

基于LEO的跨平台数字人 管家

开发团队

全能具象狮子座

目录

/CONTENTS



1

选题背景简介

2

开发计划介绍

3

技术实现详解

4

项目总结



选题背景简介

LEO模型在自然语言理解、生成和对话等方面展现出了惊人的能力，此项目旨在将LEO模型与不同平台进行整合，如何实现个性化的用户体验，打造一个功能丰富、智能高效的数字管家，为用户提供个性化、智能化的信息管理、任务执行等服务，以满足用户在日常生活和工作中的各种需求。

选题背景

随着人工智能领域各类大模型的不断涌现，为数字人提供了强大的智能支持。与此同时，显卡的加速迭代和发展使得图形处理引擎如Unity和UE5等开发软件能够处理越来越复杂的图像。在PC端，处理高质量图像已经成为可能，这意味着数字人的身体展示已经达到了高度逼真的水平。结合这两种先进技术，无疑可以创造出在各种实际场景中都能完美呈现的数字人。



选题背景



基于LLM训练的具身智能体LEO在识别、规划、推理、导航、对话和操纵等方面表现出色。作为新兴的精通3D任务的具身智能体，LEO在应用方面展现出巨大潜力，完全有能力作为数字人的强大智能核心。鉴于此，我们团队提出了“基于LEO的跨平台数字人助理”这一概念。



2

开发计划介绍

此数字人助理项目的实施分为四个主要阶段。

开发计划

第一阶段

此阶段作为项目的核心部分，主要利用LEO的API来驱动数字人，并实现高效的人机交互。

第二阶段

此阶段将开发移动应用程序(APP)，以实现移动端的交互，增加用户便利性，并确保用户移动端与PC端的数据同步。在此阶段，我们计划上线Steam等主流游戏平台，以吸引更多玩家进行测试并提供反馈，以便进行后续功能的添加和模型的微调，从而在日常生活中实现应用。

