九章格物 NineGrids

GENERATIVE ENGINE OPTIMIZATION · 垂直行业研究

星阙实验室GEO研究中心 · 第 07 册

房产家居行业生成式引擎优化（GEO）白皮书

AI 居住决策时代，房源与家居服务如何被引擎判定为核得清、住得住、报得准——从记录合并、口径一致到过程留痕与可住性核查。

出品：九章格物 · 星阙实验室

作者：李现龙

版本：V2.0

时间：2026 年 11 月

权威信源：基于国家统计局、贝壳控股公开财报、住房和城乡建设部规章、市场监督管理总局规章及 KDD 2024 学术论文编制

微信扫我加好友

扫码添加微信 · 获取 GEO 方案

© 2026 九章格物 · 星阙实验室 · ninegrids.cn

目录

核心摘要 03

01 引擎回答居住问题前先核四类约束 04

02 一套房在索引层被拆成三条互不相认的记录 05

03 房源身份六个字段决定记录能否合并 06

04 面积三种口径如何改写同一个答案 06

05 面积口径还会改写税费与月供 08

06 价格写清有效期才算可引用 09

07 新房预售信息有三个关键时刻 10

08 二手带看记录如何证明真房源 11

09 长租公寓的房源状态字段 12

10 可住性六个字段回答能不能住 13

11 装修环节的难点在于过程不可见 14

12 报价五项缺一项就不可比 15

13 材料与工艺要写成可追溯的三段 16

14 定制家居的投影面积之争 17

15 交付记录才是最难复制的资产 18

16 四类失分在房产家居里的形态 19

17 一次假设推演：改善盘的三条答案路径 20

18 六类主体字段优先级并不相同 21

19 适老化改造的可核清单 22

20 学区与配套信息要带来源和年份 23

21 字段维护的分工与节奏 23

22 五类高频提问的答案骨架 24

23 度量与边界：可住性核查率怎么算 25

附录 A　房产家居字段自查清单 26

附录 B　术语表 27

附录 C　字段写法正误对照 28

附录 D　一次字段审计的操作步骤 28

附录 E　版本与修订记录 29

参考来源 29

核心摘要

居住类提问在生成式引擎里是**约束满足题，不是偏好推荐题**。用户给出预算、通勤、楼层、电梯、学区与交付标准，引擎要做的是把这些条件逐条落到可核字段上，再判断哪个候选满足。字段对不上，答案就只能退回形容词。

本白皮书给出三条判断。第一，房源先要在索引层合并成同一实体：一套房在平台有一条记录、在官网有一条记录、在内容平台还有一条记录，三条记录的身份键各不相同，合并失败等于候选集里出现三套不同的房。第二，面积与价格是口径问题而非表述问题，2025 年全国二手房网签面积 73685 万平方米，占新房与二手合计成交面积的 45.5%[1]，存量房信息由多方维护，口径漂移是主流失分形态。第三，装修与家居服务是过程型对象，过程不留痕，内容就无法被引用为事实。

全书 23 章，按五段推进：合并实体、统一口径、补齐交易与物理字段、处理过程型对象、给出失分与度量，末段补适老化、学区时效、字段维护与提问骨架四类落地场景。

边界需要提前写明：本白皮书讨论内容与字段如何被引擎读取，不涉及平台排名机制的内部参数；涉及具体城市与项目的判断，须以当地主管部门公示为准；所引市场数据只作背景，不用于推断单个项目的价格与去化。

01 引擎回答居住问题前先核四类约束

"300 万以内在通勤 40 分钟范围、有电梯、能上学"——这类提问在传统搜索里会被拆成一串关键词，在生成式引擎里会被拆成一串约束。约束与偏好的区别在于，**约束可以判真假，偏好只能排高低。**引擎敢写下"这套满足你的要求"，前提是它把每条约束对到了一个能判真假的字段上。

四类约束决定了核查顺序。第一类是身份约束（这是哪套房），第二类是口径约束（面积和价格怎么算），第三类是物理约束（层高、朝向、噪声、车位、电梯），第四类是过程约束（谁做的、什么时候做的、有没有凭证）。前两类决定这套房能否进入候选集，后两类决定它能否被写进答案，四类构成一个逐级收窄的结构，而不是并列四条。

答案池的一半来自存量房

国家统计局 2025 年统计公报显示，全年新建商品房销售面积 88101 万平方米，二手房交易网签面积 73685 万平方米[1][8]。按面积复算，存量房占两者合计的 45.5%（73685÷(73685+88101)=45.5%）。同一份公报还给出年末新建商品房待售面积 76632 万平方米，其中住宅 40236 万平方米；以待售面积除以月均销售面积（88101÷12≈7342 万平方米），约为 10.4 个月的库存量，其中住宅待售占待售总量 52.5%。

这三组数字放在一起说明一件事：**答案池里近一半的房源，其字段不是由开发企业一次性交付，而是由经纪机构、平台、业主多方长期维护。**新房字段的源头是同一份预售方案，存量房字段的源头可能有五六个，口径不一致几乎是结构性的，而不是个别机构的疏忽。

平台侧掌握的是强字段

贝壳控股 2025 年全年总交易额 3.18 万亿元，净收入 946 亿元[2]。按两者复算，综合变现率约为 2.97%（946÷31800）。其 2025 年第四季度移动月活跃用户平均为 4380 万名[2]。这组数字的意义不在规模，而在于平台把交易环节沉淀成了结构化字段：房源编号、楼栋、房号、面积、挂牌价、成交价、带看记录都在同一套编码里。

由此形成一个反直觉的判断：**字段多不等于被引用得多。**平台内部的强字段若没有以可被外部读取的形式写出来（同一实体名、同一面积口径、同一价格时点），引擎在跨源比对时仍会判为不同对象。字段的优势只在"能被对齐"时成立。

本书的判断与边界

本白皮书的判断是：房产家居行业的 GEO 不是把文案写得更像人话，而是把对象写成可核记录。论证顺序是——先合并实体（第 02、03 章），再统一口径（第 04 至 06 章），再补齐交易与物理字段（第 07 至 10 章），再处理过程型对象（第 11 至 15 章），最后给出失分形态、推演、主体差异与度量（第 16 至 19 章）。

边界有三条。一是平台检索与排序机制不公开，凡涉及"权重""排名"的表述，本白皮书一律不写具体参数，只写可观察的因果条件。二是市场数据仅采用官方统计与上市公司公开披露，行业机构口径因定义不一，本白皮书不引用。三是装修与家居服务的合规判断以现行部门规章为准，地方另有规定的，以地方规定为准。

表 1-1 四类约束的判断对象与失效表现

约束	判断对象	失效表现
实体	这条记录指向哪套房	同套房被判为两个项目
口径	这个数字按什么定义	面积与价格换算不一致
时点	这个值什么时候有效	旧价被当作现价引用
过程	这个结论凭什么成立	有结论无记录

四类约束的作用点是不同的：前两类决定记录能不能被正确理解，后两类决定记录能不能被信任。**缺前两类，答错；缺后两类，不敢答。**这两种失败在表面上都表现为"引擎不给答案"，但修复方式完全不同。

02 一套房在索引层被拆成三条互不相认的记录

同一套房子，在索引里通常存在三条记录：平台房源卡（含编号、面积、价格、图片）、开发企业官网的楼书页（含备案名、楼栋、户型）、内容平台的探盘文章（含口语化小区名、主观描述）。三条记录讲的是同一套房，但**彼此没有共同的键**——平台的编号在官网不存在，官网的备案名在内容平台被写成推广名。

表 2-1 三条记录的字段密度对照

字段	平台房源卡	官网楼书	探盘文章
身份键	平台内部房源编号	备案名 + 楼栋	推广名（口语）
面积	建筑面积，常带小数	户型的建筑面积区间	约数，"89 平那种"
价格	挂牌价，带更新时间	一房一价表或均价	常缺失或笼统
可住性	楼层、朝向、部分图片	户型图、交付标准	体验描述为主
过程	带看记录	施工与交付节点	无

合并失败会制造三个候选

假设一套房在平台叫"××花园 3 号楼 2 单元 1802"，在官网叫"××花园（备案名：××嘉园）3# 1802"，在探盘文里叫"××花园那个 89 平的小三房"。三条记录都进了检索结果，引擎若不能判定它们是同一对象，会出现三种后果：候选集被人为放大、价格出现三个版本、答案里出现"同一个小区有两套相似房源"的错觉。

这不是表述问题，是索引结构问题。判断顺序必须是：先让三条记录能被判定为同一实体，再讨论哪条记录写得更详细。顺序颠倒的代价很直接——**写得越详细，越容易在合并失败时被当成独立房源。**

引用机制在居住领域的具体形态

生成式引擎处理提问的链路可粗分为五段：改写提问、检索候选、抽取字段、按约束判定、生成答案并给出来源。居住类的特殊之处在第四段：判定必须落到具体字段，否则引擎无法同时满足"预算内"与"通勤 40 分钟内"两个条件。

学术侧对这一机制已有量化研究。Aggarwal 等人在 KDD 2024 发表的 GEO 框架中提出，面向生成式引擎的内容优化可在其基准测试中把可见性提升最多 40%，并明确指出**优化效果因行业而异**[7]。本白皮书引用这条结论只作方向性说明：它证明"内容可优化"成立，不证明任何具体比例可以套用到房产家居，更不能换算为成交承诺。

