

Q 请输入

⌘ K

产品概述
快速入门
产品计费
开发参考

端到端实时语音大模型
语音播客大模型
语音同传大模型
语音妙记大模型
语音合成大模型
声音复刻大模型
语音识别大模型

产品简介

大模型流式语音识别 API

大模型录音文件识别标准版API

大模型录音文件极速版识别API

大模型录音文件识别闲置版

音频生成
语音识别
音视频字幕
SDK接入文档
自学习平台
控制台相关接口
相关协议

安全设置 访问控制 操作审计

API 访问密钥

藏

工具与其他

公测申请 资源管理 配额中心

伙伴控制台

待办事项

待支付	待续费	待处理工单
0	0	0
未读消息 0		

大模型流式语音识别API

简介

本文档介绍如何通过WebSocket协议实时访问大模型流式语音识别服务，主要分为端到端实时语音大模型、双向流式模式、语音识别大模型三部分。

双向流式模式使用的接口地址是 [wss://openspeech.bytedance.com/api/v3/sauc/bigmodel_async](https://openspeech/bytedance.com/api/v3/sauc/bigmodel_async)，流式输入模式使用的接口地址是 [wss://openspeech.bytedance.com/api/v3/sauc/bigmodel](https://openspeech/bytedance.com/api/v3/sauc/bigmodel)。

- 两者都是每输入一个包返回一个包，双向流式模式会尽快返回识别结果。
- 流式输入模式会在输入音频大于15s或发送最后一包（负包）后返回识别结果。
- 无论是哪种模式，单包音频大小建议在100~200ms左右，发包间隔过大会影响性能。（注：针对双向流式模式，单包为200ms大小时性能最佳）
- 流式输入模式在平均音频时长5s时，可以做到300~400ms以内返回。

双向流式模式（优化版本）接口地址：[wss://openspeech.bytedance.com/api/v3/sauc/bigmodel_async](https://openspeech/bytedance.com/api/v3/sauc/bigmodel_async)

- 该模式下，不再是每一包输入对应一包返回，只有当结果有变化时才会返回新的数据包（性能优化 rtf 和首字、尾字节延均有一定提升）
- 双向流式版本，更推荐使用双向流式模式（优化版本），性能相对更优。

鉴权

在 websocket 建连的 HTTP 请求头（Header 中）添加以下信息

Key	说明	Value 示例
X-Api-App-Key	使用火山引擎控制台获取的APP ID，可参考 控制台使用FAQ-Q1	123456789
X-Api-Access-Key	使用火山引擎控制台获取的Access Token，可参考 控制台使用FAQ-Q1	your-access-key
X-Api-Resource-Id	表示调用服务的资源信息 ID	豆包流式语音识别模型1.0 <ul style="list-style-type: none"> 小时版： <code>volc.bigsr.sauc.duration</code> 并发版： <code>volc.bigsr.sauc.concurrent</code> 豆包流式语音识别模型2.0 <ul style="list-style-type: none"> 小时版：
X-Api-Connect-Id	用于追踪当前连接的标志 ID，推荐设置UUID	

websocket 握手成功后，会返回这些 Response header。强烈建议

Key	说明
X-Api-Connect-Id	用于追踪当前调用信息的标志 ID，推荐用 UUID
X-Tt-Logid	服务端返回的 logid，建议用户获取和打印出来以备问题

3 3827手机用户 #I... 账号管理 >

账号ID : 2103539692

主账号 个人认证 >

费用中心

可用余额 ¥ 0.00 充值汇款 >

账户总览 账单详情 费用分析

发票管理

权限与安全

Q 请输入 % K

产品概述
快速入门
产品计费
开发参考

端到端实时语音大模型

语音播客大模型

语音同传大模型

语音妙记大模型

语音合成大模型

声音复刻大模型

语音识别大模型

产品简介

大模型流式语音识别
API大模型录音文件识别标
准版API大模型录音文件极速版
识别API大模型录音文件识别闲
时版

音频生成

语音识别

音视频字幕

SDK接入文档

自学习平台

控制台相关接口

相关协议

```
GET /api/v3/sauc/bigmodel
Host: openspeech/bytedance.com
X-Api-App-Key: 123456789
X-Api-Access-Key: your-access-key
X-Api-Resource-Id: volc.bigasr.sauc.duration
X-Api-Connect-Id: 随机生成的UUID

## 返回 Header
X-Tt-Logid: 202407261553070FACFE6D19421815D60
```

