中国研究生操作系统开源创新大赛

项目功能说明书

|  |  |
| --- | --- |
| **学 校：** | **国防科技大学** |
| **参 赛 队 伍：** | **银河galax** |
| **队 伍 成 员：** | **李昌建 刘鹏程 张铸玮** |
| **指 导 教 师：** | **李爱平** |
| **完 成 日 期：** | **2024年8月** |

目录

[第 1 章 绪 论 1](#_Toc174453684)

[1.1 目的与意义 1](#_Toc174453685)

[1.2 项目背景 1](#_Toc174453686)

[第 2 章 技 术 理 论 2](#_Toc174453687)

[2.1 设计思想 2](#_Toc174453688)

[2.2 技术路线 2](#_Toc174453689)

[2.3 代码原创说明 3](#_Toc174453690)

[第 3 章 软件介绍 4](#_Toc174453691)

[3.1 软件功能介绍 4](#_Toc174453692)

[3.1.1 语音识别模块功能介绍 4](#_Toc174453693)

[3.1.2 文本转写模块功能介绍 4](#_Toc174453694)

[3.1.3 音频转换模块功能介绍 5](#_Toc174453695)

[3.1.4 账号管理模块功能介绍 5](#_Toc174453696)

[3.2 软件界面 5](#_Toc174453697)

[3.2.1 语音识别模块界面 6](#_Toc174453698)

[3.2.2 文本转写模块界面 6](#_Toc174453699)

[3.2.3 音频转换模块界面 6](#_Toc174453700)

[3.2.4 账号管理模块界面 6](#_Toc174453701)

[第 4 章 软件测试 7](#_Toc174453702)

[4.1 软件测试说明 7](#_Toc174453703)

[4.2 软件测试结果 7](#_Toc174453704)

[4.2.1 语音识别模块测试 7](#_Toc174453705)

[4.2.2 文本转写模块测试 7](#_Toc174453706)

[4.2.3 音频转换模块测试 7](#_Toc174453707)

[4.2.4 账号管理模块测试 7](#_Toc174453708)

[第 5 章 实现难点说明 8](#_Toc174453709)

[第 6 章 总 结 9](#_Toc174453710)

[参 考 10](#_Toc174453711)

# 绪 论

## 目的与意义

开放麒麟 (英语：openKylin)，是由基础软硬件企业、非营利性组织、社团组织、高等院校、科研机构和个人开发者共同创立的一个开源社区，旨在以“共创”为核心、以“开源聚力、共创未来”为社区理念，在开源、自愿、平等、协作的基础上，通过开源、开放的方式与企业构建合作伙伴生态体系，共同打造桌面操作系统顶级社区，推动Linux 开源技术及其软硬件生态繁荣发展。(wiki, 2022)该应用主要面向教育、企业、媒体和司法等领域，用户可以在课堂、会议、采访或司法场合中使用它实现音频录制、实时转文字及摘要提取功能，为用户提供便捷的记录和信息处理工具，有效提高学习、工作和决策的效率。应用旨在通过AI大模型技术，帮助用户高效记录和处理信息，减少人工记录的负担，并自动提取文本中的关键内容，使用户能够快速获取和利用核心信息，提升整体信息处理效率，同时展示AI与国产操作系统相结合的创新能力。该应用基于openKylin操作系统开发，展示了国产操作系统在人工智能领域的应用潜力，丰富了其应用生态，推动了国产软件的发展；通过满足多样化的市场需求，有助于提高信息记录的准确性和处理效率，促进了社会信息化进程的发展。

## 项目背景

基于OpenKylin操作系统，开发一款音频录制及实时转文字并提取摘要的应用，具有显著的创新性和实用性。当前领域的类似软件，如Google的Speech-to-Text和IBM的Watson Speech to Text，都在音频转文字功能上表现出色，但存在一些不足。例如，这些软件在复杂环境下的噪声干扰处理能力有限，且对某些专业术语或方言的识别准确率不高。此外，现有的摘要提取功能往往需要用户手动干预来提高准确性和相关性。

