机构图标

Giggle Academy

面试报告

|  |  |  |
| --- | --- | --- |
| 文件状态：  [√] 草稿  [ ] 正式发布  [ ] 正在修改 | 文件标识： | Giggle Academy |
| 当前版本： | 1.X |
| 作 者： |  |
| 完成日期： | 2025-06-05 |

机构公开信息

版 本 历 史

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本/状态 | 作者 | 参与者 | 起止日期 | 备注 |
| V1.0 | 庄江龙 |  | 2025-06-05 | 初稿 |
|  |  |  |  |  |
|  |  |  |  |  |

目 录

目录

[1 文档介绍 6](#_Toc1222363497)

[1.1 文档目的 6](#_Toc1434726877)

[1.2 文档范围 6](#_Toc1508233223)

[1.3 读者对象 6](#_Toc2126293420)

[1.4 参考文档 6](#_Toc338140213)

[1.5 术语与缩写解释 6](#_Toc880829929)

[2 环境搭建 6](#_Toc1503837932)

[2.1 环境安装脚本 6](#_Toc1269081581)

[2.2 环境信息 7](#_Toc646549863)

[3 服务设计与开发 7](#_Toc296293621)

[3.1 功能架构 7](#_Toc1939794401)

[3.1.1 语言文本管理子系统 7](#_Toc1175252500)

[3.1.1.1 STT任务管理服务 7](#_Toc2061666041)

[3.1.1.1.1 添加任务功能 7](#_Toc772506742)

[3.1.1.1.2 查询任务详情 7](#_Toc1982166679)

[3.1.1.1.3 启动任务功能 8](#_Toc361558042)

[3.1.1.1.4 取消任务功能 8](#_Toc1474774531)

[3.1.1.1.5 任务分发策略功能 8](#_Toc279288843)

[3.1.1.2 翻译任务管理服务 8](#_Toc1755815606)

[3.1.1.2.1 添加任务功能 8](#_Toc1420096615)

[3.1.1.2.2 查询任务详情 8](#_Toc430555547)

[3.1.1.2.3 启动任务功能 8](#_Toc1474671686)

[3.1.1.2.4 取消任务功能 8](#_Toc698256575)

[3.1.1.2.5 任务执行策略 8](#_Toc1747608817)

[3.1.1.3 语言文本服务 9](#_Toc927547300)

[3.1.1.3.1 文本相识度评分 9](#_Toc703677527)

[3.1.1.3.2 语言文本持久化与检索(基于单个json文件形式) 9](#_Toc515752260)

[3.1.1.3.2.1 语言文本久化化 9](#_Toc1004234528)

[3.1.1.3.2.2 语言文本检索 11](#_Toc1095730323)

[3.1.1.3.3 语言文本持久化与检索(基于文件顺序写入与内嵌数据库形式) 11](#_Toc1267265636)

[3.1.1.3.3.1 语言文本久化化 11](#_Toc190733306)

[3.1.1.3.3.2 语言文本检索 12](#_Toc1609072618)

[3.1.2 whisper子系统 12](#_Toc421924055)

[3.2 数据架构 12](#_Toc286589991)

[3.2.1 TSS任务表（tb\_stt\_job） 12](#_Toc2059642163)

[3.2.2 翻译任务表（tb\_translate\_job） 13](#_Toc1116927548)

[3.3 系统设计 13](#_Toc1046740809)

[3.3.1 服务接口设计 13](#_Toc386740639)

[3.3.1.1 STT任务管理-查询任务 13](#_Toc1664403851)

[3.3.1.1.1 接口说明 13](#_Toc513537935)

[3.3.1.1.2 请求数据 14](#_Toc295296252)

[3.3.1.1.3 应答数据 14](#_Toc209399147)

[3.3.1.2 STT任务管理-启动任务 15](#_Toc1793249843)

[3.3.1.2.1 接口说明 15](#_Toc1364609303)