安全设置 访问控制 操作审计

API 访问密钥

工具与其他

公测申请 资源管理 配额中心

伙伴控制台

待办事项

待支付

0

未读消息

0

待续费

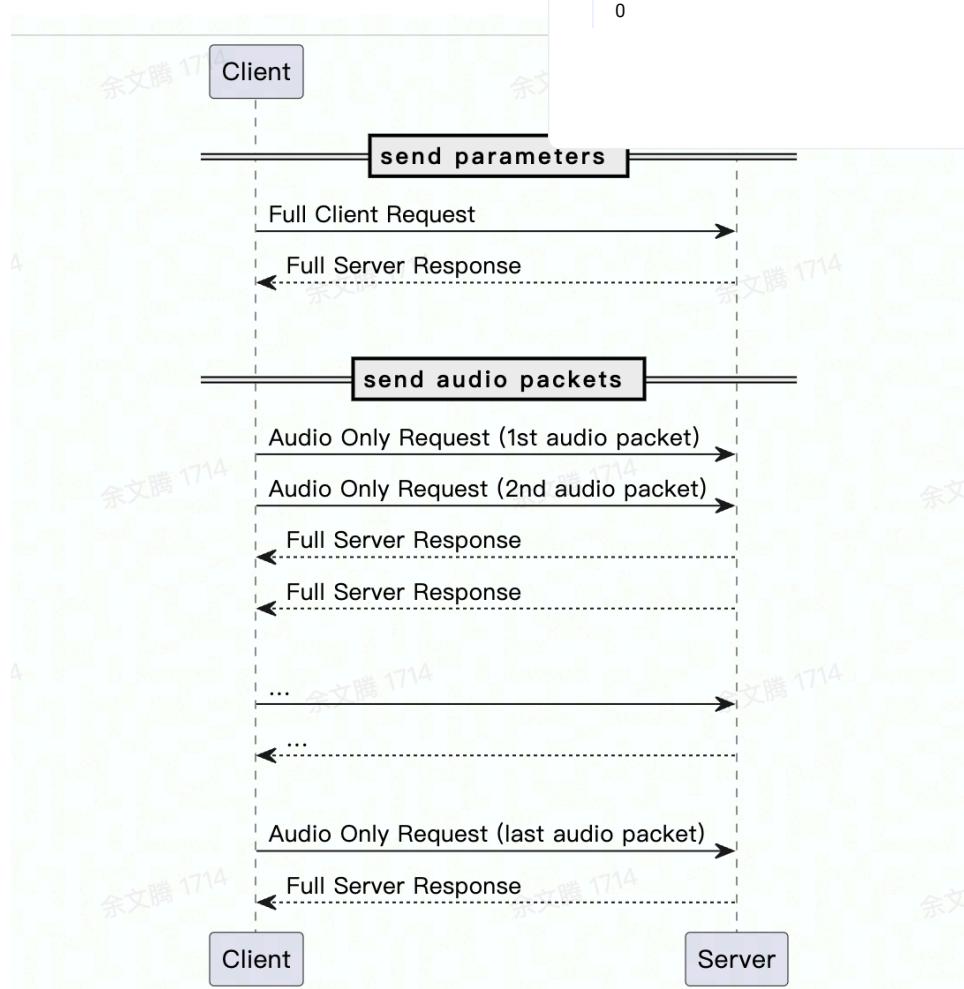
0

待处理工单

0

协议详情

交互流程



WebSocket 二进制协议

WebSocket 使用二进制协议传输数据。协议的组成由至少 4 个字节组成，其中 header 描述消息类型、序列化方式以及压缩格式等信体负载内容，依据消息类型不同 payload 内容不同。

需注意：协议中整数类型的字段都使用大端表示。

header 数据格式

Byte \ Bit	7	6	5
0	Protocol version		
1	Message type		
2	Message serialization method		
3	Reserved		

3 3827手机用户#I...

