

腾讯云智能数智人 客户端渲染 API 接入



腾讯云

【 版权声明 】

©2013–2024 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【 商标声明 】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

文档目录

客户端渲染 API 接入

概述

创建长连接通道

端渲染驱动接口

客户端渲染 API 接入概述

最近更新时间：2024-08-09 19:25:01

本文档主要描述 腾讯云智能数智人端渲染驱动aPaas 的开放接口协议。

接口调用环境

正式环境：`https://gw.tvs.qq.com`

注意事项

调用该接口前请确保下单购买了下图红框中的产品：

云智能数智人 [返回产品详情](#)

选择配置

套餐包类型	交互数智人-形象租赁 <small>本产品不包括对话服务。</small>	交互数智人-并发增购	数智人视频生成服务-小时包 <small>本产品不支持退款</small>
-------	-------------------------------------------------	------------	------------------------------------------

接口调用方式

调用每个接口时需要在接口的 query 中传入公共参数及签名。参数及签名方式说明参考以下链接：

[数智人 aPaas 接口调用方式](#)

接口接入流程

接入指引：

第1步，开通数智人平台账号，开通方式参考 [购买指南](#)；

第2步，登录平台 [数智人交互平台](#)。创建数智人项目，得到项目的 appkey、accesstoken、virtualmankey，通过 [创建长连接通道](#) 接口建立websocket连接；

第3步，通过长连接通道发送驱动请求，并获取数智人驱动下行消息。请求、返回内容见 [端渲染驱动接口](#)；

第4步，如果驱动结束，根据DriverType驱动类型，发送结束标记或断开连接。

创建长连接通道

最近更新时间：2024-08-09 19:25:01

建立 websocket 长连接通道，上行发送驱动指令、接收下行消息。

调用协议

WebSocket + JSON

WSS /v2/ws/ivh/interactdriver/interactdriverservice/driverengine

Header Content-Type: application/json;charset=utf-8

建连方式

websocket url: 参考 [数智人 aPaas 接口调用方式](#) 接口调用方式组织 url。

端渲染驱动接口

最近更新时间：2024-09-06 17:54:52

[创建长连接通道](#)后，可以通过 websocket 长连接发送文本获取驱动数据。

请求参数

参数名称	类型	必选	描述
ReqId	string	是	单次请求唯一标识，长度为32的uuid
StreamId	string	是	会话ID，用于区分多轮会话
VirtualmanProjectId	string	是	数智人项目ID，可在数智人项目中获取
InputText	string	否	请求文本内容。当DriverType为TEXT时，不可为空
SpeechParam	SpeechParam	否	定义输出音频的详细参数
DriverType	string	是	驱动类型 1.TEXT: 文本驱动; 2.CHAT: 文本对话驱动 3.STREAM_TEXT: 流式文本驱动;
ChatCommand	string	否	对话指令，默认值CHATTING 1.CHATTING: 对话 2.START_CHAT: 开启对话 3.STOP_CHAT: 中断对话
InputTextType	string	否	InputText的类型，默认为MARKDOWN 1.MARKDOWN: markdown格式，包含纯文本，可支持流式 2.SSML: ssml标准格式，不支持流式
Seq	int	否	流式文本分片的ID
IsFinal	bool	否	流式文本分片的结束标记（每一段流式文本结束必须传入结束标记）

