

设计文档

用户画像

1. 用户群体

主要面向喜欢角色扮演，以及存在对游戏或动漫的某个角色特别喜爱的群体，通过该应用实现跨次元的沟通交流。其中用户年龄段以15岁~30岁为主。其次面向用户群体为喜欢通过对话等现实情境学习新知识的群体。

2. 用户痛点

- 想要沉浸式体验，但现实中不可能和角色交流，游戏（动漫/小说等）中又受剧情约束
- 现有聊天AI太“泛化”，缺乏角色感和一致性。无法深度带入角色，存在“OOC”。
- 语音对话体验差，延迟大、不够自然
- 用户觉得通过阅读来了解某个人物或学习某个知识点太枯燥

3. 用户故事

- 小陈平时喜欢看点动漫，其中有个角色的人格魅力特别吸引他，于是他通过使用语音聊天AI跟角色实现一对一交流，在遇到选择或者是困难的时候通过跟AI对话来获得灵感以及勇气
- 小Y在闲暇的时候会看各种书籍，一次看到了《查拉图斯特拉如是说》，他觉得这书读起来太深奥了，便使用AI聊天软件问“尼采”，“上帝意思是什么意思”。于是AI使用“尼采”的思维回复他。

功能设计

1. 功能计划

- 语音聊天功能（基础对话功能，TTS语音合成输出，STT语音转文本）
- 角色扮演功能（保证AI扮演角色与人设一致）
- 角色切换功能（实现多个角色之间切换）
- 对话记忆功能（保证角色不会出现健忘现象）
- 多人聊天功能（能和朋友一起跟同一个AI聊天/或者是和多个AI聊天，类似群聊功能）
- 自定义音色（让用户有更好的体验感）
- 自定义角色（用户自己创建角色，自定义角色的性格特质）
- 聊天历史管理（让用户刷新页面后也能看到聊天记录）

2. 优先级排序

重点

- 语音聊天功能
- 角色扮演功能
- 角色切换功能
- 对话记忆功能

- 自定义音色
- 次要
- 多人聊天功能
- 自定义角色
- 聊天历史管理

3.开发功能计划

- 语音聊天功能
- 角色扮演功能
- 角色切换功能
- 对话历史管理
- 自定义音色

技术选型

1.大模型选择

DeepSeek-R1

2.模型对比

1. DeepSeek-R1

- 优点：注重推理能力，成本更低，同时开源免费，可考虑本地部署
- 缺点：容易出现服务器繁忙问题

1. OpenAI GPT-4o

- 优点：多模态（语音/文字/图片），响应快，角色扮演效果佳。
- 缺点：API 成本较高。

2. Claude 3.5 Sonnet

- 优点：角色扮演、内容创作等场景下表现优异，长文本处理能力和语境保留能力特别出色
- 缺点：该模型使用比较复杂，并且Claude公开表明锁中国国区，舆论较差

1. Qwen3

- 优点：支持多种思考模式、多语言、多模态和多能力，同时也是开源的
- 缺点：角色扮演一致性（OOC）仍需调优，可能出戏

其他技能

情感表达

能根据场景或者用户的回答表达自己的情绪，比如开心，悲伤。让人有种真的在于角色对话的感觉，而不是AI

寻找话题

娱乐型的角色可以根据用户的设定，选择是否能主动打招呼，比如早安晚安，或者是获取天气预报，再根据天气来提建议

引导思考

知识类的角色应该能够引导用户进行深层次的思考，而不只是简单的回答问题