

GEO 实测手册

AI 到底怎么推荐一个品牌

ChatGPT、豆包、DeepSeek 三家 AI 的一手实测数据
同样的问题问三家，每条回答都存了原文

硅基光哥 · aimarketinglift.com

硅基光哥

v1.0 · 2026-10-29
数据时间 2026 年 8 月至 10 月

版权与授权

书名	GEO 实测手册，AI 到底怎么推荐一个品牌
版本	v1.0
发布日期	2026-10-29
作者	硅基光哥
数据时间	2026 年 8 月至 10 月
首发地址	aimarketinglift.com

每张图和截图里都写了数据日期和出处。复测后会升版本号，最新版本放在首发地址。

按 CC BY-NC-ND 4.0 授权

可以 免费转发、分享整本手册

可以 引用里面的数字和截图，署名「硅基光哥」

不能 去掉署名、水印或图里的出处

不能 用于商业用途

不能 改编、改动内容后再发布

完整条款见 creativecommons.org/licenses/by-nc-nd/4.0/deed.zh-hans

目录

前言 这本手册的数据怎么来的	4
第 1 章 同一个品牌，三家 AI 待遇完全不同	5
正文引用飞书官网，ChatGPT 7/9，豆包 3/9，DeepSeek 0/9	
第 2 章 测一次不算数，同一天问三次	11
「国产项目管理工具推荐」问 ChatGPT 三次，飞书排第 1、第 5、第 1	
第 3 章 怎么测才算数	16
AI 从用户帖子里抽出 139 个候选问题，人工筛到 24 个正式题	
第 4 章 诊断分数不等于被引用	21
飞书网站诊断 26/75，ChatGPT 却有 7/9 道题正文引用了它的官网	
第 5 章 网站这一层，哪些值得做	26
AI 爬虫进不进得来，页面写法能不能被直接摘走	
第 6 章 站外这一层，AI 读的大多不在官网	31
DeepSeek 9 道题读了 142 条搜索结果，飞书和 Lark 官网只有 2 条	
第 7 章 被引用之后，怎么量到	35
ChatGPT 带来的 26 次访问，21 次挤在 8 月 11 日至 14 日	
第 8 章 红线	39
哪些做法已经被罚，哪些事我不做	
附录 测试问题原文、判定口径、测试条件模板	45
如果你的品牌也想测清楚	50

这本手册的数据怎么来的

手册里的数字，都是我自己在 ChatGPT、豆包和 DeepSeek 上问出来的。每条回答都存了分享链接和原文，手工逐条判一遍，再用开源的 GEO 可见度判定助手对照一遍。

案例品牌是飞书

手册只分析三家 AI 怎样回答关于飞书的问题，不评价飞书的网站和内容做得好不好。

两把尺子

严格口径只看回答正文里挂出来的引用标记。判定助手口径更宽，官网链接出现在来源列表里也算被引用。两种口径的数不能混着比，每张图都写了用的是哪一种。

测试条件

每次开新对话，能关的记忆都关掉，联网搜索打开，地区按 IP 自动判断。每章开头写了这一章的具体开关。

手册用到的几组测试

测试	日期	规模	用在
飞书 9 道题，三家各问一次	2026-10-19	27 条回答	第 1 章
同样 9 道题的 8 月测试	2026-08-27、2026-08-28	三家各 9 条	第 1 章
同一天每题问三次	2026-10-22	27 条回答	第 2 章
从用户帖子里抽测试题	2026-10-24	139 个候选，留 24 个	第 3 章
geo-audit 网站诊断	2026-10-23	飞书和自家网站	第 4 章
GA4 里 ChatGPT 带来的访问	2026-10-24	26 次会话	第 7 章

硅基光哥实测 · 2026-08-27、2026-10-24 · 手册数据台账

注 每个数字都记在数据台账里，写着测试日期、问题原文和原始记录在哪。8 月 DeepSeek 那次的具体日期没记下来，只写到月份。

怎么读 每章讲一个发现，先给数字，再给截图。图和截图下方的小字写着谁测的、哪天测的、出处是什么。



第1章

同一个品牌 三家 AI 待遇完全不同

硅基光哥 · aimarketinglift.com

同一个品牌， 三家 AI 待遇完全不同

2026-10-19，我把同样 9 道题分别问了 ChatGPT、豆包和 DeepSeek。问的都是项目管理和协作工具，看飞书在回答里会被怎样提到。每家每题问一次，一共 27 条回答，每条都存了原文。

ChatGPT	关掉记忆，开着联网搜索
---------	-------------

豆包	关掉记忆
----	------

DeepSeek	打开智能搜索，关掉深度思考
----------	---------------

地区	按 IP 自动判断
----	-----------

注 飞书是大品牌，这一章只看三家 AI 怎样回答关于它的问题，不评价飞书的网站和内容。

只测一家 AI，会得出完全相反的结论

9 道题里，回答正文引用了飞书官网的题数



硅基光哥实测 · 2026-10-19 · 三家回答原文，严格口径

只测 ChatGPT，飞书官网几乎每题都被当成依据。只测 DeepSeek，官网像是不存在。两个结论都来自真实回答。

同一批回答，换两种看法	ChatGPT	豆包	DeepSeek
来源列表里有飞书官网	9/9	4/9	2/9
正文提到了飞书	9/9	7/9	6/9

硅基光哥实测 · 2026-10-19 · 三家回答原文和来源列表

数据来源 大数字按严格口径算，只看回答正文里挂出来的引用标记。来源列表里有官网、正文却没挂，算在表格第一行。DeepSeek 的 2/9 里有 1 次是 Lark 官网。

1.1 逐题看，三家结论一致的只有一道

三家结论完全一致的只有「飞书和 Worktile 哪个更适合互联网公司」，都推荐飞书。其余 8 道，同一个问题在三家那里的位置、说法和引用都不一样。

DeepSeek 最特别。「企业级项目管理系统怎么选」里飞书被写成能力有限，「免费好用的项目管理软件有哪些」里飞书只被当成付费墙的例子。

排名怎么读

「第 1 / 8」是飞书在 8 款推荐里排第 1。豆包按场景分组，每组从 1 编号，记的是飞书所在那一组。DeepSeek 9 道题都没给带编号的推荐名单。所以排名这件事，三家没法放在一把尺子上比。

怎么用 看一个品牌在 AI 里的表现，至少三家一起测，逐题并排看，别只看一个总分。

问题原文	ChatGPT	豆包	DeepSeek
国产项目管理工具推荐	第 1 / 8 引用官网	分组第 2 / 3 引用官网	表格第 4 行
项目管理工具哪家好用	第 1 / 3 引用官网	第 1 / 10 引用官网	没出现
适合中小团队的协作工具	第 1 / 7 引用官网	第 1 / 3 引用官网	正文提到
远程办公团队用什么协作软件	第 1 / 7	第 1 / 3	正文提到
企业级项目管理系统怎么选	正文提到 引用官网	没出现	写成能力有限
OKR 管理工具哪个好	第 1 / 4	第 1 / 5	只在来源里
免费好用的项目管理软件有哪些	第 1 / 6 引用官网	没出现	当成付费墙反例
飞书和 Worktile 哪个更适合互联网公司	推荐飞书 引用官网	推荐飞书	推荐飞书
项目管理工具和即时通讯一体化的有哪些	第 2 / 4 引用官网	第 1 / 6	没出现
	正文引用了官网	正文提到	没出现
			被写成短板

硅基光哥实测 · 2026-10-19 · 三家回答原文逐条判定

1.2 三家各读了什么

ChatGPT 按问题找资料。问得宽泛就读测评和竞品写的选型指南，问题带了条件就直接去各家官网的定价页和帮助文档找答案。飞书官网 9/9 道题都被它读到了。

8/9 DeepSeek 先去知乎。9 道题里有 8 道，搜索词前面带着 site:zhihu.com，只在知乎里搜。

问题「国产项目管理工具推荐」，DeepSeek 分享页里记录的搜索词

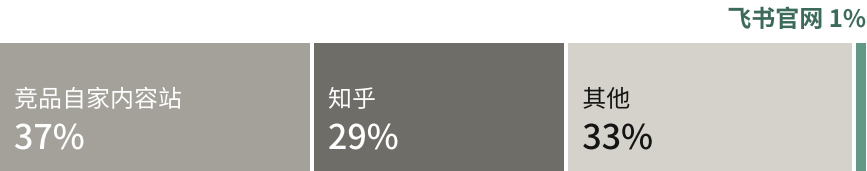
```
{
  "id": 2,
  "type": "SEARCH",
  "status": "FINISHED",
  "content": "Found 12 web pages",
  "queries": [
    {
      "query": "site:zhihu.com 2026年国产项目管理工具推荐",
    },
    {
      "query": "2026年国产项目管理软件排行榜"
    }
  ],
  "results": [
```

硅基光哥实测 · 2026-10-19 · DeepSeek 分享页数据

唯一没去知乎的是「项目管理工具和即时通讯一体化的有哪些」，那道题的回答里一款国产工具都没有。

回答前读的资料，飞书官网只占一小部分

DeepSeek 9 道题拿回 142 条搜索结果



竞品内容站 52 条，Worktile 24、ONES 22、致远互联 6。知乎 41 条。飞书和 Lark 官网 2 条。

豆包 9 道题列出 135 篇参考资料



抖音视频 36 篇。飞书官网 10 篇，其中 6 篇在「飞书和 Worktile 哪个更适合互联网公司」。其他里有厂商内容、第三方文章、研报、论文，还有法语、意大利语、西班牙语和俄语的网页。

