|  |  |
| --- | --- |
|  | **第十八届全国大学生软件创新大赛-软件设计创新赛** |
| **文档编号：SWC2025-慧议先锋** |



**慧议**

**AI Meeting**

**技术研究报告**

**Version: 1.0**



**慧议先锋**

**2025.3.3**

**All Rights Reserved**

目录

[1 问题聚焦 1](#_Toc56719301)

[1.1 问题描述 1](#_Toc56719302)

[1.2 问题抽象 1](#_Toc56719303)

[1.3 问题定位 1](#_Toc56719304)

[1.4 问题评估 1](#_Toc56719305)

[1.5 问题分解 1](#_Toc56719306)

[2 相关工作 2](#_Toc56719307)

[3 技术方案 3](#_Toc56719308)

[3.1 技术方向 3](#_Toc56719309)

[3.2 技术选择 3](#_Toc56719310)

[3.3 结果期望 3](#_Toc56719311)

[4 技术实践 4](#_Toc56719312)

[4.1 使用的开发框架及依赖的库 4](#_Toc56719313)

[4.2 技术实践过程 4](#_Toc56719314)

[5 结果验证 5](#_Toc56719315)

文档修订历史

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **序号** | **修订原因** | **版本号** | **作者** | **修订日期** | **备注** |
| **1** | **初版发布** | **1.0** | **慧议先锋** | **2025.3.3** |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

# 问题聚焦

## 问题描述

在会议场景中，利用**人工智能技术**实现会议**全流程**的智能化管理与优化，包括高效的会议**创建与组织**、精准的会议内容**记录与分析**、全面的参会者**行为洞察**以及多**语言交流**的无缝对接，以实现提升会议的效率、质量和协作效果。

## 问题抽象

将**会议组织流程**转化为基于自然语言处理的**指令解析**与**任务自动化执行**问题，例如通过 AI 创建会议，实现从自然语言指令到会议日程、邀请等一系列操作的自动生成。

把**会议内容记录**与分析抽象为**语音识别**、**信息提取**、**语义理解**与**总结概括**等技术问题，如根据不同会议内容对语音和文本内容进行**针对性处理**。

参会者行为分析可抽象为**数据挖掘**与**可视化**问题，通过对参会时长、发言频率、互动情况等数据的挖掘分析参会者的**参与度模式**。

**多语言交流**问题转化为**机器翻译技术**与**实时语音处理技术**的结合，确保不同语言背景的参会者能够实时准确地交流。

## 问题定位

**业务领域：**会议管理与协作领域，旨在提升各类会议（企业会议、学术会议、国际交流会议等）的智能化水平和效率。

**技术领域：**涉及自然语言处理、机器学习、计算机视觉、数据挖掘与分析、实时通信技术、数据可视化等多个技术领域的综合应用。

## 问题评估

**技术性：**具有较高的技术性要求，需要整合多种先进技术并解决技术融合中的兼容性、性能优化等问题。例如在智能会议笔记助手功能中，涉及到复杂的文本分类、信息提取和语义总结算法的开发与优化，在参会者行为分析功能中，涉及到将来自不同数据源和传感器的数据进行综合处理，并对融合后的数据进行进一步的处理分析和可视化。

**普适性：**该问题在当今全球化、数字化办公与交流日益频繁的背景下具有广泛的普适性。几乎所有涉及会议组织与协作的行业和场景都能受益于智能化的会议解决方案，市场需求大且持续增长。

**热度：**随着人工智能技术的发展，会议智能化是当前研究与开发的热点领域之一。众多科技企业和研究机构都在投入资源进行相关技术的研发与创新，相关技术论文和专利数量不断增加，行业关注度高。

## 问题分解

**1.5.1会议创建与组织智能化子问题：**

难度：高

描述：实现自然语言指令到会议日程、房间号生成等操作的准确转换与执行。需要构建强大的自然语言理解模型，理解用户创建会议的意图，并与会议管理系统无缝对接。

依赖关系：依赖于**自然语言处理技术**的基础模型构建与训练，同时需要与**会议数据库**和**通信模块**协同工作。

**1.5.2智能会议内容记录与分析子问题：**

1.语音识别准确性子问题：

难度：高

描述：在不同会议环境（嘈杂、多人同时发言等）下实现高精度的语音识别，准确区分不同说话人，涉及声学模型、语言模型的优化以及说话人识别技术的应用。

依赖关系：依赖于大量语音数据的收集与标注，以及深度学习算法在语音处理中的有效应用。

2.信息提取与分类子问题：

难度：中等

描述：根据不同会议形式准确提取关键信息并分类，如从报告式会议中提取关键语句、从研讨式会议中提取研究路线和专有名词等。需要开发针对不同会议类型的文本分类和信息提取算法，并结合语义理解技术。