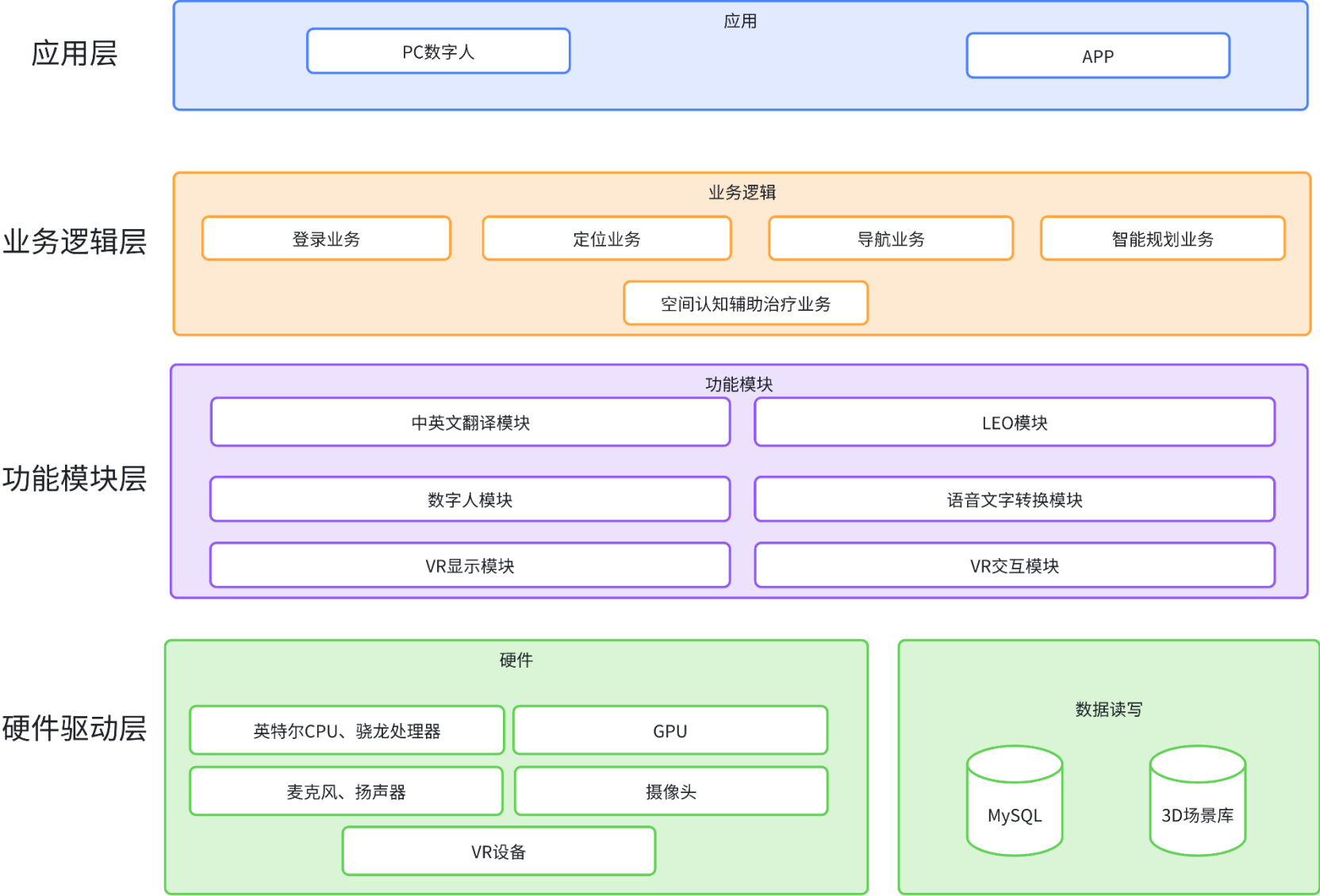
第三阶段

此阶段将根据常见的空间认知障碍康复方法，利用LEO提供的康复方法。鉴于空间认知康复与EEG信号分析的密切关系，我们将根据第二阶段收集的各类反馈对LEO进行相应的调整，从而完成在医疗领域的应用。

第四阶段

根据第三阶段的开发成果和实际应用需求，我们将尝试将系统接入AR、VR或MR设备，以实现更加丰富的场景应用。

系统架构



目标与用途

利用人工智能技术，构建一个跨平台的数字人助理，所依托的大模型名为LEO，为用户提供语音交互、备忘录管理、场景导航以及AR与VR连接等功能。其主要用途包括但不限于：

- ★ 通过手机App与LEO进行语音交互，获取反馈结果；
- ★ App提供事件记录功能，软件读取并记录；
- ★ 通过电脑端UE5搭建的数字人模型与LEO进行交互，实现与手机端事件记录互通，以及在场景图中帮助寻找物品；
- ★ 利用手机摄像头连接与AR技术，实现在找寻物品时的导航功能；
- ★ 通过VR设备与LEO进行连接交互，为大脑疾病患者提供康复训练；



移动端APP设计

01 语音交互

- 用户可以通过语音与LEO进行交互。
- App将用户的语音转换为文字，提交给LEO进行处理。
- LEO将处理结果反馈给用户。

02 自动寻物功能

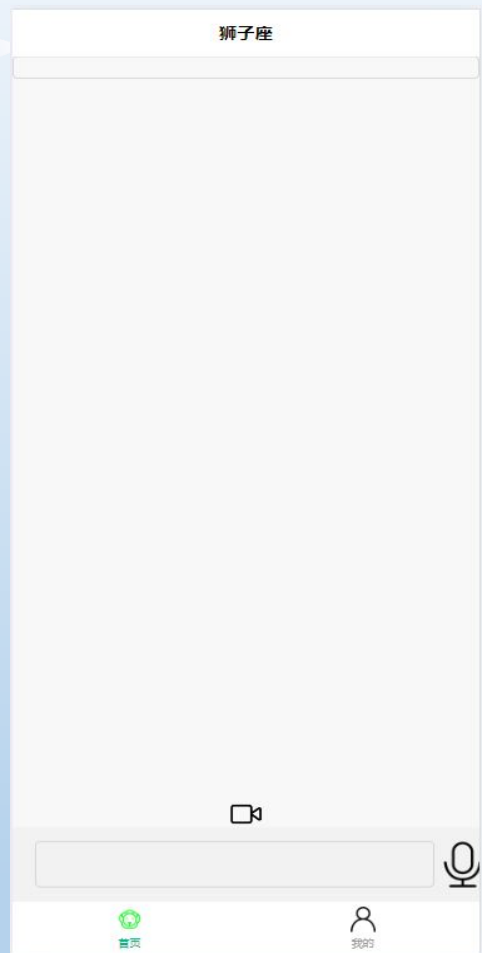
- App提供事件记录功能，用户可以在其中记录完成某一事件所需要的物品，软件读取并记录，当语音交互触发关键词时，LEO对事件所需物品按顺序指出所在位置。