账号管
理

账号ID: 2103539692

主账号 个人认证 >

费用中心

可用余额 ¥ 0.00

充值汇款 >

账户总览 账单详情

费用分析

发票管理

权限与安全

Q 请输入 K

产品概述

快速入门

产品计费

开发参考

端到端实时语音大模型

语音播客大模型

语音同传大模型

语音妙记大模型

语音合成大模型

声音复刻大模型

语音识别大模型

产品简介

大模型流式语音识别 API

大模型录音文件识别标准版API

大模型录音文件极速版识别API

大模型录音文件识别闲置版

音频生成

语音识别

音视频字幕

SDK接入文档

自学习平台

控制台相关接口

相关协议

header 字段描述

字段 (size in bits)	说明	
Protocol version (4)	将来可能会决定使用不同的协议版本，此字段是为了使客户端和服务器在版本达成共识。	only request 0b1001 - 服务端下发包含识别结果的 full server response 0b1111 - 服务端处理错误时下发的消息类型（如无效的消息格式，不支持的序列化方法等）
Header (4)	Header 大小。实际 header 大小（以字节为单位）是 header size value x 4。	0b0000 - header后4个字节不为sequence number 0b0001 - header后4个字节为sequence number且为正 0b0010 - header后4个字节不为sequence number, 仅指示此为最后一包（负包） 0b0011 - header后4个字节为sequence number且需要为负数（最后一包/负包）
Message type (4)	消息类型。	Message type 的补充信息。
Message type specific flags (4)		full client request 的 payload 序列化方法；服务器将使用与客户端相同的序列化方法。 定义 payload 的压缩方法；服务端将使用客户端的压缩方法。 保留以供将来使用，还用作填充（使整个头部总计4个字节）。
Message serialization method (4)		0b0000 - 无序列化 0b0001 - JSON 格式
Message Compression (4)		0b0000 - no compression 0b0001 - Gzip 压缩
Reserved (8)		

请求流程

建立连接

根据 WebSocket 协议本身的机制，client 会发送 HTTP GET 请求建立连接。如果需要在其中根据身份认证协议加入鉴权签名头。设置方法请参考鉴权管理。

发送 full client request

WebSocket 建立连接后，发送的第一个请求是 full client request。

31 ... 24	23 ... 16	15 ... 8	7 ... 0
Header			
Payload size (4B, unsigned int32)			
Payload			

Header：前文描述的 4 字节头。

Payload size：是按 Header 中指定压缩方式压缩 payload 后的字节数。

Payload：包含音频的元数据以及 server 所需的相关参数，一般情况下为音频流。

3 3827手机用户#I...

账号管
理

账号ID : 2103539692

主账号 个人认证 >

费用中心

可用余额 ¥ 0.00

充值汇款 >

账户总览 账单详情 费用分析

发票管理

权限与安全

文档中心 火山方舟大模型服务平台

Q 请输入 % K

- 产品概述
- 快速入门
- 产品计费
- 开发参考
 - 端到端实时语音大模型
 - 语音播客大模型
 - 语音同传大模型
 - 语音妙记大模型
 - 语音合成大模型
 - 声音复刻大模型
 - 语音识别大模型
 - 产品简介
 - 大模型流式语音识别 API
 - 大模型录音文件识别标准版API
 - 大模型录音文件极速版识别API
 - 大模型录音文件识别闲置版
- 音频生成
- 语音识别
- 音视频字幕
- SDK接入文档
- 自学习平台
- 控制台相关接口
- 相关协议

咨询电话 售前客服 售后客服

user				置	1	dict	安全设置	访问控制	操作审计
uid	用户标识	2	string				API 访问密钥		
did	设备名称	2	string				工具与其他		
platform	操作系统及 API版本号	2	string				公测申请 资源管理 配额中心		
sdk_version	sdk版本	2	string				伙伴控制台		
app_version	app 版本	2	string				待办事项		
audio	音频相关配置	1	dict				待支付 0 待续费 0 待处理工单 0 未读消息 0		
language	指定可识别的语言	2	string				英语: en-US 日语: ja-JP 印尼语: id-ID 西班牙语: es-MX 葡萄牙语: pt-BR 德语: de-DE 法语: fr-FR 韩语: ko-KR 菲律宾语: fil-PH 马来语: ms-MY 泰语: th-TH 阿拉伯语: ar-SA 例如, 如果输入音频是德语, 则此参		
format	音频容器格式	2	string	✓			pcm / wav / ogg / mp3 注意: pcm和wav内部音频流必须是		
codec	音频编码格式	2	string				raw / opus, 默认为 raw(表示pcm) 注意: 当format为ogg的时候, codec 当format为mp3的时候, codec不生		
rate	音频采样率	2	int				默认为 16000, 目前只支持16000		
bits	音频采样点位数	2	int				默认为 16, 暂只支持16bits		
channel	音频声道数	2	int				1(mono) / 2(stereo), 默认为1。		
request	请求相关配置	1	dict	✓					
model_name	模型名称	2	string						
enable_non_stream	开启二遍识别	2	bool				3 3827手机用户#I... 账号管理 > 账号ID: 2103539692		
enable_itn	启用itn	2	bool				主账号 个人认证 >		
enable_punc	启用标点	2	bool				费用中心 可用余额 ¥ 0.00 充值汇款 >		
enable_ddc	启用顺滑	2	bool				账户总览 账单详情 费用分析		