[3.3.1.2.2 请求数据 15](#_Toc2010689208)

[3.3.1.2.3 应答数据 15](#_Toc850849664)

[3.3.1.3 翻译任务管理-查询任务 15](#_Toc136697475)

[3.3.1.3.1 接口说明 15](#_Toc1814443682)

[3.3.1.3.2 请求数据 16](#_Toc1087175974)

[3.3.1.3.3 应答数据 16](#_Toc1375726342)

[3.3.1.4 翻译任务管理-启动任务 17](#_Toc2023686392)

[3.3.1.4.1 接口说明 17](#_Toc251189158)

[3.3.1.4.2 请求数据 17](#_Toc1930812151)

[3.3.1.4.3 应答数据 17](#_Toc534432040)

[3.3.1.5 语言文本管理-计算文本相识度 17](#_Toc1422684526)

[3.3.1.5.1 接口说明 17](#_Toc975902784)

[3.3.1.5.2 请求数据 18](#_Toc1665478549)

[3.3.1.5.3 应答数据 18](#_Toc1396118045)

[3.3.1.6 语言文本管理-查询打包文件 18](#_Toc1149655193)

[3.3.1.6.1 接口说明 18](#_Toc1344456692)

[3.3.1.6.2 请求数据 18](#_Toc460688710)

[3.3.1.6.3 应答数据 18](#_Toc1116601535)

[3.3.1.7 语言文本管理-快速查询 19](#_Toc2009891259)

[3.3.1.7.1 接口说明 19](#_Toc324622703)

[3.3.1.7.2 请求数据 19](#_Toc1325305941)

[3.3.1.7.3 应答数据 19](#_Toc716563703)

[3.3.1.8 Whisper服务-STT 20](#_Toc197863945)

[3.3.1.8.1 接口说明 20](#_Toc1194638059)

[3.3.1.8.2 请求数据 20](#_Toc1457241810)

[3.3.1.8.3 应答数据 20](#_Toc1959590282)

[参 考 文 献 23](#_Toc1024659182)

正文

# 文档介绍

## 文档目的

## 文档范围

## 读者对象

## 参考文档

***提示：****列出本文档的所有参考文献（可以是非正式出版物），格式如下：*

*[标识符] 作者，文献名称，出版单位（或归属单位），日期*

***例如：***

***[SPP-PROC-PP]*** *SEPG，需求开发规范，机构名称，日期*

## 术语与缩写解释

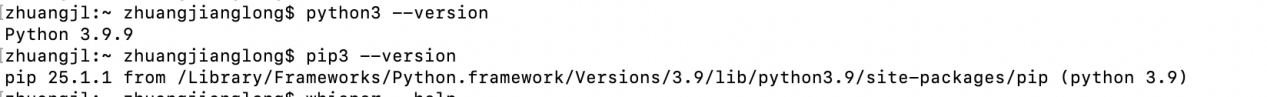
|  |  |
| --- | --- |
| **缩写、术语** | **解 释** |
|  |  |

