

2018-2026

AI 产品交互设计师 | LLM/Agent 体验设计 | 复杂系统简化专家

Portfolio Resume

BY RENZHUANGZHI



hello

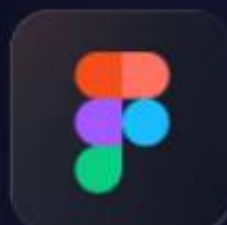
您好，很高兴可以看到我的简历，我拥有

6年 UX 设计经验 | 2年深耕 LLM/Agent 交互

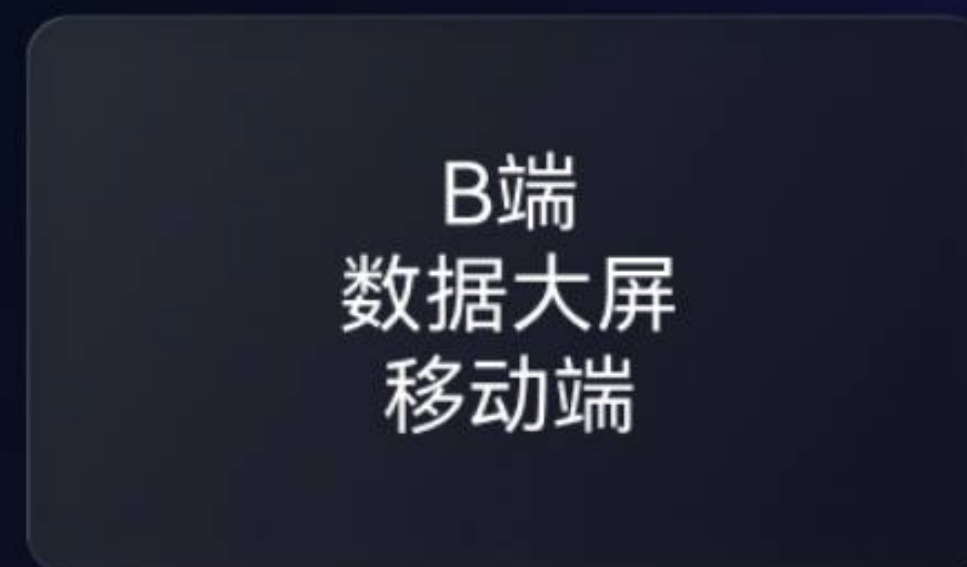
vibe coding深度探索、codex类工具熟练掌握

致力于在不确定性中创造确定性体验，让“用户想要”成为自然而然的发生

Navigation		_hello	_about me	_skills	_products
个人简历			Work Experience		
任壮志 Name 交互设计师（AI方向） Post 工作时间 Post 6年 UX 设计经验 2年深耕 LLM/Agent 交互 学历 Post 全日制本科 设计学/环艺设计 能力 Ability AI 产品交互 LLM/Agent对话设计、多模态交互 AI 反馈机制、置信度体系、人机协作范式 B端系统体验 复杂信息架构、跨角色流程优化 设计规范、组件库体系、设计资产化 用户研究 用户调研、可用性测试、行为数据分析 数据驱动迭代、设计决策 设计工具 在线协同类、氛围编码类			202512-202606 蚂蚁集团（派遣） 资深交互设计师（AI方向） • 主导模型中心 Agent 体验升级，针对复杂 AI 模型调度场景重构人机协作交互范式，优化意图理解、多轮对话及反馈机制，提升 Agent 任务完成率与开发者效率 • 独立负责风险洞察中台全域体验设计，统一机构/用户/平台三端信息架构与交互标准，打通跨角色数据流转闭环，降低协作认知成本 202305-202512 河南景信信息技术有限公司 用户体验设计师（AI方向） • 主导智慧教育 Agent 从 0 到 1 的交互体系搭建，针对模型幻觉、上下文断裂等典型痛点，设计置信度分级可视化、对话关联反馈及输出约束机制，提升模型可信度与用户掌控感，沉淀 AI 人机协作设计范式 • 搭建设计规范与组件体系，制定交互/视觉标准文档，推动设计资产化与团队协同流程标准化，显著降低沟通成本与重复造轮 • 建立数据驱动的体验优化机制，结合用户反馈、行为埋点与可用性测试，持续迭代核心交互流程，关键路径用户满意度与任务完成率有效提升 202003-202304 伊飒尔体验云科技有限公司 用户体验设计师 • 独立负责及深度参与金融、汽车、医药等领域头部产品的全链路体验设计，覆盖 C 端超级 App 与 B 端复杂业务系统 • C 端产品升级：主导中国银行 App、重庆银行 App 5.0 及广汽丰田丰云行大版本迭代，优化产品设计语言与交互框架；基于用户调研与行为数据发掘核心需求，推动体验方案高效落地 • B 端复杂系统：独立负责华润医商润药商城及供应链后台管理系统，覆盖订单、用户、商家、商品、内容、物流六大中心；从 0 到 1 搭建全局组件库与跨业务线设计规范，统一多中心交互逻辑，解决多角色长流程场景下的信息架构与交互复杂性问题		
find me in ： 964580039@qq.com		@ Renzhuangzhi phone ：137 217 18948			

Navigation		_hello	_about me	_skills (核心能力)	_products
战略与思考			逻辑/判断/思维		
用户体验策略 UX Strategy 设计思维 Design Thinking 跨部门协作 Cross-Departmental Collaboration AI 产品思维 AI Product Thinking 用户研究 User Research 产品思维 Product Mindset			人机协作范式 Human-AI Collaboration 多模态交互 Multimodal Interaction AI 安全与信任 AI Trust & Safety 对话交互逻辑 Conversational UX 模型置信度反馈 Confidence Feedback 提示词工程 Prompt		
执行与决策			技能工具		
交互设计 IXD AI 对话设计 AI Conversation Design 多模态界面设计 Multimodal UI Design 设计系统 Design Systems			氛围编码类：  在线协同类： 		
find me in : 964580039@qq.com		@ Renzhuangzhi phone : 137 217 18948			