03 AR技术支持

- 用户可以通过手机摄像头连接与AR技术，实现在找寻物品时的导航功能。

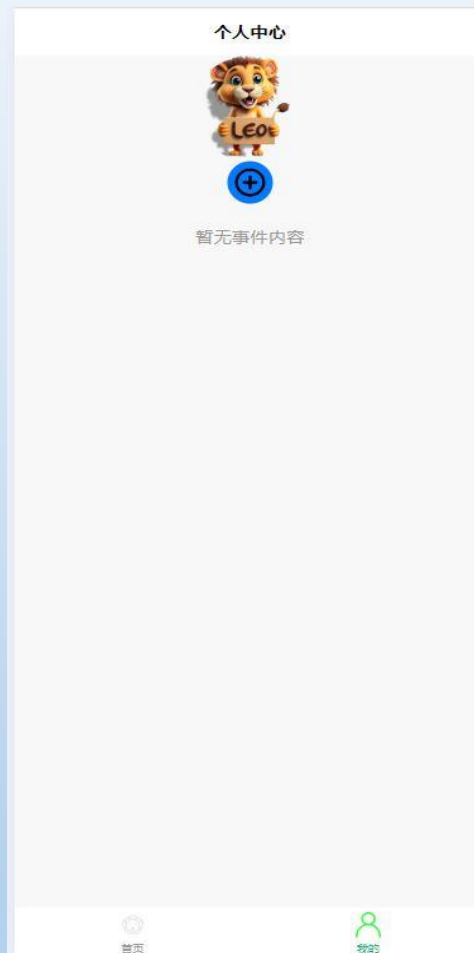
APP界面与功能实现



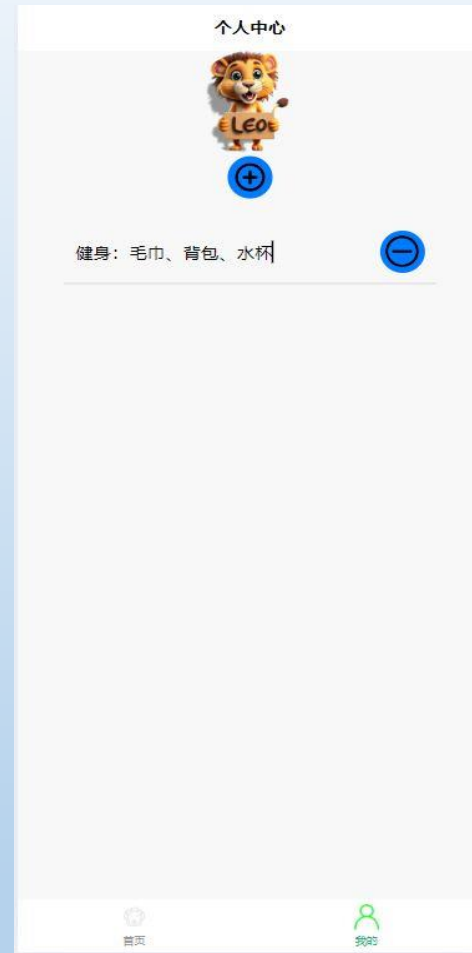
聊天框待机界面



人机交互界面以及场景展示



事件相关物品记录
待机界面

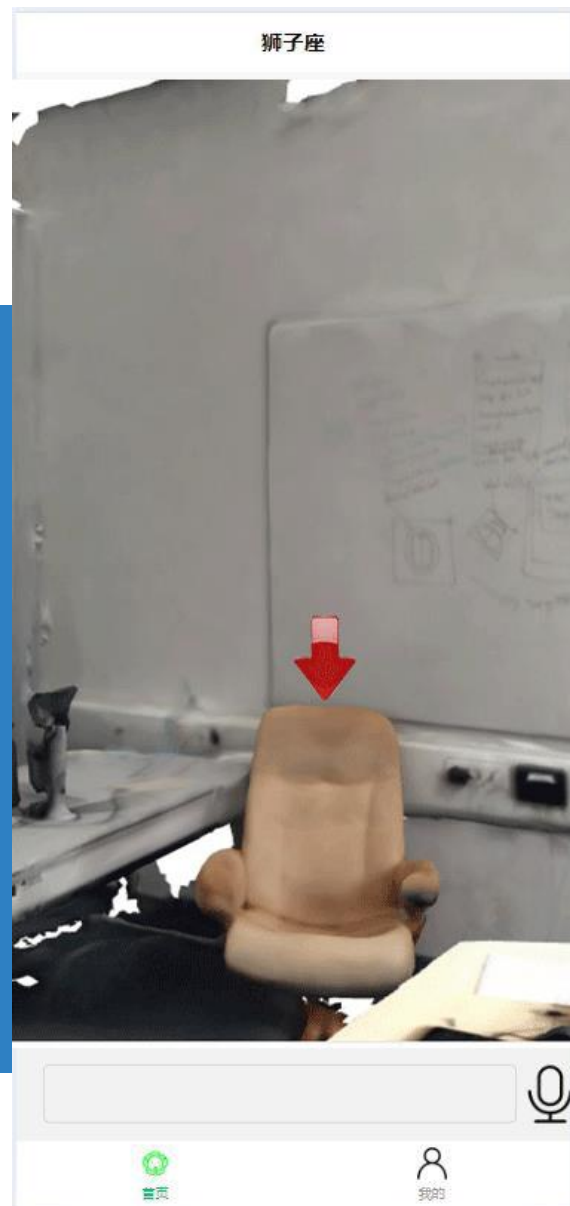


事件相关物品记录

AR技术的应用

寻物导航

当用户利用手机摄像头连接APP并应用AR技术时，他们可以体验到一种全新的导航功能，使得找寻物品变得更加直观和便捷。通过AR技术的应用，用户可以在手机屏幕上看到虚拟的导航标记，如红色箭头，指示着所寻找的物品的方向和位置。这种技术的使用不仅使得找寻物品变得更加容易，还为用户提供了一种更加沉浸式和创新的导航体验。