SpeechParam

参数名称	类型	必选	描述
------	----	----	----

TimbreKey	string	否	音色值，默认使用数智人项目配置音色。可使用的音色列表，可通过 分页查询音色列表 获取音色取值
Speed	float	否	语速: 1.0为正常语速，范围[0.5-2.0]，值为0.5时播报语速最慢，值为2.0时播报语速最快。未指定时默认取数智人项目配置的语速值
Volume	int	否	音量大小，范围[-10, 10]，对应音量大小。默认为0，代表正常音量，值越大音量越高
EmotionCategory	string	否	控制合成音频的情感，仅支持多情感音色使用，可选值参考 分页查询音色列表
EmotionIntensity	int	否	控制合成音频情感程度，取值范围为 [50,200]，只有 EmotionCategory 不为空时生效
SmartActionEnabled	bool	否	是否开启智能动作，默认不开启
SubtitleType	int	否	字幕返回的模式，按字级别还是词级别。默认按字 0: 按字; 1: 按词;
TimbreLanguage	string	否	音色语种，可选语种参考个人资产管理 API 分页查询音色列表 接口，多语种音色在合成时必须选择对应语种。

长连接下行消息

参数名称	类型	必选	描述
ReqId	string	是	单次请求ID，和入参一致
StreamId	string	是	会话ID，用于区分多轮会话，和入参一致
DriverRspType	string	是	响应类型 1.REPLY: 返回ReplyRsp，对应会话信息 2.SPEECH: 返回SpeechRsp，对应音频信息
ReplyRsp	ReplyRsp	否	会话响应，当DriverRsp为REPLY时返回
SpeechRsp	SpeechRsp	否	音频信息相应，当DriverRsp为SPEECH时返回
ErrorCode	int	是	错误码

ErrorMessage	string	是	错误信息
--------------	--------	---	------

ReplyRsp

参数名称	类型	必选	描述
ReplyType	string	是	回复语类型。 1.cloudAiGpt: 腾讯云大模型对话 2.yunxiaowei: 云小微客服对话 3.cloudAiWaiting: 首包超时等待话术 4.cloudAiTimeOut: 超时未响应话术, 会话结束 5.sensitive: 输入文本或回复语中包含敏感内容时, 返回的固定话术 6.input: InputText为纯文本或流式文本时输入的内容 7.enhanceText: 未配置对话服务时, 匹配到话术管理的内容
ReplyPro	string	否	播报内容, 包含ssml标签
ReplyDisplay	string	是	展示内容, 包含富文本标签
InteractionType	string	否	特殊消息类型
InteractionContent	string	否	特殊消息内容, 用于下发弹窗、图片等非文本类的特殊消息
Uninterrupt	bool	是	当前播报内容是否可打断
Muted	bool	是	当前播报内容, 是否关闭收音
SeqNo	int	是	子句序号, 当ReplyType为cloudAiGpt时, 正常回复语序号从1开始, 其余固定话术从0开始
ContentType	int	是	回复语内容类型 0: 未知 1: 普通字符串 2: 有序列表 3: 无序列表 4: 图片链接 5: http链接 6: 表格 8: 标题 9: SSML

TtsSupport	bool	是	当前子句是否播报
IsFinal	bool	是	是否为最后一句
IsHighLight	bool	是	是否需要高亮展示

SpeechRsp

参数名称	类型	必选	描述
Audio	string	是	base64编码的pcm音频数据
ThDim	int	是	口型维度
ThFeat	Array of float	是	口型数据
Phn	Array of [PhnInfo]	是	音素信息
Word	Array of [WordInfo]	是	分词信息
Final	bool	是	整句结束标识
SentenceFinal	bool	是	流式子句结束标识
Sampling	int	是	采样率
Action	Array of [Action]	是	动作信息
Subtitle	Array of [SubtitleInfo]	是	字幕信息
RealThType	string	是	口型参数
Expression	Array of [Expression]	是	表情信息
SeqNo	int	是	子句序号

SentenceStart	bool	是	子句开始
ThFeatFinal	bool	是	口型结束标识

PhnInfo

参数名称	类型	必选	描述
Phn	string	是	音素
Start	string	是	起始时间，单位为0.1us，该数值/10000为ms
End	string	是	结束时间，单位为0.1us，该数值/10000为ms

