

AI 搜索 RAG 与微调引用来源对比： 千问、Kimi、Gemini 63 次公开观察

DragonAI Editorial Research Team

杭州烛龙智元科技有限公司

AI Search Citation Sources for RAG vs. Fine-Tuning

63 provider observations · five fixed-query batches · six Baidu observations · 2026-07-21 to 2026-07-30

63 公开模型观察	45 五批固定查询运行	0 五批目标引用运行	0 五批品牌提及运行
---------------------	-----------------------	----------------------	----------------------

利益关系披露：DragonAI 运营 AI 产品经理课程并监测自有域名。本报告保留目标引用与品牌提及均为零的结果；发布、抓取、搜索提交、等待时长和本 PDF 本身都不构成 GEO 成功。

版本 2.0.0 · 初始发布日期 2026-07-22 · 当前证据截至 2026-07-30 · CC BY 4.0（范围见权利边界）

Technical Report No. DragonAI-AISCB-2026-07-v2.0.0

Abstract

This technical report records 63 public Qwen, Kimi and Gemini observations: nine exploratory runs, nine formal prepublication baseline runs, and 45 runs across five fixed-query batches (including historical window C). The five batches contained 552 concrete citation-URL occurrences across 194 distinct URLs and 114 hostnames. Concrete-URL coverage was 15/15 runs for Gemini, 14/15 for Qwen, and 2/15 for Kimi. DragonAI target citations and brand mentions were both zero, and the target was absent from all 18 strict-query snapshots in six read-only Baidu observations. These frozen observations do not establish answer or source quality, index coverage, long-term preference, ranking causality, or GEO success.

1. 研究问题与设计

固定无品牌元问题

“千问、Kimi、Gemini回答RAG和微调选型问题时通常引用哪些来源？请给出可复核样本、平台差异和研究限制。”

目标是观察同一问题在三种 AI 搜索表面中展示的具体引用 URL，并判断 DragonAI 研究页是否进入引用来源。研究对象是引用展示层，不是平台内部索引、答案吸收、来源质量或业务结果。

当前研究口径

证据层	观察数	当前直接结果	边界
探索样本 A	9 次运行	技术选型来源分布	不进入正式趋势
正式基线 B	9 次运行	发布前固定问题基线	仅 B v2 可作未来严格端点
五批固定查询（含 C）	45 次运行	552 次 URL；114 个主机；目标 0	均由时间盒触发；不可归因
百度六轮严格查询	18 个查询快照	目标出现 0	不是 AI 提供商运行

五批直接答案

45 次固定无品牌查询运行中，Gemini 15 / 15、千问 14 / 15、Kimi 2 / 15 次展示具体引用 URL；共记录 552 次具体 URL 出现、194 个不同 URL 和 114 个不同主机。DragonAI 目标引用和品牌提及均为 0。百度六轮、三条严格无品牌查询共 18 个快照中，目标页同样未出现。

历史正式窗口 B/C

窗口	阶段	完成运行	每平台重复	触发边界
B	发布前固定问题基线	9	3	预注册基线
C	发布后点时观察	9	3	24 小时时间盒；不可归因

纳入规则

仅纳入 API 成功、执行原生联网搜索且引用表面可复核的完成运行。具体引用 URL 必须包含 http/https、可解析主机名，并排除星号、省略号或截断说明。目标引用与品牌提及分别记录。

C 窗口及其后三批固定查询都由时间盒触发，不代表目标页已经被 Googlebot 或 Bingbot 合格抓取，也不能把跨批变化归因于页面发布、抓取、索引或改写。百度六轮只读观察也不构成排名因果或 GEO 成功。

2. 历史 B/C 汇总结果

窗口	完成运行	具体 URL 出现	不同 URL	不同主机名	目标引用	品牌提及
B	9	106	60	39	0 / 9	0 / 9
C	9	124	74	50	0 / 9	0 / 9

严格趋势

Qwen 与 Gemini 在 B/C 窗口使用相同模型，可进入同模型严格趋势：两窗合计各 6 次可比运行，目标引用率均为 0/6，变化为 0。Kimi 从 moonshot-v1-128k 迁移到 kimi-k2.6，只能把 C 视为新基线，不能并入严格趋势。

B 与 C 的引用总量和来源构成发生变化，但 DragonAI 目标引用与品牌提及都没有出现。该结果是严格无提升，不是 GEO 成功。

分平台模型可比性

平台	B 模型	C 模型	严格可比
gemini google search grounding	gemini-2.5-flash	gemini-2.5-flash	是
kimi builtin web search	moonshot-v1-128k	kimi-k2.6	否
qwen dashscope web search	qwen-plus	qwen-plus	是

3. 分平台结果

窗口 B · prepublication

平台 / 模型	逐运行引用数	URL 出现	不同 URL	不同主机名	目标引用
Gemini / gemini-2.5-flash	24 / 21 / 23	68	40	29	0 / 3
Kimi / moonshot-v1-128k	0 / 6 / 5	11	11	4	0 / 3
Qwen / qwen-plus	9 / 9 / 9	27	9	7	0 / 3

窗口 C · postpublication

平台 / 模型	逐运行引用数	URL 出现	不同 URL	不同主机名	目标引用
Gemini / gemini-2.5-flash	23 / 28 / 28	79	47	34	0 / 3
Kimi / kimi-k2.6	0 / 0 / 18	18	18	11	0 / 3
Qwen / qwen-plus	9 / 9 / 9	27	9	7	0 / 3