Q 请输入 % K

- 产品概述
- 快速入门
- 产品计费
- 开发参考
- 端到端实时语音大模型
- 语音播客大模型
- 语音同传大模型
- 语音妙记大模型
- 语音合成大模型
- 声音复刻大模型
- 语音识别大模型
- 产品简介
- 大模型流式语音识别 API
- 大模型录音文件识别标准版API
- 大模型录音文件极速版识别API
- 大模型录音文件识别闲置版
- 音频生成
- 语音识别
- 音视频字幕
- SDK接入文档
- 自学习平台
- 控制台相关接口
- 相关协议

show_utterances
顿、分句、分词信息

show_speech_rate
(仅
nostream接
口和双向流
式优化版支
持)

show_volume
(仅
nostream接
口和双向流
式优化版支
持)

enable_lid
(仅
nostream接
口和双向流
式优化版支
持)

2

bool

2

bool

2

bool

2

bool

2

bool

2

bool

API 访问密钥

工具与其他

公测申请 资源管理 配额中心

伙伴控制台

待办事项

待支付 0 待续费 0 待处理工单 0

未读消息 0

如果设为"True", 则会在additions信
支持的标签包括:

- singing_en: 英文唱歌
- singing_mand: 普通话唱歌
- singing_dia_cant: 粤语唱歌
- speech_en: 英文说话
- speech_mand: 普通话说话
- speech_dia_nan: 闽南语
- speech_dia_wuu: 吴语 (含上海)
- speech_dia_cant: 粤语说话
- speech_dia_xina: 西南官话 (含重庆)
- speech_dia_zgyu: 中原官话 (含河南)
- other_langs: 其它语种 (其它语种)
- others: 检测不出 (非语义人声和空时代表无法判断 (例如传入音频))

实际不支持识别的语种 (无识别结果)

- singing_hi: 印度语唱歌
- singing_ja: 日语唱歌
- singing_ko: 韩语唱歌
- singing_th: 泰语唱歌
- speech_hi: 印地语说话
- speech_ja: 日语说话
- speech_ko: 韩语说话
- speech_th: 泰语说话

3 3827手机用户 #I... 账号管
理 >

账号ID: 2103539692

主账号 个人认证 >

费用中心

可用余额 ¥ 0.00

充值汇款 >

账户总览 账单详情 费用分析

发票管理

权限与安全



Q 请输入 % K

产品概述

快速入门

产品计费

开发参考

端到端实时语音大模型

语音播客大模型

语音同传大模型

语音妙记大模型

语音合成大模型

声音复刻大模型

语音识别大模型

产品简介

大模型流式语音识别 API

大模型录音文件识别标准版API

大模型录音文件极速版识别API

大模型录音文件识别闲置版

音频生成

语音识别

音视频字幕

SDK接入文档

自学习平台

控制台相关接口

相关协议

enable_emotion_recognition (仅 nostream 接口和双向流式优化版支持)

启用情绪检测

2

bool

待办事项

待支付

0

未读消息

0

待续费

0

待处理工单

0

如果设为 "True", 则会在分句 additio 双向流式优化版启用此功能会默认开

enable_gender_recognition (仅 nostream 接口和双向流式优化版支持)

启用性别检测

2

bool

result_type

结果返回方式

2

string

默认为 "full", 全量返回。

设置为 "single" 则为增量结果返回, E

enable_accelerate_text

是否启动首字返回加速

2

bool

如果设为 "True", 则会尽量加速首字 默认 "False"

accelerate_score

首字返回加速率

2

int

配合 enable_accelerate_text 参数使

vad_segment_duration

语义切句的最大静音阈值

2

int

单位 ms, 默认为 3000。当静音时间

end_window_size

强制判停时间

2

int

单位 ms, 默认为 800, 最小 200。静

force_to_speech_time

强制语音时间

2

int

单位 ms, 最小 1。音频时长超过该值 推荐设置 1000, 可能会影响识别准确

sensitive_words_filter

敏感词过滤

2

string

敏感词过滤功能, 支持开启或关闭, 支示例:
system_reserved_filter // 是否使用系 filter_with_empty // 想要替换成空的 filter_with_signed // 想要替换成 * 的

enable.poi_fc

(nostream 接口 & 双向流式优化版 - 开启二遍支持)

2

bool

3 3827 手机用户 #I...