本应用的创新之处在于结合了深度学习和自然语言处理技术，通过优化音频预处理和模型训练，显著提高了在噪声环境中的识别准确率。同时，利用先进的生成式大语言模型能够实时生成更加准确和有用的文本摘要，减少了人工干预的需要。同时，设计中还充分考虑了用户体验，提供了友好的界面和简洁的操作流程，使得音频录制和文本处理变得更加高效和便捷。

。

# 技 术 理 论

## 设计思想

在软件设计过程中，为实现音频录制、实时转文字和摘要提取功能，采用了以下研究方法和开发思想：

需求分析与用户研究。首先，通过调研用户需求和市场现有产品，明确应用的核心功能和用户痛点。这一阶段的目标是确保软件设计能够满足实际需求，并解决现有软件的不足。模块化设计，软件设计中采用模块化思想，将音频录制、文字转录、摘要提取等功能分成独立模块。这样不仅便于开发和维护，也可以根据需要对某些模块进行优化或替换。用户体验设计，在软件开发中重视用户界面的设计和用户体验，确保操作简便、界面友好。通过原型设计和用户测试，及时收集反馈并进行迭代改进。持续集成与测试，在开发过程中采用持续集成（CI）和持续部署（CD）的方法，确保各模块的功能能够无缝整合，并通过自动化测试发现并解决潜在问题。优化与迭代，基于实际使用情况，持续优化算法和功能，改进系统性能，提升识别准确性和处理速度。

## 技术路线

该软件的技术路线基于前后端分离的架构，结合了现代前端框架、后端开发框架，以及大语言模型API，以实现音频录制、实时转文字及摘要提取功能。

前端部分采用Vue 3框架，Vue 3作为一种渐进式JavaScript框架，具备灵活的组件化结构和强大的响应式数据绑定能力，能够为用户提供高效、直观的操作界面 (Vue.js, 2023)。在前端，用户可以通过简单的界面进行音频录制，并查看实时生成的文字内容。Vue 3的现代化特性，如Composition API和单文件组件，使得开发过程更加模块化和易于维护，同时支持热重载功能，提升了开发效率和用户体验。

后端采用Flask框架，Flask是一个轻量级的Python Web框架，易于扩展和部署 (Flask, 2010)。Flask主要负责处理来自前端的请求，如音频数据的接收、与大语言模型API的通信、以及转录和摘要结果的返回。通过设计RESTful API，Flask能够高效管理前后端的数据交互和逻辑处理。Flask的灵活性使其适合与各种第三方服务（如语言模型API）集成，同时也可以通过扩展库来增强功能，例如实现用户身份验证和日志记录等。

在音频处理和文本生成方面，软件集成了大语言模型API，用于实现音频转文字和自动摘要提取功能。音频数据首先通过后端发送至大语言模型API，由模型进行语音识别和文本生成，然后将生成的文字内容和摘要结果返回给前端进行展示。大语言模型的强大自然语言处理能力确保了转录的高准确性和摘要的高相关性。

整体技术路线通过Vue 3提供流畅的前端体验，结合Flask实现稳健的后端逻辑，并依托大语言模型API的强大处理能力，构建出一款功能强大、用户体验优越的音频处理应用。这一技术架构具有良好的扩展性，能够根据未来需求进一步扩展功能或优化性能。

## 代码原创说明

本软件前端界面调用了element-plus开源组件库，完全应用行数，150+行。

# 软件介绍

## 软件功能介绍

该软件的总体功能旨在为用户提供高效、便捷的音频转写及文本处理服务，满足多样化的音频输入需求，并自动生成高质量的书面文本。用户可以通过上传音频文件或实时录制音频来使用该软件。音频数据输入后，软件将利用先进的语音识别技术，将语音内容转写为文本，并提供保存功能，便于后续的编辑和引用。