电脑端数字人界面设计

背景场景图

在数字人模型周围显示一个场景图，该场景图可以是用户所在的房间、办公室等场景。

场景图可以是3D环境，用户可以通过数字人模型进行导航。

主界面

电脑端显示一个数字人模型，基于UE5搭建，具有逼真的外观和动画效果。数字人模型可以根据用户的语音或文本输入进行动作和表情的变化，以模拟真实人类交互。



交互组件

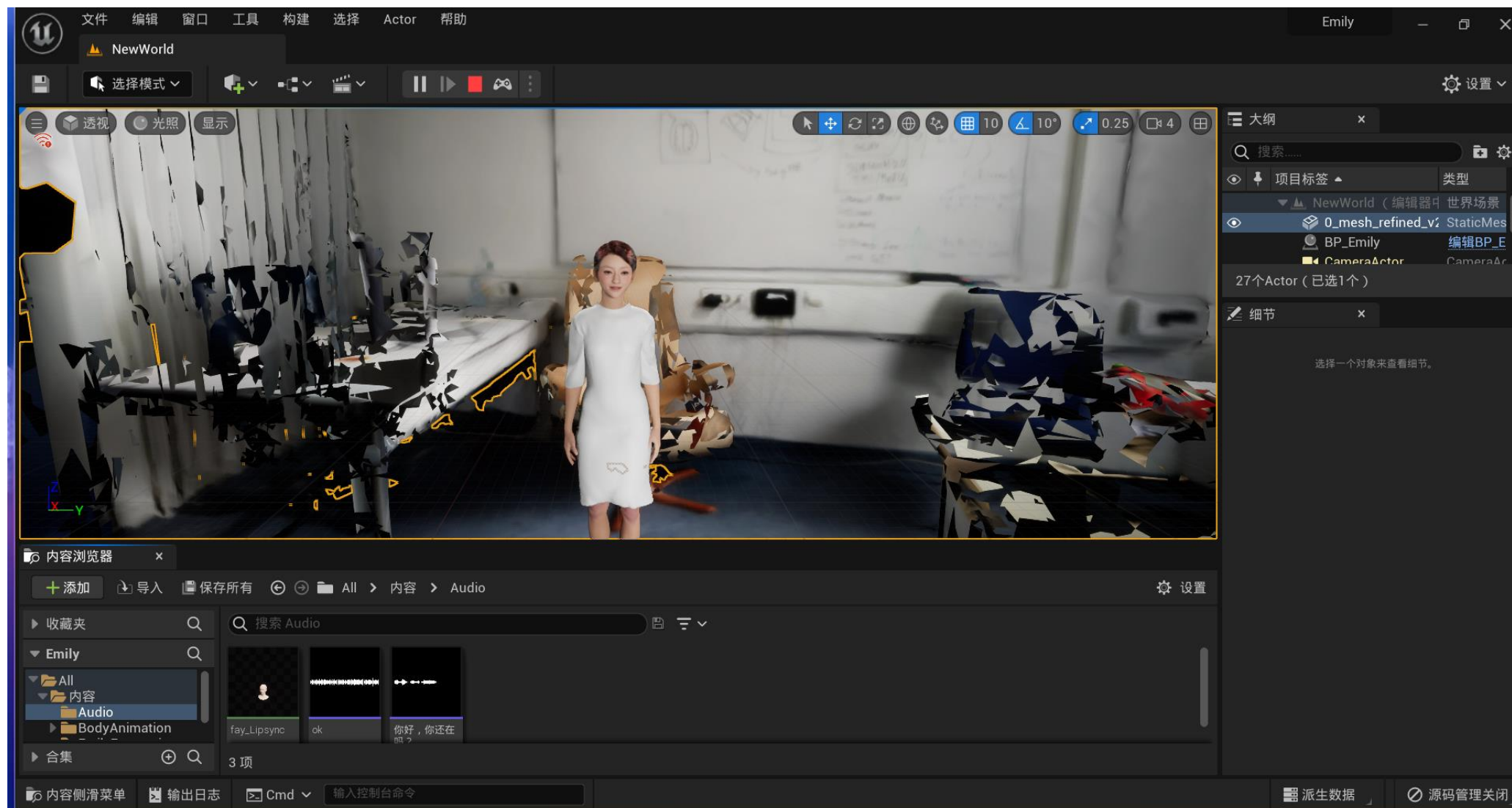
语音输入按钮：用户点击后可以开始语音输入，输入的语音会被转换为文字并发送给LEO进行处理。文字输入框：用户也可以直接在输入框中输入文字信息，然后发送给LEO进行处理。

用户可以通过语音或文本命令请求数字人模型帮助找寻指定物品。

物品搜索功能

数字人模型会根据用户的指令在场景图中搜索目标物品，并提供导航指引。

数字人界面与功能实现



数字人控制台界面

人设:

姓名: 菲菲

性别: 女

年龄: 18

出生地: gihub

生肖: 随意

星座: 随意

职业: 随意

联系方式: 11111111

喜好: 随意

Q&A文件: qa_demo.xlsx

使用面板播放: 是

声音选择: 晓晓

敏感度:

Fay Eyes

唤醒模式:

唤醒词: 你好

唤醒方式: 普通

远程音频: 未连接

tts合成:

麦克风

麦克风 (Realtek(R) Audio)

消息

使用提示: Fay可以独立使用, 启动数字人将自动对接。

开启

保存配置

数字人: 未连接

2024-05-04 22:46:22

屋子内有桌子, 椅子, 床, 电脑, 窗户。

2024-05-04 22:46:31

一直在哪儿

2024-05-04 22:46:33

我一直在这里, 随时准备帮助你。

2024-05-04 22:46:40

椅子在哪

2024-05-04 22:46:42

椅子在桌子旁边。

2024-05-04 22:46:48

发送些内容给 Fay...

