费用中心

公测申请

待支付订单

待续费项

待处理工单

退出登录

未读消息

 \equiv

0

0

0

10

Q 请输入关键字 2 文档中心 文档 备案 控制台 文档首页/豆包语音/ 豆包语音 开发参考/ 在本产品文档中搜索 Q 语音识别大模型/ 大模型流式语音识别API Q 搜索目录或文档标题 大模型流式语音识别API 产品概述 最近更新时间: 2025.08.14 15:55:55 首次发布时间: 2024.10.15 11:59:11 快速入门 □ □ 1 我的收藏 □ 有用 □ 无用 产品计费 开发参考 简介 端到端实时语音大模型 本文档介绍如何通过WebSocket协议实时访问大模型流式语音 豆包语音妙记 识别服务 (ASR), 主要包含鉴权相关、协议详情、常见问题和使 用Demo四部分。 语音播客大模型 双向流式模式使用的接口地址是 wss://openspeech.bytedance.com/api/v3/sauc/bigmodel ☐ 语音同传大模型 流式输入模式使用的接口地址是 语音合成大模型 wss://openspeech.bytedance.com/api/v3/sauc/bigmodel_n ostream₫ 声音复刻大模型 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 别到的字符,速度较快。 语音识别大模型 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 产品简介 返回识别到的结果,准确率更高。 无论是哪种模式,单包音频建议在100~200ms大小左右,不能 大模型流式语音识别API 过大或者过小, 否则均会影响性能。 (注:针对双向流式模式,单包为200ms大小时性能最优,建 大模型录音文件识别API 议双向流式模式选取200ms大小的分包) 大模型录音文件极速版识别API 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。 音频生成 语音识别 双向流式模式 (优化版本) 接口地址: 音视频字幕 wss://openspeech.bytedance.com/api/v3/sauc/bigmodel_a sync ☐ 自学习平台 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 SDK接入文档 (性能优化 rtf 和首字、尾字时延均有一定程度提升) 调用量查询接口说明 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。 QPS/并发查询接口说明 鉴权 控制台相关接口 相关协议 在 websocket 建连的 HTTP 请求头 (Header 中) 添加以下信 息 生成式模型服务专用条款

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

Neilc_5Qwx

账号ID: 2101484258

主账号 个人认证

语音播客大模型 语音同传大模型

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

音频生成

语音识别

音视频字幕

自学习平台

SDK接入文档

控制台相关接口

生成式模型服务专用条款

相关协议

调用量查询接口说明

QPS/并发查询接口说明

大模型流式语音识别API

大模型录音文件识别API

大模型录音文件极速版识别API

Q 请输入关键字 备案 控制台 文档中心 文档 识别 volc.bigasr.sauc .concurrent • 大模型流式语音 豆包语音 识别 67ee89ba-用于追踪当前连接的 X-Api-7050-4c04-Q 标志 ID、推荐设置 Connect-Id a3d7-UUID等 ac61a63499b3 产品概述 websocket 握手成功后,会返回这些 Response header。强烈 快速入门 建议记录X-Tt-Logid (logid) 作为排错线索。 产品计费 说明 Value 示例 开发参考 端到端实时语音大模型 豆包语音妙记

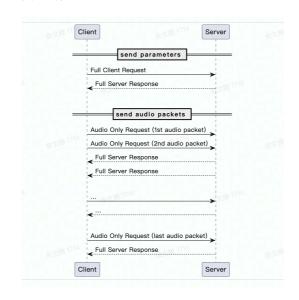
0	费用中心		=
Ç	公测申请		_
	待支付订单	0	
	待续费项	0	
	待处理工单	0	
	未读消息	10	
	退出登录		

Key	1元明	value 示例
X-Api- Connect-Id	用于追踪当前调用 信息的标志 ID,推 荐用UUID等	67ee89ba-7050- 4c04-a3d7- ac61a63499b3
X-Tt-Logid	服务端返回的 logid,建议用户获 取和打印方便定位 问题	2024072615530 70FACFE6D1942 1815D605

HTTP // 建连 HTTP 请求头示例 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.duratio X-Api-Connect-Id: 随机生成的UUID ## 返回 Header X-Tt-Logid: 202407261553070FACFE6D19421815D

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 字节的可变 header、payload size 和 payload 三部分组成,其 中 header 描述消息类型、序列化方式以及压缩格式等信息,

