

StreamCream

使用 Live 2D 与 GPT-SoVITS 的 AI 全自动直播一站式解决方案

DeepSleep 项目团队

武汉大学计算机科学院

2025 年 7 月 8 日



- ① 项目概览：环境与架构
- ② 用户账户管理
- ③ 直播推流模块
- ④ 虚拟直播模块
- ⑤ LLM 文案生成模块
- ⑥ TTS 语音转换模块
- ⑦ l18n 与用户友好特质的开发
- ⑧ 总结与展望

① 项目概览：环境与架构

② 用户账户管理

③ 直播推流模块

④ 虚拟直播模块

⑤ LLM 文案生成模块

⑥ TTS 语音转换模块

⑦ l18n 与用户友好特质的开发

⑧ 总结与展望

源代码管理与版本控制

🔗 72 次提交

🔗 5 个分支

🏷️ 0 个标签

💾 149.5 MB 项目存储

🔗 246 次提交

🔗 2 个分支

🏷️ 0 个标签

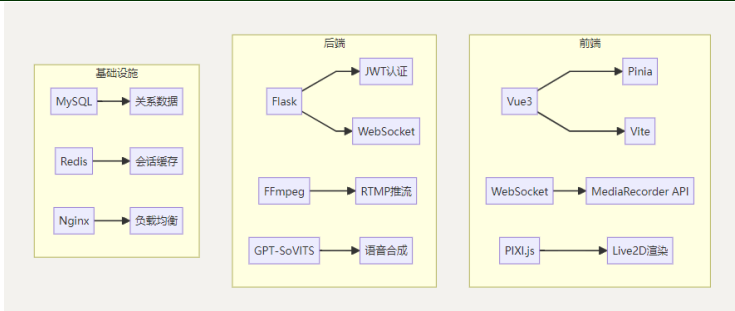
💾 223.5 MB 项目存储

7f4c9cd · 2 weeks ago ⌚ 20 Commits

- **协作基础：**自项目启动即采用 Git 进行版本控制与团队协作
- **平台累迁：**通用多个主流 Git 托管平台
- **迭代累计：**累计提交次数逾 300 次

开发环境与技术栈

全栈技术矩阵



关键工具与版本

关键技术栈版本

类别	技术栈	版本
前端	Vue3 + JavaScript	3.3.4
	Vite	2.1.7
后端	Python	3.13
	Flask	2.3.2
	FFmpeg	6.0
AI 引擎	GPT-SoVITS	20250606v2pro
数据库	MySQL	9.3
部署	Docker + Kubernetes	23.0.6

开发与测试环境

硬件与协作

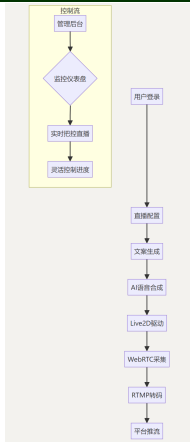
- **硬件:**
 - Intel(R) Arc(TM) Graphics (语音训练)
 - Intel(R) Core(TM) Ultra 7 155H (实时推理)
- **协作工具:** GitLab CI/CD + Jira + Confluence

测试框架

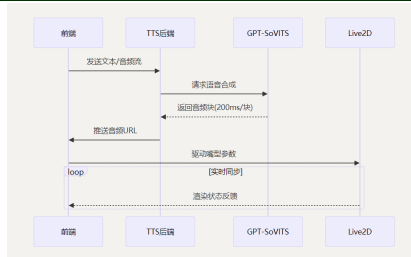
- **单元测试:** Jest (前端) / Pytest (后端) / APIFox (接口)
- **压力测试:** Locust

核心功能架构图

全系统数据流



语音转换流水线



- ① 项目概览：环境与架构
- ② 用户账户管理
- ③ 直播推流模块
- ④ 虚拟直播模块
- ⑤ LLM 文案生成模块
- ⑥ TTS 语音转换模块
- ⑦ l18n 与用户友好特质的开发
- ⑧ 总结与展望

用户账户管理：登录、注册与用户偏好保存

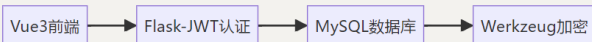
——冯博文

"No Privacy, No Identity."