依赖关系：依赖于大规模会议文本数据的训练，以及自然语言处理中的文本分类、命名实体识别等技术的成熟应用。

3.会议内容总结与概括子问题：

难度：中等

描述：对会议内容进行语义理解基础上的**总结概括**，生成有价值的会议笔记。涉及到文本摘要技术、语义分析技术的深度应用。

依赖关系：依赖于前面信息提取与分类子问题的结果，以及先进的文本摘要算法和语义理解模型。

**1.5.3参会者行为分析子问题：**

1.数据收集与整理子问题：

难度：中等

描述：准确收集参会者的发言时间、频率、互动行为等数据，并进行有效的整理和存储，需要在会议系统中设计合理的数据采集点和数据结构。

依赖关系：依赖于会议基础功能模块的数据输出，以及数据库技术的支持。

2.行为模式识别与分析子问题：

难度：高

描述：通过数据挖掘和机器学习算法识别参会者的行为模式，评估参与度。需要开发适合会议行为分析的算法模型，并进行大量数据的训练与优化。

依赖关系：依赖于数据收集与整理子问题提供的数据基础，以及机器学习算法的有效应用。

**1.5.4多语言交流子问题：**

1.实时翻译准确性子问题：

难度：高

描述：实现不同语言之间的实时准确翻译，处理语言的多样性、专业性和口语化表达。需要构建高质量的机器翻译模型，结合语言模型和翻译记忆技术。

依赖关系：依赖于大规模多语言语料库的积累与训练，以及翻译算法的优化与创新。

2.语音翻译同步子问题：

难度：较高

描述：确保语音翻译与会议语音流的同步，避免延迟和卡顿。涉及到实时语音处理技术、音频流同步技术的应用。

依赖关系：依赖于实时翻译准确性子问题的解决，以及音频处理技术的支持。

# 相关工作

1. **腾讯会议：**是一款广泛使用的视频会议软件，具备基础的会议日程创建、成员邀请、屏幕共享、音视频通话等功能，在多人视频会议的稳定性和画质方面表现较好。然而，其智能会议记录功能**相对简单**，仅能提供基本的语音转文字记录，**无法根据会议内容生成相应的热点词云图和情感分析等**便于了解会议动态与参会者状态的图表。在**参会者参与度评估**方面几乎没有涉及，实时翻译功能虽然支持多种语言，但在翻译**准确性和专业性**上还有提升空间。
2. **钉钉会议：**在企业办公场景中有较高的使用率，除了基础会议功能外，与办公协同功能有一定整合。但在智能会议记录方面，同样只是简单的语音转文字，缺乏对不同会议类型的深度理解与信息处理。参会者行为分析仅能提供一些基础的参会数据统计，如参会时长等，**不够全面深入**。对于AI创建会议、专业的**会议笔记助手**以及对会议内容的**情感分析和词云图绘制**等功能**尚未完善**。
3. **飞书会议：**注重团队协作与沟通效率，在会议互动功能上有一定特色。但在智能会议记录和分析方面，未能很好地满足根据会议形式**定制化**记录和分析的需求。**实时翻译功**能在应对复杂会议场景时的灵活性不足，且缺少对会议**整体数据**的**深度挖掘与可视化呈现**，如参会者参与度的综合分析与展示。

相比之下，本项目的AI +会议智能体开发具有显著创新性。通过AI创建会议，可根据用户指令**快速生成**个性化会议安排；智能会议笔记助手能够依据不同**会议类型精准提取**关键信息并进行**深度处理**；全面的参会者行为分析涵盖发言频率、互动情况等**多维度数据**，并结合**情感分析**和**词云图**绘制，为会议提供更深入的数据洞察。同时，在**实时翻译**功能上致力于提高**准确性**和**专业性**，更好地服务于国际会议等多语言交流场景，从而为用户提供更智能、高效、全面的会议体验，填补了现有市场竞品在这些功能方面的不足与空白。

# 技术方案

## 技术方向

本项目采用的技术涵盖多个方向，在通信与网络技术方面，使用 Zoom Video SDK 和 WebRTC 实现实时视频会议，结合 Web Audio API、Audio Worklet、WebSocket 以及 Paraformer、CT - Transformer 模型完成实时转录与音频处理；人工智能与机器学习领域，调用开源讯飞大语言模型 API 进行文本处理，运用 RoBERTa 模型开展自然语言情感分析；数据处理与分析方向，通过整合多维度会议数据实现会议参与度评估；安全技术上，利用 JWT 进行用户身份验证和权限控制；数据库技术方面，以 Firestore 实现用户信息和会议数据的存储与管理；前端开发则基于 Vue.js 框架和 Vite 构建工具，提升开发效率和界面交互性。

