

AI 智能 · 学习搭子

功能设计文档



声像科技

Sheng Xiang Ke Ji

发包方：数字马力

承接方：声像科技

日期：2025 年 10 月

需求分析

本项目旨在解决当代学习者在两大核心场景中面临的痛点：

一阶段（校园考试）：知识吸收效率低、缺乏个性化指导、学习过程孤独、难以精准评估薄弱环节。

二阶段（就业技能）：高校教学与市场需求脱节、缺乏对标岗位的能力评估、缺少持续的学习动力和规划。

基于此，我们设计了“AI 智能·学习搭子”解决方案，并设计以下的三大模块、四大系统：

可视化交互模块：数字人伴学系统

核心需求分析：

情感陪伴与激励：学习过程孤独是普遍痛点，用户需要一个能提供情感支持、缓解压力、保持学习热情的“学习搭子”。

自然、沉浸的交互体验：传统的文字问答缺乏温度。用户期望通过自然语言（尤其是语音）与学习伙伴进行流畅、低延迟的对话，获得接近真人交流的体验。

个性化形象与认同感：一个固定、无特色的形象难以建立长期陪伴关系。用户希望拥有可定制、甚至能与自己形象关联的专属学习伙伴，增强认同感和使用粘性。

解决方案：

端到端实时语音交互：采用 SenseVoice ASR + Qwen3-4B LLM + CosyVoice 2.0 TTS 的技术栈，构建平均延迟约 3 秒的实时对话流，确保交互的流畅性。

情感化智能陪伴：系统不仅理解语义，还通过 ASR 的情感识别（SER）和 LLM 的情感分析，驱动 TTS 输出不同情感语调（鼓励、耐心等），并计划联动 LAM Avatar 引擎实现表情同步，打造有温度的陪伴体验。

个性化形象生成：利用 LAM 引擎，未来支持学生上传照片生成专属 3D 数字人形象，将“学习搭子”从工具转变为有情感连接的伙伴。

后端服务模块：多智能体-知识库协作系统

核心需求分析：

知识的动态性与准确性：静态知识库无法应对知识的快速迭代。用户需要一个能整合多源信息（教材、论文、行业报告）、并能动态更新的“活”知识库。

复杂问题的深度解答：单一问答模型难以处理跨领域、深层次的问题。用户期望系统能像一个由“图书管理员”、“内容编辑”、“学业顾问”组成的专家团队，协同提供精准、易懂的答案和规划。

从“被动问答”到“主动规划”：用户不仅需要答疑，更需要系统能根据其目标和薄弱点，主动规划学习路径，实现从“授人以鱼”到“授人以渔”的转变。

解决方案：

基于 AgentUniverse 的多智能体融合架构：构建四大核心智能体——知识库检索智能体、内容生成与摘要智能体、学习规划智能体、反思与进化智能体。它们遵循 MCP/A2A 协议，通过 Service Mesh 微服务架构协同工作，形成一个可推理、可协作、可进化的动态知识网络。

活知识库与自我进化：“反思与进化智能体”通过分析交互日志和用户反馈，自动发现知识盲区并触发更新流程，确保系统“越用越懂你”。

深度协同与规划：该系统作为整个方案的“大脑”，为数字人伴学提供知识支持，为评估系统提供规划依据，真正实现了知识服务的智能化和主动化。

评估系统模块：多学科客制化评估系统 & 游戏化学习平台及评估系统

核心需求分析：

分层评估需求：

入门层（K12/兴趣导向）：用户需要无压力、高趣味性的评估方式，将学习与游戏结合，通过即时正向反馈激发内在动机。

进阶层（初高中/高校/就业导向）：用户需要精准、多维度、对标目标的评估，不仅能判断对错，更能诊断薄弱环节，并与学习目标（考试大纲/岗位技能）挂钩。

评估即学习：用户反感孤立的考试。期望评估过程本身就是学习过程的一部分，能提供即时、具体的反馈和指导，实现“以评促学”。

学习路径的无缝衔接：两大评估系统需要打通，实现从“兴趣激发”到“目标导向”的平滑过渡，覆盖用户完整的成长路径。

解决方案：

游戏化学习平台（入门层）：以 Ant Design 为 UI 基座，设计角色成长、即时奖励、挑战任务等游戏化机制。评估完全嵌入式于游戏流程中（如通过口语对话闯关），利用 ASR/NLU 技术实时评估，让学习者在“玩”中无感完成能力诊断。

