

百度收录异常排查表

按现象查原因 · 每条都给验证方法和处置动作

收录出问题的时候，最容易犯的错是「凭感觉改」。这篇按七种常见现象分类，每种现象列出可能原因、怎么验证、以及该动什么。建议先验证、再动手——顺序反了，问题会越修越多。

现象一：整站一条都没被收录

可能原因	怎么验证
robots.txt 把整站封了	浏览器直接打开 /robots.txt 看有没有 Disallow: /
页面 meta 输出 noindex	查看网页源代码，搜 noindex
后台误开了「屏蔽搜索引擎」	WordPress「设置 → 阅读」检查那个勾
服务器对蜘蛛返回异常	用站长平台的抓取诊断，看百度蜘蛛拿到的是什么
域名刚注册、无任何外部链接	正常现象，新站需要时间，先保证内容与结构没问题
有防火墙拦了蜘蛛	检查 CDN / WAF 的爬虫白名单，确认百度 IP 段放行

现象二：新发的文章迟迟不收录

可能原因	怎么验证
栏目页没有入口，成孤岛	从首页出发点三次能不能点到这篇文章
发布频率忽高忽低	一天上百篇又突然停更，会被判定为异常
正文太短或与已有内容高度重复	检查是否只是换了个说法重写
没有主动推送	资源平台的普通收录 / 快速收录配额是否用满
站点抓取频次本身很低	资源平台看「抓取频次」曲线，低就是根本原因
文章没有被写入站点地图	打开 sitemap 检查这篇的 URL 在不在

现象三：原本收录了，现在掉光

可能原因	怎么验证
站点长时间不可访问	查服务器可用性记录，连续 5xx 是很强的降权信号
被判定批量采集或低质内容	近期是否大量发布同质化内容
网站被黑，挂马或被注入垃圾页	site 命令看是否有异常页面被收录
外链批量被清理	购买的链接集中掉线，会带动排名波动
主体内容被大改	改版后 URL 结构与内容全面变化，需要重新积累

现象四：只收录首页，内页全不进

可能原因	怎么验证
内页没有任何入口链接	最常见的真实原因 —— 内页是孤岛
全站内容高度雷同	模板化生成、只换关键词的页面会被整体忽略
内页全是列表，没有正文	分类归档页只有标题串，属于典型薄页
参数 URL 过多	同一内容有大量带参数的地址，稀释抓取配额

现象五：收录了，但搜不到排名

可能原因	怎么验证
词选得太难	目标词首屏是不是全是大型平台和行业头部
页面主题不聚焦	一页想覆盖十几个不相关的词，结果哪个都不强
内容深度不够	竞争对手在讲方法，你只是在重复概念
站点整体权重低	新站承接大词，通常是排不动的
缺少点击数据	有曝光没点击，搜索引擎会降低排序

现象六：快照长期不更新

可能原因	怎么验证
抓取频次低	资源平台看抓取数据，频次低则快照自然不更新
内容长期不更新	页面本身没有变化，蜘蛛没有重抓的理由

可能原因	怎么验证
页面被判定为低优先级	在站点里的位置太深、内链太少

现象七：抓取频次上不去

可能原因	怎么验证
站点响应慢或经常超时	抓取是有成本的，慢站点会被降低配额
死链和跳转太多	蜘蛛把配额浪费在无效地址上
sitemap 里混有大量 301、404	地图应该是「干净的目标地址清单」
站点结构过深	内容被埋在四五层以下，蜘蛛爬不到
内容更新不稳定	没有新内容，蜘蛛自然就不常来

排查的正确顺序

1. 先确认蜘蛛到底有没有爬进来 —— 没有抓取，谈收录和排名都是空的。
2. 抓取正常再看收录 —— 有抓取不收，问题在内容质量或重复度。
3. 收录正常再看排名 —— 有收录没排名，问题在选词和内容深度。
4. 排名波动看站点健康度 —— 服务器、外链、改版，都可能造成整体波动。

很多人在第 3 步遇到问题，却一直去改第 1 步的东西，来回折腾几个月没有进展。

需要人看一下？

把域名、想做的关键词、以及你在资源平台看到的数据（抓取频次、收录量）发给我们，我们会按上面的顺序帮你定位一遍，给出先动哪一块的判断。

站点：<https://ccombaidu.com> | 免费获取网站诊断报告