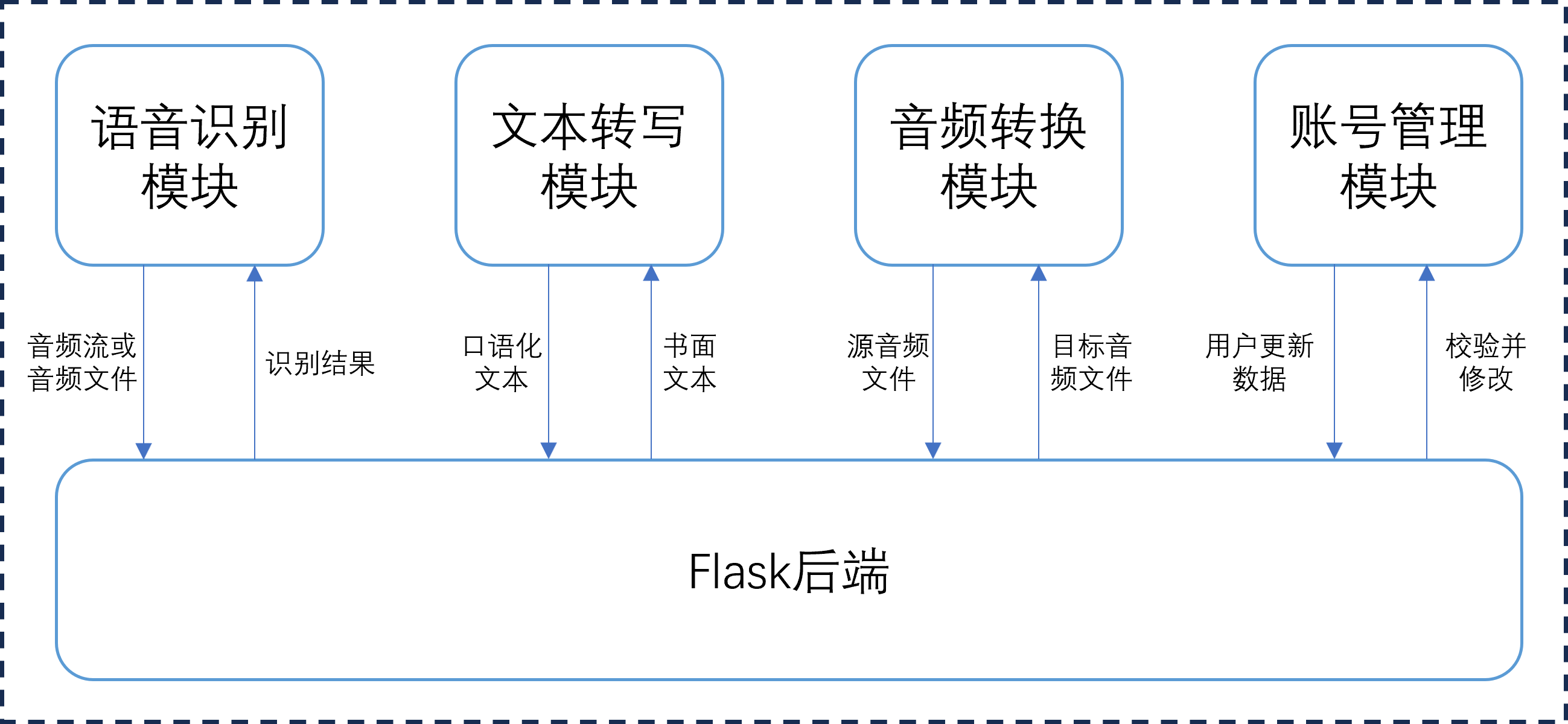
软件的核心功能模块包括以下几部分：语音识别模块、 文本转写模块、音频转换模块、账号管理模块。

图 1软件总体模块与结构

### 语音识别模块功能介绍

该模块负责处理音频输入，并利用大语言模型进行语音识别，将语音内容精确地转录为文本。无论是实时录音还是上传的音频文件，该模块都能高效处理，并支持多种音频格式。语音识别过程还考虑了复杂的语音环境，提升了转写的准确性

### 文本转写模块功能介绍

在文本转写过程中，该模块不仅能将口语化的文本转录为书面形式，还会自动调整语法结构和措辞，使输出的文本更加正式和规范。这一功能特别适用于需要将会议记录、采访内容等口语化信息转化为书面文档的场景。