X

Neilc_5Qwx 账号ID: 2101484258

主账号 个人认证

https://www.volcengine.com/docs/6561/1354869

文档中心	O 语标	俞入关键字	文档			费用中心	=
文档中心	Q IAH		, , ,		~0	公测申请	=
		0	Protocol version			待支付订单	0
豆包语音		1	Message type			待续费项	0
(2	2	Message serialization	n method		待处理工单	0
		3	Reserved			未读消息	10
产品概述	~	4	[Optional header exte	ensions]		退出登录	
快速入门	~	5	[Payload, depending	on the Message Type]			
产品计费	~	6					
开发参考	^	header 字段描	÷*				
端到端实时语音大模型	~	neader J. FX TH	<u>γ</u> υ				
豆包语音妙记	~	字段 (size in bits)	说明	值			
语音播客大模型	~		将来可能会决定使				
语音同传大模型	~	Protocol	用不同的协议版 本,因此此字段是	0b0001 - version			
语音合成大模型	~	version (4)	为了使客户端和服 务器在版本上达成	1 (目前只有该版本)			
声音复刻大模型	~		共识。				
语音识别大模型	^		Header 大小。实				
产品简介		Header (4)	际 header 大小 (以字节为单位)	0b0001 - header size = 4 (1 x 4)			
大模型流式语音识别API			是 header size value x 4。				
大模型录音文件识别API				0b0001 - 端上发			
大模型录音文件极速版识别人	\PI			送包含请求参数的			
音频生成	~			full client request 0b0010 - 端上发			
语音识别	~			送包含音频数据的 audio only			
音视频字幕	~			request			
自学习平台	~	Message	消息类型。	0b1001 - 服务端 下发包含识别结果			
SDK接入文档	~	type (4)		的 full server response			
调用量查询接口说明				0b1111 - 服务端			
QPS/并发查询接口说明				处理错误时下发的 消息类型(如无效			
控制台相关接口	~			的消息格式,不支 持的序列化方法			
相关协议	~			等)			
生成式模型服务专用条款		Message type specific flags (4)	Message type 的补充信息。	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

Neilc_5Qwx

账号ID: 2101484258

主账号 个人认证

2

3/8/24 21:30			大模型流式语音识别API豆包语音-火山引					
	文档中心	Q 请转	渝入关键字	文档	备案 控制台 N			
豆包语音					sequence number且需要为 负数(最后一包/ 负包)			
产品概述快速入门产品计费	Q		Message serialization method (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。	0b0000 - 无序列 化 0b0001 - JSON 格式			
开发参考端到端实时语音大	^		Message Compression (4)	定义 payload 的压缩方法; 服务端将使用客户端的压缩方法。	0b0000 - no compression 0b0001 - Gzip 压 缩			
豆包语音妙记 语音播客大模型 语音同传大模型	· ·		Reserved (8)	保留以供将来使 用,还用作填充 (使整个标头总计 4个字节)。				
语音合成大模型	~		请求流程					
声音复刻大模型语音识别大模型	^		建立连接					
产品简介	音识别API		请求和 server 建立	协议本身的机制,clien 立连接做协议升级。 份认证协议加入鉴权签	nt 会发送 HTTP GET 名头。设置方法请参考			

鉴权。

发送 full client request

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

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

的长度,使用**大端**表示。

Payload: 包含音频的元数据以及 server 所需的相关参数,一 般是 JSON 格式。具体的参数字段见下表:

字段	说明	层级	格式
user	用户相关配 置	1	dict
uid	用户标识	2	string
did	设备名称	2	string
platform	操作系统及 API版本号	2	string
sdk_versio n	sdk版本	2	string
ann versio			