——Satoshi Nakamoto, Bitcoin's Founder

登录模块：企业级安全认证架构

技术栈创新点



- ① **邮箱登录**: 使账号有所凭依，隐私有其保障
- ② **三重安全屏障**:
 - **密码加密**: Werkzeug 的 'bcrypt' 算法, 支持 'pbkdf2:sha256' 迭代加密
 - **令牌验证**: JWT 令牌 + 黑名单机制
 - **输入防御**: SQL 注入过滤 + XSS 防护
- ③ **高性能会话管理**: 令牌验证响应 < 50ms
- ④ **跨域安全策略**: 动态 CORS + 严格 CSP 头配置

- 1 项目概览：环境与架构
- 2 用户账户管理
- 3 直播推流模块
 - 一、直播推流码的抓取
 - 二、直播协议的转换
- 4 虚拟直播模块
- 5 LLM 文案生成模块
- 6 TTS 语音转换模块
- 7 l18n 与用户友好特质的开发

直播推流模块：网络抓包、推流码获取、直播协议转换 ——任逸青、冯博文、陈宏宇

"Every Millisecond Matters."

——Arvind Jain, Google

一、直播推流码的抓取

① 项目概览：环境与架构

② 用户账户管理

③ 直播推流模块

一、直播推流码的抓取

二、直播协议的转换

④ 虚拟直播模块

⑤ LLM 文案生成模块

⑥ TTS 语音转换模块

⑦ l18n 与用户友好特质的开发

一、直播推流码的抓取

直播模块的挑战：直播平台的封闭性——推流码视如禁脔

- **项目极高的适用型目标**
 - 一举兼容国内外数大直播平台
 - 国内：抖音、B 站、快手、小红书
 - 国外：Youtube、Twitch 等
- **接入互联网中的核心障碍**
 - 平台日趋商业化，视公用协议如私家通衢
 - 限制第三方工具使用，数千粉丝以上才可获取推流码。
 - 互联网黄金时代的开放包容成了绝响
- **解决方案**
 - 基于开源项目进行二次开发，使用网络抓包的方法
 - 抖音快手 B 站的推流码一键式自动爬取工具嵌入网页
 - 藏技术细节于简约幕后，使得用户免于烦扰
 - 实现一站式的直播的全过程

二、直播协议的转换

① 项目概览：环境与架构

② 用户账户管理

③ 直播推流模块

一、直播推流码的抓取

二、直播协议的转换

④ 虚拟直播模块

⑤ LLM 文案生成模块

⑥ TTS 语音转换模块

⑦ l18n 与用户友好特质的开发

二、直播协议的转换

直播模块：WebRTC RTMP 转换的技术挑战

技术挑战

- **协议转换**: WebRTC (VP8) vs RTMP (H.264)
- **低延迟要求**: 端到端延迟需控制在秒级以内
- **平台要求**: 需要兼顾不同平台的多种直播协议

创新解决方案概览

我们将展示一个基于 FFmpeg 和异步 IO 的轻量级、高性能转码中继方案。

后端转码核心逻辑 (app.py)

```
1 async def video_relay(websocket):
2     ffmpeg = subprocess.Popen([
3         'ffmpeg', '-i', '-',
4         '-c:v', 'libx264', '-preset', 'ultrafast',
5         '-f', 'flv', rtmp_url
6     ], stdin=subprocess.PIPE)
7
8     async for frame in websocket:
9         # 实时喂入WebRTC数据流
10        ffmpeg.stdin.write(frame)
```