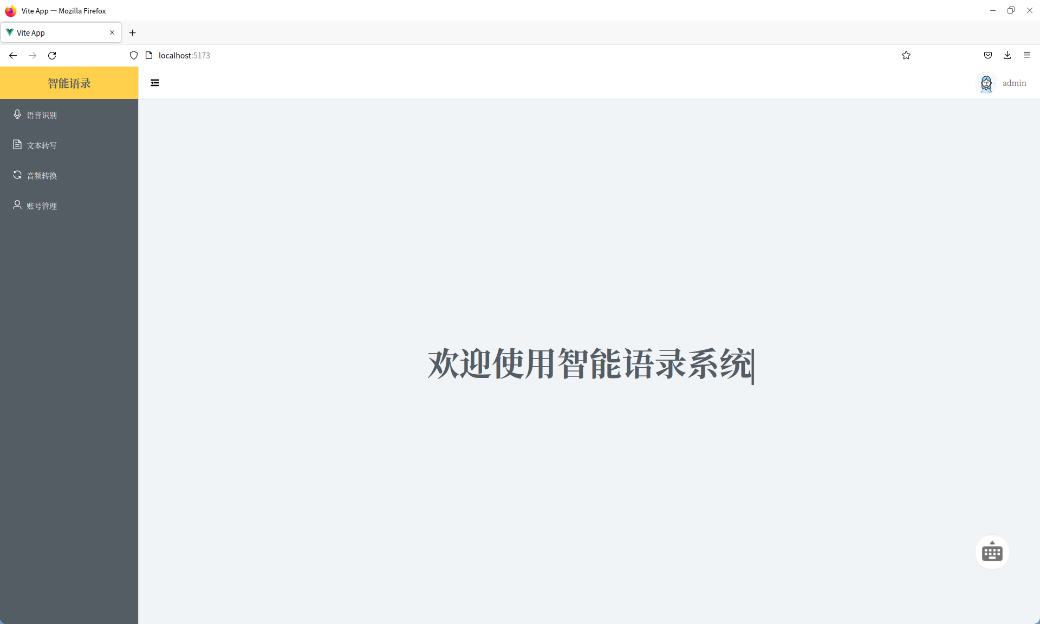
### 音频转换模块功能介绍

为了提升系统的兼容性和用户体验，该模块支持将各种格式的音频文件转换为系统支持的标准格式，从而确保音频文件能够顺利被识别和转写。

### 账号管理模块功能介绍

用户可以通过该模块进行账号注册、登录和管理。系统支持个性化设置，用户可以保存历史转写记录，并根据需求对文本进行再次编辑或导出。该模块还提供安全性保障，确保用户数据的隐私和安全。