费用中心 \equiv 公测申请 0 待支付订单 待续费项 0 待处理工单 0 未读消息 10 退出登录

X

Neilc_5Qwx

账号ID: 2101484258

主账号 个人认证

大模型录音文件识别API

音频生成

语音识别

音视频字幕

自学习平台

SDK接入文档

调用量查询接口说明

控制台相关接口

生成式模型服务专用条款

相关协议

QPS/并发查询接口说明

大模型录音文件极速版识别API

٥

	Q 请输入关键字		文档 备案	控制台 N
	audio	置	1	dict
豆包语音	format	音频容器格 式	2	string
Q	codec	音频编码格 式	2	string
产品概述	rate	音频采样率	2	int
快速入门	bits	音频采样点	2	int
产品计费		位数		
开发参考 へ	channel	音频声道数	2	int
端到端实时语音大模型 ~	request	请求相关配 置	1	dict
豆包语音妙记	model_na	模型名称	2	string
语音播客大模型	me	X = 113	_	
语音同传大模型	model_vers ion	模型版本	2	string
语音合成大模型				
声音复刻大模型	enable_non	开启二遍识	2	bool
语音识别大模型・ペーパー・ペープ	stream	别		
产品简介				
大模型流式语音识别API	enable_itn	启用itn	2	bool
大模型录音文件识别API				
大模型录音文件极速版识别API	enable_pun c	启用标点	2	bool
音频生成				
语音识别	enable_ddc	启用顺滑	2	bool
音视频字幕		输出语音停		
自学习平台	show_utter ances	顿、分句、	2	bool
SDK接入文档	ahaw anaa	分词信息		
调用量查询接口说明	show_spee ch_rate			
QPS/并发查询接口说明	(仅 nostream接	分句信息携		
控制台相关接口	口支持,双	带语速	2	bool
相关协议	向流式和二 遍识别下不			
生成式模型服务专用条款	支持)			
	show_volu me(仅 nostream接 口支持,双 向流式和二 遍识别下不 支持)	分句信息携带音量	2	bool

 费用中心

 公测申请
 0

 待支付订单
 0

 待续费项
 0

 待处理工单
 0

 未读消息
 10

Neilc_5Qwx

×



文	(档中心	Q 请输入关键字		文档 备案	控制台 N	O ₂	费用中心公测申请	≡
豆包语音		向流式和二 遍识别下不 支持)					待支付订单 待续费项	0
	Q	result_type	结果返回方 式	2	string		待处理工单 未读消息	0
产品概述	~	vad_segme nt_duration	语义切句的 最大静音阈	2	int		退出登录	
快速入门	~		值					
产品计费	~	end_windo w_size	强制判停时 间	2	int			
开发参考	^	force_to_s	强制语音时	2	int			
端到端实时语音大模	型	peech_time	间	_				
豆包语音妙记	~							
语音播客大模型	~							
语音同传大模型	~	sensitive_w ords_filter	敏感词过滤	2	string			
语音合成大模型	~	ords_liiter						
声音复刻大模型	~							
语音识别大模型	^							
产品简介		corpus	语料/干预 词等	2	dict			
大模型流式语音说	只别API	boosting_t	自学习平台					
大模型录音文件说	只别API	able_name	上设置的热 词词表名称	3	string			
大模型录音文件机	及速版识别API		自学习平台					
音频生成	~	boosting_t able_id	上设置的热 词词表id	3	string			
语音识别	~		自学习平台					
音视频字幕	~	correct_tab le_name	上设置的替 换词词表名	3	string			
自学习平台	~	ie_name	称					
SDK接入文档	~	correct_tab	自学习平台	_				
调用量查询接口说明		le_id	上设置的替 换词词表id	3	string			
QPS/并发查询接口证	说明	context	热词或者上	3	string			
控制台相关接口	~		下文					
相关协议	~							
生成式模型服务专用条款	款							
								×
							N Neilc_5Qwx 账号ID: 210148425	8
							主账号	



31 ... 24

Header

23 ... 16

15 ... 8

```
    费用中心

    公测申请

    待支付订单
    0

    待续费项
    0

    待处理工单
    0

    未读消息
    10
```