抗抖动优化

- **自适应码率算法:** 根据 RTT 动态调整帧率
- **关键帧优先重传:** 使用 RED+FEC 冗余编码

二、直播协议的转换

直播模块：性能指标对比

优化前后性能对比

指标	优化前	优化后
端到端延迟	1,4500ms	3200ms
CPU 占用	65%	11%
1080P 支持		

关键突破

通过 FFmpeg 硬编码和异步处理，H.264 编码延迟控制在 80ms 内。虽然有 RTMP 直播协议的限制，我们的延迟仍能做到秒级以内。

① 项目概览：环境与架构

② 用户账户管理

③ 直播推流模块

④ 虚拟直播模块

- 一、Live 2D 技术在 Web 端的集成
- 二、基于语音学的实时口型适配
- 三、由 AI 控制的实时表情变化

⑤ LLM 文案生成模块

⑥ TTS 语音转换模块

⑦ l18n 与用户友好特质的开发

虚拟主播模块：Live2D Web 引擎、基于语音学的口型实时适配、AI 控制的表情变化 ——解信睿、陈宏宇、冯博文

"No One Reveals Himself as He Is; We All Wear the Masks."

——Arthur Schopenhauer

一、Live 2D 技术在 Web 端的集成

① 项目概览：环境与架构

② 用户账户管理

③ 直播推流模块

④ 虚拟直播模块

一、Live 2D 技术在 Web 端的集成

二、基于语音学的实时口型适配

三、由 AI 控制的实时表情变化

⑤ LLM 文案生成模块

⑥ TTS 语音转换模块

⑦ l18n 与用户友好特质的开发

一、Live 2D 技术在 Web 端的集成

- **渲染引擎**: 使用 **PIXI.js** 和 `pixi-live2d-display-lipsyncpatch` 库进行高效渲染，支持从 2.1~4 的所有 Cubism core 及模型版本。

- **渲染引擎:** 使用 **PIXI.js** 和 `pixi-live2d-display-lipsyncpatch` 库进行高效渲染, 支持从 2.1~4 的所有 Cubism core 及模型版本。
- **展示方式:** 模型在 HTML Canvas 元素中展示, 支持动态调整分辨率以适应不同屏幕。

一、Live 2D 技术在 Web 端的集成

一、Live 2D 技术在 Web 端的集成

- **渲染引擎**: 使用 **PIXI.js** 和 `pixi-live2d-display-lipsyncpatch` 库进行高效渲染，支持从 2.1~4 的所有 Cubism core 及模型版本。
- **展示方式**: 模型在 HTML Canvas 元素中展示，支持动态调整分辨率以适应不同屏幕。
- **状态管理**: 通过 Pinia 进行状态管理，使用 `useLive2DStore` 集中存储和同步模型状态（如表情、动作等）。

一、Live 2D 技术在 Web 端的集成

- **渲染引擎**: 使用 **PIXI.js** 和 `pixi-live2d-display-lipsyncpatch` 库进行高效渲染，支持从 2.1~4 的所有 Cubism core 及模型版本。
- **展示方式**: 模型在 HTML Canvas 元素中展示，支持动态调整分辨率以适应不同屏幕。
- **状态管理**: 通过 Pinia 进行状态管理，使用 `useLive2DStore` 集中存储和同步模型状态（如表情、动作等）。
- **用户自定义与上传**: 我们支持用户通过 Zip 压缩包的形式上传其自己的 Live 2D 模型文件，系统会自动根据其 `model.json` 或 `model3.json` 确定模型版本，并调用加载方法。经由中间层统一不同版本模型骨骼的控制方式，综合官方 SDK 与我们自己实现的 SDK 实现“开发一次、全版本适用”的效果。

二、基于语音学的实时口型适配

① 项目概览：环境与架构

② 用户账户管理

③ 直播推流模块

④ 虚拟直播模块

一、Live 2D 技术在 Web 端的集成

二、基于语音学的实时口型适配

三、由 AI 控制的实时表情变化

⑤ LLM 文案生成模块

⑥ TTS 语音转换模块

⑦ l18n 与用户友好特质的开发

二、基于语音学的实时口型适配

二、基于语音学的口型实时适配

音频驱动模型核心 (Live2DModel.vue)

```
1 const analyzeAudio = () => {  
2   const dataArray = analyzer.getBytesFrequencyData();  
3   // 基于FFT的元音识别  
4   const vowel = detectVowel(dataArray);  
5   // 动态参数映射  
6   model.internalModel.eyex = vowel === 'A' ? 1 : 0;  
7 };
```

