

文档中心

Q 请输入关键字

文档 备案 控制台 N

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

退出登录

豆包语音

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

豆包语音妙记

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

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

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

QPS/并发查询接口说明

控制台相关接口

相关协议

生成式模型服务专用条款

Byte	Bit	说明	值
0		Protocol version	
1		Message type	
2		Message serialization method	
3		Reserved	
4		[Optional header extensions]	
5		[Payload, depending on the Message Type]	
6		...	

header 字段描述

字段 (size in bits)	说明	值
Protocol version (4)	将来可能会决定使用不同的协议版本，因此此字段是为了使客户端和服务端在版本上达成共识。	0b0001 - version 1 (目前只有该版本)
Header (4)	Header 大小。实际 header 大小 (以字节为单位) 是 header size value x 4。	0b0001 - header size = 4 (1 x 4)
Message type (4)	消息类型。	0b0001 - 端上发送包含请求参数的 full client request 0b0010 - 端上发送包含音频数据的 audio 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)
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

文档中心

Q 请输入关键字

文档 备案 控制台 N

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

退出登录

豆包语音

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

豆包语音妙记

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

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

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

QPS/并发查询接口说明

控制台相关接口

相关协议

生成式模型服务专用条款

sequence number且需要为负数（最后一包/负包）

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

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 compression0b0001 - Gzip 压缩

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）。

请求流程

建立连接

根据 WebSocket 协议本身的机制，client 会发送 HTTP GET 请求和 server 建立连接做协议升级。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。

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_version	sdk版本	2	string
app_version			

Neilc_5Qwx

账号ID：2101484258

主账号 个人认证

文档中心

Q 请输入关键字

文档 备案 控制台 N

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

退出登录

豆包语音

Q

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

豆包语音妙记

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

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

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

QPS/并发查询接口说明

控制台相关接口

相关协议

生成式模型服务专用条款

audio	音频格式	1	dict
format	音频容器格式	2	string
codec	音频编码格式	2	string
rate	音频采样率	2	int
bits	音频采样点位数	2	int
channel	音频声道数	2	int
request	请求相关配置	1	dict
model_name	模型名称	2	string
model_version	模型版本	2	string
enable_non_stream	开启二遍识别	2	bool
enable_itn	启用itn	2	bool
enable_punc	启用标点	2	bool
enable_ddc	启用顺滑	2	bool
show_utterances	输出语音停顿、分句、分词信息	2	bool
show_speech_rate (仅nostream接口支持，双向流式和二遍识别下不支持)	分句信息携带语速	2	bool
show_volume (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

N Neilc_5Qwx

账号ID：2101484258

主账号 个人认证

文档中心

Q 请输入关键字

文档 备案 控制台 N

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

退出登录

豆包语音

Q

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

豆包语音妙记

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

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

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

QPS/并发查询接口说明

控制台相关接口

相关协议

生成式模型服务专用条款

enable_lid (仅nostream接口支持，双向流式和二遍识别下不支持)	启用语种检测	2	bool
enable_emotion_detection (仅nostream接口支持，双向流式和二遍识别下不支持)	启用情绪检测	2	bool
enable_gen	启用性别检测	2	bool

N

Neilc_5Qwx
账号ID：2101484258

主账号 个人认证

文档中心

Q 请输入关键字

文档 备案 控制台 N

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

退出登录

豆包语音

Q

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

豆包语音妙记

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

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

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

QPS/并发查询接口说明

控制台相关接口

相关协议

生成式模型服务专用条款

向流式和二遍识别下不支持)			
result_type	结果返回方式	2	string
vad_segment_duration	语义切句的最大静音阈值	2	int
end_window_size	强制判停时间	2	int
force_to_speech_time	强制语音时间	2	int
sensitive_words_filter	敏感词过滤	2	string
corpus	语料/干预词等	2	dict
boosting_table_name	自学习平台上设置的热词词表名称	3	string
boosting_table_id	自学习平台上设置的热词词表id	3	string
correct_table_name	自学习平台上设置的替换词词表名称	3	string
correct_table_id	自学习平台上设置的替换词词表id	3	string
context	热词或者上下文	3	string

