产品起源 • 功能拆解 • 工程方案

错题本为什么存在：

一个被「抄题」撑起来的品类

需求来自哪 · 有哪些功能 · 用什么技术实现

基于 220 篇小红书笔记、554 条评论、1,097 条竞品差评

回答的问题　为什么会出现这个产品？需求来自哪里？功能怎么拆？技术上怎么做？

数据基础　小红书 220 篇 / 554 条评论；App Store 6 个竞品 / 1,097 条评论（1–2 星 146 条）

写作口径　事实标注来源；工程参数属估算，已明确标「估算」；推断已标「推断」

一、为什么会出现这个产品

一句话：因为「错题要重做一遍」这件对的事，被「把题抄下来」这件苦的活挡在了门口。

错题学习的闭环只有三步：发现错题 → 重做 → 直到会。三步里只有第一步需要工具，第二步需要时间，第三步需要诚实。纸质错题本把三步全部压给学生，而其中最耗时的「抄题」，几乎不产生任何学习价值——一道带图形的大题，抄题干、画图、抄公式，通常要 3–5 分钟（估算）；一张卷子十道错题，就是半小时。这半小时不产生理解，只产生「我有在努力」的感觉。产品化的缝隙就在这里。

1.1 三次形态迁移：每次都解决了上一个瓶颈，也带来新的瓶颈

阶段	形态	解决了什么	新的瓶颈	大致时间
一	纸质手抄 本子 + 剪刀 + 胶水	把分散的错题集中到一处	抄题耗时；图形、公式难复制； 无法搜索	长期存在
二	拍照打印 错题打印机 / 试卷宝	不用抄，拍完直接印	只能出纸：不讲解、不归类、 不排期；耗材另收费	2016–2022
三	识别 + 讲解 + 调度 App / 网页 / 小程序	讲解、知识点归类、按遗忘曲线排复习	采集在一端、使用在另一端； 提醒做不了	2023 至今

值得注意：第二次迁移解决的是「手」，第三次迁移想解决的是「脑」。但数据说明，第三次迁移里用户抱怨最多的仍然是「手」的问题——打印、导出、题图分离。也就是说，**这一代产品还没有真正把手的问题解决完，就急着去解决脑的问题。**

1.2 为什么是现在：三个前置条件同时到位

- **条件一：**手机摄像头 + 试卷切题能力可用。整卷拍下来能自动切成一道一道题，这件事 2019 年前后在工程上才稳定。
- **条件二：**大模型把「讲一道题」的边际成本压到「分」级。同样一道题，人工讲 5 分钟，模型讲 3 秒、成本不到 3 分钱（估算）。
- **条件三：**家长端已经为「学习硬件/内容」付过钱。学习机、教辅、错题打印机把付费习惯教育完了，198 元/年不再需要解释。

1.3 一个反直觉的判断：用户买的不是「整理」，是「尽头」

554 条评论里点赞最高的四条，没有一条在讨论用什么软件，全部在讨论「怎么整理」「整理到什么程度」——哪科用剪贴、哪科用荧光笔、什么时候整理、整理完怎么复习。这说明用户把这个问题理解成一套动作，而不是一个 App。

> 「我一般就是每天把错题贴在本子上，但是不做。考试前，或者是一周一个周期做一遍没做的错题。」
>
> — 小红书评论，60 赞

「贴在本子上，但是不做」——这句话是这个品类的病根，也是产品的定义。用户真正缺的不是一个更大的容器，而是一个**明确的完成状态**：什么时候做、做哪几道、做到什么程度算过。标杆产品把「连对两次才算过关」写进卖点，本质上卖的就是这个「尽头」。

> **所以这个产品到底是什么** 它不是「错题管理器」，是「错题的重做日程表 + 验收官」。收集只是入场券，排期和验收才是它存在的理由。一旦产品只做收集，就会退化成「又一个网盘」，用户用三天就忘。

二、需求来自哪

把这个品类的需求拆开，是四股力量叠在一起：考试密度提供「量」，三方角色提供「钱」，家长的可控感提供「急」，替代品价差提供「划算」。四股力量合起来，才推得动一次购买。