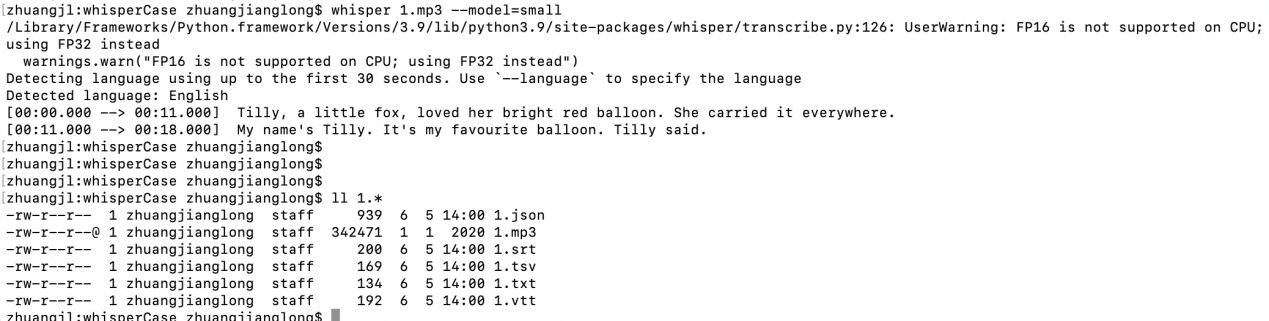
# 环境搭建

## 环境安装脚本

|  |
| --- |
| #python3.9  pip install setuptools-rust  #whisper  pip install -U openai-whisper  # ffmpeg  brew install ffmpeg |