项目类型



模型中心Agent

Agent交互



文档批量上传



用户旅行图（情绪曲线）

模型中心Agent-交互

用户场景与痛点地图



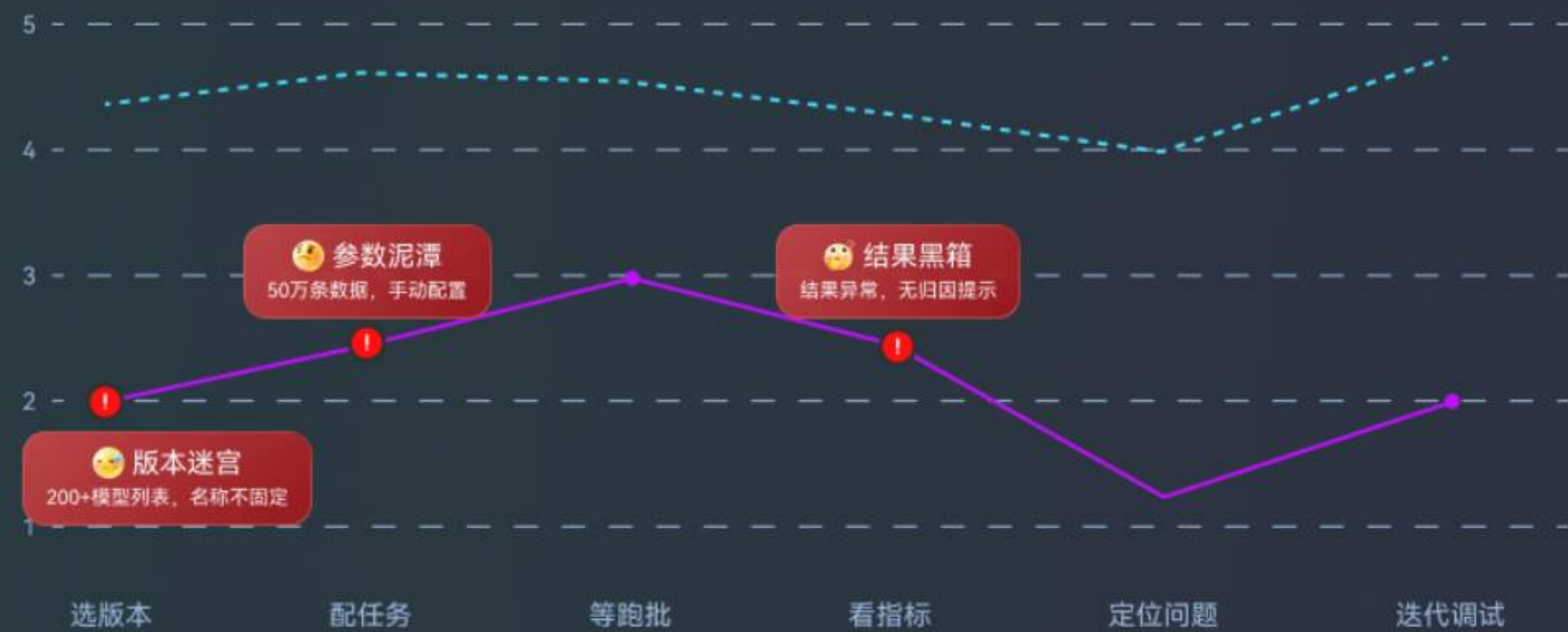
模型同学 | 算法/风控模型工程师

场景：验证本月评级模型50万条数据跨版本对比

工作流：上传测试数据 → 选择基准模型 → 配置数据跑批 → 对比指标 → 定位偏差 → 迭代模型

具体场景：小红要验证本月新的评级模型，需要同时对比上月、上上月的模型在 50 万条测试数据 上的结果表现。他打开模型中心，面对 200+ 模型版本列表，不知道哪个版本最适合本月的评级场景，只能逐个手动创建任务、复制参数、等待结果。3 小时后，他发现某个模型在 FDP 上准确率只有 60%，但系统没有提示他“这个场景下该换模型”或“该调整评分卡分段”，他只能凭经验猜测原因。