三条记录的质量差异从哪来

三条记录的产出目的不同，决定字段密度不同。平台记录为促成带看，强项是价格、图片、楼层；官网记录为合规披露，强项是备案信息、预售许可、交付标准；内容记录为获取阅读，强项是体验描述、周边配套。

三种目的各自合理，问题在于**没有任何一条记录是为"被引擎读取"而写的。**这正是 GEO 在房产家居的切入口：不是推翻原有三套内容，而是补一条能被读取的公共字段层，让三条记录通过它对齐。

表 2-2 公共字段层建议纳入的字段

层	字段	作用
身份层	城市、行政区、备案名、楼栋、房号、权属类型	让三条记录能被判为同一实体
口径层	建筑面积、套内面积、得房率、计价单位	让数字带定义，可被比对
时点层	价格生效日、有效期、状态更新日	让"现在"可判定
过程层	施工方、方案编号、验收单、工单	让"做过"可回溯

四层的排序不能颠倒。身份层不成立，后面三层写得再细也会被拆散；时点层缺失，口径层的数字就失去"当下有效"的前提。**公共字段层不是一份更长的文案，而是一张更短的对照表**——它只在内部对齐，对外仍由三条记录各自表达。

03 房源身份六个字段决定记录能否合并

合并需要键。房源身份键不能只用"小区名 + 面积"，因为同名小区、同名户型遍地都是。本白皮书给出的合并键是六个字段，六项全等才判同一实体，缺一项只能判"疑似"。

表 3-1 房源身份六字段与合并判定

字段	常见写法差异	合并时的作用
城市 + 行政区	只写城市不写区，或把新区写成老区名	排除跨区同名小区
备案名	官网用备案名，平台与内容用推广名	主键，必须以备案名为准
楼栋号	"3 号楼""3#""三期 3 栋"三种写法	把同小区不同楼栋分开
房号 / 单元	写"1802""2-1802"或不写	区分同楼层同户型的两套
建筑面积	89、89.36、89 平、89 平方米	辅助校验，不单独作主键
权属类型	商品房、经适房、回迁房混写	决定可交易性与贷款条件

说明：六字段为合并判定条件，非平台强制字段；缺项时只能给出"疑似同一房源"，不得直接断言。

备案名与推广名必须同时出现

推广名是营销用语，备案名是行政登记名称，两者在法律意义上并不等价。只写推广名的内容，在跨源合并时等同于缺主键；只写备案名的内容，在用户提问里几乎不会被命中。可行写法是**"推广名（备案名：××）"**并列置于首次出现处，其后统一用备案名。

这个动作的成本极低，收益却最直接：它让引擎在第一条记录里就拿到主键，后续所有字段都能挂上去。反过来，等到用户在答案里看到两套"同名但不同价"的房，再去补救，代价是全站改写。

权属类型为什么必须单列

权属类型决定三件事：能不能交易、能不能按揭、税费怎么算。经适房与回迁房在部分城市存在上市年限限制，若内容只写"满五唯一、随时过户"，就与权属字段冲突。**冲突字段比缺失字段更危险**：缺失会被引擎判为未知，冲突会让整条记录被降权，甚至整段不被引用。

可行写法是把权属与交易条件写在同一句，形成可判定的组合：**"权属为商品房，不动产权证已办理，可正常上市交易；契税与个税按当地现行口径计算"**。这类句子把"能不能买"这个问题回答完整，而不是留给用户自己去查。

合并之后才谈内容丰富度

合并成功只是入场。接下来的问题是口径：同一套房的面积有三个数字、价格有三种说法，合并后依然会在答案里打架。这一层留给第 04 至 06 章。

多平台同房源的编号归一

同一套房在不同平台有不同编号，是合并失败的第二种常见形态。可核做法是**在内部维护一条房源主键，并把各平台编号登记在一张对照表里**：主键由城市 + 行政区 + 小区备案名 + 楼栋 + 房号组成，平台编号只作别名，不参与实体身份。

表 3-2 房源主键与平台别名对照

项	示例值	是否参与身份判定
房源主键	某市/某区/××嘉园/3 栋/1202	是
平台 A 编号	××-1202-甲	否，仅别名
平台 B 编号	YY20260312-01	否，仅别名
经纪人工号	××××	否，属责任人字段

这样做的直接效果是：当平台 A 下架、平台 B 涨价时，主键不变，对外表达仍指向同一套房。**主键稳定，才谈得上状态同步。**反之，若以平台编号为身份，房源一旦跨平台发布，就会出现两条互不相认的记录，恰是第二章描述的那个问题。

04 面积三种口径如何改写同一个答案

一套房可以有三个面积：建筑面积、套内建筑面积、使用面积。三者不是精度差异，是定义差异。建筑面积含公摊，套内建筑面积只算套内墙体围合部分，使用面积再扣掉墙体厚度。同一套房，三个数字可能相差 8% 至 25%，取决于公摊比例与墙体厚度。

法规对这一点的要求是明确的：《中华人民共和国广告法》第二十六条规定，房地产广告中房源信息应当真实，面积应当表明为建筑面积或者套内建筑面积[3]；《商品房销售管理办法》第十四条要求商品房销售宣传广告内容必须真实、合法、科学、准确[5]。两条合起来说明一件事：**面积不是可以省略口径的描述，而是必须带定义的数量。**

同一套房三个数字的差距

图 4-1 同一套房三种面积口径的数值差距（示例口径）



示例说明：按得房率 80%（套内÷建筑面积）与墙体扣减 8.7% 推算使用面积，仅用于说明口径差异，不对应任何真实房源。

三个数字摆在一起，答案会变三次。"89 平方米三房"听起来宽松，"65 平方米使用面积的三房"听起来紧张。用户真正在意的是后者，但行业默认写前者。这中间没有说谎，只是**把定义留在了省略里。**

写法上让数字与口径同句出现

可引用的写法有三个要求：三个数字同句或同表出现、每个数字带定义、口径比例可复算。例如"建筑面积 89.0 ㎡，套内建筑面积 71.2 ㎡（得房率 80%），使用面积约 65.0 ㎡"。这样的句子在抽取时能被拆成三个带标签的量，任一提问口径都能命中。

反过来，只写"89 ㎡"的句子在抽取时只能得到一个无标签数值，引擎若把它对到"套内 89 ㎡"的提问上，就会生成一个错误答案。错误的来源不是内容假，而是**标签缺失导致的错配。**

表 4-2 面积口径的换算关系与常见误写

口径	换算关系	常见误写
建筑面积	含公摊的合同面积基准	把"套内 71"写成"89 ㎡"
套内建筑面积	建筑面积×得房率	不给得房率，直接写套内数
使用面积	套内再扣墙体厚度	与套内混用，两个词指同一个数

三种口径之间不存在"更准"的一种，只存在"是否标明"的差别。户型图上标注的是轴线尺寸，宣传页上写的是建筑面积，用户算的是使用面积——**三个数各有出处，混写就会变成三个答案。**

得房率的复算校验

得房率 = 套内建筑面积÷建筑面积，只要两个数都在，比值就能复算。这一点对内容侧很有用：**凡是给出得房率的内容，都应能被读者用两个数字验算一次。**如果 71.2÷89.0 得不出 80%，说明数据有误，需要回查图纸。

反过来，只写得房率不给两个原始数，读者也无从验算，这个百分比就成了一个无法核的修饰词。这也是本白皮书主张"三数同句"的原因：不是为了信息量，是为了让复算成为可能。

得房率并非越高越好

得房率高意味着公共空间被压缩。梯户比、单元门厅、设备平台都会影响实际体验，而这几项往往不写进得房率。因此本白皮书不建议把"高得房率"当单独卖点陈述，而应**把得房率与梯户比、公共部位配置同段写**，让读者同时判断舒适度与面积效率。

这一条也解释了为什么单一指标不能作为结论：在居住领域，几乎每个数字都有一个反向代价。只写数字不写代价，答案看似精确，实际把判断成本转移给了用户。

05 面积口径还会改写税费与月供

面积口径不只影响观感，还影响两笔钱：税费与月供。多数城市在契税、维修资金、物业费的计费基数上以建筑面积为准，而部分城市的套内计价楼盘按套内面积签约。同一个"89 ㎡"，在两种签约口径下，用户实际支付的总额不同。

这里必须给出边界：**具体计费口径由当地税务与银行执行，全国并不统一**，本白皮书不给出跨城通用的换算比例，只提出一个内容要求——凡涉及税费与月供的表述，必须写明其计算基数用的是哪一个面积口径。

表 5-1 面积口径在合同与费用中的常见对应关系

环节	常见计费基数	内容侧须写明的信息
网签合同	建筑面积或套内建筑面积（按当地规则）	合同采用哪一种口径
契税	通常按成交价，与面积档位相关	面积档位对应的税率区间
住宅专项维修资金	常按建筑面积	单价与基数
物业费	常按建筑面积	单价、基数、包含项
装修报价	建筑面积或套内面积，各家不同	报价采用哪一种口径