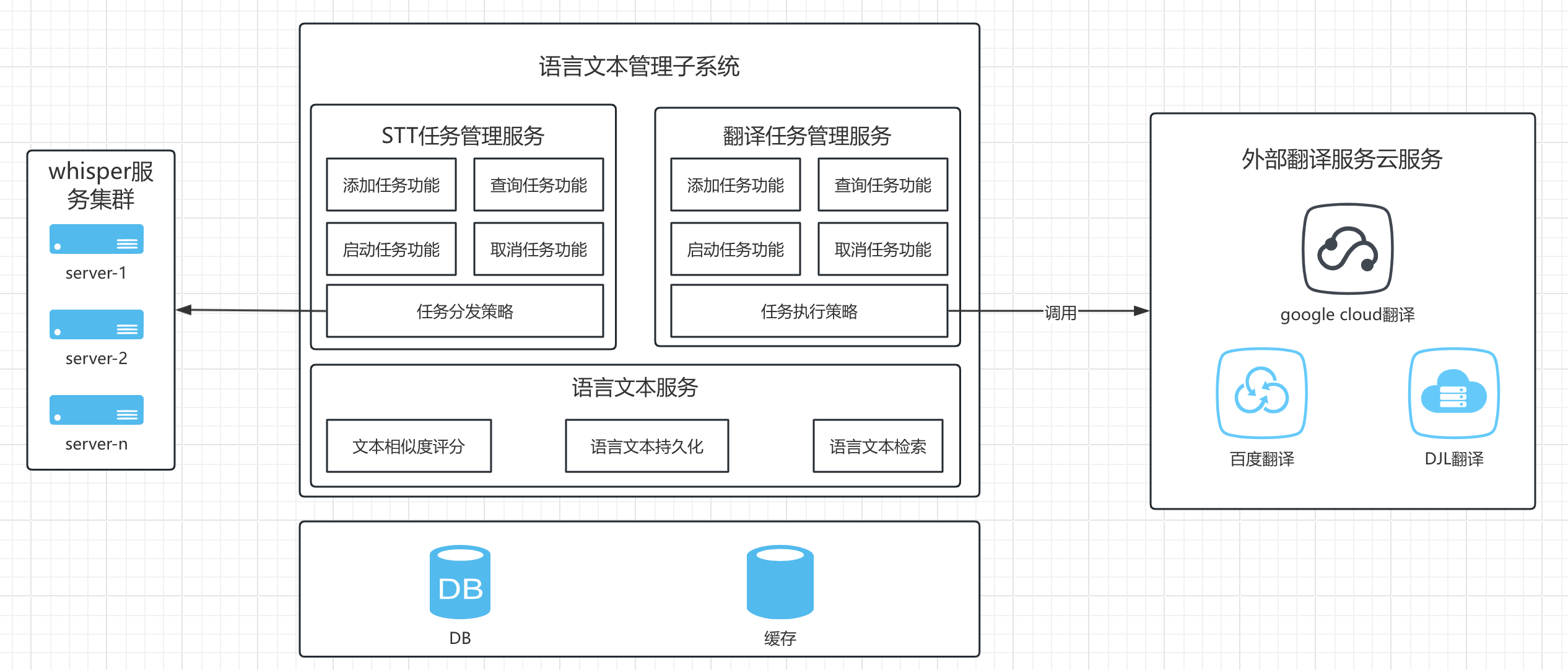
## 环境信息





# 服务设计与开发

## 功能架构



### 语言文本管理子系统

采用Java语言，使用SpringBoot、Mybatis、H2数据库实现。

#### STT任务管理服务

##### 添加任务功能

添加一个STT任务。

##### 查询任务详情

查询STT任务详情。

##### 启动任务功能

在添加完STT任务后，启动后异步执行STT任务，并更新任务结果。

##### 取消任务功能

取消STT任务，设置任务状态为取消。

##### 任务分发策略功能

根据whisper服务集群的服务实例根据负载策略进行STT任务分发。负载策略包含：

* 根据whisper服务实例链接数判断，优先分发给链接数最小的whisper服务实例；
* 根据whisper服务实例使用资源判断，优先分发给资源最空闲的whisper服务实例；

#### 翻译任务管理服务

##### 添加任务功能

添加一个文本翻译任务。

##### 查询任务详情

查询文本翻译任务详情。

##### 启动任务功能

在添加完文本翻译任务后，启动后异步执行文本翻译，并更新任务结果。

##### 取消任务功能

取消文本翻译任务，设置任务状态为取消。

##### 任务执行策略

采用工厂设计模式与策略设计模式，实现根据任务配置调用不同的翻译实现。翻译策略实现包含：

* Google cloud翻译
* 百度翻译
* DLJ翻译

#### 语言文本服务

##### 文本相识度评分

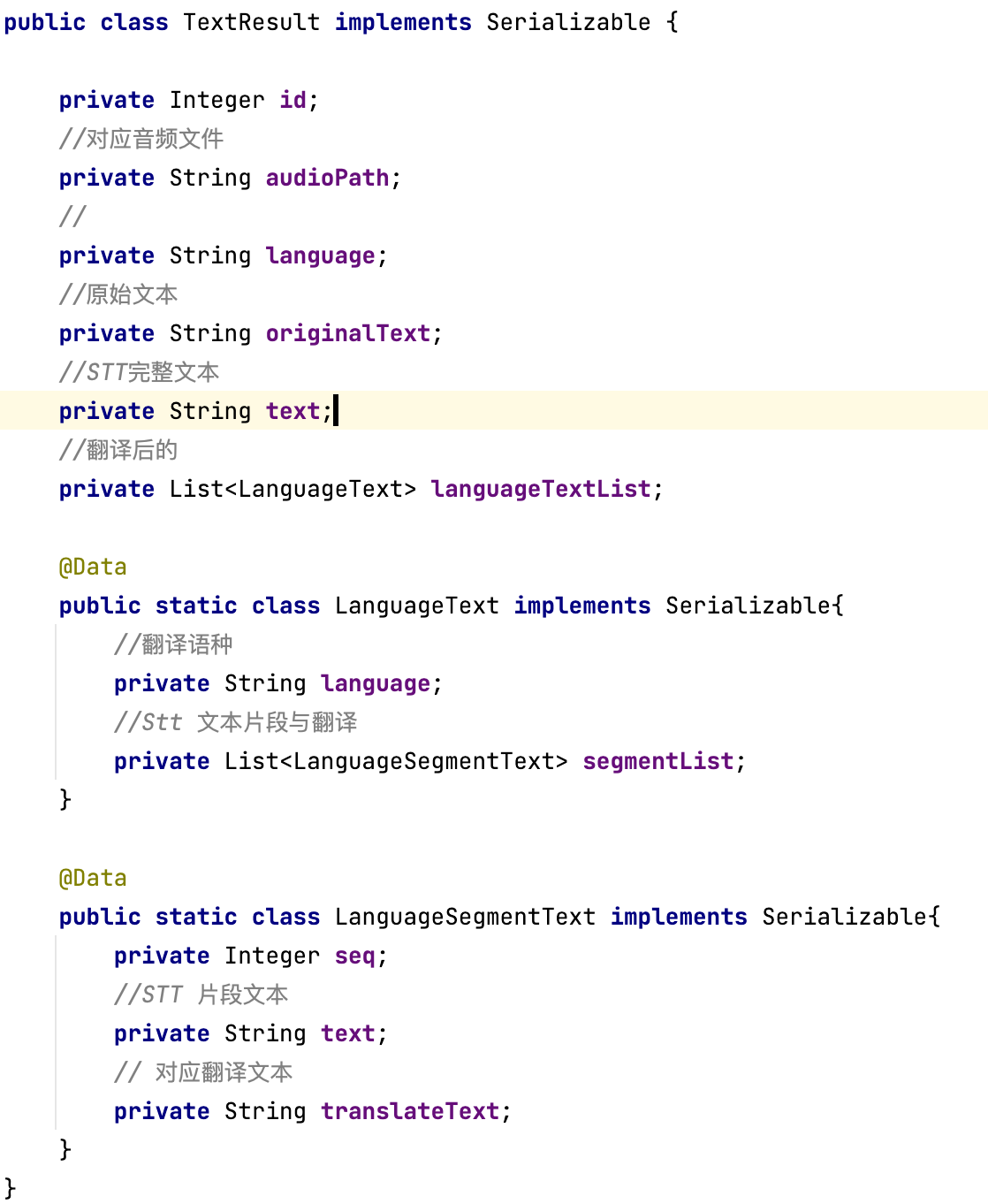
使用Apache Commons Text进行文本字符相识度计算验证原始音频⽂本信息与 STT ⽣成的文本内容的准确性校验。包含以下三种计算方式：