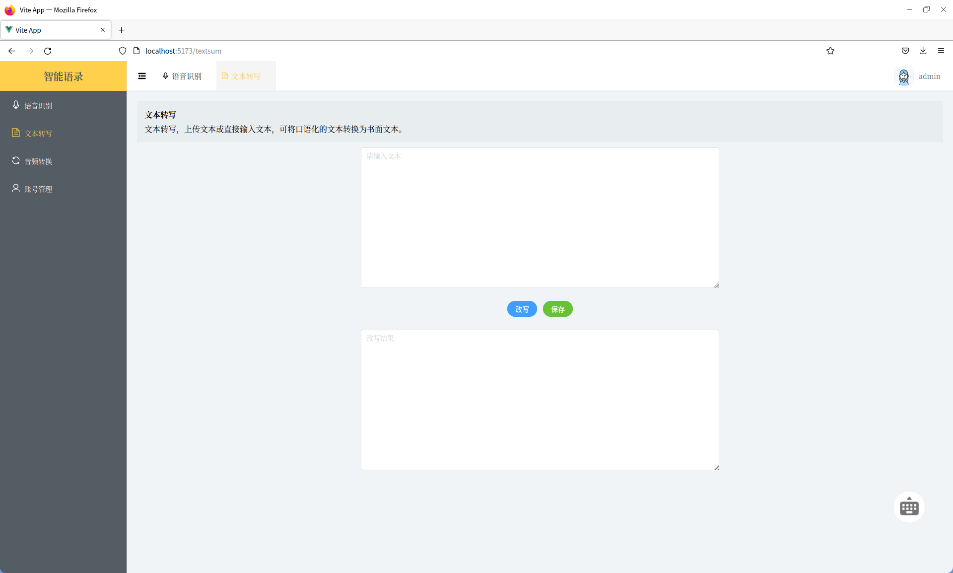
## 软件界面

在开发音频录制、实时转文字及文本摘要提取的软件过程中，我们采用了模块化设计、单一职责原则、面向接口编程以及用户中心设计等关键设计原则。模块化设计将软件划分为独立的功能模块，如语音识别、文本转写、音频转换和账号管理模块。每个模块专注于特定任务，确保功能清晰、代码可维护性强。遵循单一职责原则，每个模块只负责一项具体功能。例如，语音识别模块只处理音频转文字，文本转写模块则专注于将口语化的文本转换为正式的书面语。这样的设计减少了模块之间的耦合，使系统更具灵活性。面向接口编程的应用进一步增强了模块之间的解耦性，各模块通过定义明确的接口进行交互。这种设计方法提高了系统的可扩展性和维护性，当需要升级或替换某个模块时，其他模块无需调整。最后，用户中心设计贯穿整个开发过程，确保软件界面友好、操作简便。用户可以轻松上传音频或实时录音，快速获取准确的转写和摘要结果，从而提高工作效率。这些设计原则共同作用，构建出一个功能强大、灵活且用户体验优越的软件系统。

### 语音识别模块界面

该模块主要提供实时语音识别、音频文件识别以及录音文件和转换结果文件保存的功能。使用实时语音识别功能时，点击开始录音按钮并开始说话，识别的结果会实时显示在下面的文本框中，说话结束后点击停止按钮。如果对于识别的结果不满意可以点击重试，重新录音。使用保存音频按钮可以下载音频文件，保存文本按钮可以下载文本文件。使用音频文件识别功能时，点击或将文件拖入上传框可以上传文件，然后系统会将识别的结果显示在下面的文本框中。

### 文本转写模块界面

该模块提供文本转写功能。能够将由录音转换的口语化的文本转换为较为正式的书面文本。使用时在上面的输入框中输入需要转写的文本内容，然后点击改写按钮。转换的结果会显示在下面的文本框中，如果需要保存点击保存按钮即可。

### 音频转换模块界面

该模块提供音频文件转换功能。能够上传音频文件并进行对应的格式转换。使用时先选择需要转成的文件格式，然后点击或将文件拖入上传框中可以上传文件，转换完成后点击保存可以下载转换后的音频文件