终点不是「产品不好用」，是「产品没进入日常」——真正要解决的是第 4 到第 5 环之间那一段。

图 1　需求传导链：从考试密度到「闲置」，再回到下一次购买

这张图最关键的不是前面的链条，而是最后那个回环：**用户不是不买，是买完用三天就停。**所以问题的位置很清楚——不在「买」，在「第 4 环到第 5 环之间」。

2.1 制度源：考试密度决定错题密度

错题不是自然产生的，是考试产生的。校内周测、月考、单元测，一科一周一套卷，一套卷 20 题、正确率七成，就是 6 道错题；五科就是 30 道（估算）。这个量级决定了两件事：一是「手抄」必然不可持续，二是「按周排期」是天然合适的节奏——这也解释了为什么遗忘曲线的第一档间隔通常就是 1 天、3 天、7 天。

错题之所以值得单独做成产品，是因为它是**信号最强、噪声最低**的学习数据：它直接指向知识缺口，还自带学生的真实作答过程。相比之下，作业对错、上课听讲都是弱信号。

2.2 角色源：三方错位——老师下命令、孩子执行、家长付钱

这是所有教育产品的结构性矛盾，在错题本上尤其尖锐：老师要的是「整齐、可检查」，孩子要的是「快点弄完」，家长要的是「提分」。三个目标指向三种产品：档案柜、加速器、教练。

数据给出的答案是：**付钱的是家长，但每天打开的是孩子。**所以产品必须同时满足两种体验——给家长看「进度和薄弱点」，给孩子用「三步内完成一次重做」。任何一边缺失，都会在第 4 环断掉：只有家长视角，孩子嫌烦；只有孩子视角，家长觉得没效果、不续费。

2.3 情绪源：家长的可控感

家长能介入的学习环节其实极少：上课看不见、作业看不懂、考试插不上手。错题本是少数「能看见、能数、能打勾」的环节——所以家长买的常常不是工具，是「我在参与」的感觉。这解释了为什么评论区里最积极的是妈妈，也解释了为什么「家长报告」「周报」这类功能虽然对孩子无用，却能显著影响续费。

2.4 商业源：替代品价差

替代方案	价格带	它解决了什么	它没解决什么
纸质错题本 + 剪刀胶水	几元	集中	耗时、难复制图形、不可检索
错题打印机（硬件）	300–800 元 + 耗材	不用抄，直接出纸	不讲解、不归类、不排期
线下整理/托管服务	按小时计费	有人代做	贵、不可持续、质量不稳定
教辅书	30–60 元/本	有题可刷	不是自己的错题
错题软件（本类）	99–198 元/年	识别、讲解、排期、打印	采集麻烦、提醒缺失（当前普遍）

198 元/年落在这个区间的下沿：比一台打印机便宜，比一学期托管便宜得多。所以「贵不贵」从来不是绝对值问题，而是「和什么比」的问题——这也是为什么差评里骂的不是「198 太贵」，而是「为什么每次点开都让我再充一层」。

2.5 四股力量对产品的硬要求

力量	最强证据	对产品的硬要求
考试密度	「什么时间整理呀……课间根本不够用」（161 赞）	单次操作必须在 3 分钟内完成
三方角色	评论区大量「老师要求统一交错题本」	同时有孩子端（快）和家长端（可看进度）

力量	最强证据	对产品的硬要求
可控感	家长高频追问「怎么用」「能不能打印」	可量化进度 + 可打印的实物产出
替代品价差	差评里收费类占 32.9%，单项第一	一次说清价格，不设二次收费

三、有哪些功能

把竞品的功能拆开，其实是同一条五段链路。理解这条链路，比背功能清单有用——因为链路的断点位置，决定了产品能不能活。



图 2　错题产品的五段链路：任何一段落在「另一个设备」上，用户就会掉队

3.1 十个功能，分别挂在哪一段

段	功能	谁最需要	现在谁做好了	差异机会
采集	整卷拍照 + 自动切题	孩子（要快）	试卷宝、蜜蜂较好	多页连拍、错题勾选一步完成
采集	题干与图形绑定	所有理科用户	普遍差（差评明确骂）	同页多图同步提取为一道题
加工	错因归类（概念/计算/审题/方法）	家长（要诊断）	少数产品有，浅	归类可解释、可统计
加工	AI 讲解	孩子	豆包等免费产品已做到	不做卖点，只做兜底
加工	知识点打标 / 薄弱点图谱	家长	橙果（老师端）较好	对齐教材目录，能说人话
调度	遗忘曲线排期	孩子（要少而准）	忆题等少数	每天 3–5 道，不贪多
调度	到期提醒	家长	App 可做，纯网页做不了	订阅消息 + 微信服务号