情绪曲线



分析核心诉求

快速验证评级模型效果、对比历史版本差异、定位模型衰减原因



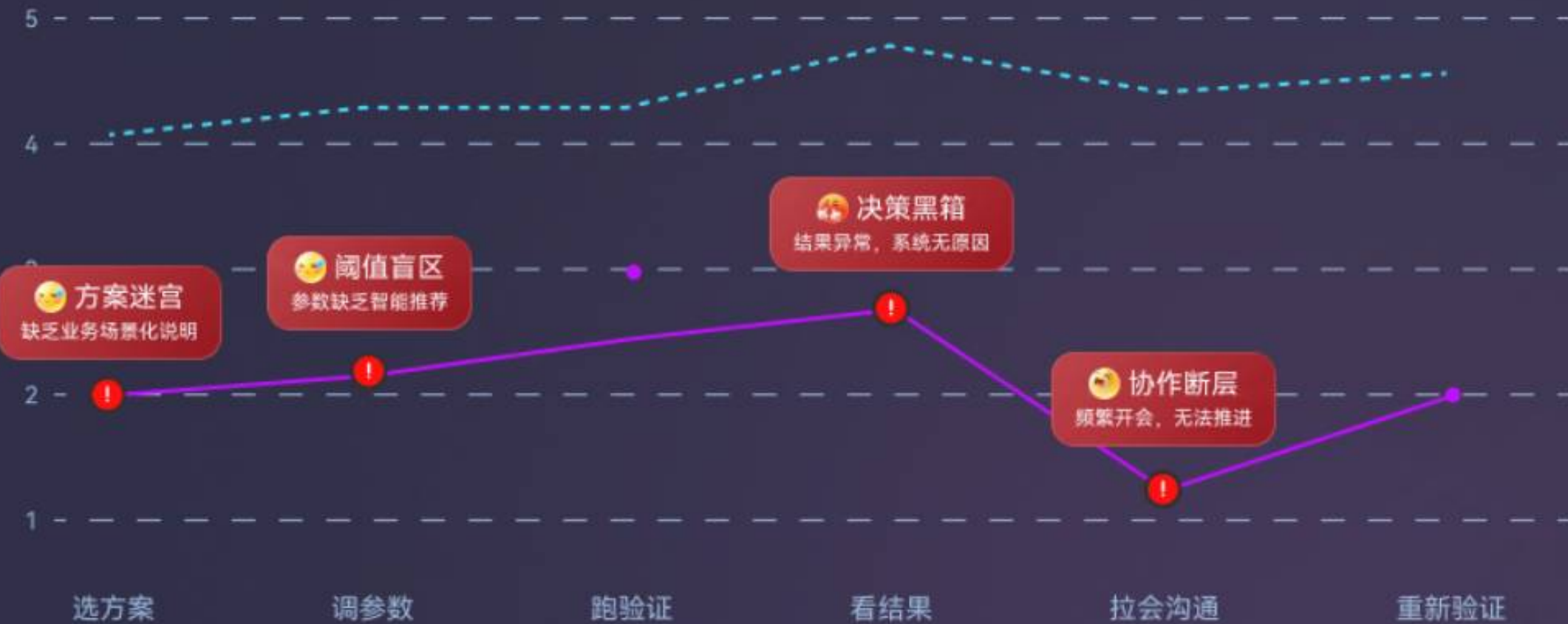
策略同学 | 风控策略/业务策略

场景：上线智能额度调整阈值配置

工作流：定义业务规则 → 选择模型方案 → 配置置信度阈值 → 跑批验证 → 上线监控 → 调优

具体场景：小明负责上线“智能额度调整”策略，需要把模型输出的风险评分自动映射到额度升降规则。他在模型中心看到“评级模型”标签下有 8 个版本，但不知道置信度阈值设 0.65 还是 0.85 更合适。他试着跑了一批数据，结果 25% 的样本被标为“无法决策”，系统没有告诉她“这是因为本月客群分布偏移，建议切换至 V3.2 版本并调整分段阈值”。她只能拉模型同学开会，来回 3 轮才定位问题。

情绪曲线



分析核心诉求

基于模型输出快速上线风控策略、置信度可控、异常可预警

设计目标与策略

压缩模型验证成本

从手动逐条创建任务、复制参数，转变为智能批量对比、自动推荐基准版本、一键生成差异报告

长耗时 → 30min

精简耗时

建立模型可信度沟通

将「黑箱概率」转化为可理解的置信度分级。当指标大幅度偏离时，不仅告知数字，更告知「为什么」以及「该怎么办」

黑箱 → 白盒

决策透明度

打通策略-模型协作链路

消除策略同学与模型同学之间的信息差。策略同学能读懂模型能力边界，模型同学能理解业务场景诉求

频繁会议确认 → 直接决策

跨团队沟通

人机协作能力边界



策略框架

01 模型类别透明化

构建模型能力画像卡片：场景标签、性能基线、版本变更、风险预警，让 200+ 版本可读懂、可搜索、可对比

场景化标签

性能基线可视化

智能搜索推荐

02 渐进式自动化

L1 建议层（手动确认）→ L2 协同层（自动 A/B 测试 + 人工审核异常）→ L3 托管层（PSI 监控 + 自动备用模型）

分层控制

A/B自动化

通查报告

03 容错与优雅降级

指标异常时给出归因与行动建议；定时走查模型参数，输出模型增益报告或提供降级决策，同步与策略同学

异动归因

自动降级

及时响应

核心交互流程 · 从用户输入到结果交付



意图澄清机制 · 当用户意图模糊时

用户输入常存在歧义时，Agent 不直接执行，通过澄清卡片主动追问，并给出 2-3 个结构化选项，降低用户认知成本，并防止模型回答幻觉

场景识别

用户说“对比模型”→ Agent 识别意图，追问：对比维度（FDP/KS/AUC）？对比范围（本月/上月/上上月）？

对比随心贷_大盘评级

✦ 已思考（用时10S）

好的，为您对比模型。请选择对比维度：

☒ KS

☐ IV

☒ 缺失率

☐ 分区数

正在对比中~

请输入内容（支持自然语言或SQL）

↑

选项卡片

不开放输入，提供结构化选项：① 快速对比（3个版本）② 深度对比（全量指标）③ 自定义对比（手动选择）

对比模型

✦ 已思考（用时10S）

好的，为您对比模型。请选择：

① 快速对比（推荐 3 个最优版本）

② 深度对比（全量指标分析）

③ 自定义对比

请输入内容

↑

分步操作

嵌套子Agent编排 + 多阶段工作流 + 人在回路（Human-in-the-loop），主Agent负责任务拆解与全局调度，将复杂任务委派给多个子Agent分阶段执行。

请对比大盘v5、v6中，正在购买奶粉的妈妈人群使用随身贷的比例

✦ 已思考（用时40S）

好的，正在为您分析 大盘v5 与 大盘v6 中「正在购买奶粉的妈妈人群」使用随身贷的比例。根据您的需求，我已自动生成以下任务流，请您确认：

1 圈选

2 关联

3 计算

4 生成

💡 资源预估：离线计算引擎，总耗时 < 1分钟。可选参数：时间范围（默认近3个月）、地区拆分（按需调整）。

修改配置

确认运行

请输入内容（支持自然语言或SQL）

↑

上下文管理 · 多轮对话状态保持与恢复

对话历史

保存用户输入、Agent 回复、用户选择，支持多轮内回溯

↓

任务状态机

记录当前任务阶段、已配置参数、待确认项，断点续传

↓

上下文恢复

检测到断裂时主动提示：「您刚才说的是...，现在是要...？」

帮我把上周最好的那个CTR模型用新数据再跑一次，对比AUC

找到你上周最好的实验是 CTR_v3.2（AUC 0.832，日期 2026-05-28）。是否用它作为基线？

确认 选其他 取消

好的，已确认基线模型。以下是训练参数汇总，请确认：

训练参数确认：

开始跑批 我自己改 取消

训练运行中，model ID: CTR_V2

● 数据校验 ● 模型特征 ● 计算结果 ● 评估 ● 报告

预计剩余时间：~ 5分钟

请输入内容（支持自然语言或SQL）

↑

三级置信度界面

基于模型输出概率与事实核查，建立分层交付策略，避免“一刀切”式回复

高置信度 (>85%)

数据完整、逻辑自洽、有明确来源支撑。Agent 直接输出结论，附带溯源卡片，用户可快速采纳。

通过Markdown格式对相关确定指标做信任度标记

溯源链接放置于下方，增加用户心理预期

对比大盘v5、v6中妈妈人群使用随身贷的比例

✦ 已思考 (用时10S)

基于完整数据验证，**大盘v6中妈妈人群使用随身贷比例为 34.2%**，较 v5 的 28.7% **↑ 5.5pp**。

🔗 数据溯源

CRM

妈妈人群标签库 · 2026-05-28 全量快照

交易

随身贷产品流水表 · 覆盖度 99.7%

验证

交叉验证通过 · 样本量 42,183

请输入内容 (支持自然语言或SQL)

↑

中置信度 (50%-85%)

部分数据缺失或存在多解性。Agent 输出结果但标注不确定性，提供“快速复核”入口，建议人工确认关键节点。

通过Markdown格式对信任度低数据指标做标记

增加人工确认节点

预测下月该人群的转化率趋势

✦ 已思考 (用时10S)

基于近 6 个月时序模型，**预测下月转化率约 38.5% ± 4.2%**，但需注意：

⚠ 不确定性来源

!

618 大促政策尚未完全落地，历史同期数据可能失真

!