说明：表内为常见做法，用于说明"必须写明基数"，不构成对任何城市规则的概括；实际以当地现行规定与合同约定为准。

这张表的价值不在于教会读者算账，而在于提示一个高频错配：**报价按套内、物业按建筑、合同按建筑，三个基数混在一页里，用户就会得到三个"总价"**。当引擎试图把这三段内容合成一段答案时，任何一段缺基数，答案就会失真。

月供表述要带首付与年限

"月供约 1.2 万元"这类句子几乎不可用，因为它缺首付比例、贷款年限、利率与是否含税费。可核写法是**"总价 380 万元，首付三成，贷款 266 万元、30 年期、按当前 LPR 加点测算，月供约××元，测算日期××"**，把测算条件写出来，句子才从结论变成可复算的推算。

还要提醒一句：利率会变，测算值必须带日期，否则半年后这条内容就成了错误信息。这与第 06 章的价格有效期是同一类问题——**凡是会随时间变化的量，都必须带时点。**

三税一费的口径来源

税费是面积口径的下游效应，也是最容易被写错的一类：税率与认定条件由地方执行，跨城市不可直接套用。可核做法是**写清适用条件而不写具体税额**："契税按首套认定、面积 90 ㎡以下或以上分档计征，具体比例以当地现行规定为准"[11]。

表 5-2 交易环节四项费用的判断依据

项目	判断依据	内容侧可写与不可写
契税	套数认定 + 面积分档	可写条件，不可写死税率
增值税及附加	持有年限	可写年限条件，税额以核定为准
个人所得税	是否满五唯一	可写认定口径，不作税额承诺
维修资金	按建筑面积计缴	可写计缴方式与单价

四项里，只有维修资金可以按面积直接算出确定数值，其余三项都依赖认定结果。**把"应缴多少"写成确定数，等于替主管部门下结论**——这既不可核，也有合规风险。

06 价格写清有效期才算可引用

价格失分比面积更常见，原因是价格会变而页面不会跟着变。一条 2024 年写的"均价 4.2 万/㎡"，到 2026 年仍挂在官网上，引擎抓到的是一个没有时点的数字，只能当作现行事实使用。

法规对此有直接要求：《房地产广告发布规定》第十条规定，房地产广告中对价格有表示的，应当清楚表示为实际的销售价格，明示价格的有效期限[4][10]。这条规定同时给出三个要素：**价格数值、计价单位、有效期**。三者缺一，价格就不可引用。

表 6-1 价格写法与可引用性判定

写法	缺什么	判定
"均价 4.2 万"	计价单位、有效期	不可引用：无法判断是元/㎡还是总价，也无时点
"总价 380 万起"	对应条件、有效期	不可引用："起"未说明对应户型、楼层与付款方式
"4.2 万元/㎡（建筑面积），2026-04-30 前有效"	—	可引用：数值、单位、口径、时点四项齐全

判定依据：《房地产广告发布规定》第十条关于实际销售价格与有效期限的要求[4]。

"起"字必须带条件

"起"是价格区间的最低端，但最低端通常对应特定户型、楼层或一次性付款。若只写"起"不写条件，用户会当成普遍价格，引擎也会当成可比价。可引用写法是把条件写进同一句：**"总价 380 万元起，对应 2 号楼 89 ㎡两房、一次性付款，2026-04-30 前有效"**。

一口价与区间价分开写

一口价是确定值，区间价是范围，混写会让答案出现"同一套房两个价格"。可行做法是同一条目不混用：要么给一口价，要么给区间并注明上下限对应条件。跨房源比较时只比较同口径价格——元/㎡建筑面积对元/㎡建筑面积。

存量房的三套价格

二手房价格不是单一数字，而是挂牌价、成交价、评估价三套。挂牌价可调，成交价滞后，评估价受银行口径影响。内容若只写"该小区均价 4.2 万"，实际上把三个口径混成了一层。

表 6-2 存量房三套价格的取用场景

价格类型	谁在用	内容侧写法要求
挂牌价	业主与经纪机构	写明挂牌日期与可谈空间
成交价	买方与税费计算	写明成交月份与统计样本量
评估价	银行与贷款审批	写明评估机构与评估时点

可核写法是**分别标注"近 90 天成交单价区间 3.8 万至 4.3 万元/㎡（建筑面积），样本 12 套，据平台成交记录"**。区间加样本量加时点，才能支撑"这个小区大概什么价"这个结论。

一房一价与楼层差的可核表达

同一项目不同楼层价差，是价格字段里最容易被含糊处理的一项。"楼层越好价格越高"是常识，但不是可核信息。可核写法是**给出价差规则与基数**：如"同户型中间楼层为标准价，每升高一层加 30 元/㎡，顶层与底层各减 200 元/㎡"。

表 6-3 一房一价表的最小字段

字段	示例	缺失后果
房号	3 栋 1202	无法定位到具体单元
面积口径与数值	建筑面积 89.0 ㎡	单价失去分母
单价与总价	4.20 万元/㎡，总价 373.8 万元	读者无法复算
生效日与有效期	2026-04-30 前有效	旧价被当作现价
优惠条件	一次性付款减 2%	"起价"与"实价"混淆

五个字段齐全，一房一价才是可核的；缺任何一项，读者只能得到"大概多少钱"。**价格的可引用性，来自能被复算与能被判定时点**，不来自数字的大小。

07 新房预售信息有三个关键时刻

新房的内容风险集中在时点。同一份预售方案，在拿证时、在开盘时、在交付时，能写的结论不同。混用三个时点的信息，是开发侧最常见的内容失分。

图 7-1 新房内容可写范围的三个时点



说明：三个时点对应的可写范围为本白皮书的判定口径，具体以各地预售管理与交付管理规定为准。

第一个时点最重要。预售许可把"能卖什么、卖多少钱"固定下来，一房一价表就是这个时点的结构化产物。**有预售许可的房源才有可引用的价格**，没有许可的价格只是意向报价，写进内容会误导用户判断可交易性。

实测面积与合同面积可能不同

交付时按实测面积结算，实测与合同约定的建筑面积可能出现误差。误差如何处理，取决于合同约定与当地规则。**内容侧不应把合同面积写成最终面积**，而应写成"合同约定建筑面积××㎡，交付以实测为准"。

这一条看似保守，实则提升可信度：它告诉用户"这个数字的来源是合同，不是实测"，用户在看到交付信息时不会产生落差。引擎读到这类句子，也能把它归入"约定值"而非"实测值"。

交付标准要逐项落到可交付物

"精装交付"不是一个可核结论。可核写法是把交付标准拆成可交付物清单：门窗品牌与型号、地板材质与规格、厨卫洁具品牌、开关面板品牌、是否含中央空调与新风。每一项都是可验收的实体，**验收时能对得上的才是标准，对不上的只是描述。**

实践中还有一个易错点：样板间与交付标准的差异。样板间常含非交付项（软装、家具、灯具升级），内容若把样板间效果当作交付效果，就等于把展示品当成了商品。可行写法是在描述样板间时同步标注"非交付项清单"。

08 二手带看记录如何证明真房源

存量房的核心问题不是描述不美，而是"这套房还在不在"。房源下架、价格调整、业主换中介，都会让页面内容与事实脱节。带看记录是唯一能证明"这套房真实存在且仍可交易"的动态证据。经纪环节的信息真实性也有专门的管理要求（《房地产经纪管理办法》）[9]，其中对房源信息发布与核验有明确规定，本白皮书只讨论由此产生的字段要求。

表 8-1 真房源的四个可核要素

要素	可核形式	缺失后的后果
房源编号	平台内可检索的唯一编号	用户无法确认真伪
最近带看时间	具体日期	无法判断是否仍在售
价格更新日期	调价记录与生效日	报价可能是数月前的
核验方式	产权核验、实地核验、业主委托	无法区分委托房源与信息房源

四个要素里，**时间字段是唯一不能被文案修饰的。**一条"3 天前带看、5 天前调价"的记录，比一段"业主诚心出售"的描述更能支撑"这套房现在能买"的结论。

委托与信息房源必须区分

有业主书面委托的房源，机构可以代为披露房屋信息并安排带看；仅有信息、没有委托的房源，披露范围受限，成交路径也不确定。内容侧若不区分这两类，用户会误以为所有挂牌都能即时签约。

可行写法是**在房源条目内标注委托状态与委托期限**，例如"业主独家委托，委托期至 2026-12-31"。这一条既是合规要求，也直接影响引擎对"可交易性"的判断。

带看记录的隐私边界

带看记录可以披露"何时带看过"，不宜披露"谁看过"。前者是房源状态，后者是个人信息。**记录摘要只写时点与结果，不写客户身份**，这条边界同样适用于第 15 章的交付与维修记录。

表 8-2 带看记录五项最小字段

字段	示例	作用
房源主键	××嘉园 3 栋 1202	锚定到唯一房源
带看时间	2026-03-12 14:00	证明时点可回溯
带看人	工号××，实名可查	证明责任可追溯
房源状态	钥匙实勘、可看	证明房源真实在售
反馈摘要	"采光好，临路一侧有噪声"	同时提供缺陷信息