Neilc_5Qwx

账号ID：2101484258

主账号 个人认证

文档中心

Q 请输入关键字

文档 备案 控制台 N

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

退出登录

豆包语音

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

豆包语音妙记

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

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

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

QPS/并发查询接口说明

控制台相关接口

相关协议

生成式模型服务专用条款

参数示例:

JSON

```
{  "user": {    "uid": "388808088185088"  },  "audio": {    "format": "wav",    "rate": 16000,    "bits": 16,    "channel": 1,    "language": "zh-CN"  },  "request": {    "model_name": "bigmodel",    "enable_itn": false,    "enable_ddc": false,    "enable_punc": false,    "corpus": {      "boosting_table_id": "通过自学习"    },    "context": {      "context_type": "dialog",      "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			

文档中心

Q 请输入关键字

文档 备案 控制台 N

费用中心
公测申请

豆包语音

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

豆包语音妙记

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

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

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

QPS/并发查询接口说明

控制台相关接口

相关协议

生成式模型服务专用条款

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
definite	是否是一个确定分句	3	bool

JSON

```
{  "audio_info": {"duration": 10000},  "result": {    "text": "这是字节跳动， 今日本头条母公司。",    "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,
```

退出登录

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

文档中心

Q 请输入关键字

文档 备案 控制台 N

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

退出登录

豆包语音

Q

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

豆包语音妙记

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

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

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

QPS/并发查询接口说明

控制台相关接口

相关协议

生成式模型服务专用条款

```
},
{
  "blank_duration": 0,
  "end_time": 1400,
  "start_time": 1200,
  "text": "节"
},
{
  "blank_duration": 0,
  "end_time": 1560,
  "start_time": 1400,
  "text": "跳"
},
{
  "blank_duration": 0,
  "end_time": 1640,
  "start_time": 1560,
  "text": "动"
}
]
},
{
  "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"
    },
    {
      "blank_duration": 0,
      "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": {
```

N Neilc_5Qwx

账号ID: 2101484258

主账号 个人认证

豆包语音

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 大小、保留数据。由于每次消息中这些字段值相同，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) (4bits)
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"

```

version:  b0001
header size:  b0001
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: 0001
header size: 0001
message type: 0010 (audio only client request)
message type specific flags: 0000 (用户设置正数
sequence number)
message serialization method: 0000 (none – raw bytes)
```

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](#)

产品简介

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

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

QPS/并发查询接口说明

控制台相关接口

相关协议

生成式模型服务专用条款

x

N

Neilc_5Qwx

账号ID: 2101484258

主账号

个人认证

Q



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



调用量查询接口说明

QPS/并发查询接口说明

生成式模型服务专用条款

payload: Gzip 压缩后的响应数据

payload: Gzip 压缩后的音频数据

payload: Gzip 压缩后的 JSON 数据

payload: 错误信息对象的 JSON 数据

错误码	含义	说明
20000000	成功	
45000001	请求参数无效	请求参数缺失必需字段 / 字段值无效 / 重复请求。
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	
		服务过载，无法处理当

[退出登录](#)

X

N

Neilc_5Qwx

账号ID: 2101484258

主账号

个人认证

文档中心

Q 请输入关键字

文档 备案 控制台 N

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

退出登录

Python:

sauc_python.zip

未知大小

↓

Go:

sauc_go.zip

未知大小

↓

Java:

java_client.zip

未知大小

↓

上一篇
产品简介

下一篇
大模型录音文件识别API

产品简介

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

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

QPS/并发查询接口说明

控制台相关接口

相关协议

生成式模型服务专用条款

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

📍钟寺广场1号楼