妈妈人群标签 5 月刚更新，样本稳定性待观察

查看模型参数

标记待复核

请输入内容 (支持自然语言或SQL)

↑

低置信度 (<50%)

信息不足、矛盾或超出知识边界。Agent 拒绝直接回答，主动请求澄清，并提供结构化选项引导用户补充信息。

通过Markdown格式对超出超出知识边界数据指标做标记

主动澄清，提供结构化选项

分析一下这个人群在竞品平台的借贷行为

✦ 已思考 (用时10S)

当前知识库**缺乏竞品平台的外部数据授权**，无法直接生成可靠结论。

请选择补充信息来源：

📄 上传第三方数据报告

🔗 接入外部数据接口

⚙ 缩小分析范围至自由数据

请输入内容 (支持自然语言或SQL)

↑

幻觉识别与优雅降级

当 Agent 检测到生成内容可能存在事实编造时，主动拦截并降级处理，而非“自信地胡说”



事实核查

RAG 检索与生成内容逐句比对
标记“无来源支撑”的陈述





优雅降级

不直接说“我不知道”，而是提供：① 已确认的事实摘要 ② 待核实的问题清单 ③ 用户可补充信息的入口





幻觉标记

对无法验证的内容使用虚线边框 + 问号角标，视觉上与确认内容区分，避免用户误信

v6 模型升级总结

1. 特征工程升级：引入 实时行为序列特征，KS 提升 0.03（已验证）

2. 模型架构：由 XGBoost 升级为 LightGBM + 深度网络融合（已验证）

3. 样本扩容：训练样本由 120w 增至 280w（内部数据确认，外部报道为 300w，待核实）

4. 业务收益：AUC 提升 0.832 → 0.861（已验证），预计年化收益增加 2400 万（模型推算，未经过财务审计）

⚠ 不确定性来源

！ 样本 300w与内部记录（280w）不符，可能来源于外部媒体报道

！ 年化收益 2400 万为模型基于历史数据推算，非财务部门确认数字

查看内部原始记录

修正

👍👎🔄

您也可以标记问题，Agent 将学习您的修正并标记

如果您需要向管理层汇报，建议采用以下版本：

v6 模型在特征工程与架构层面完成升级，核心指标 AUC 从 0.832 提升至 0.861（验证日期：2026-05-28）。训练样本扩容至 280w，较 v5 增加 133%。业务收益测算正在进行中，预计 6 月 15 日前出具财务审计版报告。

请输入内容（支持自然语言或SQL）

↑

用户修正与模型迭代闭环

将用户反馈转化为结构化数据，打通“发现错误 → 提交修正 → 验证生效 → 模型迭代”的全链路

一键纠错

对置信度中低回答，提供反馈入口，降低用户纠错成本
支持点选错误类型 + 文本框补充描述



实时修正

用户提交修正后，当前会话立即生效（RAG缓存更新），无需等待模型重训
复杂修正进入迭代队列



迭代追踪

用户可在个人中心查看自己提交的所有反馈
状态：已采纳 / 审核中 / 已合入模型
形成正向激励，促进Agent更友好

用户反馈

您指出“广告位历史CTR”的重要性不应下降，请补充说明：

数据错误

计算逻辑错误

数据版本过期

描述性歧义

其他

请输入

此反馈将同步至模型团队，并更新当前会话上下文

提交

请输入内容（支持自然语言或SQL）

↑

感谢您的反馈

修正已记录并生效

当前会话已更新：广告位历史CTR重要性修正为 0.146（↑ 0.008）
该反馈已生成 Ticket #FM-2026-0618-042
状态：已采纳 → 待合入 v3.3

修正后特征重要性 TOP 5

排名	特征名	重要性	修正标志
1	近7天点击频次	0.184	-
2	用户生命周期价值	0.152	-
3	广告位历史CTR	0.146	✓ 用户修正
4	设备品牌偏好	0.095	-
5	时段活跃度	0.087	-

迭代追踪

提交

审核

发版

合入

生效

预计 v3.3 版本（2026-06-25）自动生效 · 您将收到邮件通知

请输入内容（支持自然语言或SQL）

↑

假设验证 · 设计进化

回顾设计过程中的关键决策、意外发现与未竟之事，为后续迭代提供方向



假设验证

回顾设计初期提出的核心假设，哪些被证实、哪些被推翻、哪些需要修正



意外发现

用户行为中超出预期的模式，往往是最有价值的设计洞察来源



迭代方向

基于验证结果与用户反馈，明确下一阶段优先级最高的 3 个设计课题

三级置信度，对专家用户够用，但对新手仍显抽象

假设验证 · 被修正

原以为高/中/低三级足够清晰，但可用性测试发现：3 名非技术背景用户将「中置信度 67%」理解为「有 33% 的概率完全错误」，而非「存在部分不确定性」

后续修正：在中置信度卡片中增加解释文案，并允许用户 hover 查看详细说明

反馈闭环的进度追踪功能意外成为用户粘性来源

💡新洞察

设计反馈闭环的初衷是让修正有去处且可标记并使模型自迭代，但访谈发现，些许模型和算法同学把过于关注「修正已记录并生效」

后续方向：AI产品的用户参与度设计，可降低用户反馈结果的感知程度

低置信度场景的「结构化选项」过于封闭，限制了用户表达

设计遗憾

我在低置信度场景下提供了 3 个预设选项，意图是降低用户认知成本，但过度结构化反而扼杀了用户的真实意图表达。理想方案应该是「结构化选项 + 自由输入」的混合模式，让用户既能快速选择，也能自由描述需求。这是本次设计中最深刻的教训：不要为了降低认知成本而牺牲用户主权。

未来优化方向

个性化置信度阈值 + 跨会话记忆可视化

个性化置信度阈值

不同角色对「可接受的不确定性」容忍度不同，允许用户自定义阈值，而非统一标准

跨会话记忆可视化

当前记忆仅在单会话内有效，下一步要做到Agent从与用户对话中学习用户习惯

开放选项设计

低置信度场景的选项改为「推荐选项 + 自由输入」混合模式，修复本次最大遗憾

验证不是终点，而是下一轮设计的起点

智慧教学Agent-智教专家

产品方案



定位

作为河南省高等学校人工智能教育平台的核心交互入口，「智教专家」Agent 面向教师、校领导、省高教厅三类角色，提供自然语言驱动的教学辅助与数据决策服务。它并非外挂工具，而是深度嵌入平台“教学融合—学科建设—人才培养”全链条的智能中枢。