1. JaccardSimilarity：该算法是基于两个字符串的交集和并集来计算相似度；
2. LevenshteinDistance：该算法是基于两个字符串之间的最少编辑操作次数来计算相似度；
3. CosineSimilarity：该算法是基于两个字符串的向量空间模型来计算相似度。

##### 语言文本持久化与检索(基于单个json文件形式)

###### 语言文本久化化

根据STT任务标识，把音频原始文本信息、STT提取的文本信息和其各语言翻译文本组装生成如下json格式文件，并基于音频文件标识命名持久化。



{

"audioPath": "/Users/zhuangjianglong/py-workspace/whisperCase/1.mp3",

"id": 1,

"language": "en",

"languageTextList": [{

"language": "zh",

"segmentList": [{

"seq": 0,

"text": " Tilly, a little fox, loved her bright red balloon. She carried it everywhere.",

"translateText": " 小狐狸蒂莉特别喜欢她那只鲜红的气球。她走到哪儿都带着它。"

}, {

"seq": 1,

"text": " My name's Tilly. It's my favourite balloon. Tilly said.",

"translateText": " “我叫蒂莉。这是我最喜欢的气球。”蒂莉说。"

}]

}, {

"language": "zh-TW",

"segmentList": [{

"seq": 0,

"text": " Tilly, a little fox, loved her bright red balloon. She carried it everywhere.",

"translateText": " 小狐狸蒂莉特別喜歡她那隻鮮紅的氣球。她走到哪裡都帶著它。"

}, {

"seq": 1,

"text": " My name's Tilly. It's my favourite balloon. Tilly said.",

"translateText": " 「我叫蒂莉。這是我最喜歡的氣球。」蒂莉說。"

}]

}],

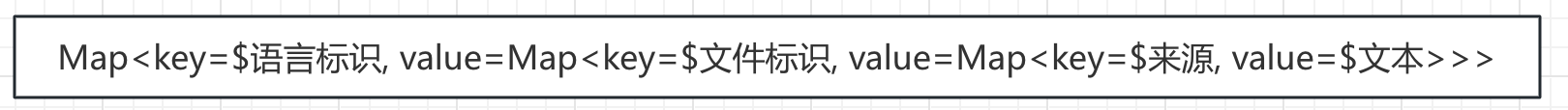
"originalText": "Tilly, a little fox, loved her bright red balloon. She carried it everywhere.\\n“It’s my favorite balloon!” Tilly said.\\n\n",

"text": " Tilly, a little fox, loved her bright red balloon. She carried it everywhere. My name's Tilly. It's my favourite balloon. Tilly said."