WordInfo

参数名称	类型	必选	描述
Phn	string	是	音素
Word	string	是	对应单词

Action

参数名称	类型	必选	描述
Pos	string	是	动作名称
Start	string	是	起始时间点，单位为0.1us，该数值/10000为ms

SubtitleInfo

参数名称	类型	必选	描述
Word	string	是	对应单词
Start	string	是	起始时间点，单位为0.1us，该数值/10000为ms
End	string	是	结束时间点，单位为0.1us，该数值/10000为ms
PosStart	string	是	文本中的起始unicode位置，注意为左闭右开形式 [PosStart,PosEnd)
PosEnd	string	是	文本中的结束unicode位置，注意为左闭右开形式 [PosStart,PosEnd)

Expression

参数名称	类型	必选	描述
Name	string	是	表情名称
Start	string	是	起始时间点，单位为0.1us，该数值/10000为ms
End	string	是	结束时间点，单位为0.1us，该数值/10000为ms
Loc	string	是	表情在文本中的unicode位置
Flag	string	是	B: 该段文本包含表情起始; I: 该段文本属于表情中间的一部分 E: 该段文本表情结束 S: 该段文本包含表情的起始、结束

请求示例

```

{
  "Header": {},
  "Payload": {
    "VirtualmanProjectId": "253b2a182d694a60bed82635b18025a2",
    "InputText": "在人工智能产业中，哪些领域的AI发展基础条件表现较  
优?",
    "ReqId": "d7aa08da33dd4a662ad5be508c5b77cf",
    "StreamId": "92597c35-3a99-415e-9bae-3124771b7749",
    "DriverType": "TEXT",
    "SpeechParam": {
      "TimbreKey": ""
    }
  }
}
    
```