二、基于语音学的口型实时适配

核心功能

- 文本到拼音转换
- 元音提取与分类
- 口型参数映射
- 实时动画驱动
- 平滑过渡处理

性能突破

- 60FPS 流畅渲染
- 嘴型检测延迟低于 40ms

二、基于语音学的实时口型适配

核心算法

元音提取算法

```
1 def extract_vowel(py):
2     """
3     从拼音中提取元音，支持复合元音
4
5     算法逻辑：
6     1. 优先检查复合元音 (ai, ao, ei, ou, an, en, ang,
7     eng 等)
8     2. 再检查基本元音 (a, o, e, i, u, ü)
9     3. 返回小写元音字符串
10    """
```

随后根据字符和元音类型计算发音时长

性能优化

前端优化

- 防重复处理：避免相同字幕在 100ms 内重复处理
- 动画帧优化：使用 `requestAnimationFrame` 确保流畅动画
- 参数缓存：缓存当前参数值，减少重复计算
- 降级处理：对不支持的参数进行优雅降级

后端优化

- 拼音缓存：缓存常用字符的拼音结果
- 批量处理：支持批量文本处理
- 错误处理：完善的异常处理机制

三、由 AI 控制的实时表情变化

① 项目概览：环境与架构

② 用户账户管理

③ 直播推流模块

④ 虚拟直播模块

一、Live 2D 技术在 Web 端的集成

二、基于语音学的实时口型适配

三、由 AI 控制的实时表情变化

⑤ LLM 文案生成模块

⑥ TTS 语音转换模块

⑦ l18n 与用户友好特质的开发

三、由 AI 控制的实时表情变化——原理概述（1/5）

核心流程：

① 表情类型总结：

首先定义一系列基础情感（开心、愤怒、惊讶、悲伤等），通过 Live2D Cubism Editor 将情感转化为模型骨骼动作。

② 文本语义嵌入：

演讲过程中对当前讲稿进行实时 Embedding，使用 BAAI/bge-large-zh-v1.5 等模型将文本转化为语义向量。

③ 表情与情感相似度计算：

计算讲稿语义向量与预定义情感/表情的相似度，确定匹配度最高的动作/表情类型。

三、由 AI 控制的实时表情变化——原理概述（2/5）

④ AI 模型生成适当的表情指令：

在讲稿生成过程中用特殊标识嵌入当前的表情动作。

例如

“哈喽，各位旅行者们！欢迎来到今天的直播间 \$\$ 开心 ##，我是你们的主播。”