多学科客制化评估系统（进阶层）：针对不同学科特性定制评估方式。例如，计算机学科基于 OpenSumi IDE 框架，提供云端编程环境和多维度代码评测；数学学科利用 Qwen-VL 识别手写公式；化学学科通过 Three.js 构建 3D 分子模型进行可视化评估。系统可对接相应行业的技能图谱，为就业导向用户提供对标企业标准的能力雷达图。

双系统协同闭环：当游戏化平台用户达到预设水平，系统自动引导其进入客制化评估系统；客制化评估的诊断结果则会触发数字人伴学系统进行主动关怀和针对性辅导，形成“学-练-测-评-优”的完整闭环。

通过四大系统协同工作，声像科技团队构建了一个覆盖“兴趣激发-知识学习-能力评估-职业规划”全链路的智能学习生态。方案深度融合蚂蚁集团技术生态，确保系统的先进性、安全性和可扩展性。

业务流程

声像科技团队就“AI 智能·学习搭子”的项目目标，设计了三种分学段循序渐进的业务流程规划：

业务流程一：小学（K12）（面向兴趣激发与入门学习）

核心目标：激发学习兴趣，培养学习习惯，为后续进阶学习打下坚实基础。

游戏化入口与角色创建：小学生用户进入“游戏化学习平台”，创建专属虚拟角色（Avatar），并可选择与蚂蚁庄园等 IP 联动的皮肤，快速建立情感连接。

玩中学的沉浸式体验：学习过程完全融入游戏。例如，在“单词冒险岛”关卡中，学生需通过口语对话（ASR/NLU 评估）与 NPC 互动来完成任务；在“句法对对碰”游戏中，通过匹配正确词组来消除方块。

即时反馈与正向激励：每完成一个微小任务，系统立即给予动画、音效等多感官即时奖励，形成强大的正向激励循环，有效解决低龄用户注意力难以集中的问题。

数据驱动的个性化难度：平台后台记录学生的发音准确度、错误类型、反应时间等数据，并动态调整后续关卡的难度，确保挑战始终与学生能力匹配，维持其学习兴趣；同时将相应数据发送给家长及教师，更好地掌握学生的学习情况。

平滑过渡至进阶学习：当学生在游戏化平台中达成预设里程碑（如完成小学英语全部课程），系统会自动推荐并引导其进入“多学科客制化评估系统”的更高阶学习路径，实现从“兴趣驱动”到“目标导向”的无缝衔接，覆盖完整的成长轨迹。

业务流程二：初高中（面向“一阶段”校园考试导向）

核心目标：解决传统教学“千人一面”的问题，提供精准的薄弱点诊断与个性化辅导，提升应试效率。

日常学习与嵌入式诊断：学生在日常学习中使用“多学科客制化评估系统”完成作业或单元测试。系统利用 Qwen-VL 识别手写数学公式，或通过 Three.js 进行化学分子结构可视化评估，实现无感化、过程性的学情诊断。

精准定位薄弱环节：系统不仅能判断对错，更能通过 LLM 智能分析，精准定位知识盲区（如“三角函数图像变换理解不清”），并生成详细的错题报告和知识点溯源。

主动干预与伴学辅导：诊断结果实时触发两大系统：

数字人伴学系统：数字人主动发起对话，以温和耐心的语气引导学生回顾错题，并进行针对性讲解。

多智能体-知识库协作系统：学习规划智能体自动调整学习计划，增加相关知识点的练习和复习任务。

情感化陪伴与动力维持：面对升学压力，数字人通过情感化语音（CosyVoice 2.0）和未来将实现的表情同步（LAM），为学生提供持续的情感支持和正向激励，有效缓解学习孤独感和焦虑情绪。