硅基光哥实测 · 2026-10-19 · DeepSeek 搜索结果、豆包参考资料，9 道题合计

注 豆包实际列出 136 篇，「远程办公团队用什么协作软件」里有 1 篇重复，去掉后按 135 篇算。

1.3 读到了，也不一定引用

「飞书和 Worktile 哪个更适合互联网公司」，豆包列出的资料里有 6 篇飞书官网，结论也推荐飞书，句末的引用标记却写着 Worktile。

一句话结论

绝大多数互联网公司优先选飞书；如果你核心诉求是重度项目 / 任务自定义、私有化部署、预算敏感，选 Worktile。

简单理解：飞书是一体化办公底座（沟通 + 文档 + 会议 + 项目 + 知识库）；Worktile 是以项目任务为核心的协作平台，IM、文档属于配套能力 Worktile。

硅基光哥实测 · 2026-10-19 · 豆包分享页

放大看句末

简单理解：飞书是一体化办公底座（沟通 + 文档 + 会议 + 项目 + 知识库）；Worktile 是以项目任务为核心的协作平台，IM、文档属于配套能力 Worktile。

硅基光哥实测 · 2026-10-19 · 豆包分享页

注 「来源里有官网」和「正文引用官网」差得不少。只数来源，会高估官网的作用。

1.4 谁写了选型指南，AI 就按谁的框架回答

「企业级项目管理系统怎么选」，DeepSeek 读了致远互联的 5 篇选型指南，5 篇都被引用。飞书项目被归进轻量协作工具一类，句末接着「能力有限」。

• 跨部门轻协同（市场/运营/专业服务）：如果项目类型杂、周期短、重在任务分配与进度可视，轻量协作工具或表格化系统（如飞书项目、Smartsheet、Asana）落地更快。但需注意，这类工具在复杂的多项目资源冲突与预算管控上能力有限。 1 5

硅基光哥实测 · 2026-10-19 · DeepSeek 分享页

放大看后两行

工具或表格化系统（如飞书项目、Smartsheet、Asana）落地目资源冲突与预算管控上能力有限。 1 5

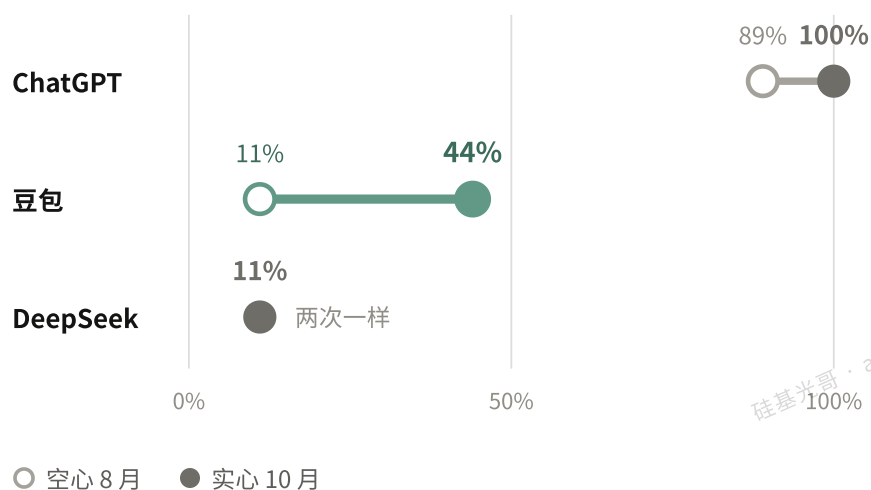
硅基光哥实测 · 2026-10-19 · DeepSeek 分享页

豆包这道题读的也是致远互联的指南，飞书没出现。「国产项目管理工具推荐」里也一样，DeepSeek 读到 ONES 的 6 篇文章，回答里 ONES 排第一。

1.5 隔两个月再问

8月底用同样9道题测过一次。豆包引用飞书官网的题从1道变成4道，DeepSeek 没变。

飞书官网被引率，判定助手口径



硅基光哥实测 · 2026-08 至 2026-10-19 · 判定助手口径

注 这里用判定助手口径，官网链接出现在正文或来源列表里都算被引用，比前面的严格口径宽，两组数不能混着比。两次都是每题只问一次。这两个月飞书、竞品和三家 AI 各自做了什么，我都看不到，所以只记变化，不猜原因。

1.6 先弄清客户用哪家 AI，再看它去哪读

ChatGPT 问得宽泛就读测评和选型指南，带了条件就去查各家官网

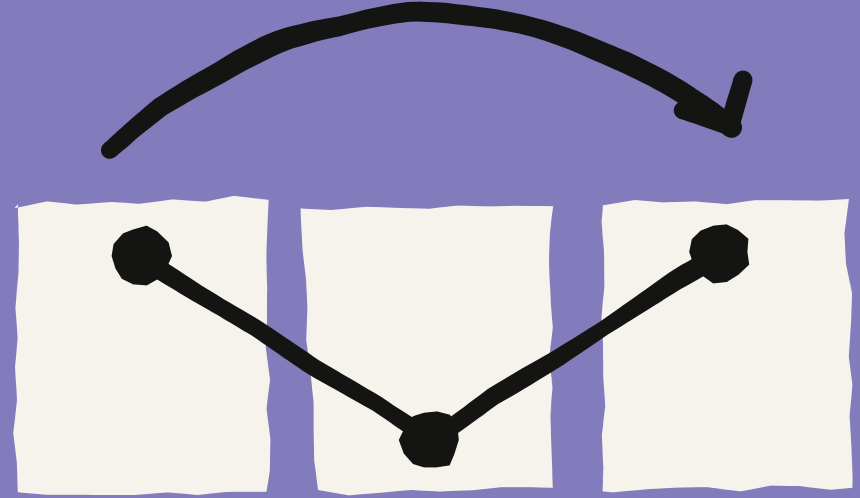
DeepSeek 先去知乎，其次是竞品自家的内容站

豆包 读得最杂，抖音视频占四分之一

在你的品类里，谁把「怎么选」写清楚了，AI 就拿谁的框架回答。这份选型指南最好由你自己来写，写在客户常用的那家 AI 会去读的地方。

怎么用 先问清客户平时用哪家 AI 查东西，再拿你的品类问题去测那一家，看它回答前读了哪些网站。

注 每题只问了一次，单题结果可能被下一次推翻，能信的是三家之间的整体差别。第 2 章讲同一天问三次会差多少。



硅基光哥 · aimarketinglift.com

第2章

测一次不算数 同一天问三次

硅基光哥 · aimarketinglift.com

测一次不算数， 同一天问三次

2026-10-22，我从第 1 章的 9 道题里挑了 3 道，同一天在三家 AI 上各问 3 次，每次都开新对话，一共 27 条回答。什么都没改，只看回答自己会差多少。

国产项目管理工具推荐	宽泛的推荐题
OKR管理工具哪个好	8 月 DeepSeek 唯一引用了飞书官网的一题
飞书和Worktile哪个更适合互联网公司	点名两家的对比题
ChatGPT	关掉记忆，默认模式
豆包	关掉记忆，联网搜索
DeepSeek	没有记忆开关，每次开新对话，打开智能搜索
地区	按 IP 自动判断

有没有被提到很稳，排第几、读了什么每次都在变

9 次回答里，正文提到飞书的次数



硅基光哥实测 · 2026-10-22 · 三家回答原文，每题问 3 次

27 次回答里，飞书只有 1 次没被提到。可同一道题问三次，排名、读的网站、搜不搜索都在变。下面一家一家看。

注 唯一没提到飞书的那次，是 DeepSeek 第 1 次回答「OKR管理工具哪个好」。同一道题第 2、3 次它又提到了。

2.1 ChatGPT，同一道题排第 1、第 5、第 1

「国产项目管理工具推荐」问三次，飞书项目两次排第一，一次排第五。只看第二次，会以为飞书掉到了中间。

飞书在推荐列表里的位置



硅基光哥实测 · 2026-10-22 · ChatGPT 分享页原文

点名对比的那道最稳。「飞书和Worktile哪个更适合互联网公司」三次都说更推荐飞书，引用区每次都有飞书官网。

注 9 次回答里，正文引用区有飞书官网的是 8/9，没有的那次正是排第五的那次。

读的网站，三次差了好几倍

「国产项目管理工具推荐」来源面板的条数，其中飞书官网



硅基光哥实测 · 2026-10-22 · ChatGPT 来源面板，按网址去重

第 2 次只列了 12 条来源，一条飞书官网都没有，引用区全是 ONES 和 Worktile 写的选型指南。飞书项目照样被推荐，只是排到了第五。

怎么用 看来源构成，要几次合起来看。单次读了多少网站，波动可以差到八九倍。

2.2 DeepSeek，提到飞书 0 次、3 次、6 次

「OKR管理工具哪个好」正文提到飞书的次数

0 3 6

第 1 次 第 2 次 第 3 次

硅基光哥实测 · 2026-10-22 · DeepSeek 分享页数据

第 1 次开头的总体建议，国内团队先看 Worktile，没有飞书

到日常项目执行，国内团队可以优先看 Worktile；

硅基光哥实测 · 2026-10-22 · DeepSeek 对话截图

第 3 次开头的总体建议，飞书 OKR 进来了

orktile；若已深度使用飞书生态，飞书 OKR 的推广阻力最小。

硅基光哥实测 · 2026-10-22 · DeepSeek 对话截图

每次都去知乎，搜的词却不一样

9/9 9 次回答，每一次的搜索词里都有 site:zhihu.com。第 1 章看到的知乎偏好，同一天重复问也很稳。

搜几次、搜什么，每次都不一样。「飞书和Worktile哪个更适合互联网公司」三次分别搜了 4、5、7 个词，后两次多了 site:feishu.cn 和 site:worktile.com，直接去两家官网找。