战略契合

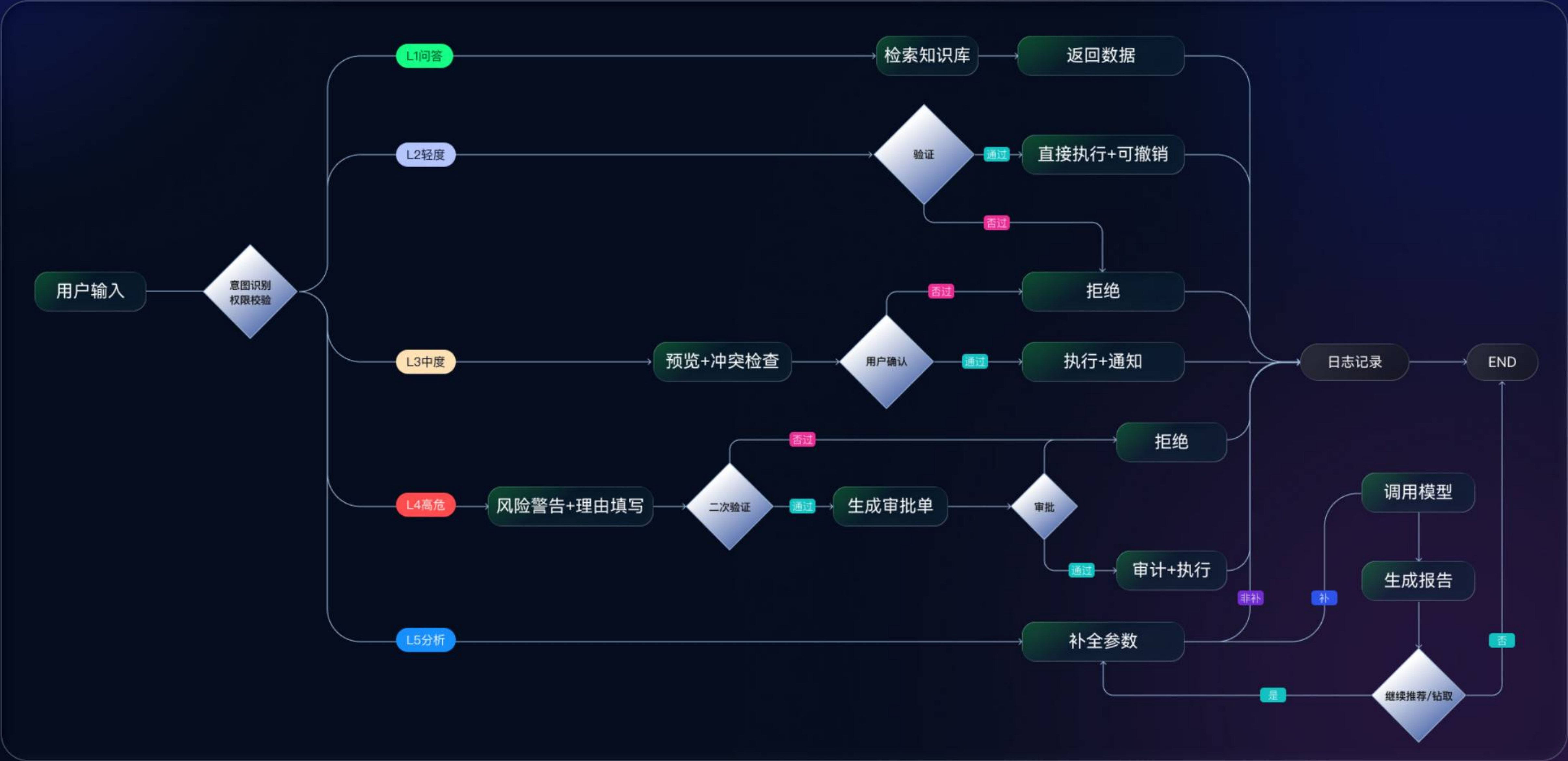
平台承载着《河南省“人工智能+教育”三年行动计划（2025—2027）》的落地使命，需同时解决三大痛点：资源碎片化（校际壁垒）、教学数字化转型阻力（教师操作负担）、产教衔接错位（人才与产业需求脱钩）。智教专家Agent 以“低门槛交互 + 分级安全管控 + 跨角色数据贯通”回应上述挑战：

- 统一资源检索与推荐：教师通过自然语言即可触达资源池中 200 余门精品课程及“基础通识—专业核心—场景应用”三级资源，打破校际壁垒。
- 教学全流程辅助：从调课、组卷、知识库开放等轻中操作，到学情诊断、学习路径优化等高阶分析，Agent 将“教学活动”“知识点线面管理”“AI仿真课”等模块串联为对话式工作流。
- 产教融合闭环：省厅/校领导可通过 Agent 跨校对比人才供给与产业需求匹配度，生成政策建议，实现“教学—实践—就业”数据贯通。



节点区域	目的	可控性体现
意图识别 + 权限校验	确保只有合法角色能进入后续流程，且从源头区分数据范围	前置拦截，避免越权操作
L1 ~ L5 分级分支	根据操作风险自动路由到不同处理逻辑，避免所有操作都走审批或都直接执行	分级清晰，规则可配置（管理员可调整L3/L4阈值）
L3 预览+二次确认	中度决策影响多人，必须让用户看到冲突结果再确认	提供“取消”选项，可逆
L4 审批+2FA+审计	高危操作强制多重验证、审批流、全量审计日志	不可逆操作有完整授权链
L5 参数反问回路	分析类需求往往参数不全，Agent反问补全后才执行，避免错误分析	形成闭环，用户可不断钻取
Monitor 异常监测	记录所有交互日志，当某类问题频繁导致用户取消或失败时，通知管理员优化	系统自我进化能力