五项中最后一项最容易被省略。只记"带看成功"的记录，只能证明有人去过；记录了客户的真实反馈，才同时证明房源存在、缺陷已知。**愿意写缺陷的记录，比只写成功的记录更可信。**

09 长租公寓的房源状态字段

租赁是居住领域里状态变化最快的场景。同一间房，一个月内可能经历"可租—已定—已租—退租—保洁—可租"六个状态。内容若只写"精装一居，拎包入住"，用户看到的可能是一间已经租出去的房。

表 9-1 租赁房源的状态字段与更新要求

字段	取值示例	更新要求
当前状态	可租 / 已定 / 已租 / 维保中	状态变化当日更新
可入住日期	具体日期	随状态联动
租期与起租	最短租期、起租日	随政策调整
押付方式	押一付三、押一付一	随合同模板调整
维修责任	家电维修、日常保洁归属	合同变更时更新

租赁内容的特殊性在于：**状态字段的价值高于描述字段。**一间房的装修描述写得多好，只要状态还是"已租"，答案就会误导用户。因此租赁侧的 GEO 首先是把状态做成可读字段，而不是把文案写得更动人。

押付方式为什么影响决策

押一付三与押一付一，对租客的现金流压力差异明显。同一间房，两种押付方式下的"首期支出"可能相差两个月租金。内容若不写押付方式，用户无法完成"我能不能租得起"这一步判断。

可核写法是**"押一付三，首期支付四个月租金，含首月租金与押金；物业费 3.8 元/㎡•月由租客承担"**。把首期支出算清楚，比强调"性价比高"有效得多。

集中式与分散式的字段差异

集中式公寓（整栋运营）有统一的房源编码与状态系统，字段更新及时；分散式房源（个人业主委托）依赖人工维护，状态滞后更常见。两类主体的字段优先级因此不同：**集中式先保状态，分散式先保委托与可看房时间。**

房源状态的三层证据

"在租"与"可租"是两个不同的字段，前者描述现状，后者描述可交易性。可核写法要求三层证据：**状态词、时点、依据。**

表 9-2 长租房源状态的三层证据

层	内容	示例
状态词	在租/空置/维修中/已预订	"空置，可租"
时点	状态更新日与最早起租日	"2026-03-18 更新，最早 2026-04-01 起租"
依据	谁维护、依什么维护	"运营主管每日巡检更新"

三层齐备后，"这套现在能不能租"就有了确定答案；缺任一层，用户只能打电话问，引擎也无法引用。**可租性不是形容词，是一个带时点的布尔值。**

10 可住性六个字段回答能不能住

身份与口径解决"这是哪套房、多少钱"，第三个问题才真正决定用户去不去看房：能不能住。可住性不是主观印象，它可以被拆成六个可核字段。

图 10-1 可住性问题到六个可核字段的分层结构



说明：字段与子问题的对应关系为本白皮书整理的判定口径，非平台官方分类。

表 10-1 可住性六字段的判定方式与缺失后果

字段	可核形式	缺失后答案会变成
净层高	数值，如"净高 2.85 米"	"层高舒适"这类无法判定的形容词
朝向与采光	"南北通透""主卧朝南"，落到具体房间	只剩户型图，无法判断日照
临路噪声	"临××路一侧，夜间有货车通行"	用户到现场才发现，形成退订
车位配比	"1:0.8，含地下产权车位数量"	"车位充足"式结论，无法核
梯户比	"两梯四户""一梯两户"	高峰等待时间无从判断
物业费标准	"3.8 元/㎡·月（建筑面积），含公摊水电"	长期成本无法纳入预算比较

六个字段的共同点是**都可被第三方核验**：层高可以量，朝向可以看，梯户比写在图纸上，物业费在合同里。正因为可核，引擎才敢把结论写进答案。这与营销语言的分界线很清楚——可核的说事实，不可核的只能说感受。

为什么噪声字段最容易被删

噪声是唯一一个"知道的人不愿写、不知道的人写不出"的字段。开发企业清楚临路一侧的价值折损，因此倾向省略；经纪机构带看时才知道，通常不写进页面。结果是引擎在回答"这套房子安静吗"时，只能给出一句"小区环境安静"，而这句结论没有任何字段支撑。

更可核的做法是把噪声写成可判定的描述：临哪条路、什么时段、什么车型、有无隔声屏障。这类写法不会美化项目，但会显著提升答案可信度，因为**引擎引用的是能判真假的句子**。

表 10-2 一套示例房源的核查率算例

字段	是否写明	计入
城市与行政区、备案名、楼栋、房号、建筑面积、权属类型	6 项已写明	6
净层高、朝向	已写明	2
临路噪声、车位配比	只写"安静""车位充足"	0

字段	是否写明	计入
梯户比、物业费	已写明	2

算例说明：已写明 10 项，应核 12 项，核查率 83.3%（10÷12）。缺失两项均为可住性字段，属"描述型"而非"数值型"，须改写为可核形式后才计入。

这个算例说明一件事：**核查率的分母固定，分子只认可核写法。**写"车位充足"与没写，在分子上完全一样——不是因为这句话无价值，而是因为它无法被验证，也就无法被引用。

可住性六个字段的排序依据

六个字段并非等权。按"影响退订概率"排序，噪声与梯户比排在前两位：噪声无法改造，梯户比无法更改；层高与朝向次之，车位与物业费再次之。这个排序不是价值判断，而是**不可逆性排序**——越不可逆的缺陷，越应该在内容里提前写明。

表 10-3 六项可住性字段的核查方法

字段	核查方法	证据形态
净层高	现场实测，或引用竣工图	数值 + 测量位置
朝向	对照户型图与指北针	文字 + 图
采光	拍摄时间 + 窗位	影像带时间戳
噪声	临路一侧 + 时段描述	文字 + 位置说明
通风	窗向组合与对流路径	文字
车位	配比 + 车位性质	数值 + 产权说明

六项的核查成本差异很大：净层高与车位可查图纸，采光与噪声必须现场，朝向与通风介于两者之间。**成本高的项更需要在内容里写清来源**，否则读者无法判断这个数字是实测还是抄来的。

11 装修环节的难点在于过程不可见

房源是结果，装修是过程。引擎可以读到"这套房子长什么样"，很难读到"这个装修是怎么做出来的"。而装修纠纷几乎都发生在过程里：材料换了、隐蔽工程没验收、工期一拖再拖。

过程之所以可核，是因为法规已经规定了必须留下的痕迹。《住宅室内装饰装修管理办法》第五条规定了禁止行为，包括未经原设计单位或具有相应资质等级的设计单位提出设计方案而变动建筑主体和承重结构[6]；第六条规定搭建构筑物、改变外立面、拆改供暖与燃气管道等行为须经批准[6]；第九条要求涉及上述内容的装修活动必须委托具有相应资质的装饰装修企业承担[6]。三条给出三个可核点：**有没有方案、有没有审批、由谁施工。**

可被引用的装修内容长什么样

把"精工细作、良心工程"换成三段可核表述：谁做的（企业名称与资质）、按什么做的（设计方案与变更单编号）、做到什么程度（隐蔽工程验收记录与竣工日期）。三段的共同点是有落点——落到了具体记录物上。

反直觉的判断是：**效果图越精美，越难被判定为事实。**效果图是渲染结果，不含施工信息；而一句"防水做了两遍，闭水试验 48 小时，2026 年 3 月 12 日验收，验收单编号××"在抽取时是一条完整事实句。

开工申报是被忽略的公开信息

该办法第十六条要求装修人或装修人与装饰装修企业，与物业管理单位签订住宅室内装饰装修管理服务协议，协议内容包含实施内容、实施期限、允许施工的时间、废弃物清运与处置等[6]。这意味着一个合规项目在物业侧本来就留有一份书面记录。

这份记录通常不公开，但可以被转述：**"已办理装修申报，施工期为 2026 年 3 月至 5 月，允许施工时段 8:00 至 18:00"**。这类信息不影响隐私，又能让"这是一个正规施工的项目"变得可核。

违规行为的处罚是硬信息

该办法第三十八条对若干行为规定了罚款，例如将没有防水要求的房间或阳台改为卫生间、厨房间，或拆除连接阳台的砖、混凝土墙体，对装修人与装饰装修企业分别处以不同额度罚款[6]。这意味着"拆阳台墙改卫生间"不是风格选择，而是有明确法律后果的行为。

内容侧的价值在于**把这类禁止项写成前置提示**，而不是等纠纷发生后再解释。用户在提问"能不能把阳台改成卫生间"时，引擎能引用的答案只有两种：法规禁止，或需要审批并委托有资质企业。写清这两种，比承诺"都能做"更可信。

同一字段对不同家庭的权重不同

可住性字段没有统一权重。带婴儿的家庭把临街噪声排在前面，有老人的家庭把电梯与门槛高差排在前面，长期出差的家庭更在意物业响应。内容侧无法替每个家庭排序，能做的是**把字段写全，让读者自己排。**

表 11-2 三类家庭的字段权重差异

家庭类型	首要字段	次要字段
有婴幼儿	临街噪声、通风采光	电梯、物业响应
有老人同住	电梯、门槛高差	采光、就近医疗
通勤为主	距离与出行方式	车位配比、物业费