这道题 DeepSeek 三次都没给明确结论，都说看你更看重一体化协作还是项目管理。ChatGPT 三次都说更推荐飞书。两家各自都很稳，只是稳在不同的结论上。

注 「国产项目管理工具推荐」三次的搜索结果里，飞书官网都是 0 条 (0/3)。这一点三次一致，比单次结论可靠。

2.3 豆包，9 次里 2 次没搜就答了

2₉ 9 次回答里有 2 次，豆包一个词都没搜，直接给了答案。

飞书和Worktile哪个更适合互联网公司	第 1 次	第 2 次	第 3 次
搜索次数	0	3	6
参考资料篇数	0	18	35

硅基光哥实测 · 2026-10-22 · 豆包分享页原文

平台本身也在变。10 月 19 日，豆包回答上方是一行「搜索几个关键词，参考多少篇资料」。到 10 月 22 日变成了「共用时多少秒」，展开是一步步的调研，有时还会画对比图。

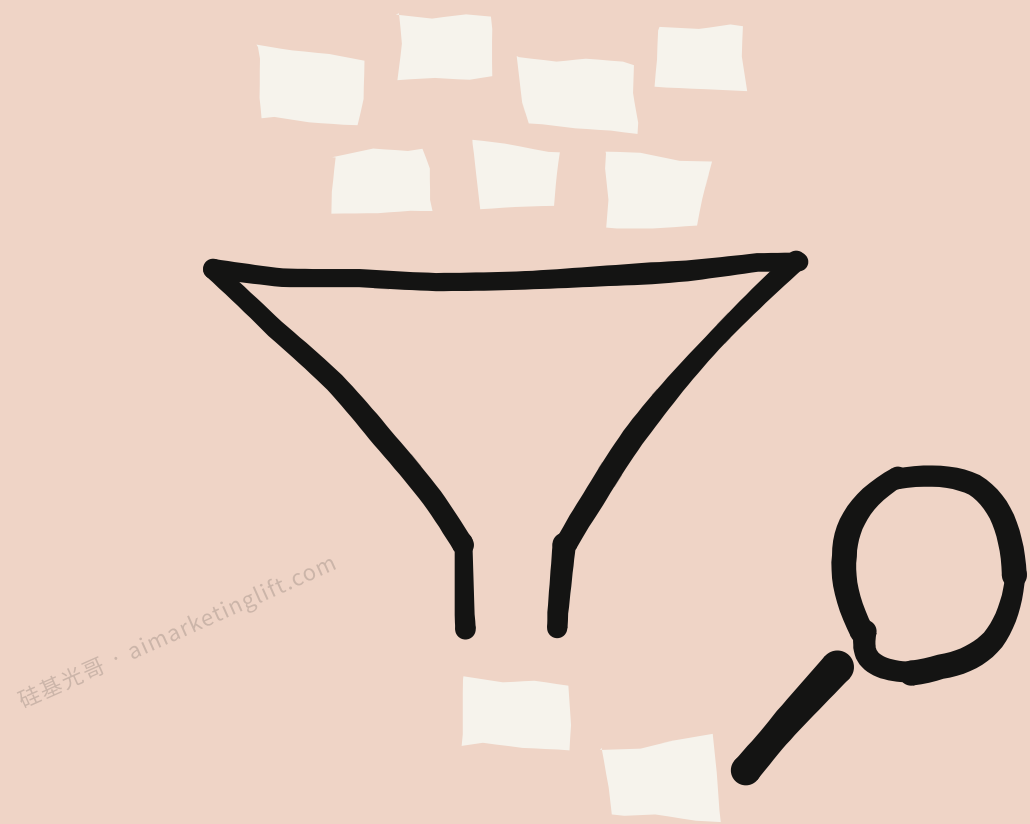
注 有一次豆包把按 IP 判断出的地区加进了搜索词。测试时地区不用手动设，但要知道它会影响搜什么。

2.4 哪些层稳，哪些层会跳

有没有提到	稳 27 次里 26 次提到飞书
对比题结论	稳 每家三次说法一致，三家之间不一样
DeepSeek 搜知乎	稳 9 次都带 site:zhihu.com
排第几	会跳 ChatGPT 同一题第 1、第 5、第 1
读了哪些网站	会跳 ChatGPT 同一题来源 107、12、27 条
搜不搜	会跳 豆包同一题搜 0、3、6 次

怎么用 每道题至少问 3 次，每次开新对话。记下日期、开关和界面的样子，平台改版也算测试条件。排名和来源只看几次合起来的结果。

注 这是一天之内、每题 3 次的结果，样本很小，只能说明波动有多大，说明不了哪家 AI 更准。



第3章

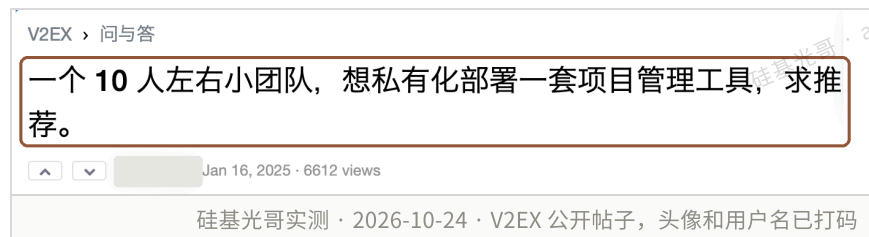
怎么测 才算数

怎么测才算数

测什么题，决定测出什么结果。判定工具本身也会判错。这一章讲三件事，题从哪来，回答怎么判，判完怎么抽查。

题从真人的问法里找

前两章的 9 道题是 AI 列出来的。2026-10-24，我重新从真人的问法里找了一遍。一路是 V2EX 的公开帖子，从 176 条搜索结果里挑出 13 个用户自己发问的帖子。另一路是百度搜索下拉，12 个种子词，每个取 10 条。

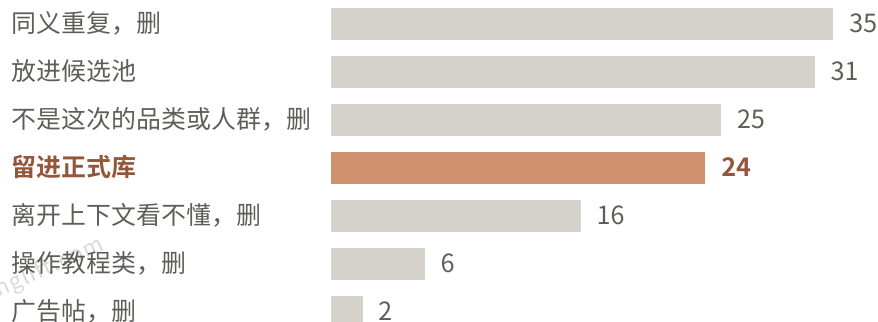


注 V2EX 上程序员多，问题偏研发和私有化部署。来源本身就有偏，选题时要知道。

3.1 AI 抽出 139 个问题，人工留下 24 个

我把帖子、回复和下拉词交给 AI，让它抽出用户的问题，并标明来历。AI 交出 139 条，原话 68、改写 65、自己补的 6。我再逐条审，每条定一个去向。

139 条的去向



硅基光哥实测 · 2026-10-24 · AI 抽取结果的人工审核

怎么用 题库宁可小，也要每条都是真人会这么问。候选池里的题不删，以后有了更多来源再看。

AI 抽原话，哪里靠得住，哪里不行

68/68	标成原话的 68 条，用脚本去掉标点后回原文里找，一字不差。这一步 AI 做得对。
35	标成改写的条目里，有 35 条把百度下拉词改成了完整问句。意思没错，可用户在搜索框里打的就是那几个词。
12	还有 12 条，把回复里别人说的话改成了问题，标成改写。原文里没人这么问过。

硅基光哥实测 · 2026-10-24 · AI 抽取结果的人工审核

回复里写的	AI 改成的问题
企业微信可以加微信，这个无敌	企业微信能加个人微信吗？
经常有人会说飞书贵很多	飞书是不是比钉钉贵很多？

注 进正式库的题一律换回原句，用 V2EX 的帖子标题或百度下拉的原词。

24 个正式题，按组分开算

组	题数	一个原句
办公协作，不带品牌	4	协作办公软件哪个好
办公协作，带品牌	6	飞书, 企业微信, 钉钉 选谁？
项目管理，不带品牌	6	一个 10 人左右小团队，想私有化部署一套项目管理工具，求推荐。
项目管理，带品牌	2	飞书项目管理收费标准
多维表格和知识库	4	多维表格与excel的区别是什么
品牌事实	2	飞书和钉钉是一个公司吗

硅基光哥实测 · 2026-10-24 · 正式题库，原句照抄

办公协作和项目管理要分开测，带品牌名的题和不带的也要分开算分。带品牌的题里飞书本来就会出现，混在一起算会把出现率抬高。

怎么用 品牌事实那组单独记，看 AI 说得准不准，比如飞书是哪家公司的。

3.2 每条回答判成五种结果之一

判定助手按下面的顺序看，前一种成立就不再往下看。正文点名了，就不再看是否也被引用。

正文提及	品牌名直接出现在回答正文里
仅被引	官网被当作来源引用，正文没写品牌名
仅检索到	只出现在搜索或来源里，正文和引用都没用到
无引用数据	回答里没有链接，也没有来源，只能确认正文没点名
未出现	有来源数据，哪一层都没有品牌