### 账号管理模块界面

该模块主要提供用户信息修改功能，包括用户名和密码等，用户通过输入新的用户名和密码，点击提交按钮进行修改。

# 软件测试

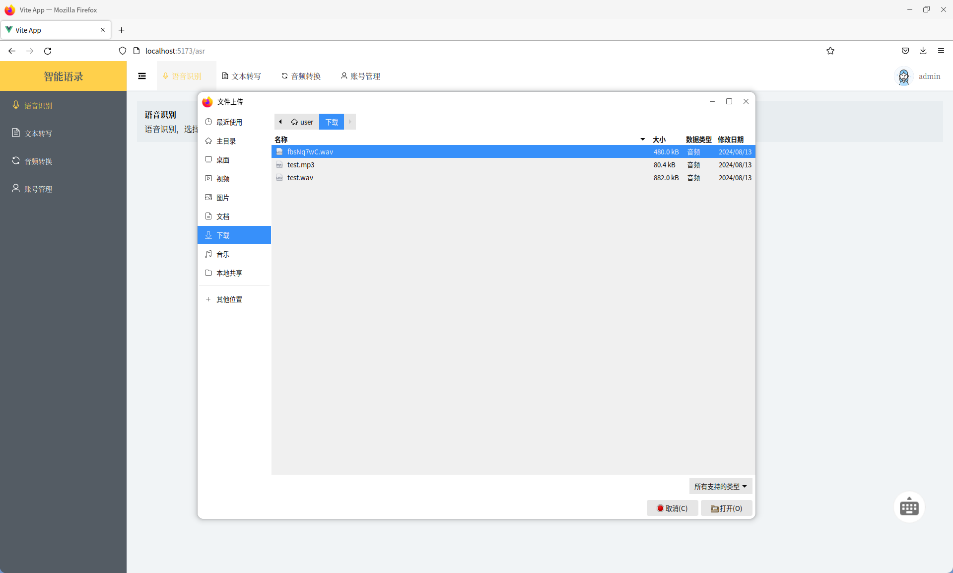
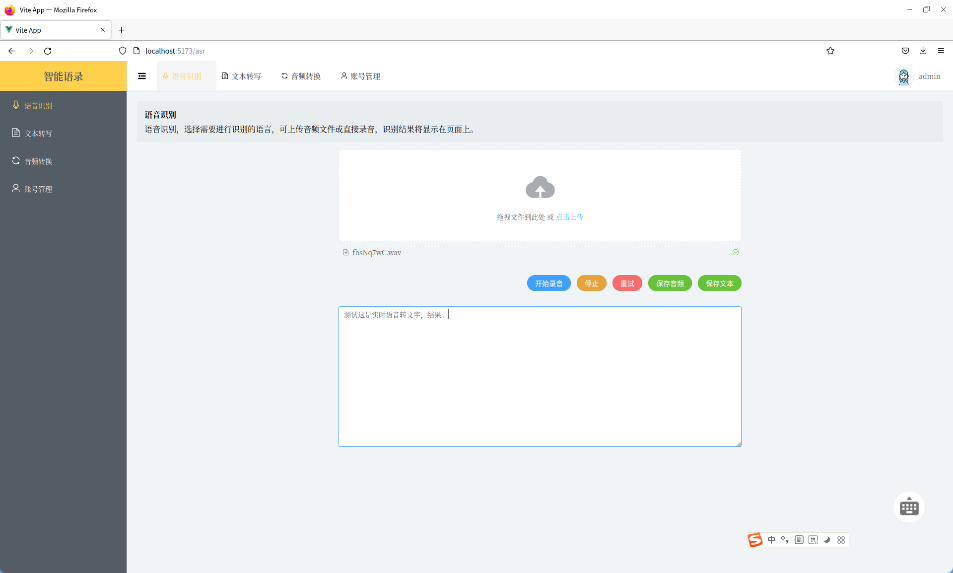
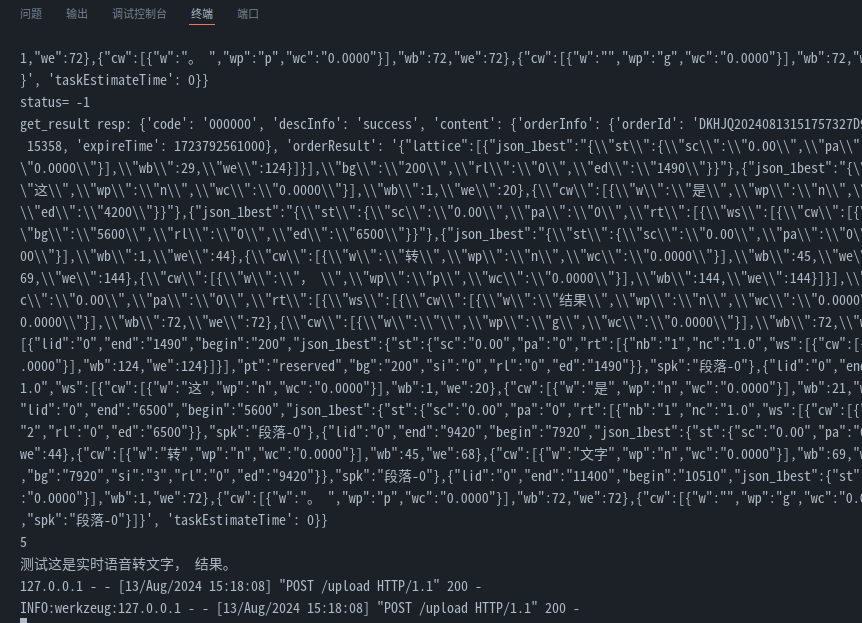
## 软件测试说明

为了确保该音频录制、实时转文字及文本摘要提取软件的功能正常工作，我们设计了一套全面的软件测试方案，我们通过模拟实际使用场景，测试语音识别、文本转写、音频转换和账号管理模块的基本功能是否正确，这些测试确保每个模块能够在实际的使用场景中正常工作。