⑤ 解析标记与 Live2D 模型控制：

前端解析 AI 输出文本，提取表情标记并通过 Live2D Cubism SDK 控制模型做出对应表情。

三、由 AI 控制的实时表情变化

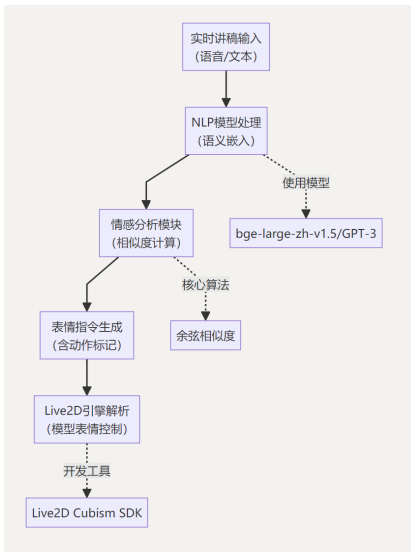
三、由 AI 控制的实时表情变化——原理概述（3/5）

关键技术栈

- 自然语言处理（NLP）模型：BGE、BERT 等
- 情感分析与分类模型
- 语义向量计算（余弦相似度等）
- 动态表情生成引擎（Live2D）

三、由 AI 控制的实时表情变化

三、由 AI 控制的实时表情变化——系统架构（4/5）



三、由 AI 控制的实时表情变化——系统架构（5/5）

流程说明

1. 输入：实时输入当前 LLM 生成的讲稿
2. 处理：NLP 模型嵌入语义向量并进行情感分析
3. 匹配：将语义向量与预定义表情库求相似度匹配
4. 输出：生成控制指令驱动 Live2D 模型表情变化

① 项目概览：环境与架构

② 用户账户管理

③ 直播推流模块

④ 虚拟直播模块

⑤ LLM 文案生成模块

一、全面的 AI 服务提供商与模型的适配

二、人类可控的 AI 直播文案生成与即时性的优化

⑥ TTS 语音转换模块

⑦ l18n 与用户友好特质的开发

LLM 文案生成模块：人类的可控性，多服务商适配与完全即时性的优化

——任逸青、冯博文、陈宏宇

"Can Machines Think?"

——Alan Turing, The Father of Artificial Intelligence

一、全面的 AI 服务提供商与模型的适配

① 项目概览：环境与架构

② 用户账户管理

③ 直播推流模块

④ 虚拟直播模块

⑤ LLM 文案生成模块

一、全面的 AI 服务提供商与模型的适配

二、人类可控的 AI 直播文案生成与即时性的优化

⑥ TTS 语音转换模块

⑦ l18n 与用户友好特质的开发

一、全面的 AI 服务提供商与模型的适配

覆盖全球主流 AI 服务提供商

- **国际头部厂商：**OpenAI、Anthropic、Google
- **国内领先平台：**硅基流动、火山方舟、腾讯云
- **新兴技术力量：**DeepSeek、SORUX 等
- **灵活扩展支持：**自定义 API 接入能力

全谱系模型适配矩阵

- **OpenAI 系列：**GPT-4o、GPT-4o-mini、GPT-3.5 Turbo 等
- **Anthropic 家族：**Claude 3/3.5/3.7 系列 (Opus/Sonnet/Haiku)、Thinking 增强模型
- **Google Gemini：**2.5 Pro/Flash (含推理优化版)、预览版模型
- **Deepseek 家族：**DeepSeek V3-1216、DeepSeek V3-0324、DeepSeek R1 等等
- **专项能力模型：**Grok-3 (深度搜索/推理)、O1/O3 系列、MJ Chat 等

核心优势

- 无需重复开发适配层，降低跨平台集成成本
- 支持模型版本无缝切换，灵活应对业务场景变化
- 兼顾国际前沿技术与国内自主研发，全球化与本地化并重
- 覆盖从联网搜索、文案生成到表情控制的全流程矩阵，满足系统对于大模型多种多元能力的需求

二、人类可控的 AI 直播文案生成与即时性的优化

① 项目概览：环境与架构

② 用户账户管理

③ 直播推流模块

④ 虚拟直播模块

⑤ LLM 文案生成模块

一、全面的 AI 服务提供商与模型的适配

二、人类可控的 AI 直播文案生成与即时性的优化

⑥ TTS 语音转换模块

⑦ I18n 与用户友好特质的开发

二、人类可控的 AI 直播文案生成与即时性的优化

文案生成模块：AI 驱动的直播流程

三步式生成流程

为实现 AI 全流程直播，我们设计了人机协作的文案生成模式，确保内容质量与人类的可控性：

- 1 **主题输入**：用户设定直播核心主题。
- 2 **大纲生成**：AI 或用户基于主题快速构建直播框架。
- 3 **内容填充与调整**：AI 生成具体文案，用户可随时调整大纲与内容，掌控全局。