注 「无引用数据」占比高，多半是复制回答时漏了来源，不是 AI 真没引用。前言里说的两把尺子，管的是「被引用」这一层怎么算。

3.3 选什么题，决定测出什么

2026-10-24，我从正式题库里挑了 5 道，三家 AI 各问 3 次，一共 45 条回答。

3 次里正文提到飞书的次数	ChatGPT	豆包	DeepSeek
协作办公软件哪个好	3/3	3/3	3/3
项目管理软件推荐	3/3	3/3	3/3
一个 10 人左右小团队，想私有化部署一套项目管理工具，求推荐。	0/3	0/3	0/3
飞书和钉钉哪个收费高	3/3	3/3	3/3
钉钉和飞书哪个不容易封号	3/3	3/3	3/3

硅基光哥实测 · 2026-10-24 · 三家回答原文，每题问 3 次

4 道题三家 9 次全都提到飞书。换成那道带具体条件的真人问法，三家 9 次一次都没提。只拿宽泛的题去测，会以为飞书哪里都在。

怎么用 题库里要有宽泛的推荐题，也要有带人数、部署方式、预算这类条件的真人问法。

3.4 判定工具也要人工抽查

10月19日那组测试，ChatGPT 每条回答都先用一句话把推荐的几款点一遍，再展开列表。判定助手因此把9条都判成不在编号名单里（0/9），人工看有7/9进了编号名单。

判定助手对「国产项目管理工具推荐」的判定



被 ChatGPT 正文点名了

品牌名直接出现在回答正文里，这是最强的可见度。

出现位置

在列表里，但不在编号项内（共8项）

硅基光哥实测 · 2026-10-19 · GEO 可见度判定助手

同一天豆包换了复制格式，判定助手分不清正文和参考资料，把可见率算成89%，人工重算是78%。

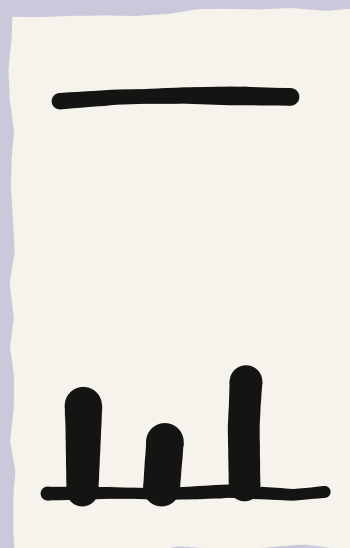
注 判定助手2026-07-10上线第一版，每接一家AI、每次对方改界面都要跟着改，到2026-10-19还在改。代码从558行长到2,065行。

45 条回答，人工又抽出这几处

排名	DeepSeek 一条回答，判定助手判飞书第3，原文加粗小标题里飞书排第2
看哪张列表	ChatGPT 开头和第一张表都把飞书项目放第一，后面的编号列表里排第二，汇报时要写明按哪张算
价格	「飞书和钉钉哪个收费高」，DeepSeek 3次里2/3把飞书价格写错，ChatGPT 和豆包6/6写对。判定助手不判事实
冒名网站	ChatGPT「协作办公软件哪个好用」第2、3次来源里有21个和25个冒名网站。判定助手按域名判，没算成官网；人只看标题会被骗

怎么用 每批结果挑一部分回原文核对，排名、官网、事实三样都看。AI改了界面，先拿旧的回答样本把判定工具复测一遍。

注 价格对的是飞书官网版本对比页，2026年10月24日取。



硅基光哥 · aimarketinglift.com

第4章

诊断分数 不等于被引用

硅基光哥 · aimarketinglift.com

诊断分数 不等于被引用

geo-audit 是我开源的网站诊断工具，从技术、内容写法、实体信号和站外存在感四项打分。2026-10-23，我用同一个版本给飞书官网和自家网站各跑了一遍。

geo-audit 四项得分	飞书官网	自家网站
技术可访问性	11/20	18/20
内容可摘取性	3/30	11/30
实体与权威信号	12/25	19/25
第三方存在感	unknown	unknown
总分	26/75	48/75

硅基光哥实测 · 2026-10-23 · geo-audit 诊断报告

注 第三方存在感工具查不到，记 unknown，不计入总分。

4.1 诊断只有 26 分的官网，ChatGPT 照样引用

26/75

飞书官网的诊断分数

7/9

ChatGPT 正文引用飞书官网的题数

硅基光哥实测 · 2026-10-19、2026-10-23 · geo-audit 报告、ChatGPT 回答原文

诊断说飞书官网「能访问，但被引用的条件偏弱」，内容可摘取性只有 3/30。可第 1 章那组测试里，ChatGPT 9 道题有 7 道在正文里引用了飞书官网。

维度	得分	置信度	判断
技术可访问性	11/20	●	●
内容可摘取性	3/30	●	●
实体与权威信号	12/25	●	●
第三方存在感	unknown/25	—	—
总分	**26/75**		**Leve

硅基光哥实测 · 2026-10-23 · geo-audit 终端输出

注 7/9 是严格口径，只算正文里的引用标记。按判定助手口径算是 9/9。

4.2 工具看到的，不等于 AI 看到的

geo-audit 抽查了飞书的版本对比页。直接抓下来的网页里只有一个标题，没有小标题，也没有正文段落。页面内容是浏览器运行脚本以后才画出来的。

Title: 版本对比 | 飞书
H1: 0 H2: 0 H3: 0

硅基光哥实测 · 2026-10-23 · geo-audit 页面抽查输出，照抄

按工具的看法，这一页对 AI 几乎是空的，内容可摘取性因此很低。

怎么用 诊断报告说某一页「读不到」，先拿一道相关问题去问 AI，看它回答里有没有这一页的内容，再决定要不要改。

ChatGPT 却写出了这一页上的内容

「免费好用的项目管理软件有哪些」，ChatGPT 把飞书排第一，写了免费版最多 100 用户，句末的引用指向飞书官网的版本对比页。

1. 飞书 (Feishu) 国内团队首选
中文体验好，集成任务管理、文档、聊天、会议和多维表格
免费版支持最多100用户，包含多维表格；部分高级能力需付费。[飞书 +1](#)
[官网](#)

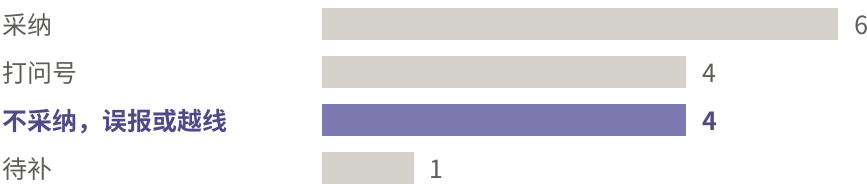
硅基光哥实测 · 2026-10-19 · ChatGPT 分享页

注 ChatGPT 怎样拿到这一页的内容，从外面看不到。能确定的只有一点，工具抓到的和 AI 读到的不是同一份东西。

4.3 诊断报告要逐条批注

自家网站那份报告，我逐条批注，每条问三件事。有没有官方或实测依据，和自己的实测对不对得上，改了以后怎么验证。

15 条建议的批注结果



硅基光哥实测 · 2026-10-21 · geo-audit 报告人工批注

误报	robots.txt 第一段已经放行所有爬虫，报告却说 10 个 AI 爬虫都没提到
误报	首页没有问答，工具把 4 个案例折叠框当成了 FAQ
越线	建议免费工具页补评分标记。没有真实评分，补上就是编评分，第 8 章会讲

修好工具，网站一个字没改，分数变了

4 条不采纳的问题都在工具里修掉了。修好以后同一个网站再跑一遍，总分从 52/75 变成 48/75，内容可摘取性从 15/30 变成 11/30。

维度	上次	本次	变化
技术可访问性	18/20	18/20	±0
内容可摘取性	15/30	11/30	-4 ↓
实体与权威信号	19/25	19/25	±0

硅基光哥实测 · 2026-10-23 · geo-audit 复测得分趋势

少的 4 分，是旧版把那 4 个折叠框算成了 FAQ。两次的 robots.txt 我逐行比过，完全一样。

注 分数掉了，不一定是网站变差了。先看工具版本和口径有没有变。

4.4 同一个网站，同一天跑出 17 分

26 17 26

第一次

晚上再跑

改了抽样以后

硅基光哥实测 · 2026-10-23 · geo-audit 三次运行输出

2026-10-23 晚上再给飞书官网跑一次，总分成了 17/75，实体与权威信号从 12/25 掉到 3/25。网站没变，工具口径也没变。

原因在抽样。这一次有一个子 sitemap 没读到，页面数从 12128 变成 9925，工具抽到的文章页换成了一篇没有作者、没有日期的新闻。当晚改了抽样方式再跑，回到 26/75。