这张表对内容写作的指导意义是：**不要预设买家的优先级**。一旦预设，就会把次要字段写成卖点、把关键缺陷省掉，答案的可信度随之下降。写全六项，再由读者排序，是更稳的做法。

12 报价五项缺一项就不可比

装修报价是房产家居领域最容易被截断的内容。一句"899 元/㎡ 全包"在传播上极有效，在判定上几乎无效，因为它没有说明计价面积口径、包含项、不含项、工期与保修。

法规侧已有对应要求。该办法第二十四条规定，装修人与装饰装修企业应当签订书面合同，合同内容包括房屋间数与建筑面积、装饰装修的项目方式规格质量要求与验收方式、开工竣工时间、保修内容与期限、工程价格与计价支付方式[6]。其中至少三项直接决定报价可比性。

表 12-1 报价五项与不可比的具体后果

要素	合规依据	缺失后的后果
计价面积口径	建筑面积或套内面积须写明	同一报价在两种口径下相差 20% 以上
包含项	项目、方式、规格、质量要求	"全包"实际不含主材或水电改造
不含项	同上，须反向列明	开工后增项，预算失控

要素	合规依据	缺失后的后果
工期与延期责任	开工、竣工时间	无法比较交付确定性
保修内容与期限	保修内容与期限	长期成本不可估

合规依据：《住宅室内装饰装修管理办法》第二十四条关于装修合同应包含内容的规定[6]。

五项齐全的报价才构成可比值。可引用写法为：**"899 元/㎡（按建筑面积计），含主材、辅材、人工与管理费，不含拆改、水电改造与家电，工期 60 个工作日，延期按日千分之三计，保修 2 年、防水 5 年"**。这一句里每个数字都能被单独抽取。

表 12-2 三种报价形态的可比性差异

形态	常见表述	可比性
平米包	"××元/㎡，含基础施工"	口径未定则不可比
套餐包	"三室两厅全包 ××万"	需按面积与项目逐项折算才可比
清单报价	按项目逐条列价	可比性最高，核对成本也最高

三种形态各有用处，但**不能混在同一比较里**。平米包对应面积，套餐包对应户型，清单报价对应项目。用户问"哪家便宜"时，引擎若把三种形态的价格放在一起排序，得出的结论必然失真。

为什么"不含项"比"包含项"更重要

包含项是卖方主动披露，不含项是买方事后才发现。引擎在回答"这个报价贵不贵"时，真正需要的是完整边界。只写包含项的报价，会让答案把一个不完整的价格当成完整价格，最终伤害的是转化后的信任。

由此给出一个可判定的比值：**报价完备度 = 含项数 ÷（含项数 + 不含项数）**。这个比值不需要精确到小数点，只需要回答"边界是否写全"。

团购价与限时价的时点问题

"限时 899 元/㎡"若不写活动期，半年后仍会被引用为现行报价。**凡带促销属性的价格，必须写截止日期与适用条件**，这与第 06 章的规则同源：价格是变量，不是常量。

13 材料与工艺要写成可追溯的三段

材料信息决定装修内容能否被当成事实引用。三段式是成本最低的写法：品牌与型号（可查）、执行标准与检测（可核）、施工做法（可复述）。

图 13-1 装修过程五个节点对应的留痕物



说明：五个节点对应《住宅室内装饰装修管理办法》第二十四条所列合同要素中的项目、质量验收与保修内容[6]。

型号比品牌重要

"使用××品牌乳胶漆"只说明品牌，不说明型号。同一品牌下不同型号的耐擦洗次数、遮盖力、环保指标差异很大。可核写法是品牌＋型号＋执行标准：**"××品牌××型号内墙乳胶漆，执行 GB/T 9756，提供该批次检测报告编号"**。型号与标准都能被第三方查到，句子才成为事实。

环保等级不能写成"零甲醛"

人造板及其制品的甲醛释放限量有国家标准可依（GB 18580），检测以报告形式出具。把"零甲醛"当结论陈述，既不符合检测口径，也会造成预期落差。可核写法是**"板材甲醛释放量按 GB 18580 检测，检测报告编号××，检测日期××"**，把结论交给报告，而不是交给形容词。

施工做法要能被复述

"工艺精湛"无法复述，"防水两遍、闭水 48 小时、墙面找平两遍打底"可以复述。能被复述的句子才可能被引擎转述进答案。这也是本白皮书对"内容可核"的最低要求：**写完一段工艺描述后，让别人照着这段话说一遍，说不出来就是不可核。**

还有一个常被忽略的细节：材料变更。**变更必须留痕，否则"用料一致"这个结论就是假的。**可行做法是在内容里写出"如遇断货，替换需业主书面确认并注明替代型号"，把变更规则前置，而不是事后解释。

工期与延期责任怎么写才可核

工期是家居家装报价里第二高频争议项。可核写法不是"约 45 天完工"，而是**给起点、给口径、给延期处理**："自开工日计 45 个工作日，不含主材到场等待；因甲方变更或主材延误顺延；超期按每日合同额 1‰ 支付违约金"。

三个要件缺一不可。没有起点，45 天从哪天算是分歧来源；没有口径，工作日与自然日会被混用；没有延期处理，工期条款就只是一句愿望。**写入内容侧的工期，应当是合同条款的摘要，而不是宣传口径的概括。**

表 13-2 材料与工艺三段的追溯链

段	要写什么	可回溯到
品牌与型号	品牌、型号、规格	采购合同与进场单
执行标准	标准编号与检测报告号	检测机构报告
施工做法	遍数、顺序、验收方式	施工记录与验收单

三段之间是递进关系：只写品牌，读者知道用了什么；加上标准，读者知道是否合规；加上做法，读者才知道做得对不对。**三段的差别不在信息量，在能否支撑结论。**

14 定制家居的投影面积之争

定制柜体的计价口径比装修报价更隐蔽。同一组柜子，可以按投影面积（宽×高）、展开面积（所有板材面积之和）、延米（按长度）三种方式计价，三种算法下的单价差异极大，而消费者通常只看到"××元/㎡"这一个数字。

表 14-1 三种计价口径的适用与风险

口径	计算方式	内容侧必须写明
投影面积	柜体宽×高	是否含柜内层板、五金、门板

口径	计算方式	内容侧必须写明
展开面积	所有板材面积之和	板材厚度、柜体结构
延米	按柜体长度计价	台面、地柜、吊柜的比例

同一组柜子按投影计价与按展开计价，最终的"元/㎡"单价可以相差一半以上，因为两者的分母不同。**分母不同，单价就不可比**——这是定制家居最典型的可比性陷阱，也是用户投诉的高发点。

五金与层板是最常被省略的部分

柜体报价常含柜体与门板，不含五金（铰链、滑轨、拉篮）与增加层板。用户按投影面积算出的预算，在加上五金与层板后可能上浮两成。**可核写法是把五金品牌、数量、是否包含写明**，例如"含××品牌铰链 12 只、三节滑轨 3 副，增加层板按××元/块计"。

交付周期是定制家居的第二决策点

定制柜体从测量到安装通常需要数周，工期受板材、产能与物流影响。内容若不写交付周期，用户无法把它与装修工期对齐，最终形成"柜子等不到、装修停工"的常见问题。可核写法是"**下单至安装完成约 35 个工作日，含两次上门测量与一次复尺**"。

15 交付记录才是最难复制的资产

房产家居行业最容易被模仿的是文案，最难被模仿的是记录。交付验收单、隐蔽工程影像、维修工单，这三类记录构成服务方的履约档案，也是引擎在回答"这家公司靠不靠谱"时唯一能抓到的硬证据。

行业结构正在往这个方向走。贝壳控股 2025 年净收入中，非房产交易业务占比提升至 41%[2]，家装与租赁成为增长来源。这个变化的含义是：交易平台正把一次性撮合延伸为长期服务，而长期服务的核心资产就是过程记录。

记录的价值在可回放，不在好评

好评可以被运营，工单时间戳不能。一条"2026 年 3 月 12 日上门处理漏水，更换角阀，客户签字"的记录，包含时间、动作、结果三个可核要素，比一百条五星好评更能支撑"响应及时"这个结论。

由此形成判断：**内容侧应把记录摘要当作常规内容形态，而不是把客户评价当作唯一证据。**记录摘要不需要披露隐私，只需保留时间、部位、动作、结果四项。

表 15-1 三类交付记录的最小字段

记录	最小字段	可支撑的结论
交付验收单	验收日期、房间、验收项、双方签字	交付时状态已确认
隐蔽工程影像	拍摄日期、部位、施工阶段	隐蔽部分做法可回溯
维修工单	报修日期、上门日期、处理动作、结果	响应时效与处理能力

边界要强调：记录摘要可以公开，原始凭证不宜公开。公开的是**可核要素与编号**，不是客户信息与合同原件。这既是合规要求，也是内容安全的底线。

记录如何转成内容

记录本身不是内容，需要一次转写：把工单里的"更换角阀"转成"这类问题的常见原因是角阀老化，处理方式是更换并做水压测试"。转写之后，记录从个案变成可复用的知识，才可能被引擎引用。**不转写，记录永远只是一份档案。**