总体流程图

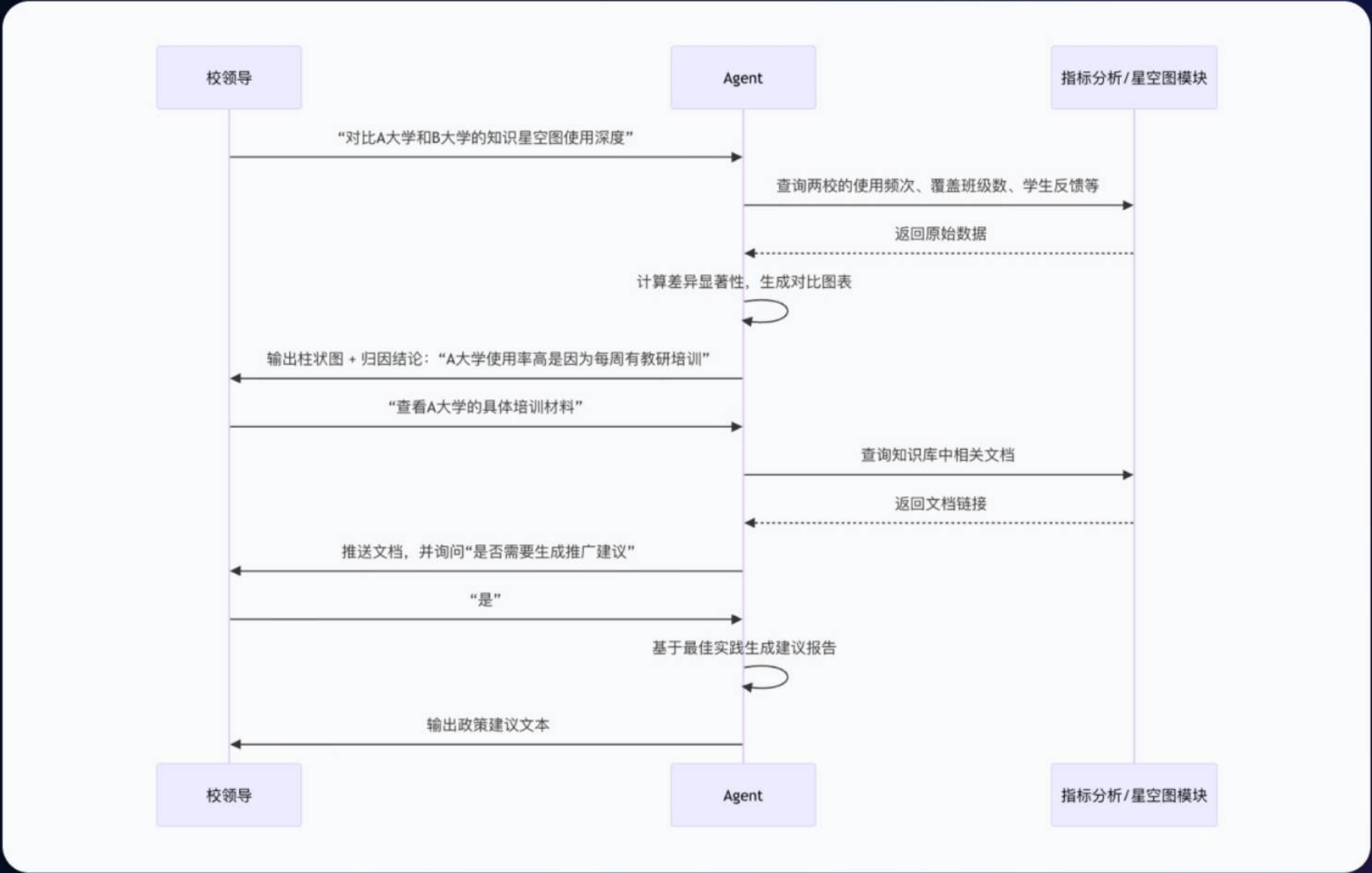


权限泳道图

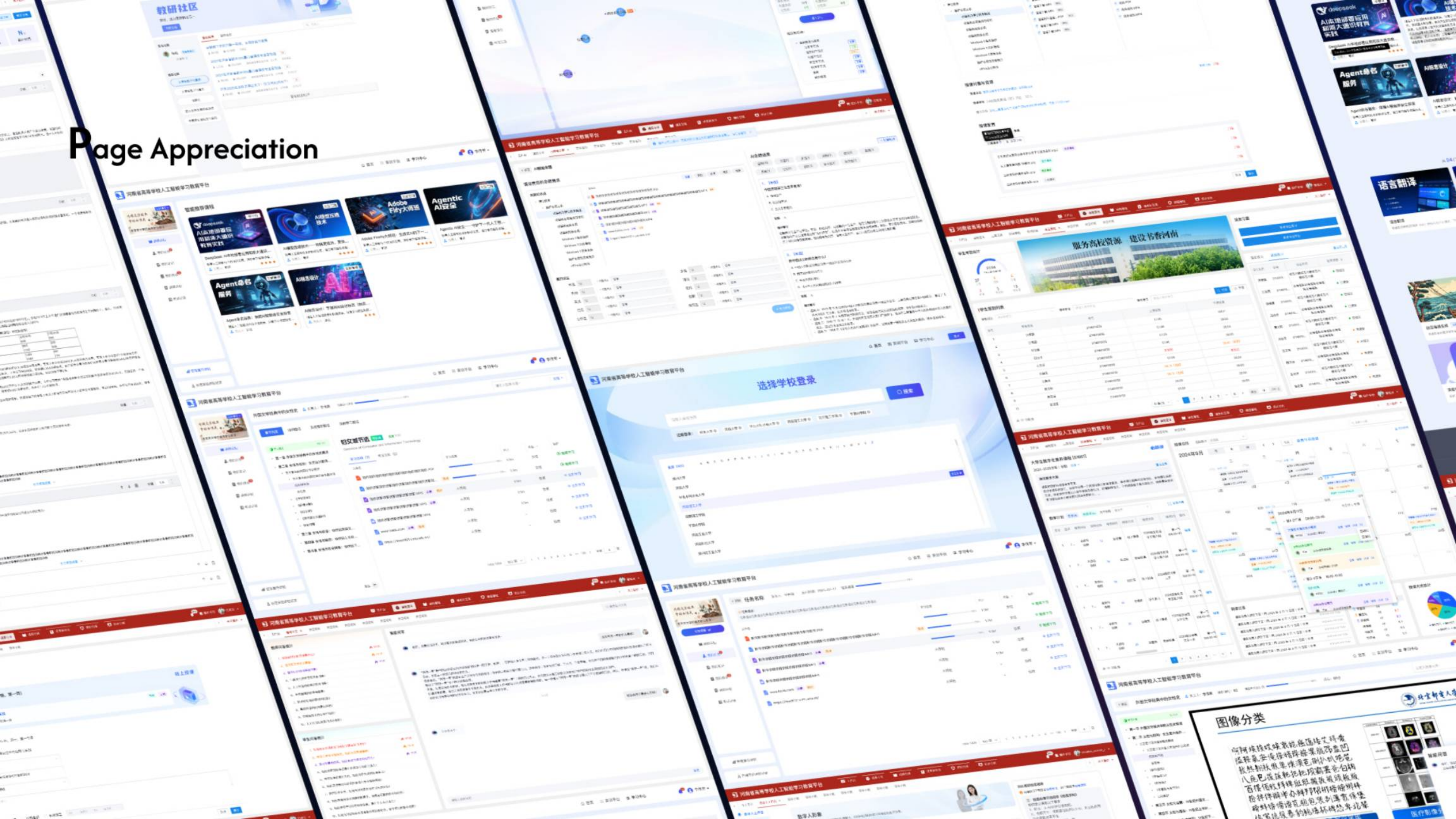
L3+L4场景



L5智能分析场景



Page Appreciation



B端、数据大屏、移动端、HMI

Platform

面向案件链条相关人员

复杂逻辑简单化

方便易用

清晰的节点管理

Platform

Website

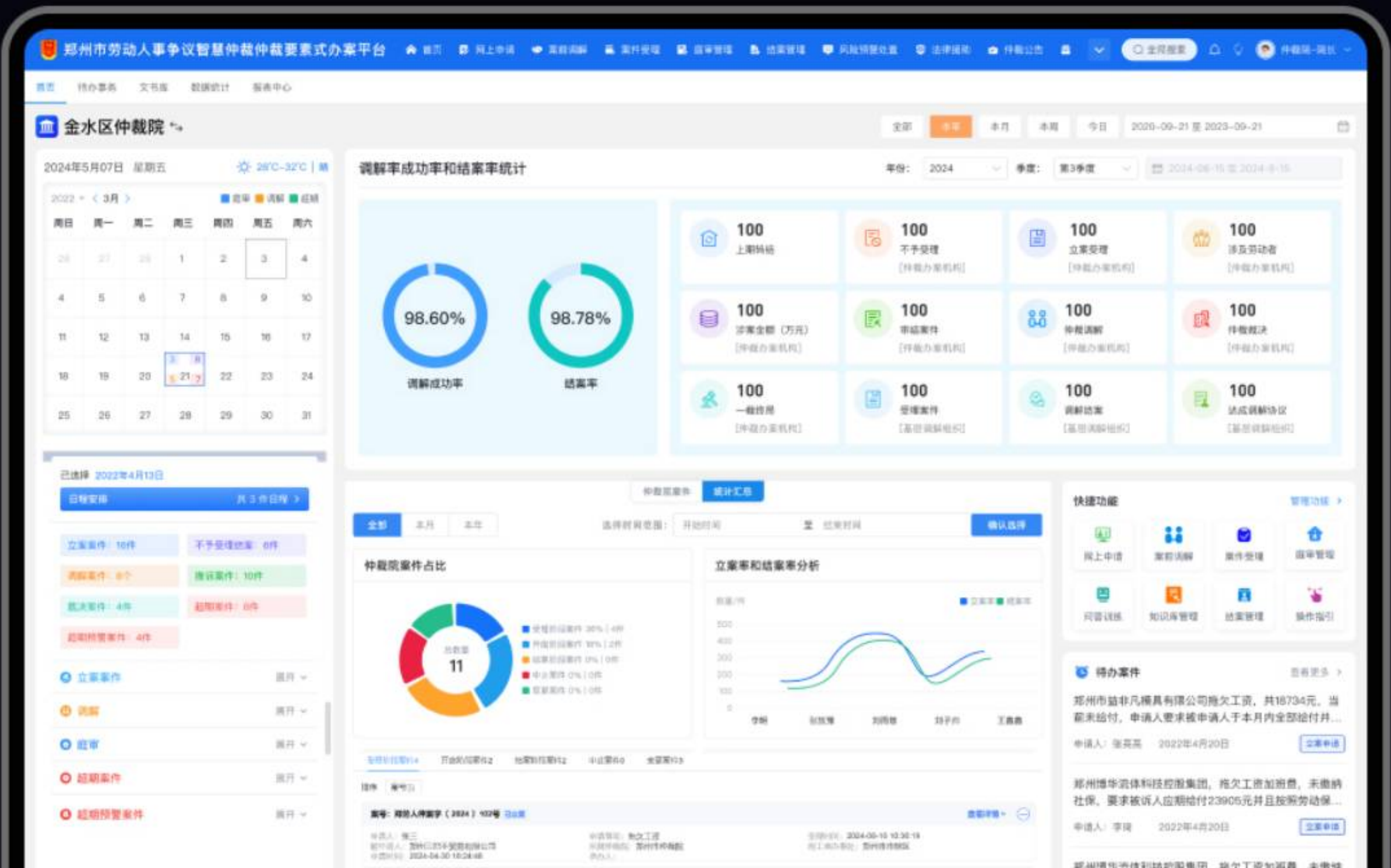