## 技术选择

### 视频会议 SDK 和 WebRTC 技术

为了实现实时视频会议功能，本项目使用了 **Zoom Video SDK** 和 **WebRTC** 技术。Zoom Video SDK 提供了一个强大的开发框架，使开发者能够快速构建具有视频和音频通信功能的应用程序。WebRTC（Web Real-Time Communication）是实现浏览器与浏览器之间实时通信的开源技术，支持音频、视频通话和文件共享。在本项目中，WebRTC 用于处理视频流的传输，并确保了低延迟的视频体验。

### 调用开源大语言模型API

使用 RESTful API 调用开源**讯飞大语言模型**服务，对文本进行摘要生成、问答等处理，通过集成这些API，实现文本数据的深度理解与生成。并且在前端与后端通过 API 请求进行交互，保证请求的高效性与稳定性。

### **自然语言处理（NLP）技术**

运用情感分析技术，借助由 Facebook AI 开发的基于BERT 的改进版情感分析**RoBERTa 模型**，实时获取会议记录文本的情感分析结果，并以八种不同语调分类，同时绘制情感分布条形图、饼图和雷达图，以直观展示会议中的情感信息。

### 实时转录与音频处理技术

本项目结合 Web Audio API 和 Audio Worklet，实现了音频流的实时处理。利用 WebSocket 技术将音频数据传输至基于对于大规模语料库时的处理有较快的推理速度和出色的识别性能**的Paraformer 模型**构建的转录服务，确保了音频数据能准确且迅速地被转录，然后将转录出的文本结果使用CT-Transformer模型进行标点预测和不流畅性检测，最终实现了低延迟的音频数据处理和即时字幕展示。

### 会议参与度评估计算方法

本项目通过整合会议中的各类数据，如参会者的进入离开时间、发言时长、视频时长等，将其分成行为、认知和情感参与度三个维度，分别采用不同的方法进行计算，最后得到综合的参与度得分，并生成可导出的数据统计表格，实现全面的会议参与度评估，帮助用户深入分析会议效果和参与者的贡献，为人事考核或团队管理提供有力依据。

### 安全性与身份验证：JWT

在项目中，通过 **JWT** (JSON Web Token) 技术实现用户的身份验证和权限控制。JWT 提供了一种无状态的认证方式，用户登录后可以通过 JWT 获取访问令牌，并在后续请求中使用该令牌进行身份验证，确保了会议的安全性和用户的隐私保护。

### 数据存储与用户管理：Firestore

为了实现用户信息和会议数据的存储与管理，本项目使用了 Firestore 作为数据库从而实现实时存储和更新用户的会话历史、转录文本以及会议记录等信息，使得用户的行为（如加入会议、音频转录等）能够及时更新到云端数据库，并为系统的其他部分提供实时反馈。

### 前端框架：Vue.js 和 Vite

本项目的前端界面基于 Vue.js 和 Vite 构建。Vue.js 是一款渐进式 JavaScript 框架。在本项目中，Vue.js 提供了易于使用的双向绑定、响应式数据流和组件化结构，从而实现了会议创建和加入表单的动态更新。Vite 是一个现代化的构建工具，提供了更快的热更新和开发体验，在本项目中，Vite 加速了开发流程，提升了编译速度。

## 结果期望

**1.会议创建与组织：**用户能够通过自然语言指令快速、准确地创建会议，系统自动生成详细的会议日程、唯一的房间号，并高效地邀请参会成员。会议创建的成功率达到90%以上，整个创建流程在1分钟内完成。

**2.智能会议内容记录与分析：**语音识别准确率在不同会议环境下平均达到85%以上，能够准确区分不同说话人，信息提取的准确率根据不同会议类型达到80%-90%以上，会议内容总结概括能够准确反映会议的核心要点和关键信息，生成的**会议笔记**具有较高的**可读性**和**实用性**，能够有效辅助参会者回顾会议内容和进行后续工作。

**3.参会者行为分析：**能够全面准确地收集参会者的行为数据，数据收集的完整性达到85%以上。通过行为模式识别与分析，为每个参会者生成详细的参与度报告，报告中的参与度评估指标具有较高的可信度和有效性，能够为会议组织者提供有价值的决策依据，例如根据参与度情况调整会议流程、激励参会者积极参与等。

**4.多语言交流：**实时语言翻译功能能够支持至少5种主流语言的互译，翻译的准确率在日常会议交流场景下达到85%以上，在专业领域会议中达到70% 以上。语音翻译同步延迟控制在3秒以内，确保参会者能够流畅地进行多语言交流，无明显卡顿和误解。

# 技术实践

## 使用的开发框架及依赖的库

## 技术实践过程

# 结果验证