表 15-2 从记录到内容的三步转写

步骤	动作	产出示例
摘要素	取时间、部位、动作、结果	"2026 年 3 月，卫生间角阀老化漏水，更换并做水压测试"
归类型	把个案归入常见问题	"卫生间渗漏的三类常见原因"
给边界	写清适用条件与不适用情形	"适用于角阀与软管老化，不适用于防水层破损"

三步之后，一段工单能支撑一条知识条目；十段同类工单能支撑一个判断，例如"该小区交付 5 年以上，角阀老化是高频报修项"。**这类判断只有做过服务的一方写得出来**，也正是最难被复制的内容资产。

16 四类失分在房产家居里的形态

把前文问题归并，房产家居的失分收敛为四类：实体散列、口径漂移、过程无痕、记录缺失。四类不是并列关系，而是四道依次收窄的闸门——每过一道，候选集都会缩小。

图 16-1 四道闸门对候选集的收窄（示例推演值）



图注：四道闸门的通过比例为本白皮书按字段完整性推演的示例值，非平台实测数据，仅用于说明收窄关系。

表 16-1 四类失分的表现、修复动作与代价

类型	典型表现	修复动作	代价
实体散列	同名小区三种写法，房源被拆成多条	统一备案名 + 楼栋 + 房号	低，一次改表即可
口径漂移	面积不带定义、价格不带时点	补口径与有效期	中，需回改全部页面
过程无痕	装修只有效果图，没有验收与检测	按节点补记录	高，需重建流程
记录缺失	交付与维修无凭证	建立工单与验收留档	最高，需改组织结构

修复顺序必须从低代价到高代价。前三类的修正在内容侧完成，第四类要在流程侧完成。若顺序倒过来，先做记录体系而实体仍是散的，记录本身也会被拆成多条，等于做了两遍。

> **反直觉的一点**：四类失分里，对答案影响最大的不是"写得少"，而是"写得多但口径不同"。散列与漂移都发生在信息丰富的页面上，信息越全、口径越不统一，错配的机会越多。

表 16-2 三道闸门的自检问题

闸门	自检问题	答"否"时先改什么
第一道：能不能合并	换一台设备搜同一小区，结果是同一套房吗	备案名与楼栋房号
第二道：能不能对齐	两个页面写的面积能互相验算吗	面积与得房率同句
第三道：能不能回溯	结论能否指向一份可出示的记录	验收单与工单摘要

三个问题都可以由非技术人员回答，也都能在一小时内验证。**能被快速验证的检查，才会被反复执行**——这是三道闸门设计成问句而非指标的原因。

17 一次假设推演：改善盘的三条答案路径

为把前面的判定串起来，这里做一次假设推演。推演对象是一个虚构的改善型楼盘与两项虚构的家装服务，**不对应任何真实主体**，所有数值均为示例口径。

表 17-1 三条路径的字段状态与答案形态

路径	字段状态	引擎能给出的答案
A 只写卖点	无备案名、无面积口径、无价格时点	无法判定是否满足约束，通常不进入候选
B 字段齐全但口径混乱	六字段齐，面积三口径混写，价格无有效期	进入候选，但两个候选价格互相矛盾，答案自相冲突
C 字段与口径一致	六字段齐、三面积同句、价格带有效期、装修有记录	可写成"满足预算与通勤条件、使用面积约 65 ㎡、2026-04-30 前 4.2 万元/㎡"的完整句

三条路径的差别不在文案水平，而在**每次提问都能被同一条记录接住。**路径 A 的问题是记录不可判，路径 B 的问题是记录自相冲突，二者都不是"写得不好"，而是"写得不一致"。

推演还暴露一个容易被忽略的点：路径 B 的页面通常看起来更专业——有图、有参数、有对比表。这恰恰说明，**内容完备度与可判定度是两个指标，前者可以靠人力堆出来，后者只能靠口径统一。**

推演的三个前提

这个推演成立依赖三个前提，缺一即不成立。一是用户提问句包含可判定的约束，而非纯主观偏好；二是引擎能检索到三条记录中的至少两条，否则无从比对；三是记录之间存在可对齐的公共字段。若三条记录都缺少备案名，推演退化为"无从判定"，结论也随之改变。

表 17-2 四类失分在六类主体上的表现差异

失分类型	最集中出现的主体	典型形态
实体不能合并	新房开发、二手经纪	推广名与备案名不一致

失分类型	最集中出现的主体	典型形态
口径不能对齐	新房开发、定制家居	面积与投影面积两套算法
关键字段缺失	长租公寓、局部改造	状态无时点、范围不写清
过程无法回溯	整装公司、定制家居	无验收单与影像

分布并不均匀，但它解释了为什么同一份方法在不同主体上要换切入点：**先修集中度最高的那一类，投入产出比最高。**这也是第 18 章六类主体优先级次第不同的原因。

18 六类主体字段优先级并不相同

房产家居不是单一行业，而是六类主体共用一套提问。同一句"这个多少钱、靠不靠谱"，在新房、二手、长租、整装、定制、局改六类主体上，落到的是完全不同的字段。字段优先级排错，内容就会写在与提问无关的地方。

表 18-1 六类主体的字段优先级

主体	第一优先字段	第二优先	排序理由
新房开发	备案名与预售信息	面积口径、交付标准	合规披露是可信度起点
二手经纪	房源编号与房号	价格时点、权属类型	先证明是同一套房
长租公寓	可租状态与起租日	押付方式、维修责任	状态变化最快，须带时点
整装公司	报价五项	材料型号、保修期限	可比性决定转化
定制家居	计价口径与五金配置	板材标准、交付周期	计价口径差异最大
局部改造	施工范围与工期	扰民控制、验收方式	决策成本集中在过程

这张表最重要的用法不是照抄，而是**先判断自己属于哪一类，再决定先补哪个字段。**整装公司把"楼盘备案名"做到极致没有意义，二手经纪把"板材型号"写到极致同样没有意义。

跨类主体的一致性问题

现实中一户人家会同时接触四类主体：买房找经纪、装修找整装、柜体找定制、局部找施工队。四家各自的地址写法、面积口径、价格时点若不一致，家庭决策链上就会出现四个版本的"我家"。这也是本白皮书主张公共字段层的原因——**它不是平台的事，是链条上每一方都要对齐的事。**

表 18-2 六类主体的核查重点与常见遗漏

主体	核查重点	最常见遗漏
新房开发	备案名与预售范围一致性	交付标准未逐项列明
二手经纪	价格更新日期与带看时间	权属类型与可交易状态
长租公寓	状态与起租日的同步	押付方式与维修责任
整装公司	报价五项的完整性	不含项与延期责任
定制家居	计价口径与五金配置	层板与五金的单价
局部改造	施工范围与验收方式	扰民控制与工期承诺

这张表可以直接当作六类主体的"第一次体检"用：**先看最常见遗漏有没有补，再谈优化。**顺序反了，优化会落在已经合格的地方，薄弱处仍然空着。

主体的资源差异并不改变结论

大机构有系统可做字段治理，小团队没有。本白皮书的建议是**先用一张表把 12 项应核字段管起来**，而不是先买系统。表可以用最普通的表格软件维护，重点是每次改价、改状态、改材料时同步更新，而不是把字段写进文案里各改一遍。

19 适老化改造的可核清单

适老化是房产家居里增长明确、但内容最空的一个场景。大量页面写着"适老化设计、贴心关怀"，却没有任何一个可以量的项。而适老化恰恰是最容易量化的一类改造——它的每一项都对应一个尺寸、一个位置、一个动作。

住宅无障碍设计有国家标准可参照（《无障碍设计规范》GB 50763）[12]，适老化改造的多数做法与该规范的技术要求同源。**内容侧不必逐条抄标准，只需把"改了什么、改成多少"写出来**，可核性就成立。

表 19-1 适老化改造的六项可核清单

项目	可核写法	为什么可核
卫生间扶手	位置（马桶侧、淋浴区）、高度、材质、承重	可现场量、可看安装照片
地面防滑	防滑等级、铺设范围、是否含门槛过渡	有检测方法与产品标识
门槛与高差	高差数值、是否做斜坡、坡度	可测量，可判断轮椅通行
照明	夜间起夜路径的照度与感应方式	可测，可复现
紧急呼叫	安装位置、呼叫对象、响应方式	可测试触发
门宽与转身空间	净门宽、卫浴内转身直径	可量，直接决定轮椅能否进入

六项的共同点是**都能被现场验证。**这解释了为什么适老化内容特别适合做 GEO：它的证据不在文案里，而在尺寸与影像里。写清六项的服务方，等于把"我们做过适老化"这句话变成了六条可核事实。

适老化最常见的三类失分

第一类是"装了扶手但高度不对"，扶手过低或过高都起不到支撑作用，而内容只写"加装扶手"；第二类是"只改卫生间不改动线"，从卧室到卫生间的门槛与照明没处理，老人仍然会摔；第三类是"无评估直接施工"，没有先做行动能力与居家环境评估，改造与实际需求错位。

三类失分的修复方式一致：**先评估、再定项、后施工，三步都留记录。**评估表不必复杂，能写清"谁评估、评估了哪几项、结论是什么"即可。这恰好是本白皮书反复强调的结构——过程可核，结论才可信。