账号管
理

专

账号 ID: 2103539692

主账号 个人认证 >

费用中心

可用余额 ¥ 0.00

充值汇款 >

账户总览 账单详情

费用分析

发票管理

权限与安全

文档中心 火山方舟大模型服务平台 文档 备案 控制台 3 安全设置 访问控制 操作审计

Q 请输入 K

产品概述 快速入门 产品计费 开发参考

端到端实时语音大模型 语音播客大模型 语音同传大模型 语音妙记大模型 语音合成大模型 声音复刻大模型 语音识别大模型

产品简介 大模型流式语音识别 API 大模型录音文件识别标准版API 大模型录音文件极速版识别API 大模型录音文件识别临时版

音频生成 语音识别 音视频字幕 SDK接入文档 自学习平台 控制台相关接口 相关协议

enable_music_fc (nostream 接口&双向流式优化版-开启二遍支持)

corpus 语料/干预词等

boosting_table_name 自学习平台上设置的热词词表名称

boosting_table_id 自学习平台上设置的热词词表id

correct_table_name 自学习平台上设置的替换词词表名称

correct_table_id 自学习平台上设置的替换词词表id

context 热词或者上下文

参数示例:

JSON

```
{ "user": { "uid": "388808088185088" }, "audio": { "format": "wav", "rate": 16000, "bits": 16, "channel": 1, } }
```

API 访问密钥 工具与其他 公测申请 资源管理 配额中心 伙伴控制台 待办事项 待支付 0 未读消息 0 待续费 0 待处理工单 0 替换词功能和设置方法可以参考文档 替换词功能和设置方法可以参考文档 1. 热词直连 (优先级高于传热词表) "context": { "hotwords": [{ "word": "热词" }] } 2. 上下文, 限制800 tokens及20轮 context_data字段按照从新到旧的顺序包流式语音识别模型2.0, 支持将通过理解图像内容, 帮助模型更精准图片限制传入1张, 大小: 500k以内 SQL 上下文: 可以加入对话历史、聊天历史 a. 对话历史: 把最近几轮的对话记录 b. 聊天所在bot信息: 如"我在和 c. 个性化信息: "我当前在北京市 d. 业务场景信息: "当前是中国平安 {"context_type": "text", "context_data": [{ "text": "热词", "image_url": null }] } 账号管理 账号ID: 2103539692 主账号 个人认证 >

费用中心 可用余额 ¥ 0.00 充值汇款 > 账户总览 账单详情 费用分析 发票管理 权限与安全

咨询电话 售前客服 售后客服

Q 请输入

⌘ K

产品概述

快速入门

产品计费

开发参考

端到端实时语音大模型

语音播客大模型

语音同传大模型

语音妙记大模型

语音合成大模型

声音复刻大模型

语音识别大模型

产品简介

大模型流式语音识别
API大模型录音文件识别标
准版API大模型录音文件极速版
识别API大模型录音文件识别闲
时版

音频生成

语音识别

音视频字幕

SDK接入文档

自学习平台

控制台相关接口

相关协议

```

    "model_name": "bigmodel",
    "enable_itn": false,
    "enable_ddc": false,
    "enable_punc": false,
    "corpus": {
        "boosting_table_id": "通过自学习平台",
        "context": [
            {
                "context_type": "dialog_ct",
                "context_data": [
                    {"text": "text1"},
                    {"text": "text2"},
                    {"text": "text3"},
                    {"text": "text4"},
                    ...
                ]
            }
        ]
    }
}

```

发送 audio only request

Client 发送 full client request 后，再发送包含音频数据的 audio-only client request。音频应采用 full client request 中指定的格式（音频格式、编解码器、采样率、声道）。格式如下：