段	功能	谁最需要	现在谁做好了	差异机会
验收	变式题 / 举一反三	家长（要检验）	普遍质量差（差评 5 条）	同知识点、同难度、可判分
验收	连对两次才算过关	家长	标杆产品有	状态机化，可见进度
输出	A4 排版 + PDF 打印	家长（要纸）	普遍差（差评 27.4%）	答题区留白、题图不分离、不限题数

3.2 功能分三层：必须有、差异化、可以不做

分层依据不是「重要不重要」，而是「缺了会不会死」。判据来自 146 条 1–2 星差评的结构：导出/还原 17.1% + 打印/排版 10.3% = 27.4%，说明纸面链路缺失会直接导致差评；而「举一反三」只有 5 条（3.4%），说明变式题目前做得太差，用户还没形成期待。

层	功能	判断依据
必须有	拍照采集、自动切题、错题入库、按曲线排期、到期提醒、PDF 导出打印	缺任一项，产品退化成网盘或相册；打印缺失已被差评验证
差异化	题图绑定排版、答题区留白、连对两次验收、家长周报	竞品普遍没做好，且直接对应差评原文；做了就是口碑点
可以不做	AI 视频讲解、社区/排行榜、题库商城、错题社交	免费产品已覆盖；与「每天 3 分钟」的核心场景冲突；成本高、留存低

> **一个反常识的取舍**　「AI 讲解」不该是卖点。豆包爱学免费、47.9 万评分、4.55 分，评论区已经在说「这个 ai 不就能做了」。把讲解当卖点，等于和免费产品正面竞争；把「排期 + 验收 + 打印」当卖点，才是别人没做的。

四、如何实现

实现这件事的难点不在算法，而在**分工**：哪个动作发生在哪个设备上。下面按「端 → 云 → 纸」的顺序给方案，关键参数标了估算口径。

4.1 总体架构与技术选型

层	选型	为什么
客户端	微信小程序（主）+ H5 兜底	免安装、能发订阅消息；个人主体无教育类目，需企业资质
接入层	FastAPI / Node BFF	上传、鉴权、限流；与异步任务解耦
异步任务	Redis 队列 + Worker	OCR/切题/讲解耗时 3–20 秒，必须异步，前端轮询或推送
OCR / 切题	版面分析模型（PP-Structure 类）+ 题目检测框	整卷切题的核心；商用 OCR 可作为冷启动
公式识别	多模态大模型直接转 LaTeX	比专用公式 OCR 更省事，准确率已够用（估算）
讲解	两个不同厂商的模型交叉生成	单模型会自相矛盾，交叉可拦截大部分错解
存储	PostgreSQL + 对象存储	结构化关系 + 原图/切片
调度	定时任务（每日 06:00）+ 推送	按「今日到期」生成任务，一次算好，不做实时计算
导出	服务端 HTML → PDF（WeasyPrint / 无头浏览器）	排版规则用 CSS 控制，改版不改代码

4.2 采集与切题：把「抄题」这一步彻底删掉

流程：拍照（可多页）→ 透视校正 → 版面分析（文本/图/表/公式分块）→ 题目检测 → 题干与图形绑定 → 家长勾选 → 入库。其中「题干与图形绑定」是最容易被跳过、也最容易挨骂的一步。

> 「几乎每张卷子都会有几道题的图集合在同一行里，出现题图分离的情况……没法对同一页上分在两个地方的题和图同步提取为一道题。」
>
> — 试卷宝用户，1 星

这条差评给出了明确的工程要求：**同一页上分离的题与图，要能合并成一道题。**实现上不是纯视觉问题，而是「检测框 + 空间关系 + 语义归属」三步：先检测所有块，再用纵向距离与编号（如「图 3」）建立候选关系，最后让模型判断归属。做不好这一步，后面所有功能都白搭。