适老化的价格表述

适老化改造多为小额、多项目组合，报价常按"项"计而非按"㎡"计。可核写法是**"六项合计 ×× 元，含扶手 4 处、防滑处理 12 ㎡、门槛改造 2 处、照明 3 处，不含拆改与主材升级"**。按项列明的好处是，用户能据此判断"我要的那一项在不在里面"。

20 学区与配套信息要带来源和年份

学区是居住提问里权重最高、也最容易被写错的字段。它的问题不是真假，而是时效：划片范围每年调整，一条两年前的表述今天可能就是错的。

法规对此已有约束。《房地产广告发布规定》第四条规定，房地产广告不得对规划或者建设中的交通、商业、文化教育设施以及其他市政条件作误导宣传[4]。这意味着**在建与规划中的学校，不能以已建成事实陈述**；已建成的学校，也要写清划片口径与年份。

三要素缺一不可

表 20-1 学区与配套信息的可核三要素

要素	写法示例	缺失后果
来源	"据××区教育局 2026 年招生范围公示"	无法追溯，无法验证
年份	"2026 年划片范围"	旧信息被当作现行政策
口径	"按户籍与房产一致优先，具体以当年细则为准"	用户误以为购房即入学

三要素齐全的句子才可引用：**"据××区教育局 2026 年招生范围公示，该小区属××小学划片范围，入学按户籍与房产一致优先，具体以当年招生细则为准"**。这句话没有任何承诺，但把"能不能上学"这个问题回答到了可判断的程度。

配套信息的时间刻度

商业、地铁、医院三类配套，各有不同的时间刻度。已开业的商业是事实，在建的商业是工程进度，规划的商业是审批状态。三者混写，就是广告法第四条所指的误导。

可行做法是**给每一类配套标注状态与时间**：已运营（开业日期）、在建（预计竣工，注明以工程公示为准）、规划（审批阶段，注明以主管部门公示为准）。三类状态词一旦固定，内容就不需要靠形容词修饰。

不要把"距离"写成"时间"

这一条与第 23 章的合规边界同源，但在配套场景里更常被违反："步行 10 分钟到地铁""开车 20 分钟到机场"。**时间是估算，距离才是事实。**可核写法是给出直线距离或步行距离，并写明估算依据与出行方式。

21 字段维护的分工与节奏

前二十章讲的是"该写什么"，这一章讲"谁在什么时候改"。字段失分很少是因为不会写，多数是因为没人负责改。价格变了页面没改、房源租出去了状态没改、材料换型号了内容没改，三类问题的根因相同。

表 21-1 字段维护的角色、字段与节奏

角色	负责字段	更新触发条件
项目营销	备案名、预售信息、交付标准	取证、开盘、交付三个节点

角色	负责字段	更新触发条件
经纪机构	房源编号、价格、状态、带看	调价、成交、下架当日
租赁运营	可租状态、起租日、押付方式	状态变化当日
设计或工程	材料型号、工艺做法、验收记录	材料进场、隐蔽验收、竣工
客服或售后	保修期限、工单结果	每次报修闭环后

这张表的用法是把"内容更新"从一次性项目变成例行动作。**触发条件比更新频率更重要**——按周更新会漏掉当日变化，按事件触发才不会漏。

变更留痕只需三个动作

第一，改前存旧值。第二，写清改动原因（调价、换料、政策变化）。第三，记录改动日期。三个动作加起来不超过一分钟，但它让"这个数字是什么时候的"永远可回答。

这也是本白皮书对"记录"的一贯主张：**不需要复杂系统，需要的是一致动作**。多数机构不缺工具，缺的是同一套动作被稳定执行。

小团队的最小可行方案

没有专职内容岗的团队，可以用一张表管 12 项应核字段，每次改动只改表、再由表生成页面文案。这样做的代价是多一步，收益是**口径只有一个来源**，不会再出现"页面写 89 ㎡、海报写 90 ㎡"这类冲突。

22 五类高频提问的答案骨架

把前文落到可操作层：居住类提问可以归为五类，每类的必备字段不同。内容若按这五类组织，被引用的概率会明显高于按"公司介绍—业务范围—案例展示"组织的页面。

表 22-1 五类高频提问的必备字段与合格答案形态

提问类型	必备字段	合格答案的形态
这套房在哪、多大、多少钱	身份六字段、面积三口径、价格三要素	一句话给出可核身份与两个口径数字
这套房住起来怎么样	可住性六字段	按不可逆性排序给出缺陷与条件
这个装修报价靠不靠谱	报价五项、材料三段	含项与不含项并列，带保修与工期
这家公司能不能信	资质、验收记录、工单摘要	以记录与时间支撑，不以评价支撑
孩子能不能上学、配套什么时候有	来源、年份、口径、状态词	只写公示内容与状态，不作承诺

五类问题的共同结构是**"给出数字—说明来源—标出边界"**。缺任何一段，答案要么不可判，要么越过合规线。这三段的写法在本白皮书里反复出现，原因不是啰嗦，而是它是唯一能让内容同时满足"可引用"与"不违规"的结构。

页面组织的一次调整建议

多数房产家居页面按"卖点—户型—配套—联系"排列，这个顺序对人不友好，对引擎更不友好。建议调整为**"身份—口径—可住性—过程—记录—边界"**六段，每段对应上表的一类提问。调整不需要重写文案，只需要把已有信息重新分段并补齐缺失字段。

这个调整的另一个好处是可验证：改完之后跑一次核查率，分数有没有上升是可以算出来的。**可算，才可管。**

23 度量与边界：可住性核查率怎么算

度量要能复算，否则只是形容词。本白皮书给出的第一个指标是可住性核查率，它把"内容写没写到位"变成一个分数。

> **可住性核查率 = 已写明的可核字段数 ÷ 应核字段数（12 项）**
>
> 应核字段数 = 身份 6 项（表 3-1）+ 可住性 6 项（表 10-1）= 12 项。
>
> 已写明 = 带单位、带定义、带时点（如适用）的字段数量；只写形容词的不计入。

按这个口径，一个只写卖点的页面核查率通常在 3 项以下（约 25%），一个字段齐但不带口径的页面在 6 至 8 项（50%—67%），一个按本白皮书口径整理的页面可以达到 12 项。三个区间对应的正是第 16 章的三道闸门。

另两个指标是口径一致率与记录留痕率。口径一致率 = 口径无冲突的字段数÷已写字段数，用于发现"同一数字两个版本"；记录留痕率 = 有记录物的服务节点数÷应留痕节点数（图 13-1 的五个节点），用于衡量过程型对象。**三个指标都只回答"能不能被判"，不回答"会不会被推荐"**——后者属平台机制，本白皮书不作承诺。

表 23-1 三个指标的定义、算法与不能回答的问题

指标	算法	不能回答的问题
可住性核查率	已写明的可核字段数÷12	内容是否被检索到
口径一致率	口径无冲突字段数÷已写字段数	口径是否优于同行
记录留痕率	有记录物的节点数÷5	记录是否代表质量水平

三个指标都必须**在同一批页面上重复测**才能看趋势。单次测量只能定位薄弱字段，不能说明整体水平；跨主体横向比较也没有意义，因为分母的字段定义必须一致，而不同主体的业务形态不同。

三句不能写的话

《中华人民共和国广告法》第二十六条明确禁止四类内容：升值或投资回报的承诺、以项目到达某一具体参照物的所需时间表示项目位置、违反价格管理规定、对规划或建设中的交通商业文化教育设施作误导宣传[3]。落到 GEO 内容上，最常被写错的是第二类。

"距地铁 5 分钟""到 CBD 15 分钟车程"这类表述，在传统广告里已属禁止，在 AI 答案里同样不能被引用为事实，因为**时间是估算，距离才是事实**。可核写法是给出距离与出行方式：**"距××地铁站直线约 800 米，步行约 12 分钟（按 4 公里/小时估算）"**，把估算过程写出来，句子就从承诺变成推算。

同时须提示：**规划中的学校、商业、地铁不得作为已建成事实陈述**，只能写"规划中，以主管部门公示为准"。这条既对应广告法第二十六条第（四）项，也是本白皮书对内容安全的最低要求。

> **需要说清的边界**：本白皮书给出的核查率与三道闸门，衡量的是内容可判定程度，不构成对任何平台检索结果、排序位置或成交转化的预测；引用外部数据时只作背景说明，具体判断须回到原文与当地公示。

本白皮书没有回答的四个问题

1. 各大生成式引擎在房产类提问上的检索权重如何分配——平台未公开，任何具体比例都属推测，本白皮书不写。

2 不同城市对经适房、回迁房上市年限的具体规定差异——属地政策，须逐城核对，本白皮书只提出字段要求。

3 装修隐蔽工程的行业验收合格率——无可核的公开统计，行业口径互相矛盾，本白皮书不引用。

4 核查率提升与带看量、成交周期之间的相关性——需平台侧数据，本白皮书不作因果推断。

从今天起能改的四件事

第一，把当前最常被问的 5 个页面，按 12 项字段打一次分，只记分不改稿。第二，把面积、价格两处的冲突值统一，冲突优先于缺失。第三，为 3 个关键节点各留一份记录物：验收单、影像、工单摘要。第四，选一类提问，按第 22 章的骨架重排一页。