简化界面

清晰明确的流程进度

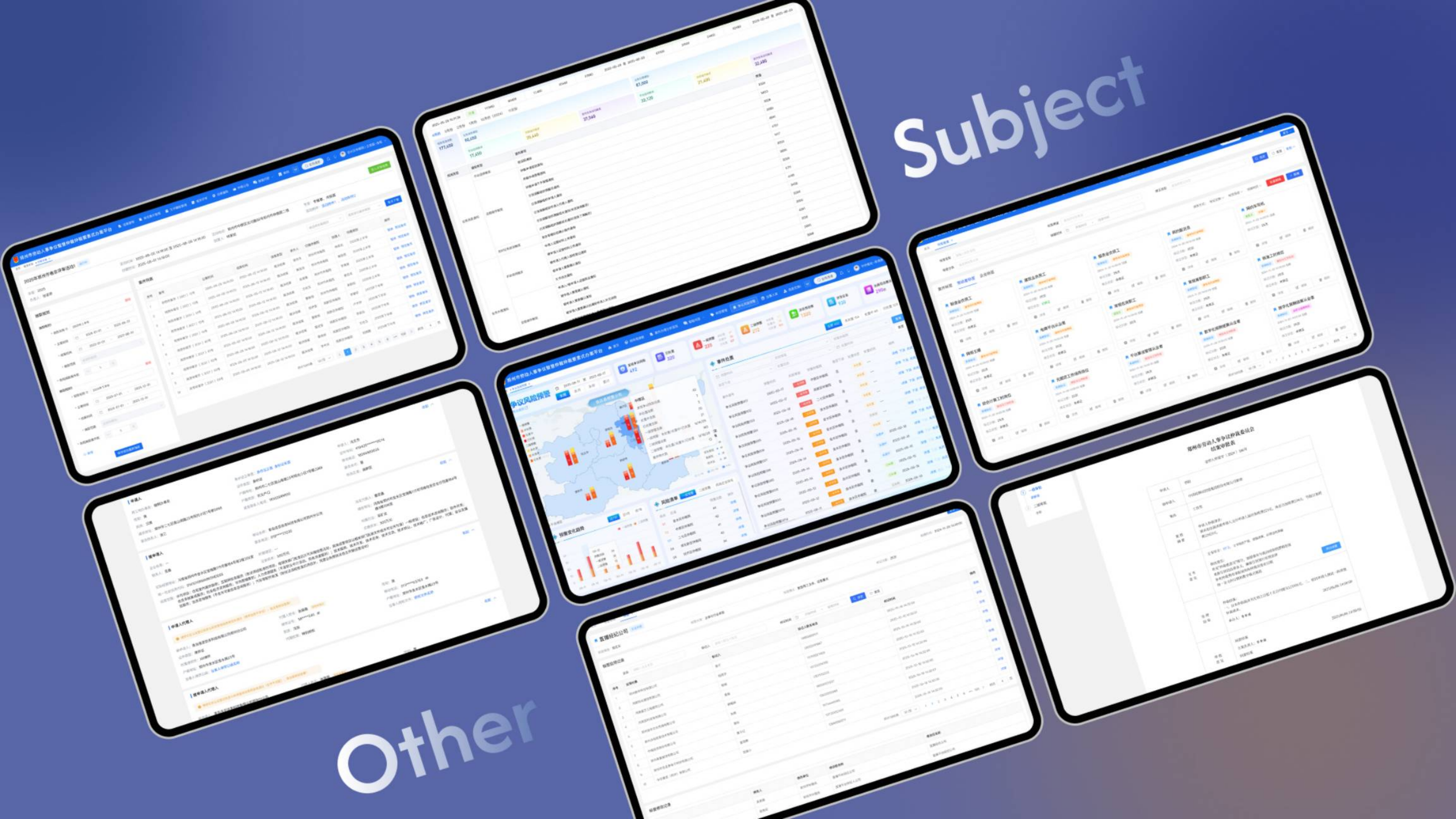
操作直观

Website

面向前端用户群体



Subject



Other

可视化系统 & 帆软

河南开放大学

低代码平台

内网部署

超视距大屏

本项目以帆软 FineReport 为技术底座，依托 FineBI 搭建可视化平台，通过低代码开发与基础函数配置构建页面动态数据交互能力，实现硬件数据的动态感知、实时传输与转译处理，最终达成页面数据的全量及时更新。