返回示例

```

//DriverRspType为REPLY
{
  "Header": {
    "RequestID": "fe0e4c13f2a34cb69b2475d8483f28de",
    "SessionID": "gza802cc9317231084402578413",
    "DialogID": "",
    "Code": 0,
    
```

```
    "Message": ""
  },
  "Payload": {
    "DriverRspType": "REPLY",
    "ErrorCode": 0,
    "ErrorMessage": "",
    "ReplyRsp": {
      "ContentType": 1,
      "InteractionContent": "",
      "InteractionType": "",
      "IsFinal": true,
      "IsHighLight": true,
      "Muted": false,
      "ReplyDisplay": "哪些领域的AI发展基",
      "ReplyPro": "\u003cspeak\u003e哪些领域的AI发展基础条件表现较优?
\u003c/speak\u003e",
      "ReplyType": "input",
      "SeqNo": 2,
      "TtsSupport": true,
      "UninterrId": "fe0e4c13f2a34cb69b2475d8483f28de",
      "SpeechRsp": {
        "Action": [],
        "Audio": "",
        "Expression": [],
        "Final": false,
        "Phn": [],
        "RealThType": "",
        "Sampling": 0,
        "SentenceFinal": false,
        "SentenceStart": false,
        "SeqNo": 0,
        "Subtitle": [],
        "ThDim": 0,
        "ThFeat": [],
        "ThFeatFinal": false,
        "Word": []
      },
    },
    "StreamId": "92597c35-3a99-415e-9bae-3124771b7749"
  }
}
}
}

//DriverRspType为SPEECH
```

```
{
  "Header": {
    "RequestID": "fe0e4c13f2a34cb69b2475d8483f28de",
    "SessionID": "gza802cc9317231084402578413",
    "DialogID": "",
    "Code": 0,
    "Message": ""
  },
  "Payload": {
    "DriverRspType": "SPEECH",
    "ErrorCode": 0,
    "ErrorMessage": "",
    "ReplyRsp": {
      "ContentType": 0,
      "InteractionContent": "",
      "InteractionType": "",
      "IsFinal": false,
      "IsHighLight": false,
      "Muted": false,
      "ReplyDisplay": "",
      "ReplyPro": "",
      "ReplyType": "",
      "SeqNo": 0,
      "TtsSupport": false,
      "Uninterrupt": false
    },
    "ReqId": "fe0e4c13f2a34cb69b2475d8483f28de",
    "SpeechRsp": {
      "Action": [],
      "Audio": "", //内容过长，不展示
      "Expression": [],
      "Final": false,
      "Phn": [
        {
          "End": "200000",
          "Phn": "sil0",
          "Start": "0"
        },
        {
          "End": "1100000",
          "Phn": "z4",
          "Start": "200000"
        }
      ],
    }
  }
}
```

```
{
  "End": "2700000",
  "Phn": "ai4",
  "Start": "1100000"
},
{
  "End": "3800000",
  "Phn": "r2",
  "Start": "2700000"
},
{
  "End": "5100000",
  "Phn": "en2",
  "Start": "3800000"
},
{
  "End": "5800000",
  "Phn": "g1",
  "Start": "5100000"
},
{
  "End": "7300000",
  "Phn": "ong1",
  "Start": "5800000"
},
{
  "End": "8100000",
  "Phn": "zh4",
  "Start": "7300000"
},
{
  "End": "9000000",
  "Phn": "iii4",
  "Start": "8100000"
},
{
  "End": "9800000",
  "Phn": "n2",
  "Start": "9000000"
},
{
  "End": "11300000",
  "Phn": "eng2",
```

```
    "Start": "9800000",
  },
  {
    "End": "12800000",
    "Phn": "ch3",
    "Start": "11300000"
  },
  {
    "End": "14000000",
    "Phn": "an3",
    "Start": "12800000"
  },
  {
    "End": "16000000",
    "Phn": "ie4",
    "Start": "14000000"
  },
  {
    "End": "17100000",
    "Phn": "zh1",
    "Start": "16000000"
  },
  {
    "End": "19200000",
    "Phn": "ong1",
    "Start": "17100000"
  },
  {
    "End": "24200000",
    "Phn": "sil0",
    "Start": "19200000"
  }
],
"RealThType": "3D_standard",
"Sampling": 24000,
"SentenceFinal": false,
"SentenceStart": true,
"SeqNo": 1,
"Subtitle": [
  {
    "End": "2700000",
    "PosEnd": "1",
    "PosStart": "0",
```

```
"Start": "200000",
"Word": "在"
},
{
  "End": "5100000",
  "PosEnd": "2",
  "PosStart": "1",
  "Start": "2700000",
  "Word": "人"
},
{
  "End": "7300000",
  "PosEnd": "3",
  "PosStart": "2",
  "Start": "5100000",
  "Word": "工"
},
{
  "End": "9000000",
  "PosEnd": "4",
  "PosStart": "3",
  "Start": "7300000",
  "Word": "智"
},
{
  "End": "11300000",
  "PosEnd": "5",
  "PosStart": "4",
  "Start": "9000000",
  "Word": "能"
},
{
  "End": "14000000",
  "PosEnd": "6",
  "PosStart": "5",
  "Start": "11300000",
  "Word": "产"
},
{
  "End": "16000000",
  "PosEnd": "7",
  "PosStart": "6",
  "Start": "14000000",
```



```
    "Word": "业"
  },
  {
    "End": "19200000",
    "PosEnd": "9",
    "PosStart": "7",
    "Sta": "16000000",
    "Word": "中, "
  }
],
"ThDim": 52,
"ThFeat": [], //内容过长, 不展示
"ThFeatFinal": false,
"Word": [
  {
    "Phn": "z-ai4",
    "Word": "在"
  },
  {
    "Phn": "r-en2|g-ng1",
    "Word": "人工"
  },
  {
    "Phn": "zh-iii4|n-eng2",
    "Word": "智能"
  },
  {
    "Phn": "ch-an3|ie4",
    "Word": "产业"
  },
  {
    "Phn": "zh-ong1",
    "Word": "中"
  }
]
},
"StreamId": "925999-415e-9bae-3124771b7749"
}
```