二、人类可控的 AI 直播文案生成与即时性的优化

文案生成模块：核心功能架构

模块交互逻辑

- **输入处理**: 用户输入主题，通过 v-model 触发 generateOutline(), 进入加载状态。
- **提纲生成**: 调用 AI 接口，成功后解析响应，生成带唯一 ID 的章节对象。
- **内容生成**: 采用三级体系（主生成、预生成、按需生成）填充章节内容，确保直播流畅。

二、人类可控的 AI 直播文案生成与即时性的优化

文案生成模块：核心功能架构

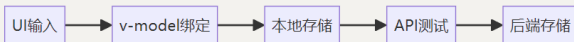
提纲生成 API 示例 (JavaScript)

```
1 // 构造 API 请求
2 callOpenAI(为主题 "${topic}" 生成直播提纲...)
3
4 // 响应处理
5 response.split('\n')
6 .map(line => ({ id: uuidv4(), title: line, content: ''
  })))
```

二、人类可控的 AI 直播文案生成与即时性的优化

文案生成模块：配置系统与数据流

配置数据流设计



关键校验与保存逻辑

- **连接测试**: 检查必填项，构造测试请求，密码字段脱敏显示。
- **双模保存**: 支持带提示的即时保存与后台静默保存，保存前会自动执行连接测试。

二、人类可控的 AI 直播文案生成与即时性的优化

文案生成模块：算法细节与触发条件

智能分句算法 (JavaScript)

```
1 function splitIntoSentences(text) {
2   // 智能分句逻辑：
3   // 1. 按标点分割为数组
4   // 2. 奇偶项重组，保留标点
5   // 3. 末尾项特殊处理并过滤空句
6 }
```

预生成触发条件

触发场景	判断条件	执行方法
首句播放	sentenceIndex === 0	generateNextBlockContent(+1)
拖拽结束	currentBlockIndex 变化	检查后续章节空内容
标题修改	oldTitle !== newTitle	generateContentForSpecificBlock()

文案生成模块：健壮性与性能优化

异常处理体系

- **错误分类处理：**
 - API 错误：HTTP 状态码分析、响应体验证、重试机制（最多 3 次）。
 - 本地错误：LocalStorage 读写异常、用户输入验证、组件卸载清理。
- **恢复机制：**
 - 缓存最后有效配置，失败时自动回滚。
 - 使用错误边界组件（Error Boundary）包裹，防止应用崩溃。

二、人类可控的 AI 直播文案生成与即时性的优化

文案生成模块：健壮性与性能优化

性能优化策略

- **内存管理**: 分块加载文案内容，定时清理不再使用的音频缓存。
- **渲染优化**: 采用虚拟滚动列表处理长篇大纲，避免 DOM 臃肿。
- **网络优化**: 合并部分 API 请求，预加载关键资源。

- 1 项目概览：环境与架构
- 2 用户账户管理
- 3 直播推流模块
- 4 虚拟直播模块
- 5 LLM 文案生成模块
- 6 TTS 语音转换模块
- 7 l18n 与用户友好特质的开发
- 8 总结与展望

TTS 语音转换模块：推理、训练与开发者 API

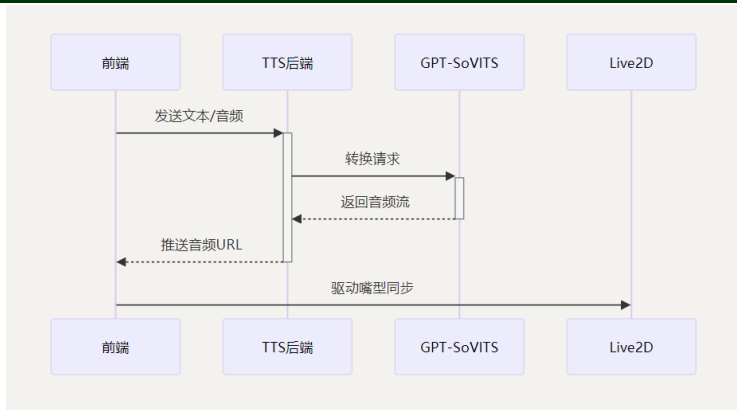
——陈宏宇

"Stop Trying to Reinvent the Wheel."