Other Subject

超视距可视化大屏

内网部署

超视距大屏





Navigation

_hello

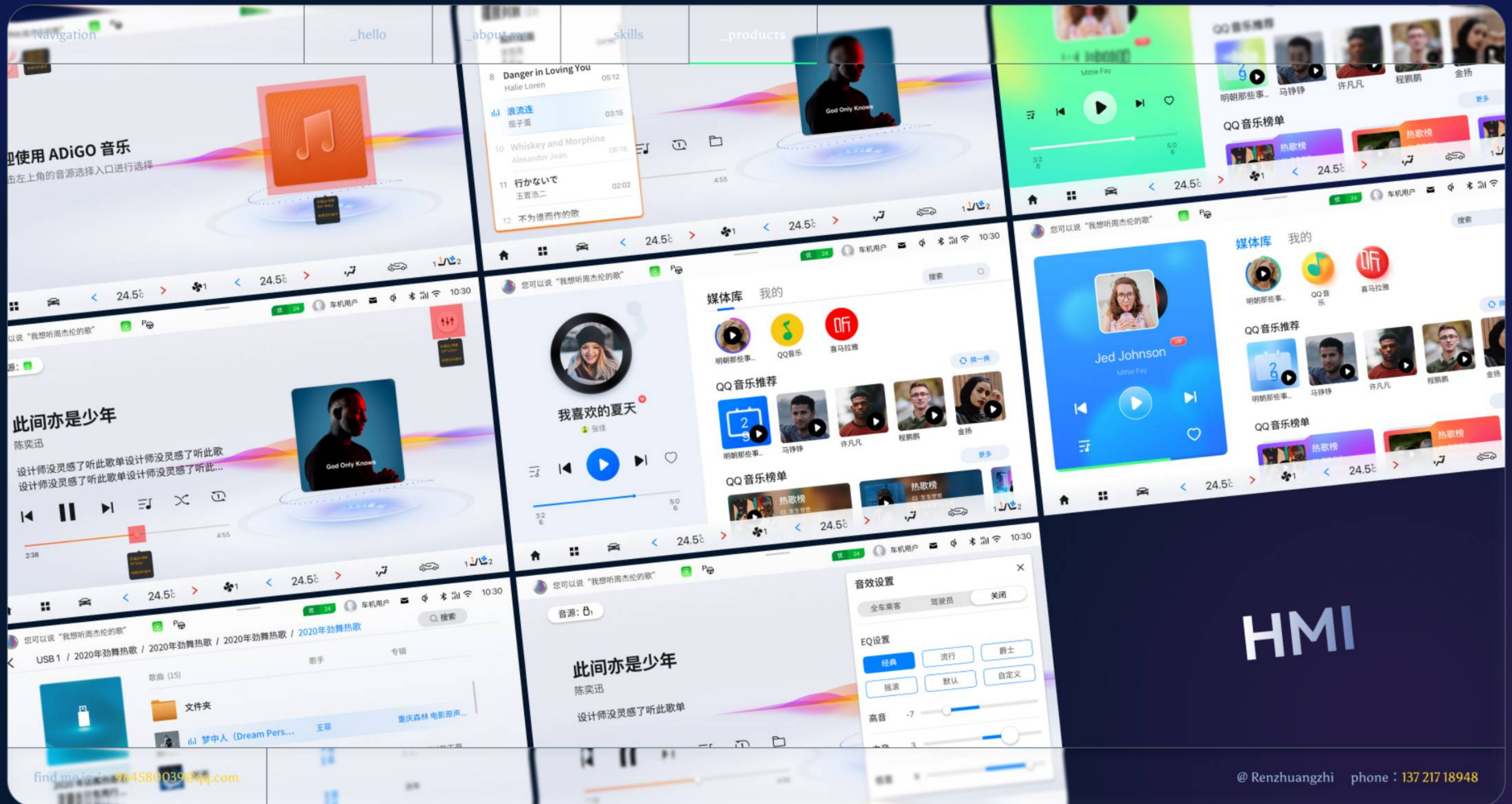
_about me

_products

重庆银行APP

964580039@qq.com

@ Renzhuangzhi phone : 137 217 18948



HMI

期待您的回复

我的联系方式：155 440 85444 | 137 217 18948

Thank You