},

###### 语言文本检索

基于Hash算法快速寻址特点，采用三级Map缓存生成语言文本信息，提供语言文本快速检索查询。



##### 语言文本持久化与检索(基于文件顺序写入与内嵌数据库形式)

###### 语言文本久化化

借鉴MQ对消息持久化机制，采用顺序写入机制将文本信息写入到textlog文件中，每个textlog文件按照分片方式，固定大小为1G，总长度20位，标记当前文件的物理偏移量，高位用0补齐，文件命名如下：

00000000000000000000 第一个文件，物理偏移量为0

00000000001073741824 第二个文件，物理偏移量为1G

00000000002147483648 第三个文件，物理偏移量为2G

其中，每个文件的存储内容格式如下：



###### 语言文本检索

采用内嵌数据库，入SQLite、Derby等内置在用户端上，按照offset、语言类型、音频标识、来源纬度建表并设置复合索引，将textlog的索引信息写入到数据表中。查询时，根据语言类型、音频标识、来源纬度获取对应offset值，根据offset计算textlog文件获取对应语言文本信息。

### whisper子系统

采用python语言，使用whisper、flask框架实现。

## 数据架构



### TSS任务表（tb\_stt\_job）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **类型** | **必填** | **描述** | **约束** | **备注** |
| id | int | Y | 主键 | PK | 主键 |
| audio\_id | int | N | 音频文件标识 |  |  |
| audio\_path | varchar(1024) | Y | 音频文件路径 |  | 冗余 |
| original\_text\_path | varchar(1024) | N | audio文件对应原始文本文件路径 |  | 冗余 |
| type | TINYINT | Y | TSS方式 |  | 1-whisper;2- |
| status | TINYINT | Y | 状态 |  | 0-成功；1-新建；2-执行中；4-取消；3-whisper异常；9-未知错误 |
| retry\_count | TINYINT | N | 重试次数 |  |  |
| text\_path | varchar(1024) | Y | TSS文本文件地址 |  |  |

### 翻译任务表（tb\_translate\_job）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **类型** | **必填** | **描述** | **约束** | **备注** |
| id | int | Y | 主键 | PK | 主键 |
| stt\_job\_id | int | N | TSS任务标识 | FK |  |
| type | TINYINT | Y | TSS方式 |  | 1-whisper;2- |
| target\_language | char(8) | Y | 目标语言 |  | en,zh\_cn,ja |
| status | TINYINT | Y | 状态 |  | 0-成功；1-新建；2-执行中；4-取消；3-whisper异常；9-未知错误 |
| retry\_count | TINYINT | N | 重试次数 |  |  |
| result\_path | varchar(1024) | Y | 翻译文本文件地址 |  |  |

## 系统设计

### 服务接口设计

#### STT任务管理-查询任务

##### 接口说明

|  |  |
| --- | --- |
|  | **内容** |
| 接口地址 | http://ip:port/stt/{id} |
| 接口调用方式 | get |
| 接口描述 | 查询任务详情接口 |

##### 请求数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

##### 应答数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

{

"code": "0000",

"msg": "Success",

"data": {

"id": 1,

"audioPath": "/Users/zhuangjianglong/py-workspace/whisperCase/1.mp3",

"language": "en",

"originalText": null,

"text": " Tilly, a little fox, loved her bright red balloon. She carried it everywhere. My name's Tilly. It's my favourite balloon. Tilly said.",