31 ... 24	23 ... 16	15 ... 8	7 ... 0
Header			
Payload size (4B, unsigned int32)			
Payload			

Payload 是使用指定压缩方法，压缩音频数据后的内容。可以多次发送 audio only request 请求，例如在流式语音识别中如果每次发送 100ms 的音频数据，那么 audio only request 中的 Payload 就是 100ms 的音频数据。

full server response

Client 发送的 full client request 和 audio only request，服务端都会返回 full server response。格式如下：

31 ... 24	23 ... 16	15 ... 8	7 ... 0
Header			
Sequence			
Payload size (4B, unsigned int32)			
Payload			

Payload 内容是包含识别结果的 JSON 格式，字段说明如下：

字段	说明	层级	格式
result	识别结果	1	list
text	整个音频的识别结果文本	2	string
utterances	识别结果语音分句信息	2	list
text	utterance级的文本内容	3	string
start_time	起始时间 (毫秒)	3	int
end_time	结束时间 (毫秒)	3	int

3 3827手机用户#I... 账号管
理

账号ID : 2103539692

主账号 个人认证 >

费用中心

可用余额 ¥ 0.00 充值汇款 >

账户总览 账单详情 费用分析

发票管理

权限与安全

Q 请输入 % K

- 产品概述
- 快速入门
- 产品计费
- 开发参考
 - 端到端实时语音大模型
 - 语音播客大模型
 - 语音同传大模型
 - 语音妙记大模型
 - 语音合成大模型
 - 声音复刻大模型
 - 语音识别大模型
- 产品简介
- 大模型流式语音识别 API
 - 大模型录音文件识别标准版API
 - 大模型录音文件极速版识别API
 - 大模型录音文件识别闲置版
- 音频生成
- 语音识别
- 音视频字幕
- SDK接入文档
- 自学习平台
- 控制台相关接口
- 相关协议

definite 分句 3 bool

JSON

```
{
  "audio_info": {"duration": 10000},
  "result": [
    {
      "text": "这是字节跳动，今日头条母公司。",
      "utterances": [
        {
          "definite": true,
          "end_time": 1705,
          "start_time": 0,
          "text": "这是字节跳动，",
          "words": [
            {
              "blank_duration": 0,
              "end_time": 860,
              "start_time": 740,
              "text": "这"
            },
            {
              "blank_duration": 0,
              "end_time": 1020,
              "start_time": 860,
              "text": "是"
            },
            {
              "blank_duration": 0,
              "end_time": 1200,
              "start_time": 1020,
              "text": "字"
            },
            {
              "blank_duration": 0,
              "end_time": 1400,
              "start_time": 1200,
              "text": "节"
            }
          ]
        },
        {
          "definite": true,
          "end_time": 3696,
          "start_time": 2110,
          "text": "今日头条母公司。",
          "words": [
            {
              "blank_duration": 0,
              "end_time": 3070,
              "start_time": 2910,
              "text": "今"
            },
            {
              "blank_duration": 0,
              "end_time": 3230,
              "start_time": 3070,
              "text": "日"
            },
            {
              "blank_duration": 0,
              "end_time": 3390,
              "start_time": 3230,
              "text": "头"
            }
          ]
        }
      ]
    }
  ]
}
```

安全设置 访问控制 操作审计

API 访问密钥

工具与其他

公测申请 资源管理 配额中心

伙伴控制台

待办事项

待支付	待续费	待处理工单
0	0	0
未读消息 0		

3 3827手机用户#I... 账号管理 >
账号ID: 2103539692

主账号 个人认证 >

费用中心

可用余额 ¥ 0.00 充值汇款 >

账户总览 账单详情 费用分析

发票管理

权限与安全

Q 请输入 % K

产品概述

快速入门

产品计费

开发参考

端到端实时语音大模型

语音播客大模型

语音同传大模型

语音妙记大模型

语音合成大模型

声音复刻大模型

语音识别大模型

产品简介

大模型流式语音识别
API大模型录音文件识别标
准版API大模型录音文件极速版
识别API大模型录音文件识别闲
时版

音频生成

语音识别

音视频字幕

SDK接入文档

自学习平台

控制台相关接口

相关协议

```

        "end_time": 3550,
        "start_time": 3390,
        "text": "条"
    },
    {
        "blank_duration": 0,
        "end_time": 3670,
        "start_time": 3550,
        "text": "母"
    },
    {
        "blank_duration": 0,
        "end_time": 3696,
        "start_time": 3670,
        "text": "公"
    },
    {
        "blank_duration": 0,
        "end_time": 3696,
        "start_time": 3696,
        "text": "司"
    }
]
},
"audio_info": {
    "duration": 3696
}
}