Fay

数字人界面与功能实现



VR居家空间认知康复

VR空间认知游戏 个性化定制

病人大脑受到心脑血管疾病等影响后无法实现正常的空间认知功能。研究表明，利用VR游戏可有效的对病人的空间认知功能实现恢复。

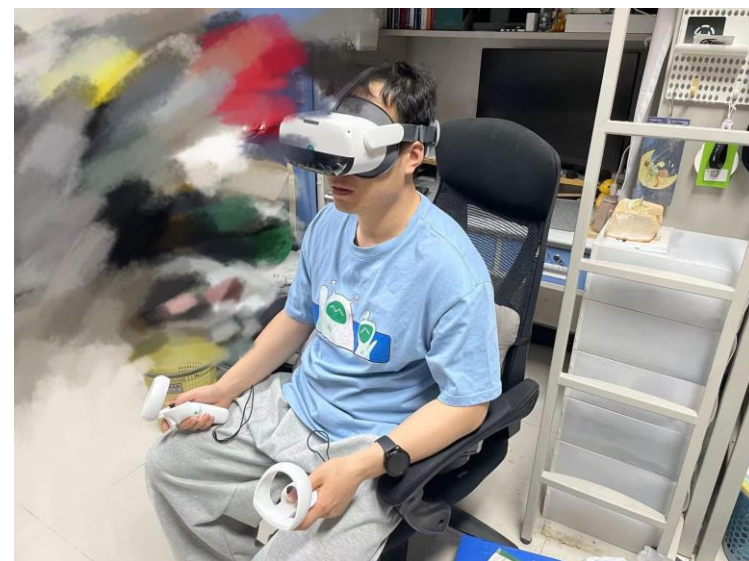
游戏设计

VR设备内随机生成要进行训练的场景，LEO随机指出某一物品所在空间位置，病人根据其所给信息在VR环境之中进行查找。



项目所用VR设备为PICO

负责人李X测试VR功能





技术实现详解

此部分包括技术实现、部署环境、数据存储以及测试方法。

01



技术实现

手机App：
UNI-APP、
Hbuilder、
Javascript、Html5
、vue框架、CSS

电脑端数字人界面
：Unreal engine 5

语音识别与处理：
语音转文字Azure

02



数据存储

用户对话记录：
存储在本地数据库中，用于备忘录功能的实现。

场景导航数据：
存储在电脑端，用于数字人助理在场景中导航。

03



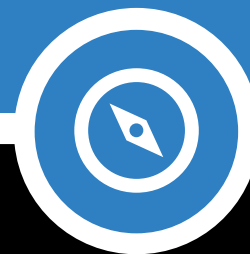
部署环境

手机环境：
Android；

PC环境：
Windows10、
Windows11；

VR设备为PICO。

04



测试方法

- ①手机App功能测试；
- ②电脑端数字人功能测试；
- ③手机摄像头连接与AR技术导航功能测试；
- ④VR设备连接与交互功能测试；
- ⑤兼容性测试；



项目总结

此部分包含项目小结、团队介绍、未来规划





技术难点

在开发过程中遇到技术难点，包括语音识别的准确率不高、AR技术与手机摄像头的兼容性问题等，通过持续跟踪技术发展，及时应用新的技术和算法，共同解决技术难题。



通过项目的实施，我们成功地开发了一个功能丰富、性能优越的智能助理软件，为用户提供了便捷、高效的数字化生活体验，在未来继续不断优化和完善软件功能。

在项目开发过程中，充分考虑了用户的需求和体验，通过对功能的详细设计和测试，确保了软件的稳定性和可靠性。同时，我们还积极采用了先进的技术，如语音识别、AR技术和VR技术，提升了用户的交互体验和功能性。



项目小结

基于LEO的跨平台数字人管家是一个集成了语音交互、事件相关物品记录、AR技术导航以及VR设备连接等多种功能的智能助理软件。通过手机App和电脑端数字人界面，用户可以与LEO进行交互，并且数字人界面与手机App备忘录功能实现了互通。此外，该软件还利用手机摄像头连接与AR技术，提供了方便的物品导航功能，并且支持通过VR设备与LEO进行交互，为用户提供了全方位的智能助理服务。