另外两个工程细节：一是**去重**（同一道题被重复拍），用题干文本指纹 + 图形感知哈希；二是**断点续传**，一次传 5 页在弱网下必然失败，必须支持分片。

4.3 讲解与判分：宁可说「不确定」，也不要给错答案

家长对 AI 讲解的容忍度是零：错一次，永久流失。所以工程上要做三件事——

- ① **交叉验证：**同一道题让两个不同厂商的模型各解一次，答案不一致就不展示「标准答案」，改展示「参考解法（待确认）」。
- ② **机械校验：**数值题把答案回代原式验算；选择题校验选项分布；几何题校验角度/长度是否自洽。
- ③ **置信度门槛：**识别置信度低、题干残缺、图文不匹配时，直接标记「这题没拍清」，让用户重拍，而不是硬讲。

> 「首先极速解题和深度解题给我两个完全不同的答案。」
>
> — 豆包爱学用户，1 星

这条差评是整份数据里对「AI 讲解」最致命的一击：不是讲错，是**自相矛盾。**用户一旦发现同一道题能出两个答案，就再也不敢让孩子跟着学。

4.4 数据模型：把「题目资产」和「错题关系」分开

表	关键字段	说明
students	id, grade, subject_ids, guardian_consent_at	14 岁以下必须记录监护人同意时间
papers	id, student_id, images[], source, parsed_at	一次上传对应一张试卷
questions	id, stem_latex, figures[], answer, kp_ids[], fingerprint	题目资产，跨学生复用（去重省成本）
mistakes	id, student_id, question_id, reason, created_at, state	学生与题目的错题关系

表	关键字段	说明
attempts	id, mistake_id, answer, correct, elapsed_ms, at	每次重做，验收的唯一依据
review_tasks	id, mistake_id, due_at, interval, ease, streak	调度状态，见 4.5
exports	id, student_id, question_ids[], pdf_url, pages	打印产物，缓存复用

把 questions 独立出来的好处有两个：同一道题被多个学生拍到只解析一次（成本直接摊薄），以及可以做「这道题全网错率」这类后续数据产品。

4.5 调度：SM-2 简化版 + 「连对两次」状态机

不需要上复杂的间隔重复算法。教育场景里真正有效的只有两点：间隔随熟练度放大，以及一个明确的「过关」条件。

```
# 每次重做后调用；state: new -> learning -> passed
def on_attempt(task, correct):
    task.streak = task.streak + 1 if correct else 0
    if task.streak >= 2:
        task.state = "passed"
        task.due_at = None          # 不再排期，只在考前「回忆测」里抽查
        return
    task.state = "learning"
    task.ease = clamp(task.ease + (0.1 if correct else -0.2), 1.3, 2.8)
    if task.streak == 0:
        task.interval = 1           # 错了，明天再来
    else:
        task.interval = round(task.interval * task.ease)
    task.due_at = today + task.interval
```

每日任务生成规则：取 due_at <= 今天 的题，按「逾期天数 × 历史错误次数」排序，**每天最多 5 道**。上限是产品规则不是技术限制——超过 8 道，用户会直接放弃整件事（这与「课间十分钟」的现实约束一致）。

4.6 打印链路：纸是终点，不是备份

A4 排版有四条硬规则，每一条都对应一条真实差评：

- ① **题图同页**：题干与图形必须在同一页，禁止跨页断开。对应差评「题图分离」「导出后题目乱排」。
- ② **答题区留白**：填空题预留横线区，解答题预留约 1/3 页空白。对应差评「不给填空题留答题区域」。
- ③ **不限量**：导出不设题数上限，按需分册。对应差评「500 题最多只能下载 100 题」。
- ④ **成组输出**：同一知识点/同一错因的题尽量排在一起，方便「专题突破」。

实现上建议「HTML + CSS 打印样式 → PDF」：排版规则写进 CSS（page-break-inside: avoid、min-height），改版不用改代码；同时保证 PDF 可被家用打印机直接双面打印。

4.7 成本与定价：99 元/学期能不能覆盖

项目	单次成本（估算）	一个学生一月的量	月成本（估算）
整卷 OCR + 切题	0.02–0.05 元/页	约 8 页	0.2–0.4 元
题目讲解	0.01–0.03 元/题	约 30 题	0.3–0.9 元
变式题生成	0.02 元/题	约 10 题	0.2 元
判分（拍照）	0.01 元/次	约 40 次	0.4 元
存储与带宽	—	—	0.2 元
合计	—	—	1.3–2.1 元/月