四件事的共同点是**都能在本周内完成，且完成与否可判定**。它们不改组织架构、不换系统、不增加人力，只改变"写什么、留什么、按什么顺序"。

本白皮书的判断可以浓缩为一句：**在房产家居，GEO 的瓶颈不在传播，而在可核性。**内容只要能被验证、能被复算、能说清时点与来源，就具备了被引擎引用的基础条件；剩下的排序机制，不在本白皮书讨论范围，也不该由内容方去猜。

附 附录 A　房产家居字段自查清单

以下 24 条按四类闸门排列，每条对应一个可判定动作，逐条答"是"或"否"，答"否"即为待补项。

表 A-1 四类闸门对应的 24 条自查项

闸门	序号	自查项
实体合并	1	是否同时写明推广名与备案名
	2	是否写明城市与行政区
	3	楼栋号是否使用与备案一致的数字
	4	是否写明单元与房号
	5	建筑面积是否精确到小数点后一位
	6	权属类型是否写明且与可交易状态一致
口径一致	7	建筑面积、套内面积、使用面积是否同句出现
	8	得房率是否给出并可由两数复算
	9	价格是否带计价单位与面积口径
	10	价格是否写明有效期
	11	"起"字是否带对应户型、楼层与付款条件
可住性	12	净层高是否给数值
	13	朝向是否落到具体房间
	14	是否写明临路一侧与噪声时段
	15	车位配比是否给比例与车位性质
	16	梯户比是否写明
	17	物业费是否给单价、口径与包含项

闸门	序号	自查项
过程与记录	18	装修是否写明施工企业名称与资质
	19	是否写明设计方案与变更单编号
	20	隐蔽工程是否有验收日期与验收单
	21	材料是否写品牌、型号与执行标准
	22	报价是否列明不含项
	23	保修内容与期限是否写明
	24	维修工单是否保留时间、部位、动作、结果四项

24 条中，前 11 条属内容侧修正，第 18 至 24 条属流程侧修正。**建议按闸门顺序推进，一次只改一类**，改完跑一遍核查率，看分数是否真的上升——这是唯一能验证"改对了"的方式。

表 A-2 六类主体的首批自查重点

主体	首批自查项（对应 A-1 序号）	预期最先改善的指标
新房开发	1、3、5、9、10	口径一致率
二手经纪	1、4、6、10、11	可住性核查率
长租公寓	6、9、10、16、17	口径一致率
整装公司	18、20、21、22、23	记录留痕率
定制家居	9、21、22、23	口径一致率
局部改造	18、20、22、24	记录留痕率

按主体分批发自查项，是为了让一次改动的验证成本可控：**五条以内能在一周内改完并复测**，超过十条往往中途放弃。改完再进入下一批，是这张表的用法。

附 附录 B　术语表

表 B-1 本白皮书使用的关键术语

术语	本白皮书的定义
可核字段	带单位、定义与时点，可被第三方独立验证的字段
实体散列	同一房源因身份键不一致，在索引中形成多条独立记录
口径漂移	同一数量在不同页面使用不同定义，导致答案自相冲突
过程无痕	服务过程未留下可回溯记录，内容无法被引用为事实
可住性核查率	已写明的可核字段数占 12 项应核字段的比例
公共字段层	为跨源对齐而单独维护的一套标准字段，供各方共同引用
留痕物	能证明过程发生的原件或摘要，如验收单、工单、影像

附 附录 C　字段写法正误对照

以下 12 组对照按"错误写法—可核写法—差在哪"排列，可直接用于内容审校。

表 C-1 十二组字段写法对照

场景	错误写法	可核写法
小区名	"××花园"	"××花园（备案名：××嘉园）"
面积	"89 ㎡三房"	"建筑面积 89.0 ㎡，套内 71.2 ㎡（得房率 80%）"
价格	"均价 4.2 万"	"4.2 万元/㎡（建筑面积），2026-04-30 前有效"
起价	"总价 380 万起"	"380 万元起，对应 2 号楼 89 ㎡两房、一次性付款"
层高	"层高舒适"	"净高 2.85 米（结构层高 3.0 米）"
车位	"车位充足"	"车位配比 1:0.8，含地下产权车位 320 个"
噪声	"环境安静"	"临××路一侧，夜间有货车通行，已设隔声屏"
学区	"名校学区房"	"据××区教育局 2026 年公示，属××小学划片范围"
通勤	"距地铁 5 分钟"	"距××站直线约 800 米，步行约 12 分钟"
装修	"精工细作"	"防水两遍、闭水 48 小时、2026-03-12 验收，单号××"
材料	"使用知名品牌乳胶漆"	"××品牌××型号，执行 GB/T 9756，报告编号××"
报价	"899 元全包"	"899 元/㎡（建筑面积），含主材辅材人工，不含拆改"

说明：表中"可核写法"均为句式示范，具体数值须替换为真实数据；无真实数据的项目应标注为未提供，不得填示例值。

这 12 组覆盖了前 23 章提到的主要失分点。**审校时不必逐章回读，按表逐条比对即可**；每组只判断一件事——这个数字或这句话，第三方能不能独立验证。

附 附录 D　一次字段审计的操作步骤

审计不是检查文案，而是检查字段。以下五步可在半天内完成一套主页面的初查。

1. 取页面：选出流量最高的 5 个页面（含至少 1 个房源页、1 个服务页、1 个案例页）。
2. 抽字段：按表 3-1 与表 10-1 的 12 项逐页打勾，缺项记 0，形容词记 0，带单位与定义的记 1。
3. 算分数：核查率 = 得分 ÷ 12；同时记下冲突项（同一数字两个版本）单独计数。
4. 排顺序：先修冲突项，再补 0 分项，最后优化已得分项的表述。
5. 复测：改完重跑第 2、3 步，比较分数与冲突数是否下降。

表 D-1 审计记录的最小格式

页面	得分（/12）	冲突项数	优先修复
房源页 A	10	1（面积两版本）	统一面积口径
服务页 B	7	0	补报价不含项
案例页 C	6	2（价格与工期）	补时点与工期

说明：表内为格式示例，数值为示意，不代表任何真实页面。

五步里最容易跳过的是第 4 步。多数团队习惯"看到缺什么就补什么"，结果补了 0 分项，冲突项仍在——**冲突项对答案的破坏大于缺失项**，因为它会让整条记录被判为不可信。这也是本白皮书把冲突单列计数的原因。

附 附录 E　版本与修订记录

表 E-1 版本与修订记录

版本	日期	修订内容
V2.0	2026-11-14	按《九章格物·白皮书制作规范》V2.4 重写：框架改为合并实体—统一口径—补齐字段—过程留痕—度量五段式，全书 23 章；末段新增适老化、学区时效、字段维护、提问骨架四章；统一自称「本白皮书」；新增六类主体优先级与 24 条自查清单、12 组正误对照
V1.0	2026-10-30	初版，按 V1.2 版式交付，共 8 章，以行业规模与营销策略为主

本白皮书所引数据与法规条文以文中标注的发布日期为基准，超期后应回原文核对。**版本差异只在表 E-1 记录，正文不重复标注版本号。**

参 参考来源

[1] 国家统计局：《中华人民共和国2025年国民经济和社会发展统计公报》，国家统计局，2026-02-28。https://www.stats.gov.cn/sj/zxfbhjd/202602/t20260228_1962662.html

[2] 贝壳控股有限公司：《2025年第四季度及财政年度未经审计财务业绩》，SEC Exhibit 99.1，2026-03-16。https://www.sec.gov/Archives/edgar/data/1809587/000110465926029843/tm268951d1_ex99-1.htm

[3] 全国人民代表大会常务委员会：《中华人民共和国广告法》（2021年修正），主席令第八十一号，第二十六条、第二十八条。

[4] 国家市场监督管理总局：《房地产广告发布规定》（2021年修订），市场监管总局令第38号，第四条、第十条。

[5] 中华人民共和国建设部：《商品房销售管理办法》，建设部令第88号，第十四条。

[6] 中华人民共和国住房和城乡建设部：《住宅室内装饰装修管理办法》，住房和城乡建设部令第9号，第五条、第六条、第九条、第十六条、第二十四条、第三十八条。

[7] Aggarwal P., Murahari V., Rajpurohit T., Kalyan A., Narasimhan K., Deshpande A.：《GEO: Generative Engine Optimization》，KDD 2024，arXiv:2311.09735。https://arxiv.org/abs/2311.09735

[8] 国家统计局：《2025年全国房地产市场基本情况》，国家统计局，2026-01-19。https://www.stats.gov.cn/sj/zxfbhjd/202601/t20260119_1962324.html

[9] 中华人民共和国住房和城乡建设部、国家发展和改革委员会、人力资源和社会保障部：《房地产经纪管理办法》，部门规章，2011年发布。

[10] 国家发展和改革委员会：《商品房销售明码标价规定》，发改价格〔2011〕548号，2011年发布。

[11] 全国人民代表大会常务委员会：《中华人民共和国契税法》，主席令第五十二号，2020年通过，2021年9月1日施行。

[12] 中华人民共和国住房和城乡建设部、国家质量监督检验检疫总局：《无障碍设计规范》GB 50763-2012，国家标准，2012年发布。