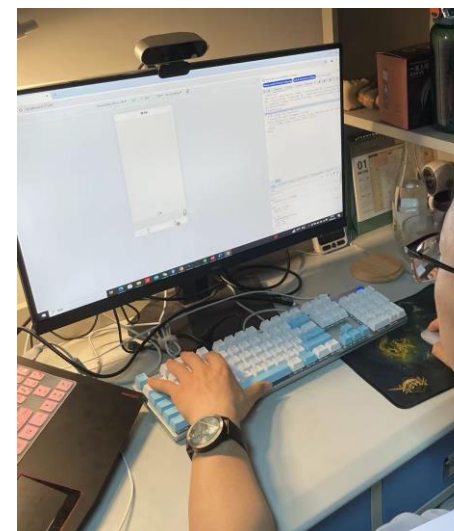
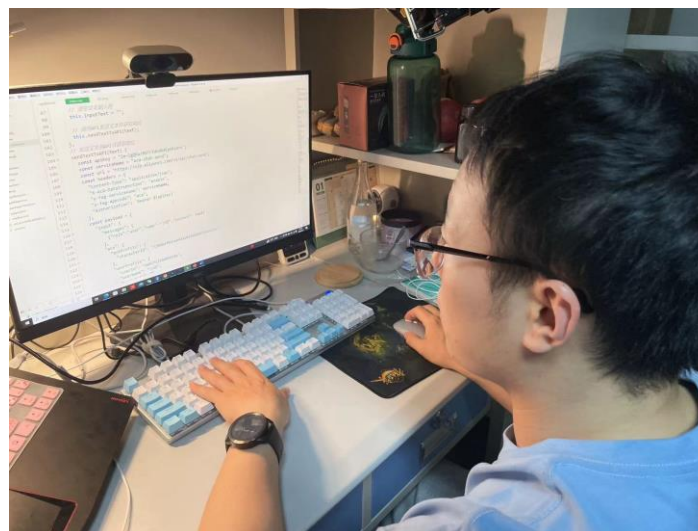
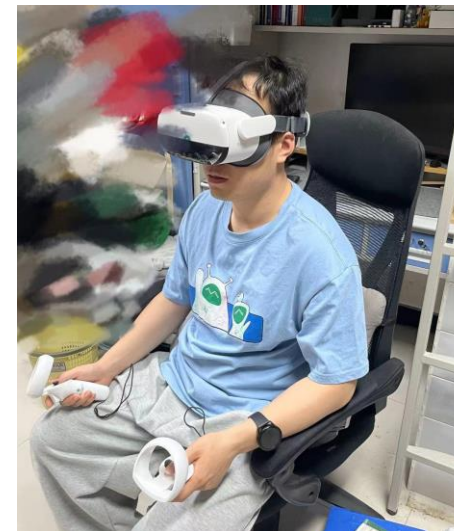
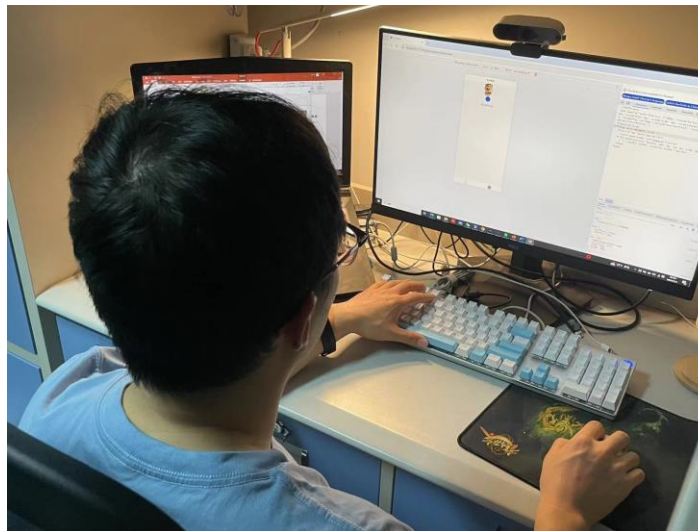
负责人介绍



【项目负责人】 李X

主要工作:

- ☑ 移动端APP开发;
- ☑ VR、AR与LEO结合设计;
- ☑ 服务器部署指导;
- ☑ 数字人制作规划;
- ☑ 项目整体规划调度。



团队介绍及分工



牛 X X

ue5数字人模型搭建
python控制台的搭建
相关API调用



张 X X

云服务器部署 leo



聂 X X

云服务器部署 leo



段 X X

ue5数字人模型搭建
相关API调用



沈 X X

ue5数字人模型搭建
word版计划书撰写



侯 X X

行业调研与PPT展示方
案制作



王 X

行业调研与PPT展示方
案制作

未来规划

01

增加智能推荐功能

根据用户的偏好和历史记录推荐相关信息或服务，提升用户体验。例如，根据用户的运动习惯和健康数据推荐适合的健身计划或营养餐单。

02

日程管理功能

让用户可以通过LEO助理安排和管理日常事务，包括会议安排、提醒事项等，提高用户的工作效率和生活质量。

01

技术升级方面

引入更智能的推荐算法，提高智能推荐的准确性和个性化程度。

02

用户体验改进

持续优化界面设计和交互体验，根据使用反馈不断优化产品功能和性能。



感谢评委老师的指导

完成时间

2024.5.4

软件开发设计文档

1. 软件概述

软件名称: 基于 LEO 的跨平台数字人管家

版本号: 1.0

开发团队: 全能具象狮子座

发布日期: 2024.5.4

2. 目标与用途

该软件旨在利用人工智能技术, 构建一个跨平台的数字人管家, 所依托的大模型名为 LEO, 为用户提供语音交互、备忘录管理、场景导航以及 AR 与 VR 连接等功能。其主要用途包括但不限于:

- 通过手机 App 与 LEO 进行语音交互, 获取反馈结果。
- 通过电脑端 UE5 搭建的数字人模型与 LEO 进行交互, 实现与手机端事件县官物品栏的互通, 以及在场景图中帮助寻找物品。
- 利用手机摄像头连接与 AR 技术, 实现在找寻物品时的导航功能。
- 通过 VR 设备与 LEO 进行连接交互, 为大脑疾病患者提供康复训练。