按 99 元/学期（约 5 个月）算，单生毛利约 90 元，扣掉支付通道与客服仍成立。但有两个前提：**不做无限量的视频讲解**（视频成本高一个量级），以及**题目资产跨学生复用**（否则 OCR 成本会翻几倍）。这也说明：积分制不是成本所迫，是定价策略——把「不确定的成本」转嫁给用户，代价是信任。

4.8 分阶段落地

阶段	范围	验收标准
第 1 期（4 周）	小程序拍照 → 切题 → 勾选 → 入库 → 排期 → 打印 PDF	一整套流程 3 分钟内走完；PDF 可直接打印
第 2 期（4 周）	订阅消息提醒 + 重做拍照判分 + 连对两次状态机	每日推 3–5 道，连续两周不中断
第 3 期	知识点图谱 + 家长周报 + 考前回忆测	家长能说出「孩子哪块最弱」
暂不做	社区、题库商城、AI 视频讲解、多端实时同步	—

五、结论与风险

第一句：这个产品存在的理由只有一个——把「抄题」删掉，把「重做」排进日程。

它不是一个「错题管理器」。收集是入场券，排期和验收才是产品本身。任何只做收集的版本，都会退化成网盘。

第二句：功能重心不在「讲」，在「排」和「验」。

AI 讲解已经免费，而且会自相矛盾；排期（今天该做哪几道）和验收（做到什么程度算过）没有免费替代品，也没有竞品做好。这两个功能加起来，才构成「尽头」。

第三句：实现上最难的不是 OCR，也不是大模型，是让采集发生在手机上、让输出回到纸上。

技术都是现成的：切题有开源模型、讲解有 API、排版有 CSS。真正决定成败的是分工——如果采集要用户「传到电脑」，如果打印要用户「自己想办法」，链路就会在第 4 环断掉。

五个必须提前想清楚的风险

风险	具体表现	应对
未成年人信息合规	14 岁以下需监护人同意；试卷可能含姓名、学校、班级	最小必要采集、默认打码、可一键删除、留存期限明示
AI 错解	一次错答案 = 家长永久流失	双模型交叉、机械校验、低置信度不展示
打印体验	PDF 是口碑入口，也是差评重灾区（27.4%）	把排版当核心功能做，而不是导出附件
渠道依赖	老师推荐能带来爆发，也能因政策一夜清零	同时做家长自然流量，不把增长押在单一渠道
巨头免费	豆包/作业帮把讲解做成免费入口	不打讲解战，打「排期 + 验收 + 纸」的差异

最后一条判断：这个品类的机会不在「做得更聪明」，而在「做得更省事、更有头有尾」。用户已经把答案写在评论区了——他们要的是一套动作，不是一堆功能。

附录：数据与口径

A. 本文引用的数据

数据	数值	来源
小红书笔记 / 评论	220 篇 / 554 条（18 篇有评论明细）	collector/cli.mjs 采集，data/out/xhs.db
App Store 评论	6 个 App / 1,097 条，1–2 星 146 条（13.3%）	collector/appstore.py，Apple 公开 RSS
差评问题分布	收费 32.9%、功能缺失 21.2%、导出 17.1%、识别 15.8%、打印 10.3%	analysis/analyze.py 关键词归类
标杆笔记	564 赞 / 1,010 藏 / 110 条评论	手工整理，用于案例对照

B. 口径说明

- **估算：**工程参数（单次操作 3 分钟、抄题 3–5 分钟、成本 1.3–2.1 元/月）均为估算，用于判断量级，不作报价依据。
- **推断：**「需求来自四股力量」「第 4 到第 5 环是断点」属分析推断，依据是评论与差评的分布，不是统计检验结论。
- **样本：**小红书评论仅覆盖 18 篇高互动笔记，用于判断「讨论强度」，不用于推断「全人群比例」。

C. 复现

```
node collector/cli.mjs run --keywords 错题本,错题打印机,错题整理 --max 80 --notes 25

python collector/appstore.py run --keywords 错题本,试卷宝 --top 3 --pages 5

python analysis/normalize.py && python analysis/analyze.py

python report/make_diagram.py && python report/build_report_v3.py
```