引用数量不能直接解释为答案质量、来源权威或平台优劣。不同平台的搜索触发、展示机制、账号、地区、模型和解析表面可能不同。

4. 主要来源主机名（描述性）

窗口 / 平台	主机名	引用出现
B / Gemini	medium.com	7
	cloud.google.com	5
	docs.cloud.google.com	4
B / Kimi	arxiv.org	6
	mp.weixin.qq.com	3
	beta.openai.com	1
B / Qwen	zhuanlan.zhihu.com	9
	cloud.tencent.com	3
	developer.cloud.tencent.com	3

窗口 / 平台	主机名	引用出现
C / Gemini	medium.com	10
	docs.cloud.google.com	6
	arxiv.org	4
C / Kimi	arxiv.org	7
	platform.moonshot.cn	2
	36kr.com	1
C / Qwen	zhuanlan.zhihu.com	9
	cloud.tencent.com	3
	developer.cloud.tencent.com	3

表中仅列每个平台窗口内引用出现次数最高的三个精确主机名。主机名相同不代表具体页面相同；出现次数更高也不代表内容更准确、长期更稳定或更值得推荐。

5. 测量定义

变量	定义	不能推出
targetCitationRuns	具体引用 URL 命中预登记 DragonAI 目标的完成运行数	内容被吸收、推荐或产生转化
brandMentionRuns	回答正文命中预登记品牌别名的完成运行数	目标网页被引用
concreteUrlOccurrences	规范化后具体引用 URL 的跨运行出现总次数	不同 URL 数或来源质量
distinctHostnames	规范化引用 URL 中不同精确主机名数量	平台索引或算法相似度

证据阶梯

页面可访问 → 合格抓取 → 产品索引或可见性 → 搜索/检索触发 → 目标引用 → 内容吸收 → 用户结果。上一层不自动证明下一层。本报告只测量可复核的引用展示与品牌提及。

失败与占位符

完成但无具体 URL 的运行仍保留在分母中。星号、省略号或截断展示只记录为丢弃占位符，不计入引用、不同 URL 或竞争域。C 窗口保留 2 条被丢弃占位符。

6. 研究限制与权利边界

研究限制

- 正式基准只有一个固定中文无品牌问题、三个提供商和每窗口每家三次运行。
- Kimi 模型迁移，因此不计入 B 到 C 的同模型严格趋势。
- C 由 24 小时时间盒触发，不能归因于目标页发布、抓取、索引或改写。
- 引用 URL 出现不证明内容被答案吸收、来源正确、排名稳定或产生业务结果。
- 模型、搜索索引、地区、账号、工具和展示机制可能变化。

权利边界

CC BY 4.0 适用于 DragonAI 的观察设计、逐运行引用 URL 记录、聚合、字段定义与本报告文字。完整模型回答、提供商 API 原始载荷、第三方页面内容、产品名称与商标不在许可授予范围内。引用 URL 的记录不构成复制、许可、准确性确认或推荐。

关系披露

DragonAI 运营 AI 产品经理课程并监测自有域名。Qwen、Kimi、Gemini 及相关模型名称属于各自权利人；本报告不暗示合作、认可或背书。

7. References / 引用、数据与完整性

建议引用

DragonAI Editorial Research Team. AI Search Citation Sources for RAG vs. Fine-Tuning: 63 Public Qwen, Kimi, and Gemini Observations. Version 2.0.0. DragonAI, 2026.

公开研究页

<https://course.dragonai.tech/research/ai-search-rag-finetuning-citations-2026-07>

历史 B/C 权威快照

<https://course.dragonai.tech/datasets/ai-search-citation-benchmark-snapshot-2026-07-22.json>

五批逐运行引用数据

<https://course.dragonai.tech/datasets/ai-search-citation-benchmark-five-batch-run-level-2026-07-30.json>

百度六轮严格查询持续性数据

<https://course.dragonai.tech/datasets/ai-search-rag-baidu-serp-six-observation-persistence-2026-07-30.json>

完整研究存档

<https://course.dragonai.tech/datasets/ai-search-citation-benchmark-archive-2026-07-22.zip>

本 PDF 的固定公开 URL

<https://course.dragonai.tech/datasets/ai-search-citation-benchmark-report-2026-07-22.pdf>

完整性对象	SHA-256
历史 B/C 快照	3dedb418d0eecebd440d046d886ad6237d7d513ec8fabfecb981b2ec0bba9a83
五批逐运行 JSON	eb006cecfda3e2c9a01a03d940f606c4ca084d7aa591f5a158f11c5bfa465b7c
百度六轮持续性 JSON	52465a6983b8d5e30a2d6bf5204064e21568dd7307043021adcb41e4ccb30eef
研究存档 ZIP	见公开 SHA256SUMS ; ZIP 包含本 PDF, 不在 PDF 内嵌自引用摘要

外部仓库状态：manual_pending。当前没有 DOI、外部仓库记录或独立反向链接；只有经过真人审批、授权账号发布并被独立复核的公开记录，才可进入外部权威证据。

生成说明

本 PDF 由确定性生成器直接读取版本 1.1.0 历史 B/C 快照、五批逐运行引用数据、百度六轮严格查询持续性数据、MANIFEST 与研究存档交接文件。它不调用 AI 模型，不补写缺失结果，也不改变权威 JSON/CSV 数据。