```

文档 备案 控制台 3 安全设置 访问控制 操作审计

API 访问密钥

工具与其他

公测申请 资源管理 配额中心

伙伴控制台

待办事项

待支付 0

待续费 0

待处理工单 0

未读消息 0

Error message from server

当 server 发现无法解决的二进制/传输协议问题时，将发送 Error message from server 消息（例如，client 以 server 不支持的序列化格式发送消息）。格式如下：

31 ... 24	23 ... 16	15 ... 8	7 ... 0
Header			
Error message code (4B, unsigned int32)			
Error message size (4B, unsigned int32)			
Error message (UTF8 string)			

Header：前文描述的 4 字节头。

Error message code：错误码，使用大端表示。

Error message size：错误信息长度，使用大端表示。

Error message：错误信息。

示例

示例：客户发送 3 个请求

下面的 message flow 会发送多次消息，每个消息都带有版本、header 等字段。由于这些字段在所有消息中都相同，所以有些消息中这些字段省略了。

Message flow:

client 发送 "Full client request"

```

version: b0001 (4 bits)
header size: b0001 (4 bits)
message type: b0001 (Full client request) (4bits)
message type specific flags: b0000 (use_specific_pos_sequence)
message serialization method: b0001 (JSON) (4 bits)
message compression: b0001 (Gzip) (4bits)
reserved data: 0x00 (1 byte)
payload size = Gzip 压缩后的长度
payload: json 格式的请求字段经过 Gzip 压缩后的数据
server 响应 "Full server response"

```

3 3827手机用户#I...

账号管
理

账号ID : 2103539692

主账号 个人认证 >

费用中心

可用余额 ¥ 0.00

充值汇款 >

账户总览 账单详情 费用分析

发票管理

权限与安全

文档
中心

火山方舟大模型服务平台

文档 备案 控制台

3

9

立即入驻

安全设置 访问控制 操作审计

≡

Q 请输入

⌘ K

产品概述 ▼
 快速入门 ▼
 产品计费 ▼
 开发参考 ▲

端到端实时语音大模型▼
 语音播客大模型 ▼
 语音同传大模型 ▼
 语音妙记大模型 ▼
 语音合成大模型 ▼
 声音复刻大模型 ▼
 语音识别大模型 ▲

产品简介

大模型流式语音识别 API

大模型录音文件识别标准版API

大模型录音文件极速版识别API

大模型录音文件识别闲置版

音频生成 ▼
 语音识别 ▼
 音视频字幕 ▼
 SDK接入文档 ▼
 自学习平台 ▼
 控制台相关接口 ▼

相关协议 ▼

message type: **b1001** (Full server response)
 message type specific flags: **b0001** (none)
 message serialization method: **b0001** (JSON 和请求相同)
 message compression: **b0001** (Gzip 和请求相同)
 reserved data: **0x00**
 sequence: 0x00 0x00 0x00 0x01 (4 byte) sequence=1
 payload size = Gzip 压缩后数据的长度
 payload: Gzip 压缩后的响应数据

client 发送包含第一包音频数据的 "Audio only client request"
 version: **b0001**
 header size: **b0001**
 message type: **b0010** (audio only client request)
 message type specific flags: **b0000** (用户设置正数 sequence)
 message serialization method: **b0000** (none – raw bytes)
 message compression: **b0001** (Gzip)
 reserved data: **0x00**
 payload size = Gzip 压缩后的音频长度
 payload: 音频数据经过 Gzip 压缩后的数据

server 响应 "Full server response"
 message type: **0b1001** – Full server response
 message specific flags: **0b0001** (none)
 message serialization: **0b0001** (JSON, 和请求相同)
 message compression: **0b0001** (Gzip, 和请求相同)
 reserved data: **0x00**
 sequence data: 0x00 0x00 0x00 0x02 (4 byte) sequence=2
 payload size = Gzip 压缩后数据的长度
 payload: Gzip 压缩后的响应数据