## 软件测试结果

### 语音识别模块测试

实时语音转文字测试，开始后以及点击下载的识别结果与下载结果如图所示，以及后端的http请求响应结果如图所示。

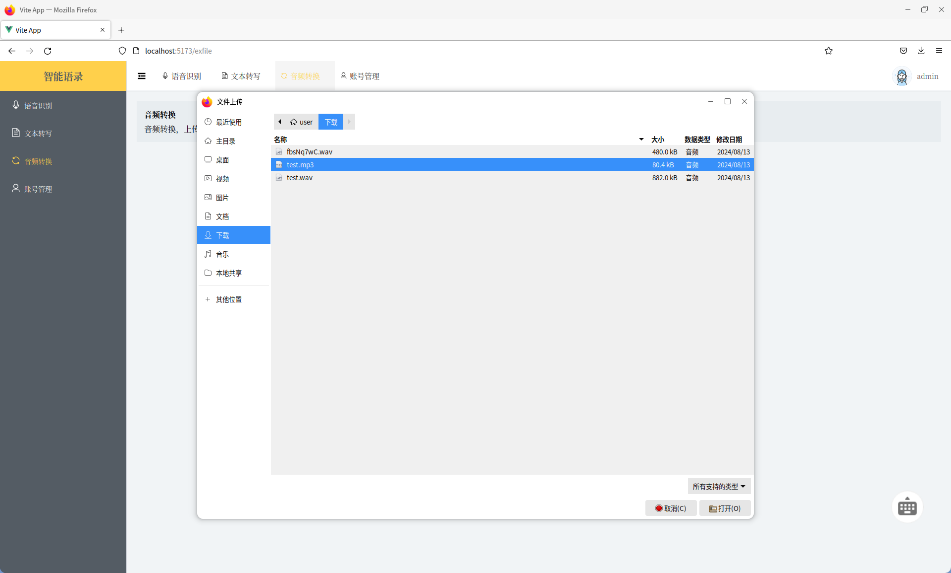


音频文件识别测试，打开文件上传列表，以及识别结果如图所示。后端http请求响应结果如图所示。

### 文本转写模块测试

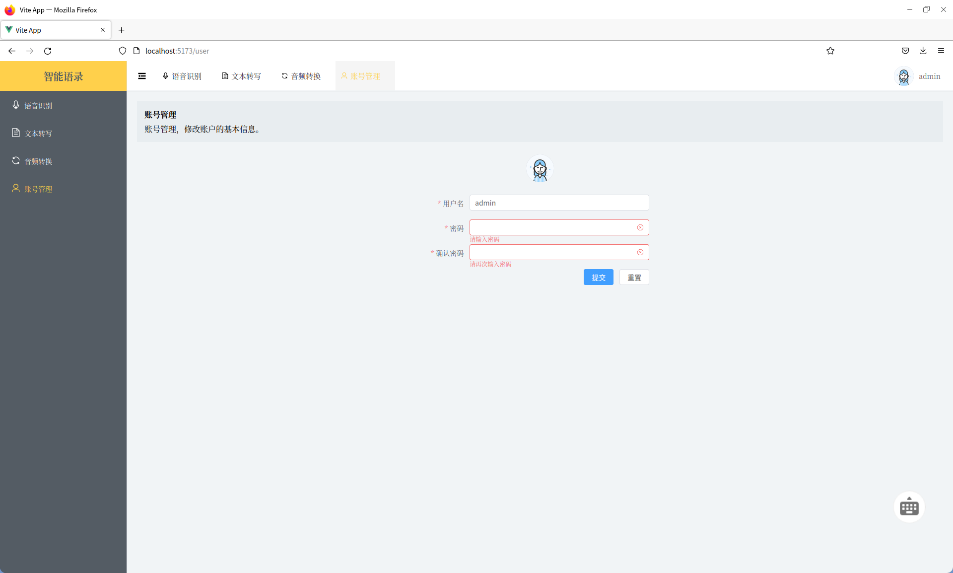
文本转写，转写结果如图所示，以及后端的http请求响应结果如图所示。

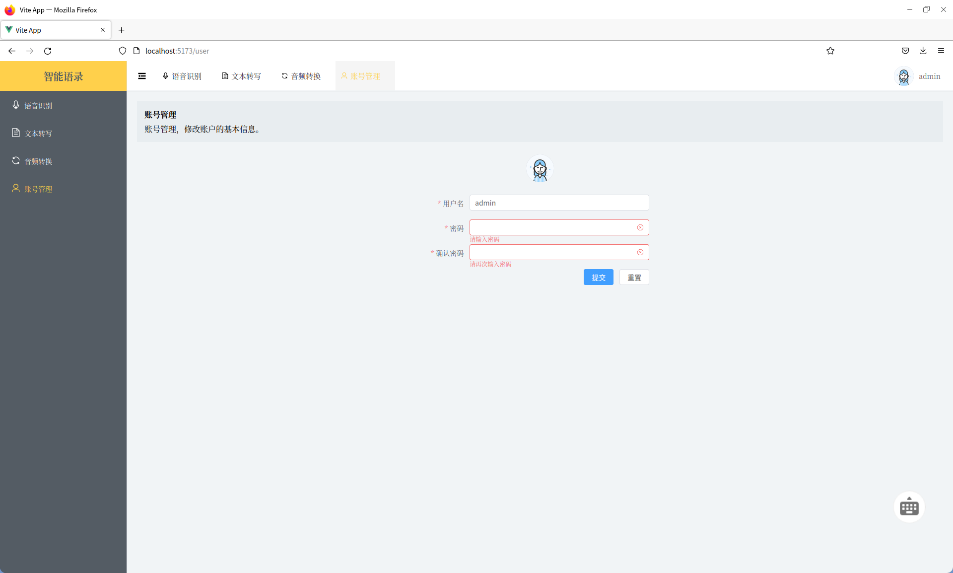
### 音频转换模块测试

音频文件转换，打开上传列表和完成转换的下载结果如图所示，后端http请求响应结果如图所示

### 账号管理模块测试

该模块开发并未完全完成，目前暂时只是实现了前端的界面和前端的表单校验





# 实现难点说明

略

# 总 结

项目伊始，我们进行了详细的需求分析，明确了软件的核心功能。我们确定了软件需要支持音频文件上传和实时录音，并能够高效地将音频内容转写为文本，同时提供自动摘要功能。在此基础上，我们制定了详细的技术方案，采用了前后端分离的架构，前端使用Vue 3，后端采用Flask，并集成了大语言模型API以实现关键的语音识别和文本处理功能。在开发过程中，我们遵循了模块化设计、单一职责原则、面向接口编程以及用户中心设计等核心原则。这些设计原则帮助我们构建了一个结构清晰、功能明确的软件系统。我们通过定义明确的模块和接口，确保了系统的可扩展性和可维护性。同时，我们能够在开发的各个阶段及时发现并修复问题，保证了软件的稳定性和功能的完整性。在项目的测试阶段，我们指定了详细的测试方案，使软件能够在实际的使用场景中正常工作。