——DRY Principle

语音转换模块：低延迟 AI 语音流水线

系统架构核心



语音转换模块：GPT-SoVITS 深度集成

GPT-SoVITS 流式处理 (API_v2.py)

```
1 # API_v2.py 流式响应
2 def generate_stream():
3     while chunk := get_audio_chunk():
4         # 200ms/块 实时推送
5         yield chunk
```

集成亮点

- **动态模型加载**: 支持 V1~V4 所有模型切换
- **流式处理**: 实现低延迟语音合成
- **情感迁移**: 通过 Prosody Embedding 实现语气情感控制

语音转换模块：前端优化与性能

前端优化策略

- **预载入机制**：提前加载这一整章的语音，降低首字延迟
- **自适应缓冲**：根据网络抖动动态调整缓冲池 (150-500ms)

关键性能数据

- 中文合成速度: **0.8CE 实时** (i7-12700H)
- 端到端延迟: **220ms** (文本输入 音频输出)
- 资源占用: **<2GB RAM**

① 项目概览：环境与架构

② 用户账户管理

③ 直播推流模块

④ 虚拟直播模块

⑤ LLM 文案生成模块

⑥ TTS 语音转换模块

⑦ l18n 与用户友好特质的开发

一、中日英三语的国际化

二、系统主题的自定义与帮助、文档页面

I18n 与用户友好特质的开发：中日英三语的国际化、系统主题的自定义与帮助、文档页面

——解侑睿、任逸青

“全球化是社会生产力和科学技术发展的客观要求和必然结果。”

——江泽民

一、中日英三语的国际化

- 1 项目概览：环境与架构
- 2 用户账户管理
- 3 直播推流模块
- 4 虚拟直播模块
- 5 LLM 文案生成模块
- 6 TTS 语音转换模块
- 7 I18n 与用户友好特质的开发
 - 一、中日英三语的国际化
 - 二、系统主题的定义与

一、中日英三语的国际化

多语言支持与国际化设计

- **三语支持**: 系统支持中、日、英三种语言，确保各类用户能够轻松使用并操作。
- **动态切换**: 通过国际化框架（i18n.js 等）动态加载语言包，用户可在设置中自由切换语言。
- **字符适配**: 处理多种语言字符集，对非拉丁字符集的全面兼容。
- **本地化内容**: 根据不同地区用户习惯对界面元素、提示文字和操作行为进行定制化展示。

语言包设计

- **中央语言管理:** 所有文案、提示信息统一存储, 易于扩展与更新。
- **多语言支持插件:** 使用 Vue i18n 插件与后台接口配合, 根据用户区域自动加载对应的语言包。
- **自动检测:** 通过浏览器语言设置与 IP 地理位置自动识别用户语言。

- ## 二、系统主题的自定义与帮助、文档页面

系统主题自定义与帮助文档

- **主题自定义**: 用户可以根据自己的偏好选择系统的主题样式，支持多种颜色主题与组件样式。
- **个性化设置**: 提供可调节的界面元素，如组件位置、界面颜色等，满足不同用户的可用性需求。
- **帮助文档**: 在线帮助文档提供详细的系统功能介绍与操作指南，并根据用户选择的语言自动显示对应文档内容。
- **FAQ 与社区支持**: 集成 FAQ 模块与社区支持，用户可以快速找到常见问题的解决方案。
- **多语言文档**: 针对不同语言的用户，提供中、日、英等语言的文档版本，确保跨国用户都能顺利获取帮助。

- ① 项目概览：环境与架构
- ② 用户账户管理
- ③ 直播推流模块
- ④ 虚拟直播模块
- ⑤ LLM 文案生成模块
- ⑥ TTS 语音转换模块
- ⑦ l18n 与用户友好特质的开发
- ⑧ 总结与展望**