效果追踪与目标达成：系统持续追踪学生在薄弱知识点上的进步情况，形成动态的学习效果报告，帮助学生和教师清晰把握备考进度，直至顺利通过考试。

业务流程三：高校（面向“二阶段”就业技能导向）

核心目标：弥合校园学习与就业需求的鸿沟，提供技能评估与职业规划指导。

入口与目标设定：大学生用户注册后，明确学习目标（如“成为 Java 后端开发工程师”），系统自动关联对应相关企业的行业技能图谱。

客制化能力评估：用户进入“多学科客制化评估系统”，完成对标岗位的实战任务（如在客制化系统的 IDE 中完成一个微服务模块开发）。系统将从代码正确性、算法效率、工程规范等多维度进行精细化评估，生成能力雷达图。

智能规划与知识协同：评估结果同步至“多智能体-知识库协作系统”。学习规划智能体结合岗位要求与用户薄弱点（如“数据库设计能力不足”），生成个性化学习路径；知识库检索智能体与内容生成智能体协同，推送精选学习资源和实战项目。

数字人伴学与情感激励：在整个学习过程中，数字人伴学系统作为“学业伙伴”提供 7x24 小时答疑。当用户在评估中受挫时，系统会主动触发数字人进行情感关怀和错因分析，缓解学习压力。

闭环与就业衔接：用户完成学习路径后，可再次进行评估，验证能力提升效果。系统最终输出一份对标企业标准的综合能力报告，并提供用于求职的简历提升指导，实现从“校园”到“职场”的无缝衔接。

用户交互原型

用户交互原型：小学（K12）——游戏化学习与评估系统

设计目标：打造一个色彩鲜艳、交互直观、充满乐趣的沉浸式学习入口，让低龄用户“玩中学”。

界面风格：采用 Ant Design 的基础组件并进行高度定制化的设计，配色方案以明亮、活泼的糖果色系为主，图标采用圆润、卡通化的设计，符合儿童审美。

核心交互流程：

角色创建页：用户进入后，首先看到一个 3D 角色编辑器。用户可以选择基础形象，并通过“金币”（初始赠送）兑换与蚂蚁集团 IP 联动的皮肤（如小鸡帽子、树叶披风）。此页面利用 Ant Design 的 Modal 和 Tabs 组件组织选项。

主游戏大厅：以岛屿地图形式呈现，每个岛屿代表一个学习主题（如“单词冒险岛”、“句法对对碰”）。岛屿图标采用 Ant Design 的 Avatar 组件，并通过 Badge 徽章组件显示任务完成度。点击岛屿进入关卡选择。

“单词冒险岛”关卡：界面分为三部分：顶部显示任务目标（如“和船长对话，拿到宝藏地图”），中部为 3D 场景（使用 Three.js），底部为麦克风按钮。学生点击麦克风，说出指定句子（如“Where is the map?”），并利用 ASR/NLU 实时评估。成功后，场景中播放庆祝动画（宝箱打开），并弹出 Ant Design 的 Notification 组件，显示“+10 XP!”。

数据看板（家长/教师端）：在家长端 App 中，通过 Ant Design 的 Progress 进度条和 Chart 图表组件，以可视化方式展示孩子的学习时长、发音准确率趋势、高频错误词等数据，便于家长掌握学习情况。