4.5 诊断能做什么，不能做什么

能做	查技术问题，爬虫能不能进、页面结构全不全、标记有没有写错
----	------------------------------

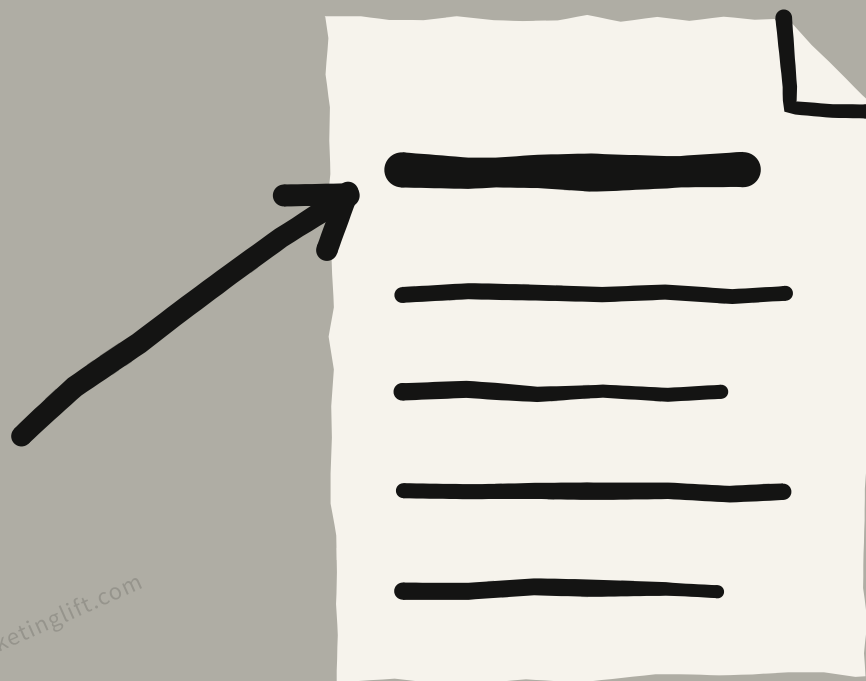
能做	给一份待办，按影响大小排好
----	---------------

不能做	判断哪家 AI 会不会引用你，这只能拿问题去测
-----	-------------------------

不能做	替你判断建议对不对，报告每条都要人批注
-----	---------------------

怎么用 有实测数据时，把诊断分数和实测结果并排放，分数低但被引用的页面先别动。每次诊断记下工具版本，分数跳了先查是不是抽样变了。

注 飞书只是一个案例。手册不评价飞书的网站做得好不好，分数只用来说明诊断和实测会对不上。



第 5 章

网站这一层 哪些值得做

网站这一层， 哪些值得做

网站上能做的事很多。这一章按我的实测和查得到原文的外部资料，排一个先后。

5.1 先确认 AI 爬虫进得来

robots.txt 第一段写 **User-agent: *** 加 **Allow: /**，所有爬虫就都放行了，不用给每个 AI 爬虫单独写一行。网站托管在 Cloudflare 这类平台上的，还要去后台看一眼 AI 爬虫的设置，平台的默认值会改，以后台显示为准。

注 别拿 AI 的回答判断它能不能读你的网站。我让 ChatGPT 总结一个不存在的网址，它先说打开是 404，接着照样写了一长篇要点。

我尝试打开了你发的文章链接，但该页面目前返回 **404 (页面未找到)**，暂时无法直接读取原文。 Internal Error

不过，从文章标题 GEO Pricing Guide 2026 来看，主题应该是 **2026 年 GEO (生成式引擎优化) 服务定价、成本构成和投入回报**。我先帮你梳理这个主题最值得关注的核心要点，并把行业参考信息与原文内容区分开来，避免把其他文章的观点误当成这篇文章的结论。

一、核心要点：2026 年 GEO 服务到底应该怎么定价？

硅基光哥实测 · 2026-10-25 · ChatGPT 对话截图

被拦下的请求，多半是扫描器在找密钥

2026-10-16 在 Cloudflare 后台看 AI 爬虫这一栏，返回 4xx 的路径有 144 个，翻下来全是 .env、密钥、配置文件这类地址。这是扫描器顶着爬虫的名字来探测，和 AI 读不读你的文章没有关系。

路径	主机
/config/env/aws_credentials.env	aimark
/.docker/config.json	aimark
/.env.production	aimark
/@fs/src/.env	aimark
/admin%2F.env	aimark
硅基光哥实测 · 2026-10-16 · Cloudflare AI 爬虫控制台	

怎么用 看 AI 爬虫记录时，先把被拦的和放行的分开。被拦的大多不用管，放行的那部分才说明 AI 来读了哪些页。

5.2 llms.txt 和专门的标记，作用有限

LLMS.txt files and other "special" markup: You don't need to create new machine readable files, AI text files, markup, or Markdown to appear in Google Search (including its generative AI capabilities), as Google Search itself doesn't use them. Note that Google may discover, crawl, and index [many kinds of files](#) in addition to HTML on a website: this doesn't mean that the file is treated in a special way.

硅基光哥实测 · 2026-07-10 · Google 搜索中心，页面更新于 2026-07-10

Google 的指南写得很直接，想出现在 Google 搜索和它的生成式 AI 功能里，不需要新建 AI 文本文件、专门的标记或 Markdown，Google 搜索本身不用它们。Ahrefs 统计，97% 的 llms.txt 在 2026 年 5 月一次访问都没有。

怎么用 已经有 llms.txt 的不用删，也不用再花时间。这个结论只说到 Google 和 Ahrefs 统计的范围，别的 AI 用不用，要自己看访问记录。

97% of those files received zero traffic in May 2026.
Nothing fetched them at all.

硅基光哥实测 · 2026-05 · Ahrefs 统计

5.3 结论往前放

- 44.2% of citations come from the first 30% of content.
- 31.1% come from the middle (30–70%).
- 24.7% come from the final third, with a sharp drop near the footer.

硅基光哥实测 · 2026-10-22 · Search Engine Land 报道

Search Engine Land 报道了 Kevin Indig 的一项分析，18,012 条 ChatGPT 引用里，44.2% 来自页面前 30%，越往后越少。

自家一篇文章，进了 Google AI 概览

2026-10-17 在 Google 搜「AI营销到底是什么」，我那篇《AI营销到底是什么》是 AI 概览来源面板的第 1 张卡片，同一天截了两次，结果一样。这篇文章第一屏就写了 AI 营销主要干两件事。



注 这只是一个案例，一天里看了两次。能说明写法有用的，主要还是上一页那份统计。

5.4 用 AI 写的内容，数字要人核

我让豆包写一份产品对比页。它写的 39/99/299/599 元/月，其实是腾讯云模型调用套餐的价格，产品官网的订阅价是 99/199/999 元。

后来我做了一个核数字的小工具。第一版跑这份稿子，查了 56 条，找到 35 条，要人打开看 18 条。人工核出的 4 处错，工具一处也没标出来。



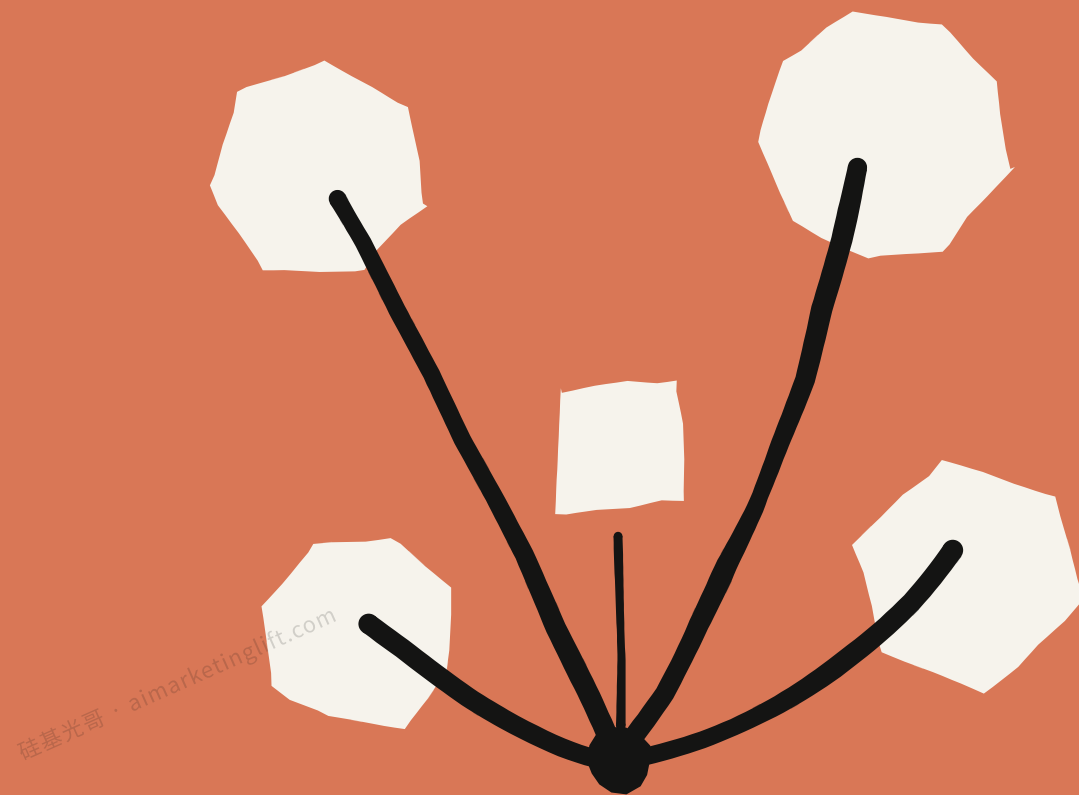
怎么用 工具说「找到了」，只说明页面上有这个数，不说明数用对了地方。价格、人数、日期这几类，发布前逐条对原页面。

5.5 网站这一层的先后

先做	robots.txt 放行，托管平台后台看一眼 AI 爬虫设置
先做	结论写在开头，一段只讲一件事，能单独拿出来引用
先做	价格、版本、人数这些事实写清楚，放在 AI 能直接读到的正文里
别多花时间	llms.txt 和专门给 AI 的标记
别全信	诊断报告和核数字工具的结论，都要人再看一遍

注 第 1 章看到，AI 读的大多不在你的官网。网站这一层做好，是让 AI 读到你的时候能读对，下一章讲官网以外的地方。

数据来源 Google 和 Ahrefs、Search Engine Land 是外部资料，原文截图都附在本章。其余是我自己网站的记录。



第6章

站外这一层 AI 读的大多不在官网

站外这一层， AI 读的大多不在官网

第 1 章那组测试，三家 AI 回答前读的资料，飞书官网只占一小部分。

AI	读得最多的	飞书官网
DeepSeek	竞品内容站 37%，知乎 29%	1%
豆包	抖音视频 27%	7%
ChatGPT	测评和竞品写的选型指南	9/9 题读到