"languageTextList": [{

"language": "zh\_cn",

"segmentList": [{

"seq": 0,

"text": " Tilly, a little fox, loved her bright red balloon. She carried it everywhere.",

"translateText": " 小狐狸蒂莉特别喜欢她那只鲜红的气球。她走到哪儿都带着它。"

}, {

"seq": 1,

"text": " My name's Tilly. It's my favourite balloon. Tilly said.",

"translateText": " “我叫蒂莉。这是我最喜欢的气球。”蒂莉说。"

}]

}]

}

}

#### STT任务管理-启动任务

##### 接口说明

|  |  |
| --- | --- |
|  | **内容** |
| 接口地址 | http://ip:port/stt/start/{id} |
| 接口调用方式 | get |
| 接口描述 | 启动任务接口 |

##### 请求数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

##### 应答数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

{

"code": "0000",

"msg": "Success"

}

#### 翻译任务管理-查询任务

##### 接口说明

|  |  |
| --- | --- |
|  | **内容** |
| 接口地址 | http://ip:port/translate/{id} |
| 接口调用方式 | get |
| 接口描述 | 查询任务详情接口 |

##### 请求数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

##### 应答数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

{

"code": "0000",

"msg": "Success",

"data": {

"id": 1,

"type": 1,

"sttJobId": 1,

"sttText": "{\n \"language\": \"en\",\n \"segments\": [\n {\n \"avg\_logprob\": -0.24758162865271935,\n \"compression\_ratio\": 1.209090909090909,\n \"end\": 11.0,\n \"id\": 0,\n \"no\_speech\_prob\": 0.3544903099536896,\n \"seek\": 0,\n \"start\": 0.0,\n \"temperature\": 0.0,\n \"text\": \" Tilly, a little fox, loved her bright red balloon. She carried it everywhere.\",\n \"tokens\": [\n 50364,\n 314,\n 6917,\n 11,\n 257,\n 707,\n 21026,\n 11,\n 4333,\n 720,\n 4730,\n 2182,\n 16994,\n 13,\n 1240,\n 9094,\n 309,\n 5315,\n 13,\n 50914\n ]\n },\n {\n \"avg\_logprob\": -0.24758162865271935,\n \"compression\_ratio\": 1.209090909090909,\n \"end\": 18.0,\n \"id\": 1,\n \"no\_speech\_prob\": 0.3544903099536896,\n \"seek\": 0,\n \"start\": 11.0,\n \"temperature\": 0.0,\n \"text\": \" My name's Tilly. It's my favourite balloon. Tilly said.\",\n \"tokens\": [\n 50914,\n 1222,\n 1315,\n 311,\n 314,\n 6917,\n 13,\n 467,\n 311,\n 452,\n 10696,\n 16994,\n 13,\n 314,\n 6917,\n 848,\n 13,\n 51264\n ]\n }\n ],\n \"text\": \" Tilly, a little fox, loved her bright red balloon. She carried it everywhere. My name's Tilly. It's my favourite balloon. Tilly said.\"\n}\n",

"targetLanguage": "zh\_cn",

"targetLanguagelist": null,

"status": 0,

"retryCount": 3,

"resultText": "[{\"seq\":0,\"sourceLanguage\":\"en\",\"text\":\" 小狐狸蒂莉特别喜欢她那只鲜红的气球。她走到哪儿都带着它。\"},{\"seq\":1,\"sourceLanguage\":\"en\",\"text\":\" “我叫蒂莉。这是我最喜欢的气球。”蒂莉说。\"}]"

}

}

#### 翻译任务管理-启动任务

##### 接口说明

|  |  |
| --- | --- |
|  | **内容** |
| 接口地址 | http://ip:port/translate/start/{id} |
| 接口调用方式 | get |
| 接口描述 | 查询任务详情接口 |

##### 请求数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

##### 应答数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

{

"code": "0000",

"msg": "Success"