总的来说，项目的整体开发并未完全完成，目前只实现了软件的基础功能。未实现功能展望：1、音频文件优化，使用降噪算法去除音频文件中的噪声增强识别的准确性并提高录音的质量。2、支持更多格式的音频文件转换，目前的项目中只支持wav和mp3两种格式的音频文件，在后续的计划中还将更新对更多音频文件的支持。3、提供文本摘要和总结功能：目前对文本的处理仅限于转写功能，预期能够实现对文本进行摘要和总结。4、实现完整的账号管理模块，接入用户数据库实现登录功能。

此次参赛过程中，我学到了很多宝贵的知识和技能，也积累了丰富的项目经验。首先，在技术层面，我深入掌握了前后端分离架构的开发方法，以及如何将复杂的AI模型集成到实际应用中。同时，我也深刻体会到软件开发中的设计原则和最佳实践的重要性，它们不仅提高了开发效率，还显著提升了代码质量和系统的可维护性。另外，团队合作和项目管理经验也是此次参赛的一大收获。在项目中，我们通过合理分工与高效沟通，顺利完成了各项任务。这让我认识到，成功的软件项目离不开良好的团队合作和科学的项目管理。总的来说，此次参赛不仅让我在技术上得到了提升，还增强了我在软件设计、开发、测试以及项目管理方面的综合能力。这些经验和收获将对我未来的学习和职业发展产生深远的影响。

参 考

维基百科. 开放麒麟. <https://zh.wikipedia.org/wiki/>开放麒麟.[2022-07-22]

Vue.js. 什么是Vue3？.https://cn.vuejs.org/guide/introduction.[2023-12-21]

Flask. Introduction of Flask. <https://flask.palletsprojects.com/en/3.0.x/>. [2010-3-5]