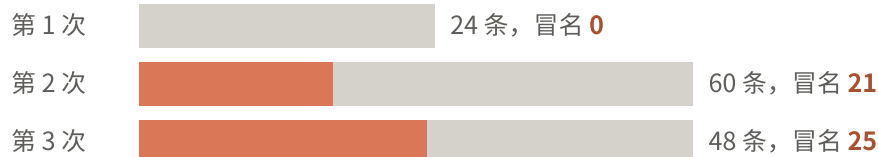
硅基光哥实测 · 2026-10-19 · 三家来源列表，9 道题合计

注 ChatGPT 读官网读得最多，可它在宽泛的推荐题里，大量引用的是竞品写的选型指南。

6.1 换一批题再看，规律还在

第 3 章那 45 条回答里，DeepSeek 10/15 次自己给搜索词加了 site:zhihu.com。

「协作办公软件哪个好用」ChatGPT 来源面板，其中冒名网站



硅基光哥实测 · 2026-10-24 · ChatGPT 来源面板，按网址去重

冒名网站是域名仿飞书、钉钉、企业微信，又自称官网的网站。第 2、3 次的来源里各有 21 个和 25 个，豆包同一批 15 次一个都没有。

怎么用 看来源时按域名看，别只看标题。标题写着「官网」的，不一定是官网。

6.2 发在 AI 偏爱的平台上，很快就会被引用

10 AI-search platforms, we analyze 17,211 citation instances over 6,356 unique source domains and find: (1) citations concentrate in platform-specific sources, with top-20 domains capturing 20.5–70.8% of per-platform citations, and 15 of 22 tested publication platforms tied to cited source domains had low or medium barriers for both account set-up and posting; (2) in our experiments, ordinary publication on preferred platforms changed what entered AI-search outputs: 8 of 10 platforms cited a fabricated concept within seven days, and one high-preference-platform article had greater citation impact than over 20 matched low-preference posts; and (3) this path is commercially available: a \$14 GEO purchase produced 13 public posts, and one AI-search platform cited GEO-posted content with our designed markers within one hour. These results reveal and measure a low-barrier publication path into AI-search ci-

硅基光哥实测 · 2026-10 · 复旦大学论文 arXiv 2610.11932 摘要

Results. Across the five matched pairs, the high-preference side produced the first citation or marker signal in every paired first round. The best low-preference case became visible only after 20 posts; it was rarely cited in RQ1, appearing only three times (twice by Kimi and once by Wenxin). The weakest case was self-hosted WordPress: even 40 posts received no citation within one week. We report these as per-topic thresholds, not a general conversion rule, because topic identity, indexing latency, stance assignment, and platform pair were not fully crossed.

硅基光哥实测 · 2026-10 · 复旦大学论文 arXiv 2610.11932 实验结果

复旦大学一篇论文看了 10 个 AI 搜索平台的 17,211 条引用。每家平台前 20 个域名占了它 20.5% 至 70.8% 的引用，豆包引字节系网站占 25.9%，文心引百度系占 24.4%。他们花 14 美元 买了一次 GEO 服务，发出 13 篇稿子，1 小时内就有一家 AI 搜索引用了。

配对实验里，发在 AI 偏爱的平台上，第一轮就有引用信号。不偏爱的平台，最好的一组要发到 20 篇才看得到。自建的 WordPress 网站发了 40 篇，一周内都没被引用。

注 论文说明这些是按话题测出来的门槛，不能当通用的换算。它还说明这条路门槛很低，第 8 章讲平台和监管怎么管。

6.3 数一数你的客户用的那家 AI 在读哪里

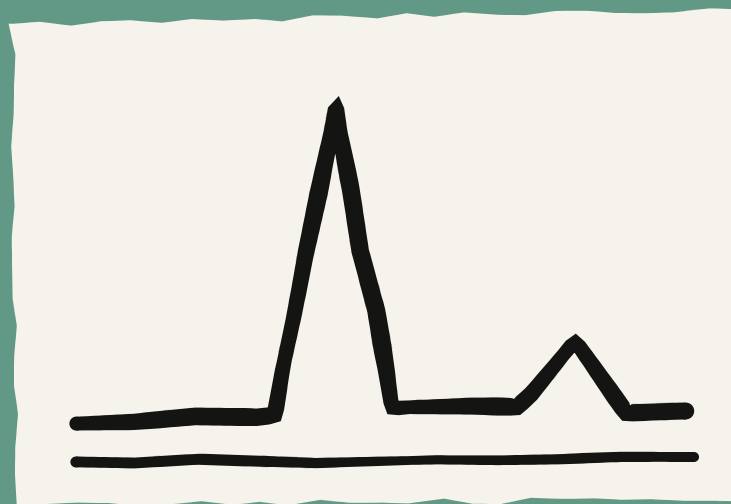
第 1 步	从第 3 章的题库里挑 5 道真人问法，宽泛的和带条件的都要有
第 2 步	在客户常用的那家 AI 上，每题问 3 次，每次开新对话
第 3 步	把每次的来源列表存下来，按域名归类，去掉重复
第 4 步	数出现最多的前 10 个网站，标出竞品自家的、平台自家的和冒名的
第 5 步	看你在这些网站上有没有内容，没有的就是要补的地方

发在哪里，比发多少更要紧

DeepSeek 偏知乎，豆包偏抖音和字节系网站，ChatGPT 偏测评和选型指南。论文里偏爱的平台一篇就有信号，不偏爱的平台发几十篇也可能没用。

注 这是一个方向，依据是上面这篇论文和我自己的几组测试，还没有更多数据验证。

怎么用 发内容时用真实身份，写真实的经验和数据。伪造「第三方」测评、批量造号发稿，都在第 8 章的红线里。



硅基光哥 · aimarketinglift.com

第7章

被引用之后 怎么量到

硅基光哥 · aimarketinglift.com

被引用之后， 怎么量到

这一章只讲去哪看、怎么算。下面六个信号，我都在自己的网站上试过。

Bing	AI Performance 报告，看哪些页面被 Copilot 引用、被引用几次
GA4	按会话来源看 chatgpt.com 等 AI 带来的访问
爬虫记录	Cloudflare 这类后台里，AI 爬虫来读了哪些页
Google	Search Console 里 AI 功能带来的展示和点击
品牌搜索	品牌名被主动搜索的次数有没有变化
问客户	咨询时直接问一句，从哪知道你的

7.1 被引用了，不等于有人点进来

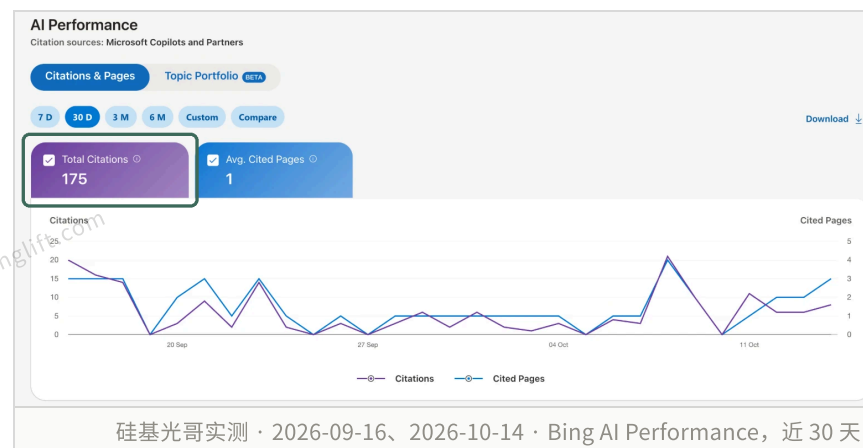
175

Bing 报告里被 Copilot 引用的次数

0

同期 GA4 里来自 Copilot 的访问

硅基光哥实测 · 2026-09-16、2026-10-14 · Bing Webmaster Tools、GA4



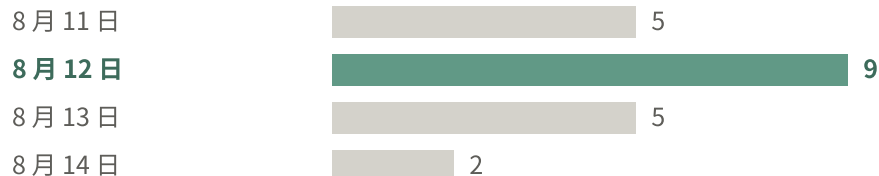
硅基光哥实测 · 2026-09-16、2026-10-14 · Bing AI Performance, 近 30 天