3. 功能概述

3.1 移动端 App 功能

3.1.1 语音交互

- 用户可以通过语音与 LEO 进行交互。
- App 将用户的语音转换为文字, 提交给 LEO 进行处理。
- LEO 将处理结果反馈给用户。

3.1.2 自动寻物功能

- App 提供事件记录功能, 用户可以在其中记录完成某一事件所需要的物品, 软件读取并记录, 当语音交互触发关键词时, LEO 对事件所需物品按顺序指出所在位置。

3.1.3 AR 技术支持

- 用户可以通过手机摄像头连接与 AR 技术, 实现在找寻物品时的导航功能。

3.2 电脑端数字人界面功能

3.2.1 与 LEO 交互

- 用户可以通过电脑端数字人模型与 LEO 进行语音交互。
- 电脑端数字人将用户的语音转换为文字, 提交给 LEO 处理, 并显示处理结果。

3.2.2 与移动端自动寻物列表互通

- 电脑端数字人可以与手机 App 上的自动寻物列表进行互通。
- 用户在电脑端与数字人对话时, 可以查看并编辑移动端自动寻物列表。

3.2.3 场景导航

- 电脑端数字人模型背景为所设场景图。
- 用户可以请求数字人助理在场景图中帮助寻找指定物品, 并提供导航指引。

3.3 VR 居家空间认知康复

3.3.1 VR 空间认知游戏个性化定制

- 病人大脑受到心脑血管疾病等影响后无法实现正常的空间认知功能。研究表明, 利用 VR 游戏可有效的对病人的空间认知功能实现恢复。

4. 技术实现

- 手机 App：UNI-APP、Hbuilder、Javascript、Html5、vue 框架、CSS
- 电脑端 UE5 数字人界面：Unreal engine 5
- 语音识别与处理：微软 Azure 语音转文字处理。

5. 数据存储

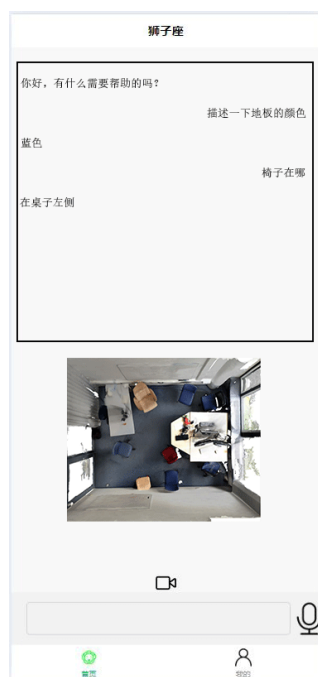
- 用户对话记录：存储在本地数据库中，用于聊天框功能的实现。
- 场景导航数据：存储在电脑端，用于数字人助理在场景中导航。

6. 用户界面设计

6.1 手机 App 界面设计：

6.1.1. 对话页面

- **对话框：**
 - 用户输入框：用户可在此处输入语音或文字进行与 LEO 的交互。
 - LEO 反馈框：显示 LEO 返回的文字信息或语音回复。
 - 对话记录：显示用户与 LEO 的对话历史记录，包括用户发送的消息和 LEO 的回复。



- **AR 寻物：**
 - LEO 根据客户提出物品查找功能，调用手机摄像头，根据场景利用 AR 技术指出物品所在位置。



6.1.2. 个人页面

- **自动寻物列表：**

- 显示用户保存的自动寻物条目。
- 每个备忘录条目包括对话内容和创建时间。
- 点击备忘录条目可以查看详细内容。



6.2 电脑端数字人界面设计：

6.2.1 主界面：

- 显示一个数字人模型，基于 UE5 搭建，具有逼真的外观和动画效果。
- 数字人模型可以根据用户的语音或文本输入进行动作和表情的变化，以模拟真实人类交互。

6.2.2 交互组件：

- 语音输入按钮：用户点击后可以开始语音输入，输入的语音会被转换为文字并发送给 LEO 进行处理。
- 文字输入框：用户也可以直接在输入框中输入文字信息，然后发送给 LEO 进

行处理。

6.2.3 背景场景图：

- 在数字人模型周围显示一个场景图，该场景图可以是用户所在的房间、办公室等场景。
- 场景图可以是 3D 环境，用户可以通过数字人模型进行导航。

6.2.4 物品搜索功能：

- 用户可以通过语音或文本命令请求数字人模型帮助找寻指定物品。
- 数字人模型会根据用户的指令在场景图中搜索目标物品，并提供导航指引。

6.3 VR 居家空间认知康复

6.3.1 VR 空间认知游戏

- 游戏设计：
 - VR 设备内随机生成要进行训练的场景，LEO 随机指出某一物品所在空间位置，病人根据其所给信息在 VR 环境之中进行查找。



7. 部署与测试

- 部署环境：手机环境 Android；PC 环境 Windows10、Windows11；VR 设备为 PICO
- 测试方法：

7.1. 手机 App 功能测试

7.1.1 语音交互

- 测试目标： 确保用户可以通过语音与 LEO 进行交互，并且语音能够准确地转换为文字提交给 LEO，收到正确的反馈结果。
- 测试步骤：
 - 启动手机 App，进入对话功能页面。
 - 点击语音输入按钮，说出一个测试指令。
 - 确认语音能够被准确识别并转换为文字，并提交给 LEO。
 - 验证 LEO 返回的反馈结果是否准确。

