

捷通华声 灵云公有云能力平台 HTTP-TTS 接口 开发手册

灵云科技 源自清华 服务全球

北京捷通华声语音技术有限公司

Beijing InfoQuick SinoVoice Speech Technology Corp.

2014 年 6 月

目 录

1. 概述.....	1
2. HTTP 头.....	1
3. TTS 访问接口.....	2
3.1. TTS 合成请求接口（同步）	2
3.2. 访问	2
3.3. 服务器应答	2
3.4. 示例	2
3.5. x-task-config 参数说明.....	3

1. 概述

灵云HTTP-TTS接口，客户端通过HTTP-TTS接口使用POST方式与灵云平台进行通讯，返回数据为XML格式。

2. HTTP 头

客户端POST请求时HTTP头信息，必须带上如下：

顺序	名称	值（例子）	含义	备注
1	x-app-key	ac5d5431	应用标识	必选，由捷通公司提供
2	x-sdk-version	3.6	sdk 版本号	必选，使用 http 接口时，可以指定为 3.6
3	x-request-date	2014-6-18 10:10:11	请求时间	必选，每个任务的请求时间不能重复
4	x-task-config	capkey=asr.cloud.freetalk, audioformat=alaw	任务参数信息	为 name=value 形式，多个参数以逗号隔开 capkey 为必选 splitsentence 为可选，默认为 yes
5	x-session-key	18EA9384AA34	请求数据签名	必选 x-session-key 生成算法说明： x-session-key =md5（x-request-date + devkey）
6	x-udid	101:1234567890	设备标识	必选，如使用设备取设备标识号，如不使用设备设置为例子中的默认值
6	x-tid		保留字段	保留字段，不用设置
7	x-eid		保留字段	保留字段，不用设置

注：头的内容必须为英文。

devkey是应用开发者密钥，在注册开发者用户时由开发者网站统一分配，用作请求数据签名；

3. TTS 访问接口

3.1. TTS 合成请求接口（同步）

本接口在提交合成本时，同步在 http 请求的回应消息中返回语音数据。

3.2. 访问

(POST) [ServiceUrl]/tts/SynthText

说明：包体中传输文本数据，包体最大 1024 字节，文本内容仅支持 utf8 字符编码。

需要以下 HTTP 头：x-app-key、x-sdk-version、x-request-date、x-task-config、x-session-key

3.3. 服务器应答

应答格式：XML文件

XML的元素（Element）及含义

顺序	名称	值（例子）	含义	备注
1	ResCode	Success/FAILED	完毕/处理失败	必选
2	ResMessage	Success/FAILED	完毕/处理失败	必选
3	ErrorNo	4	返回错误码	失败时必选
4	Result_token	123456	结果令牌	成功必选
5	Data_Len	4096	语音数据长度	携带语音数据时，成功必须

3.4. 示例

```
<?xml version="1.0" encoding="utf-8" ?>
```

```
<ResponseInfo>
```

```
    <ResCode> Success </ResCode>
```

```
    <ResMessage> Success </ResMessage>
```

```
    <ErrorNo>0</ErrorNo>
```

```
    < Result_token>123456</Result_token>
```

```
    <Data_Len>4096</ Data_Len >
```

</ResponseInfo>

[语音数据块.....]

说明：携带的语音数据接在</ResponseInfo>标记后）。

3.5. x-task-config 参数说明

以下是有效参数：

✧ capkey（语音合成能力名称，必选）

Capkey	发音人	语种	支持 tagmode
tts.cloud.wangjing	wangjing	中文女声	none,auto,s3ml
tts.cloud.haobo	haobo	中文男声	none,auto,s3ml
tts.cloud.xiaokun	xiaokun	中文女声	none,auto,s3ml
tts.cloud.nancy	nancy	英语女声	none
tts.cloud.hui	hui	中文女声	none ,vtml
tts.cloud.liang	liang	中文男声	none,vtml
tts.cloud.xiaojie	xiaojie	中文粤语女声	none,auto,s3ml
tts.cloud.julie	julie	英语女声	none,vtml
tts.cloud.kate	kate	英语女声	none,vtml
tts.cloud.paul	paul	英语男声	none,vtml
tts.cloud.yumi	yumi	韩语女声	none,vtml
tts.cloud.junwoo	junwoo	韩语男声	none,vtml
tts.cloud.milena	milena	俄罗斯女声	none,ssml
tts.cloud.audrey-ml	audrey-ml	法语女声	none,ssml
tts.cloud.thomas	thomas	法语男声	none,ssml
tts.cloud.sebastien	sebastien	法语男声	none,ssml