}

#### 语言文本管理-计算文本相识度

##### 接口说明

|  |  |
| --- | --- |
|  | **内容** |
| 接口地址 | http://ip:port/textRecognition/{type}/{id} |
| 接口调用方式 | get |
| 接口描述 | 查询文本相似度接口 |

##### 请求数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

##### 应答数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

{

"code": "0000",

"msg": "Success",

"data": 0.8646

}

#### 语言文本管理-查询打包文件

##### 接口说明

|  |  |
| --- | --- |
|  | **内容** |
| 接口地址 | http://ip:port/textresult/{id} |
| 接口调用方式 | get |
| 接口描述 | 查询文本打包文件接口 |

##### 请求数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

##### 应答数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

{

"code": "0000",

"msg": "Success",

"data": {

"id": 1,

"audioPath": "/Users/zhuangjianglong/py-workspace/whisperCase/1.mp3",

"language": "en",

"originalText": null,

"text": " Tilly, a little fox, loved her bright red balloon. She carried it everywhere. My name's Tilly. It's my favourite balloon. Tilly said.",

"languageTextList": [{

"language": "zh\_cn",

"segmentList": [{

"seq": 0,

"text": " Tilly, a little fox, loved her bright red balloon. She carried it everywhere.",

"translateText": " 小狐狸蒂莉特别喜欢她那只鲜红的气球。她走到哪儿都带着它。"

}, {

"seq": 1,

"text": " My name's Tilly. It's my favourite balloon. Tilly said.",

"translateText": " “我叫蒂莉。这是我最喜欢的气球。”蒂莉说。"

}]

}]

}

}

#### 语言文本管理-快速查询

##### 接口说明

|  |  |
| --- | --- |
|  | **内容** |
| 接口地址 | http://ip:port/stt/query/{lan}/{id}/{ori} |
| 接口调用方式 | get |
| 接口描述 | 快速接口 |

##### 请求数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

##### 应答数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

{

"code": "0000",

"msg": "Success",

"data": ""

}

#### Whisper服务-STT

##### 接口说明

|  |  |
| --- | --- |
|  | **内容** |
| 接口地址 | http://ip:port/whisper/stt |
| 接口调用方式 | post |
| 接口描述 | 快速接口 |

##### 请求数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

{

"path": "/Users/zhuangjianglong/py-workspace/whisperCase/1.mp3",

"model": "small"

}

##### 应答数据

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **选项** | **数据类型** | **长度** | **示例值或说明** |
|  |  |  |  |  |  |

{

"code": "0000",

"msg": "Success",

"data": {

"language": "en",

"segments": [{

"avg\_logprob": -0.24758162865271935,

"compression\_ratio": 1.209090909090909,

"end": 11.0,

"id": 0,

"no\_speech\_prob": 0.3544903099536896,

"seek": 0,

"start": 0.0,

"temperature": 0.0,

"text": " Tilly, a little fox, loved her bright red balloon. She carried it everywhere.",

"tokens": [

50364,

314,

6917,

11,

257,

707,

21026,

11,

4333,

720,

4730,

2182,

16994,

13,

1240,

9094,

309,

5315,

13,

50914

]

},

{

"avg\_logprob": -0.24758162865271935,

"compression\_ratio": 1.209090909090909,

"end": 18.0,

"id": 1,

"no\_speech\_prob": 0.3544903099536896,

"seek": 0,

"start": 11.0,

"temperature": 0.0,

"text": " My name's Tilly. It's my favourite balloon. Tilly said.",

"tokens": [

50914,

1222,

1315,

311,

314,

6917,

13,

467,

311,

452,

10696,

16994,

13,

314,

6917,

848,

13,

51264

]

}

],

"text": " Tilly, a little fox, loved her bright red balloon. She carried it everywhere. My name's Tilly. It's my favourite balloon. Tilly said."

}

}

**参 考 文 献**

[1] https://github.com/openai/whisper

[2]

[3]

[4]

[5]

[6]

[7]

[8]

[9]