7.1.2 事件相关物品功能

- 测试目标： 确保用户可以在手机 App 事件相关物品栏中记录，并且事件相关物品

栏能够准确记录。

- **测试步骤：**
 - 在对话页面中进行一次对话交互。
 - 确认对话内容设计的事件相关物品均在对话框中被准确指示出位置。

7.2. 电脑端数字人界面功能测试

7.2.1 与 LEO 交互

- **测试目标：** 确保用户可以通过电脑端数字人模型与 LEO 进行语音交互，并正确接收 LEO 的反馈结果。
- **测试步骤：**
 - 启动电脑端数字人界面，进入对话模式。
 - 通过语音输入或文本输入与 LEO 进行交互。
 - 确认语音或文本被准确提交给 LEO，并收到正确的反馈结果。

7.2.2 与手机端备忘录互通

- **测试目标：** 确保电脑端数字人可以与手机端 App 事件相关物品栏信息实现互通，并能正确读取和更新事件相关物品栏内容。
- **测试步骤：**
 - 在手机 App 中添加或修改事件相关物品栏内容。
 - 在电脑端数字人界面中验证事件相关物品是否能正确同步并显示最新的事件相关物品内容。

7.2.3 场景导航

- **测试目标：** 确保电脑端数字人能够正确使用场景图进行导航，并成功帮助用户寻找指定物品。
- **测试步骤：**
 - 在电脑端数字人界面中发出寻找物品的指令。
 - 确认数字人模型能够在场景图中正确导航，并找到指定物品的位置。

7.3. 手机摄像头连接与 AR 技术导航功能测试

7.3.1 导航功能

- **测试目标：** 确保用户可以利用手机摄像头连接与 AR 技术实现在找寻物品时的导航功能。
- **测试步骤：**
 - 打开导航功能页面，并允许手机使用摄像头。
 - 对手机摄像头进行扫描，并观察是否能成功识别场景和指示物品。
 - 移动手机，确认导航指示是否正确显示。

7.4. VR 设备连接与交互功能测试

7.4.1 连接功能

- **测试目标：** 确保用户可以通过 VR 设备与 LEO 进行连接交互。
- **测试步骤：**
 - 启动 VR 设备并连接到软件。
 - 进行 VR 设备与 LEO 的交互操作，如语音输入等。
 - 确认与 LEO 的交互是否正常，能否收到正确的反馈结果。

7.5. 兼容性测试

7.5.1 Windows11 兼容性测试

- **测试目标：** 确保软件在 Windows11 操作系统下正常运行。
- **测试步骤：**

- 在 Windows11 系统上安装并运行软件。
- 验证软件的各项功能是否正常工作，包括语音交互、事件相关物品栏管理和数字人界面交互等。

7.5.2 安卓兼容性测试

- **测试目标：** 确保手机 App 在安卓操作系统上正常运行。
- **测试步骤：**
 - 在安卓手机上安装并运行手机 App。
 - 验证手机 App 的各项功能是否正常工作，包括语音交互和事件相关物品栏管理等。
 - 以上是针对基于 LEO 的跨平台数字人管家的功能测试计划，通过执行这些测试，可以确保软件在不同平台上的稳定性和可靠性。

8. 可扩展性与未来规划

8.1 增加更多功能

未来版本可以考虑增加智能推荐功能，在 LEO 原有参数的情况下进行训练微调。根据用户的偏好和历史记录推荐相关信息或服务，提升用户体验。例如，根据用户的运动习惯和健康数据推荐适合的健身计划或营养餐单，指出实现以上活动相关物品所处位置的导航。

另外，可以引入日程管理功能，让用户可以通过 LEO 助理安排和管理日常事务，包括会议安排、提醒事项等，提高用户的工作效率和生活质量。

8.2 发展规划

未来版本的发展规划包括技术升级和用户体验改进：

- 技术升级方面，可以考虑引入更先进的自然语言处理技术，提升 LEO 的语音识别和对话处理能力，以及引入更智能的推荐算法，提高智能推荐的准确性和个性化程度。
- 用户体验改进方面，可以持续优化界面设计和交互体验，根据用户反馈不断优化产品功能和性能，提高用户满意度。

9. 风险与问题

9.1 技术难点

在开发过程中会遇到技术难点，语音识别的准确率不高、AR 技术与手机摄像头的兼容性问题等。解决方案可以是持续跟踪技术发展，及时应用新的技术和算法。

9.2 用户隐私保护

由于涉及到用户的个人信息和对话记录，隐私保护是一个重要的问题。解决方案可以是加强数据加密和权限控制，确保用户数据安全；制定严格的隐私政策，明确用户数据的使用和保护规定；并且进行定期的安全审查和漏洞修复，及时应对潜在的安全风险。