tts.cloud.anna	anna	德语女声	none,ssml
tts.cloud.steffi	steffi	德语女声	none,ssml
tts.cloud.yannick	yannick	德语男声	none,ssml
tts.cloud.violeta	violeta	西班牙语女声	none,vtml
tts.cloud.diego	diego	西班牙语男声	none,ssml
tts.cloud.vera	vera	葡萄牙语女声	none,ssml
tts.cloud.joana	joana	葡萄牙语女声	none,ssml
tts.cloud.maged	maged	阿拉伯语男声	none,ssml
tts.cloud.kyoko	kyoko	日语女声	none,ssml
tts.cloud.misaki	misaki	日语女声	none,vtml
tts.cloud.narae	narae	韩语女声	none,ssml
tts.cloud.alice-ml	alice-ml	意大利女声	none,ssml
tts.cloud.narisa	narisa	泰语女声	none,ssml
tts.cloud.damayanti	damayanti	印尼语女声	none,ssml

◇ domain（领域），取值范围及含义（仅部分发音人支持）：

有效参数	参数说明
finance	金融证券
weather	天气预报
queue	排队叫号
insurance	保险
voyage	航空
revenue	税务

elecpower	电力
message	短信

- ✧ pitch（合成语音的基频）：按照0~9分为10级，缺省值为5。数字越大基频越高。
- ✧ volume（合成语音的音量）：按照0~9分为10级，缺省值为9。数字越大音量越大。
- ✧ speed（合成语音的语速）：按照0~9分为10级，缺省值为5。数字越大语速越快。
- ✧ puncmode（标点符号读法）：取值范围及含义：

有效参数	参数说明
off	不合成标点符号，自动判断回车换行是否分隔符
on	合成标点符号，自动判断回车换行是否分隔符
off_rtn	不读符号，强制将回车换行作为分隔符
on_rtn	读符号，强制将回车换行作为分隔符

- ✧ digitmode（数字阅读方式），取值范围及含义：

有效参数	参数说明
auto_number	自动方式，即根据上下文自动选择电报方式或数目方式合成，如果无法判断，则使用数目方式。
telegram	电报方式，即逐个数字合成。例如：1997 读成“一九九七”。
number	数目方式，即连续的数字合成。例如：1997 读成“一千九百九十七”。
auto_telegram	自动方式，即根据上下文自动选择电报方式或数目方式合成，如果无法判断，则使用电报方式。

- ✧ engmode（英文阅读方式），取值范围及含义：

有效参数	参数说明
auto	自动判断，可能按字母、自录音词汇或者 SAPI 合成。
english	强制所有英文用字母或自录音词汇合成。
letter	强制所有英文按照单字母方式合成。

✧ tagmode（标注处理方式），各capkey可选值参见capkey列表，取值范围及含义：

有效参数	参数说明
none	文本中无任何标记，所有标记将会被视为文本读出
auto	自动判断标记类型
vtml	仅处理 VTML 标记
ssml	仅处理 SSML 标记，例如 <voice gender="female" />
s3ml	处理 S3ML 标记

✧ voicestyle（朗读风格），取值范围及含义：

有效参数	参数说明
normal	抑扬顿挫
plain	平稳庄重

✧ audioformat（音频格式，必选）

有效参数	参数说明
auto	WAV 文件格式，语音采样率和每采样的位数由系统内部决定
vox6k4bit	6KHz，4Bit VOX 格式，也即 Dialogic ADPCM 格式
vox8k4bit	8KHz，4Bit VOX 格式，也即 Dialogic ADPCM 格式
alaw8k8bit	8KHz，8Bit ALaw 格式
ulaw8k8bit	8KHz，8Bit uLaw 格式
pcm8k8bit	8KHz，8Bit PCM 格式
pcm8k16bit	8KHz，16Bit PCM 格式
pcm16k8bit	16KHz，8Bit PCM 格式
pcm16k16bit	16KHz，16Bit PCM 格式

pcm11k8bit	11KHz, 8Bit PCM 格式
pcm11k16bit	11KHz, 16Bit PCM 格式
mp3	mp3 64kbps 格式
mp3_24	mp3 24kbps 格式
mp3_16	mp3 16kbps 格式