注 175 是 2026-09-16、2026-10-14 按天加总的数。GA4 同一段时间里，会话来源含 copilot 的一次都没有。

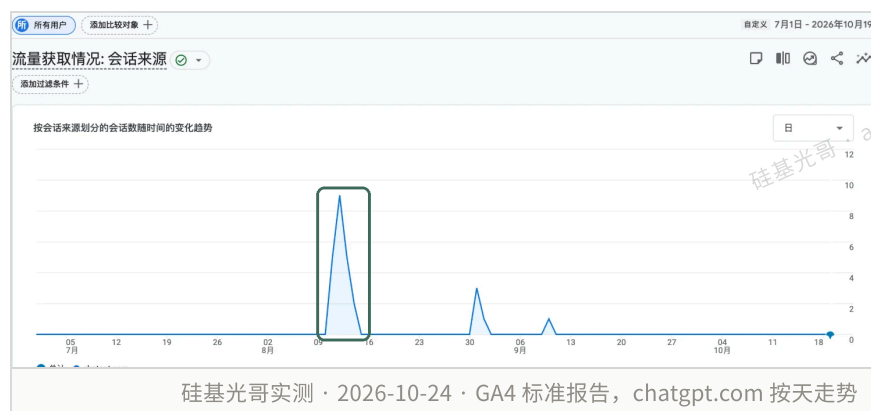
7.2 ChatGPT 带来的访问，会扎堆

21

GA4 按日期导出的表里，chatgpt.com 一共带来 26 次会话，21 次挤在 8 月 11 日至 14 日，9 月 10 日以后一次都没有。



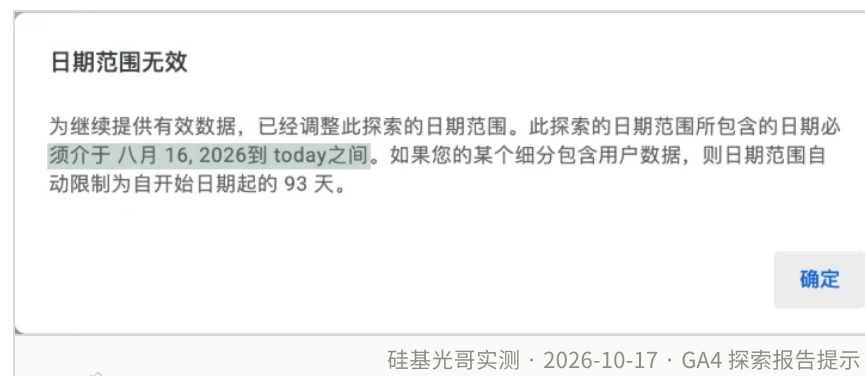
硅基光哥实测 · 2026-10-24 · GA4 会话来源 × 日期导出表



注 同一时期 GA4 的来源汇总表是 25 次，和按日期的表差 1 次。两张表的数不混在一句里用。

7.3 探索明细能查多久，要看保留设置

2026-10-17 在 GA4 探索报告里选 7 月的日期，弹出提示说日期必须在 2026-08-16 之后。



口径说明 截图反映当时这个资源可查的日期范围，不是 GA4 的统一上限。标准 GA4 的用户与事件数据保留期可选 2 或 14 个月，影响探索明细，不影响标准汇总报告。[Google 官方说明](#)

怎么用 先检查资源的数据保留设置，再按月导出来源与日期明细。延长保留期不能恢复已删除的数据。

7.4 让 AI 读后台数据，要回原表核

我把 7 张 GA4 和 Bing 的导出表交给 ChatGPT。它算的几十个数里，错了 1 处，把渠道组表的 2,398 当成了来源汇总表的数，来源汇总表其实是 2,397。



来源汇总表	2,397
渠道组汇总表	2,398
来源×日期明细求和	2,419

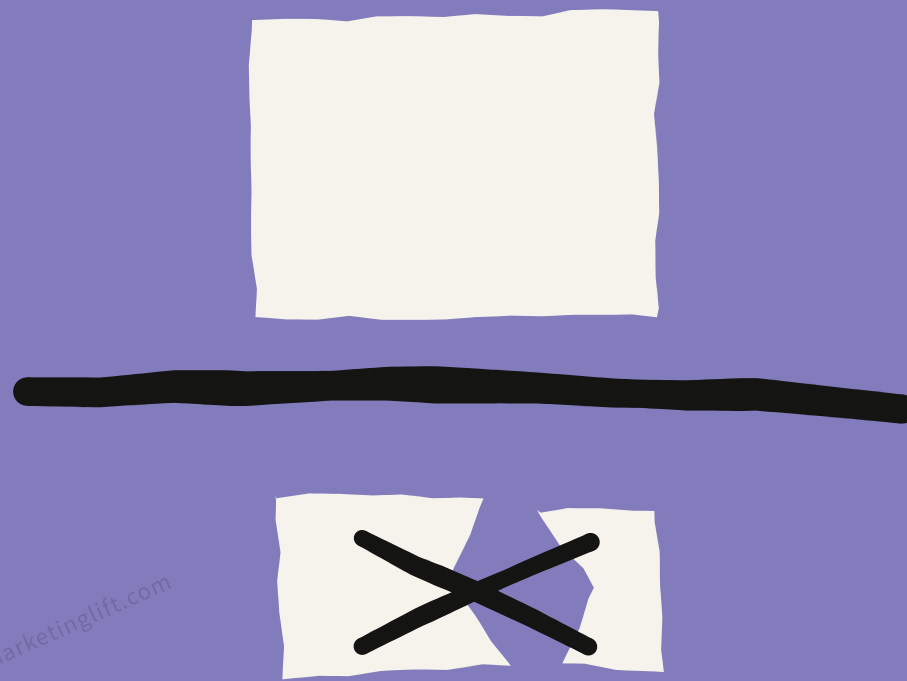
硅基光哥实测 · 2026-10-24 · ChatGPT 回答截图

豆包这次一个数都没算错，但它看到按日期的表少了 2 天，就说导出不完整。我回 GA4 查了，7 月 8 日那天本来就是 0 次会话。

7.5 量的时候记住这几条

分开算	被引用和带来访问是两件事，各看各的报告
写清口径	同一个数在不同报告里会差几次，写明出自哪张表、哪段日期
按月存档	检查探索明细的保留期限，按月导出备份
回原表核	AI 读出来的数，抽几个回导出表核一遍

注 这里的数都来自我自己的网站，量很小。它们说明的是去哪看、怎么算，说明不了别的网站会怎样。



第 8 章

红线

红线

第 6 章看到，往 AI 里塞内容的门槛很低。这一章讲哪些做法已经被罚，行业和 AI 公司怎么写进了规则，以及我自己不做的事。

8.1 已经被罚的做法

据一篇报道，到 9 月已经有三起 GEO 相关的行政处罚。三起的做法不一样，共同点是编造数据，让 AI 去读。

注 三起都出自同一篇报道，处罚决定书原文我没找到，手册不写企业名。

哪里	做了什么	罚款
北京朝阳区 6 月	服务商在自家官网虚构评分 99.6 分、市场占有率 49%	5 万元
广州越秀区 9 月 4 日公布	给两家律所批量写排行榜文章，虚构 187 起案件和满意度评分	12 万元
深圳市	用「GEO 分析系统」探测大模型偏好，编造行业评分发到多个社交平台	报道没写

硅基光哥实测 · 2026-06、2026-09-04 · 新闻报道截图，三家市场监管部门

北京那起，监管部门认为这些行为干扰了 AI 搜索产品的正常运行。虚构的评分写在服务商自家官网上，等 AI 来读。

报道原文

今年6月，北京市朝阳区市场监管局也通报了一起类似案件。一家GEO服务商在官网虚构自身评分99.6分、市场占有率49%、续费率99%等数据。监管部门认为，这些行为干扰了AI搜索产品的正常运行，对其处以5万元罚款。

硅基光哥实测 · 2026-06 · 新闻报道截图，北京朝阳区

GEO被监管处罚已不是首次。9月4日，广州市越秀区市场监管局公布的一起案件显示，一家GEO公司为两家律师事务所批量生成《2026年广州刑事案件代理专业企业TOP5》等文章，虚构187起案件、9.6分满意度、92.5分综合得分等数据，最终被罚款12万元。

硅基光哥实测 · 2026-09-04 · 新闻报道截图，广州越秀区

注 报道里写的是「今年6月」和「9月4日」，报道本身的发布日期没有存下来。

怎么用 自家官网上的评分、市场占有率、案例数，每一个都要能拿出来来源。AI会把它当成事实转述出去。

8.2 行业标准写进去了

7.1.4 GEO 服务主体和委托主体不应在宣传推广、商务沟通、服务合同、服务方案、交付报告或者验收安排中作出或者要求作出不当效果承诺。不当效果承诺包括但不限于：

a) 结果兜底类：“保证首位推荐”“保证第一提及”“100%收录”“保证固定推荐位置”“指定回答内容”“保证品牌近期真实负面不被提及”等保证特定结果的承诺；

硅基光哥实测 · 2026-10-13 · T/CAAAD 046-2026 《生成式引擎优化服务要求》第 7.1.4 条

10.2.2 不应伪造主体身份、证明材料、内容来源、监测结果或者服务效果，不应利用未经授权的第三方资源实施 GEO 服务，不应通过虚假评价、误导性比较、信息投毒、提示词注入或者其他不正当方式影响请求方判断、损害竞争主体合法权益或者干扰生成式引擎正常运行。

硅基光哥实测 · 2026-10-13 · T/CAAAD 046-2026 第 10.2.2 条

中国广告协会 2026-10-13 发布了两项 GEO 团体标准，2026-10-23 实施。服务标准不许承诺「保证首位推荐」这类结果，不许伪造内容来源和监测结果，也不许用虚假评价和信息投毒。

监测标准第 12.4 条写明，拿单次查询、单张截图当主要证据的，不能判定为通过。

注 这一条和第 2 章是一回事。测一次不算数，标准里也这么写了。

AI 公司的政策也写进去了

Manipulate the sources from which search engines or AI systems draw answers by seeding them with content that **misrepresents its origin, authorship, or independence** (e.g., networks of sites posing as unaffiliated sources corroborating the same claims)

硅基光哥实测 · 2026-11-12 · Anthropic 使用政策, 2026-11-12 生效

Anthropic 的使用政策禁止用伪装来源、作者或独立性的内容，去操纵搜索引擎和 AI 读取答案的来源，举的例子是一批假装互不相关、却在互相印证同一说法的网站。这条 2026-11-12 生效。

注 复旦论文里那种一小时就被引用的路子，正是这几条规则要管的。能做到，不等于可以做。

8.3 我不做的事

不编数据	评分、案例数、满意度，没有真实来源就不写
不造信源	不批量注册账号、不伪造「第三方」测评
不装用户	不冒充用户去问答平台发帖、回答
不承诺结果	不保证 AI 排第一、不保证被提到
汇报不造假	每个数字写明日期和口径，单次截图不当结论
不自建词条	不自己给自己建百科词条，这是我的做法，不是规定