client 发送包含最后一包音频数据 (通过 message type specific flags) 的 "Audio-only client request",
 message type: **b0010** (audio only client request)
 message type specific flags: **b0010** (最后一包音频请求)
 message serialization method: **b0000** (none – raw bytes)
 message compression: **b0001** (Gzip)
 reserved data: **0x00**
 payload size = Gzip 压缩后的音频长度
 payload: Gzip 压缩后的音频数据

server 响应 "Full server response" – 最终回应及处理结果
 message type: **b1001** (Full server response)
 message type specific flags: **b0011** (最后一包音频结果)
 message serialization method: **b0001** (JSON)
 message compression: **b0001** (Gzip)
 reserved data: **0x00**
 sequence data: **0x00 0x00 0x00 0x03** (4byte) sequence=3
 payload size = Gzip 压缩后的 JSON 长度
 payload: Gzip 压缩后的 JSON 数据

如处理过程中出现错误信息，可能有以下错误帧的返回

message type: **b1111** (error response)
 message type specific flags: **b0000** (none)
 message serialization method: **b0001** (JSON)
 message compression: **b0000** (none)
 reserved data: **0x00**
 Error code data: **0x2A 0x0D 0x0A2 0xff** (4byte) 错误码
 payload size = 错误信息对象的 JSON 长度
 payload: 错误信息对象的 JSON 数据

错误码



3 3827手机用户#I... 账号管理 >

账号ID: 2103539692

主账号 个人认证 >

费用中心

可用余额 ¥ 0.00 充值汇款 >

账户总览 账单详情 费用分析

发票管理

权限与安全

文档
中心

火山方舟大模型服务平台

文档 备案 控制台

3



安全设置

访问控制

操作审计



Q 请输入

⌘ K

产品概述

快速入门

产品计费

开发参考

端到端实时语音大模型

语音播客大模型

语音同传大模型

语音妙记大模型

语音合成大模型

声音复刻大模型

语音识别大模型

产品简介

大模型流式语音识别
API大模型录音文件识别标
准版API大模型录音文件极速版
识别API大模型录音文件识别闲
时版

音频生成

语音识别

音视频字幕

SDK接入文档

自学习平台

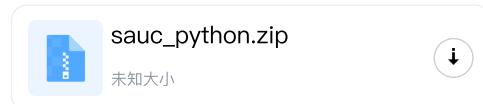
控制台相关接口

相关协议

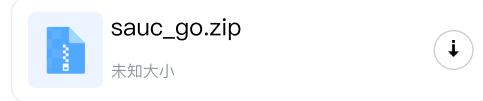
45000001	请求参数无效	请求参数缺失必需字段 /
45000002	空音频	
45000081	等包超时	
45000151	音频格式不正确	
550xxxxx	服务内部处理错误	
55000031	服务器繁忙	服务过载，无法处理当前

Demo

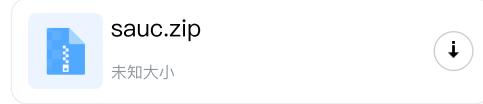
Python:



Go:



Java:



最近更新时间: 2026.01.08 19:49:13

这个页面对您有帮助吗?



< 上一篇产品简介

大模型录音文件识别标准版API 下一篇 >



极速服务应答

秒级应答为业务保驾护航



客户价值为先

从服务价值到创造客户价值



全方位安全保障

打造一朵“透明可信”的云



极速服务应答



客户价值为先



全方位安全保障

[豆包语音](#)[文档指南](#) 请输入

⌘ K

号院大钟寺广场1号楼

产品概述 ▼

快速入门 ▼

产品计费 ▼

开发参考 ^

[服务条款](#) | [隐私政策](#) | [更多协议](#)

端到端实时语音大模型▼

版权所有 代理域名注册服务机构：新网数码 商中在线 京公网安备11010802032137号 京ICP备20018813号-3 营业执照

语音播客大模型 ▼

;经营许可证京B2-20202418, A2.B1.B2-20202637 网络文化经营许可证：京网文（2023）4872-140号

语音同传大模型 ▼

语音妙记大模型 ▼

语音合成大模型 ▼

声音复刻大模型 ▼

语音识别大模型 ^

产品简介

**大模型流式语音识别
API**大模型录音文件识别标
准版API大模型录音文件极速版
识别API大模型录音文件识别闲
时版

音频生成 ▼

语音识别 ▼

音视频字幕 ▼

SDK接入文档 ▼

自学习平台 ▼

控制台相关接口 ▼

相关协议 ▼