Neilc_5Qwx 账号ID: 2101484258

主账号 个人认证

X

文档中心	Q 请输入关键字	(/ ~ · J · · ·	文档 备第	K 控制台 N	O _o	费用中心 公测申请		≡
豆包语音	Payload					待支付订单	0	
	Payload 内容是	是包含识别结果的] JSON 格式,	字段说明如下:		待续费项	0	
Q	字段	说明	层级	格式		待处理工单 未读消息	0	
产品概述	result	识别结果	1	list		退出登录		
快速入门		整个音频的						
产品计费	text	识别结果文本	2	string				
开发参考へ		'						
端到端实时语音大模型 ~	uttoronoo	识别结果语	0	liet				
豆包语音妙记	utterances	音分句信息	2	list				
语音播客大模型								
语音同传大模型		utterance						
语音合成大模型	text	级的文本内 容	3	string				
声音复刻大模型		т.						
语音识别大模型								
产品简介	start_time	起始时间 (毫秒)	3	int				
大模型流式语音识别API		(=12)						
大模型录音文件识别API								
大模型录音文件极速版识别API	and time	结束时间	3	int				
音频生成	end_time	(毫秒)	3	IIIL				
语音识别								
音视频字幕		日本日 △						
自学习平台	definite	是否是一个确定分句	3	bool				
SDK接入文档								
调用量查询接口说明	3000							
QPS/并发查询接口说明	JSON {			의 ① 다				
控制台相关接口	"resul							
相关协议		ext": <mark>"这是字</mark> tterances":		头条母公司。 ",				
生成式模型服务专用条款		"words": [{ "blank "end_t	: 1705, e": 0, X是字节跳动, _duration": ime": 860, _time": 740	0,				
		}, {					×	
		"end_t	_duration": ime": 1020, _time": 860			N Neilc_5Qwx 账号ID: 210148		
62 0 0						主账号 个人	认证	

```
费用中心
                             Q 请输入关键字
                                                                                 2
                                                                                                               \equiv
              文档中心
                                                         文档 备案 控制台
                                                                                      公测申请
                                               },
                                               {
                                                                                      待支付订单
                                                                                                          0
                                                 "blank_duration": 0,
豆包语音
                                                 "end_time": 1400,
                                                                                      待续费项
                                                                                                           0
                                                 "start_time": 1200,
                                                 "text": "节"
                                                                                      待处理工单
                                                                                                           0
                       Q
                                               },
                                                                                      未读消息
                                                                                                          10
                                               {
                                                 "blank_duration": 0,
产品概述
                                                 "end time": 1560,
                                                                                              退出登录
                                                 "start_time": 1400,
快速入门
                                                 "text": "跳"
                                               },
产品计费
                                               {
                                                 "blank_duration": 0,
开发参考
                                                 "end_time": 1640,
                                                 "start_time": 1560,
                                                 "text": "动"
  端到端实时语音大模型
                                               }
                                             ]
  豆包语音妙记
                                           },
  语音播客大模型
                                             "definite": true,
                                             "end_time": 3696,
  语音同传大模型
                                             "start_time": 2110,
                                             "text": "今日头条母公司。",
  语音合成大模型
                                             "words": [
                                               {
  声音复刻大模型
                                                 "blank_duration": 0,
                                                 "end_time": 3070,
  语音识别大模型
                                                 "start_time": 2910,
                                                 "text": "今"
    产品简介
                                               },
                                               {
    大模型流式语音识别API
                                                 "blank_duration": 0,
                                                 "end_time": 3230,
    大模型录音文件识别API
                                                 "start_time": 3070,
                                                 "text": "∃"
    大模型录音文件极速版识别API
                                               },
  音频生成
                                                 "blank_duration": 0,
                                                 "end_time": 3390,
  语音识别
                                                 "start_time": 3230,
                                                 "text": "头"
  音视频字幕
                                               },
                                               {
  自学习平台
                                                 "blank_duration": 0,
                                                 "end_time": 3550,
  SDK接入文档
                                                 "start_time": 3390,
                                                 "text": "条"
  调用量查询接口说明
                                               },
                                               {
  QPS/并发查询接口说明
                                                 "blank_duration": 0,
                                                 "end_time": 3670,
  控制台相关接口
                                                 "start_time": 3550,
                                                 "text": "母"
相关协议
                                               },
生成式模型服务专用条款
                                                 "blank_duration": 0,
                                                 "end_time": 3696,
                                                 "start_time": 3670,
                                                 "text": "公"
                                               },
                                                 "blank_duration": 0,
                                                 "end_time": 3696,
                                                 "start_time": 3696,
                                                 "text": "司"
                                               }
                                                                                                         X
                                             ]
                                           }
                                                                                           Neilc_5Qwx
                                         ]
                                                                                            账号ID: 2101484258
                                       },
                                                                                            主账号 个人认证
```