我自己的工具也差点越线

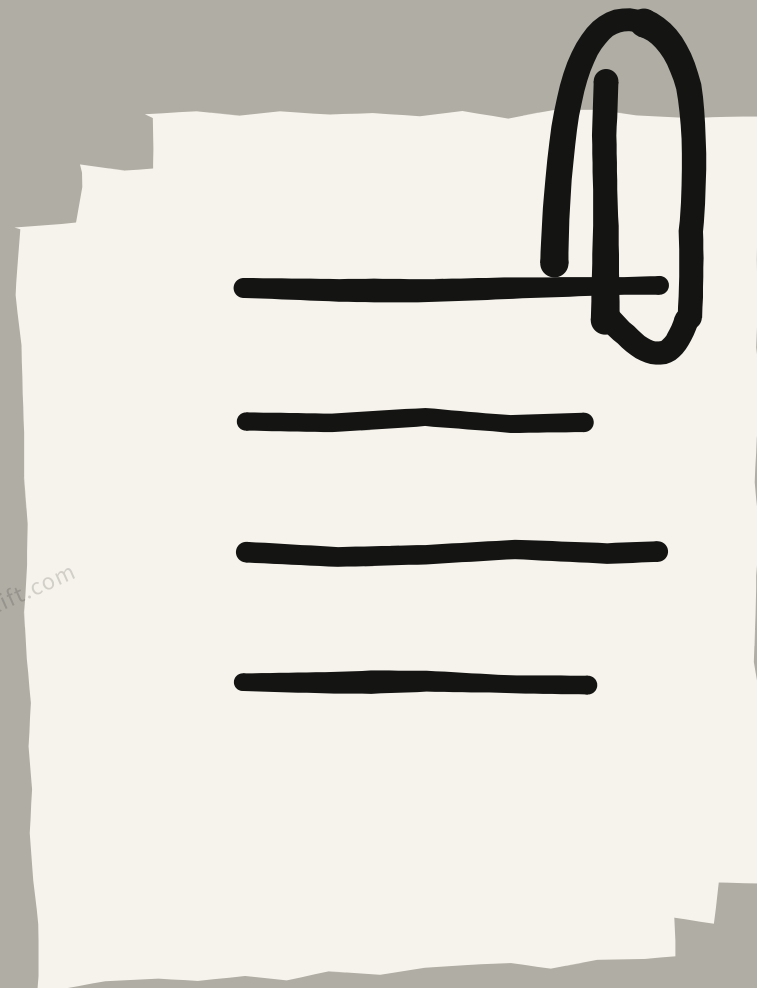
2026-10-21，geo-audit 给我网站的工具页报「缺 AggregateRating」，也就是缺评分标记。免费工具没有真实评分，照着补上就等于编评分。

第 4 章人工批注时把这条判成越线，后来改了工具，不再默认建议补评分，模板里的示例评分也删了。

怎么用 诊断报告、AI 写的方案里出现「补评分」「建词条」「多平台同步发布」，先停下来问一句，有没有真实的依据。

附录

附录



测试问题原文

逐字照抄，标点和空格都和当时输入的一样。同一道题前后两次的空格不一样，照原样记，这也是附录 C 要记「问题原文」的原因。

飞书 A 组 9 题

2026-10-19，三家各问一次，用在第 1、4 章。题目是 AI 列的。

A1 国产项目管理工具推荐

A2 项目管理工具哪家好用

A3 适合中小团队的协作工具

A4 远程办公团队用什么协作软件

A5 企业级项目管理系统怎么选

A6 OKR 管理工具哪个好

A7 免费好用的项目管理软件有哪些

A8 飞书和 Worktile 哪个更适合互联网公司

A9 项目管理工具和即时通讯一体化的有哪些

同日重复测 3 题

2026-10-22，三家各问 3 次，用在第 2 章。从 A 组里挑的。

A1 国产项目管理工具推荐

A6 OKR 管理工具哪个好

A8 飞书和 Worktile 哪个更适合互联网公司

批量测 5 题

2026-10-24，三家各问 3 次，用在第 3、6 章。从下一页的正式题库里挑的，题号沿用题库编号。

A1 协作办公软件哪个好用

B4 项目管理软件推荐

B2 一个 10 人左右小团队，想私有化部署一套项目管理工具，求推荐。

A8 飞书和钉钉哪个收费高

A9 钉钉和飞书哪个不容易封号

注 批量测的 A1、A8 和飞书 A 组的 A1、A8 不是同一道题，看题目原文区分。

正式题库 24 题

2026-10-24 人工审核后留下，用在第 3 章。括号里是出处。

A1	协作办公软件哪个好用	百度下拉
A2	团队协作软件哪个好用	百度下拉
A3	团队协作软件推荐	百度下拉
A4	协同办公软件推荐	百度下拉
A5	飞书, 企业微信, 钉钉 选谁?	V2EX 标题
A6	飞书和钉钉哪个好用	百度下拉
A7	飞书和企业微信哪个好	百度下拉
A8	飞书和钉钉哪个收费高	百度下拉
A9	钉钉和飞书哪个不容易封号	百度下拉
A10	飞书和企业微信打通	百度下拉
B1	公司有很多项目并行，有什么好的项目管理软件	V2EX 标题
B2	一个 10 人左右小团队，想私有化部署一套项目管理工具，求推荐。	V2EX 标题

B3	大家都用什么项目管理软件?	V2EX 标题
B4	项目管理软件推荐	百度下拉
B5	项目管理软件免费版	百度下拉
B6	小团队项目管理工具	百度下拉
B7	飞书项目管理收费标准	百度下拉
B8	飞书项目管理甘特图	百度下拉
C1	金山文档的多维表格可以迁移到飞书多维表格吗?	V2EX 标题
C2	有没有类似飞书多维表格的开源项目啊	V2EX 标题
C3	多维表格与excel的区别是什么	百度下拉
C4	飞书知识库中的图片如何免费添加水印	V2EX 标题
D1	飞书是哪家公司的	百度下拉
D2	飞书和钉钉是一个公司吗	百度下拉

注 A 组是办公协作，B 组是项目管理，前 10 题和前 8 题里带品牌名的和不带的分开算分。C 组是多维表格和知识库，D 组看 AI 说的品牌事实准不准。

判定口径

GEO 可见度判定助手按顺序往下查，命中第一种就停。

正文提及	品牌名直接出现在回答正文里
仅被引	正文没点名，官网链接被当作正文里的引用
仅检索到	正文和引用都没有，搜索或来源层里出现过
无引用数据	回答里没有链接，也没有填来源，只能确认正文没点名
未出现	正文、引用、来源里都没有

注 负面提及要人工勾选，工具会把它算成正文提及。

两把尺子

指标	判定助手口径	严格口径
被引率	官网链接出现在正文或来源列表里就算	正文里有指向官网的引用标记才算
检索命中率	来源标题、搜索词里出现品牌名或官网域名就算	来源列表里有官网页面才算

硅基光哥实测 · 2026-10-19 · 判定助手源码和 10 月核对记录

同一组回答，两把尺子量出来不一样。第 1 章 DeepSeek 的官网被引，严格口径是 0/9，判定助手口径是 11%。

注 「仅被引」这个状态只看正文，「被引率」这个指标连来源列表一起算，名字像，算法不一样。排名从 2026-10-19 起按品牌落在哪张列表的第几条算，和 8 月的算法不同，不能直接比。

测试条件记录模板

每次测试照着记一遍。右边一栏写的是哪次测试让我知道这一项必须记。

测试日期	有一次计划表里的日期栏一直空着，只能靠收到链接的记录补
提问时间	分享链接里只能推出分享时间，推不出提问时间
账号状态	登录与否两次都没记，现在没法回头确认
记忆开关	ChatGPT、豆包能关就关，截图留证；DeepSeek 没有这个开关
联网和深度思考	DeepSeek 记下智能搜索开、深度思考关
每次新对话	不在上一次的对话里接着问，不点重新生成
问题原文	同一道题两次的空格不一样，逐字照抄
模型和界面	记下回答上方显示的思考时间、搜索次数、参考篇数

界面改版	豆包 10-19 到 10-22 换了界面，复制格式也变了
来源展开	ChatGPT 来源会折叠，要点开到底；有一次没有来源面板
链接和题号	ChatGPT 分享页不显示提问，按顺序发链接才对得上
存档方式	DeepSeek 分享页打不开，改用数据接口存原文
次数和轮次	每题问几次，是按题问完还是按轮跑
判定设置	品牌名、官网域名、竞品名单、口径和工具版本
负面提及	DeepSeek 把飞书当付费墙的例子，工具会算成正文提及
地区	按 IP 自动判断，不手动设置

注 前面几章的测试，账号状态、提问时间这两项没有记全。手册里凡是用到的数字，都没有依赖这两项。

如果你的品牌 也想在 AI 里测清楚

这本手册里的测试，用的是两个开源工具加一套固定的记录办法。想先自己试，可以从这两个工具开始，都在 GitHub 的 jiguang9 下面。

geo-judge GEO 可见度判定助手。把 AI 的回答贴进去，判断品牌有没有被提到、官网有没有被引用

geo-audit 检查网站的技术和结构，给出一份待办清单

想一起做，我提供这四项

AI 营销培训

实战陪跑

营销 workflow 搭建

GEO 服务



硅基光哥

专注 AI Marketing、出海与一人公司实践。
公众号、B 站、小红书都叫这个名字，扫码关注
公众号。

本手册按 CC BY-NC-ND 4.0 授权。可以转发，转发要署名「硅基光哥」，不能商用，不能改编。首发 aimarketinglift.com