系统核心价值总结

我们实现了什么？

- 实现了高安全度的用户登录-注册-偏好保存系统，将之模块高度解耦，标准化了其端口。

核心突破

本系统将 AI 技术与实时流媒体深度融合，打造了从内容生成到语音转换到实时虚拟主播的全链路一站式解决方案。

系统核心价值总结

我们实现了什么？

- 实现了高安全度的用户登录-注册-偏好保存系统，将之模块高度解耦，标准化了其端口。
- 从头实现了 **WebRTC 到 RTMP** 的转码方案，攻克协议壁垒，在 RTMP 的协议限制下仍能做到秒级延迟。

核心突破

本系统将 AI 技术与实时流媒体深度融合，打造了从内容生成到语音转换到实时虚拟主播的全链路一站式解决方案。

系统核心价值总结

我们实现了什么？

- 实现了高安全度的用户登录-注册-偏好保存系统，将之模块高度解耦，标准化了其端口。
- 从头实现了 **WebRTC 到 RTMP** 的转码方案，攻克协议壁垒，在 RTMP 的协议限制下仍能做到秒级延迟。
- 实现了全版本 Live2D 模型的 Web 应用——从十年前的 Cubism core2.1 到现今的 Cubism core4.0。

核心突破

本系统将 AI 技术与实时流媒体深度融合，打造了从内容生成到语音转换到实时虚拟主播的全链路一站式解决方案。

系统核心价值总结

我们实现了什么？

- 实现了高安全度的用户登录-注册-偏好保存系统，将之模块高度解耦，标准化了其端口。
- 从头实现了 **WebRTC 到 RTMP** 的转码方案，攻克协议壁垒，在 RTMP 的协议限制下仍能做到秒级延迟。
- 实现了全版本 Live2D 模型的 Web 应用——从十年前的 Cubism core2.1 到现如今的 Cubism core4.0。
- 实现了网页端的适于直播形式的 AI 文案生成解决方案

核心突破

本系统将 AI 技术与实时流媒体深度融合，打造了从内容生成到语音转换到实时虚拟主播的全链路一站式解决方案。

系统核心价值总结

我们实现了什么？

- 实现了高安全度的用户登录-注册-偏好保存系统，将之模块高度解耦，标准化了其端口。
- 从头实现了 **WebRTC 到 RTMP** 的转码方案，攻克协议壁垒，在 RTMP 的协议限制下仍能做到秒级延迟。
- 实现了全版本 Live2D 模型的 Web 应用——从十年前的 Cubism core2.1 到现如今的 Cubism core4.0。
- 实现了网页端的适于直播形式的 AI 文案生成解决方案
- 实现了基于 GPT-SoVITS 的高度灵活的语音 TTS 转换方案。

核心突破

本系统将 AI 技术与实时流媒体深度融合，打造了从内容生成到语音转换到实时虚拟主播的全链路一站式解决方案。

系统核心价值总结

我们实现了什么？

- 实现了高安全度的用户登录-注册-偏好保存系统，将之模块高度解耦，标准化了其端口。
- 从头实现了 **WebRTC 到 RTMP** 的转码方案，攻克协议壁垒，在 RTMP 的协议限制下仍能做到秒级延迟。
- 实现了全版本 Live2D 模型的 Web 应用——从十年前的 Cubism core2.1 到现如今的 Cubism core4.0。
- 实现了网页端的适于直播形式的 AI 文案生成解决方案
- 实现了基于 GPT-SoVITS 的高度灵活的语音 TTS 转换方案。
- 达成了 **语音-Live2D-推流**全流程一站式的解决方案。

核心突破

本系统将 AI 技术与实时流媒体深度融合，打造了从内容生成到语音转换到实时虚拟主播的全链路一站式解决方案。

我们的愿景

打造下一世代的一站式全自动 AI 虚拟直播平台

由而使得机械的灵魂与人类的心脏同频共振

致谢

感谢您的聆听！



deepsleep



STREAMCREAM

我们是 DeepSleep 项目团队，
期待与您一同探索 AI 领域的无垠边界！