费用中心

Q 请输入关键字 2 文档中心 文档 备案 控制台 \equiv 公测申请 待支付订单 0 Error message from server 豆包语音 待续费项 0 当 server 发现无法解决的二进制/传输协议问题时,将发送 待处理工单 Λ Error message from server 消息 (例如, client 以 server 不 0 支持的序列化格式发送消息)。格式如下: 未读消息 10 31 ... 24 23 ... 16 15 ... 8 7 ... 0 产品概述 退出登录 Header 快速入门 Error message code (4B, unsigned int32) 产品计费 Error message size (4B, unsigned int32) 开发参考 Error message (UTF8 string) 端到端实时语音大模型 Header: 前文描述的 4 字节头。 豆包语音妙记 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。 语音播客大模型 Error message: 错误信息。 语音同传大模型 示例 语音合成大模型 示例: 客户发送 3 个请求 声音复刻大模型 下面的 message flow 会发送多次消息,每个消息都带有版 语音识别大模型 本、header 大小、保留数据。由于每次消息中这些字段值相 产品简介 同, 所以有些消息中这些字段省略了。 Message flow: 大模型流式语音识别API client 发送 "Full client request" 大模型录音文件识别API version: b0001 (4 bits) header size: b0001 (4 bits) 大模型录音文件极速版识别API message type: b0001 (Full client request) (4bits) 音频生成 message type specific flags: b0000 (use_specific_pos_sequence) (4bits) 语音识别 message serialization method: b0001 (JSON) (4 bits) 音视频字幕 message compression: b0001 (Gzip) (4bits) reserved data: 0x00 (1 byte) 自学习平台 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 SDK接入文档 server 响应 "Full server response" 调用量查询接口说明 version: b0001 QPS/并发查询接口说明 header size: b0001 控制台相关接口 message type: b1001 (Full server response) message type specific flags: b0001 (none) 相关协议 message serialization method: b0001 (JSON 和请求相同) message compression: b0001 (Gzip 和请求相同) 生成式模型服务专用条款 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 X message type: b0010 (audio only client request) Neilc_5Qwx message type specific flags: b0000 (用户设置正数 账号ID: 2101484258 sequence number) message serialization method: haaaa (none - raw hytes) 主账号 个人认证

2

Q 请输入关键字 文档中心 文档 备案 控制台 server 响应 "Full server response" 豆包语音 message type: 0b1001 - Full server response message specific flags: 0b0001 (none) message serialization: 0b0001 (JSON, 和请求相同) 0 message compression 0b0001 (Gzip, 和请求相同) reserved data: 0x00 sequence data: 0x00 0x00 0x00 0x02 (4 byte) sequence=2 产品概述 payload size = Gzip 压缩后数据的长度 快速入门 payload: Gzip 压缩后的响应数据 产品计费 client 发送包含最后一包音频数据(通过 message type specific flags) 的 "Audio-only client request", 开发参考 message type: b0010 (audio only client request) 端到端实时语音大模型 message type specific flags: b0010 (最后一包音频请求) message serialization method: b0000 (none - raw bytes) 豆包语音妙记 message compression: b0001 (Gzip) 语音播客大模型 reserved data: 0x00 payload size = Gzip 压缩后的音频长度 语音同传大模型 payload: Gzip 压缩后的音频数据 语音合成大模型 server 响应 "Full server response" - 最终回应及处理结果 声音复刻大模型 message type: b1001 (Full server response) message type specific flags: b0011 (最后一包音频结果) 语音识别大模型 message serialization method: b0001 (JSON) 产品简介 message compression: b0001 (Gzip) reserved data: 0x00 大模型流式语音识别API sequence data: 0x00 0x00 0x00 0x03 (4byte) sequence=3 大模型录音文件识别API payload size = Gzip 压缩后的 JSON 长度 payload: Gzip 压缩后的 JSON 数据 大模型录音文件极速版识别API 如处理过程中出现错误信息,可能有以下错误帧的返回 音频生成 message type: b1111 (error response) 语音识别 message type specific flags: b0000 (none) message serialization method: b0001 (JSON) 音视频字幕 message compression: b0000 (none) reserved data: 0x00 自学习平台 Error code data: 0x2A 0x0D 0x0A2 0xff (4byte) 错误码 SDK接入文档 payload size = 错误信息对象的 JSON 长度 payload: 错误信息对象的 JSON 数据 调用量查询接口说明 错误码 QPS/并发查询接口说明 控制台相关接口 错误码 含义 说明 相关协议 20000000 成功 生成式模型服务专用条款 请求参数缺失必需字段 / 请求参数无效 45000001 字段值无效 / 重复请 求。 45000002 空音频 45000081 等包超时

音频格式不正

服务内部处理

服务过载, 无法处理当

确

错误

45000151

550xxxxx

 费用中心

 公测申请

 待支付订单
 0

 待续费项
 0

 待处理工单
 0

 未读消息
 10

 \equiv

×

Neilc_5Qwx

账号ID: 2101484258

主账号 个人认证



服务条款 隐私政策 更多协议

所有 代理域名注册服务机构: 新网数码 商中在线 ❷ 京公网安备11010802032137号 京ICP备20018813号-3 营业执照 午可证京B2-20202418, A2.B1.B2-20202637 网络文化经营许可证: 京网文 (2023) 4872-140号

69





业务咨询