用户交互原型：初高中——多学科客制化评估系统

设计目标：提供专业、清晰、高效的评估与诊断界面，帮助学生和教师聚焦核心问题。

界面风格：采用基于 Ant Design 的主题设计，布局简洁、信息密度高，强调内容的可读性和操作的便捷性。整体风格与蚂蚁系产品保持一致，降低用户学习成本。

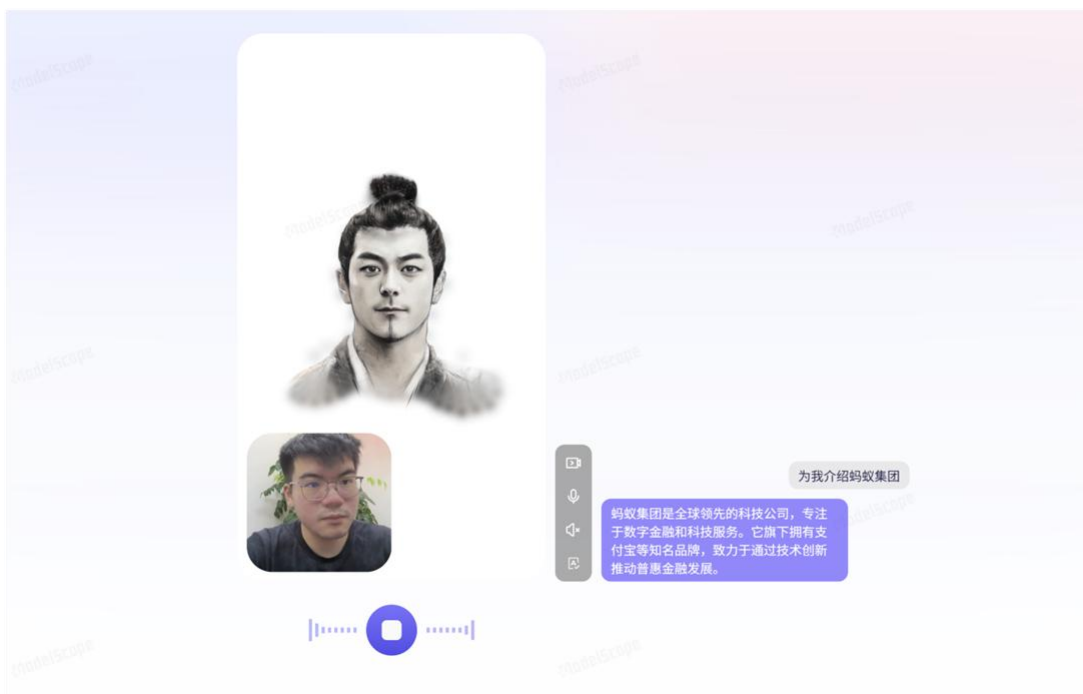
核心交互流程：

学科选择与任务面板：主界面左侧为学科导航栏（数学、化学等），右侧为任务列表。每个任务卡片（Card）清晰标注任务类型（作业/单元测试）和截止时间。

数学手写评估界面：学生点击数学任务后，进入一个类似在线答题卡的界面。学生可在画布上手写公式。提交后，系统调用 Qwen-VL 进行识别和逻辑分析。评估结果页分为三栏：左侧原题与学生手写稿（高亮错误步骤），中间为 LLM 生成的详细错因分析（如“此处应使用诱导公式，而非直接代入”），右侧为关联知识点卡片（可点击跳转复习）。

化学 3D 分子评估界面：在化学任务中，学生需在 Three.js 构建的 3D 空间中，拖拽原子构建指定分子。系统实时评估键角、键长是否符合标准。评估完成后，界面会展示标准模型与学生模型的对比，并用 Ant Design 的 Steps 组件引导学生回顾构建过程中的关键步骤。

错题本与数字人介入：所有错题自动归集到“错题本”。当学生查看错题本时，页面右下角会弹出数字人头像，并主动发起对话：“这道题看起来有点难，需要我帮你再讲一遍吗？”，实现无缝的伴学衔接。



用户交互原型：高校——就业导向的客制化评估与规划系统

设计目标：模拟真实工作环境，提供对标产业标准的专业评估和职业发展指导。

界面风格：采用 Ant Design 的中台/后台解决方案，界面专业、沉稳，功能模块划分清晰，体现客制化应用的专业感。

核心交互流程：

职业目标设定页：新用户注册后，首先进入目标设定向导。通过 Ant Design 的 Cascader 组件，用户选择目标岗位（如“后端开发工程师”→“Java 方向”）。系统自动加载相关关联企业要求的行业技能图谱。

客制化 IDE 评估环境：用户点击“开始评估”后，直接进入基于 OpenSumi 框架构建的云端 IDE。界面包含标准的代码编辑区、终端、调试面板等。任务描述将以 Ant Design 的 Descriptions 组件清晰列出需求，用户在此环境中完成编码任务。

多维度评估报告：提交代码后，系统在定制 IDE 的侧边栏生成评估报告。报告采用 Ant Design 的 Tabs 和 Table 组件，分页展示用户的代码的单元测试，代码质量及能力雷达图。

个性化学习路径：报告页底部提供“下一步行动”建议。点击后，跳转至学习路径规划页。该页面以 Ant Design 的 Timeline 组件展示动态生成的学习计划，包含推荐的课程、项目实战和复习资料，均由“多智能体-知识库协作系统”进行智能推送。

