ALIBABA CLOUD

Alibaba Cloud

智能语音交互

语音合成

文档版本：20220415

阿里云

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。

2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。

3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。

4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。

5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。

6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	危险 重置操作将丢失用户配置数据。
警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	警告 重启操作将导致业务中断，恢复业务时间约十分钟。
注意	用于警示信息、补充说明等，是用户必须了解的内容。	注意 权重设置为0，该服务器不会再接受新请求。
说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击**设置**> **网络**> **设置网络类型**。
粗体	表示按键、菜单、页面名称等UI元素。	在**结果确认**页面，单击**确定**。
`Courier字体`	命令或代码。	执行 `cd /d C:/window` 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	`bae log list --instanceid` *Instance_ID*
[] 或者 [a|b]	表示可选项，至多选择一个。	`ipconfig [-all|-t]`
{} 或者 {a|b}	表示必选项，至多选择一个。	`switch {active|stand}`

目录

1.接口说明 ---------- 05

1.接口说明

语音合成为您提供将输入文本合成为语音二进制数据的功能。

功能介绍

- 支持输出PCM、WAV和MP3编码格式数据。
- 支持设置语速、语调和音量。
- 支持设置不同类型的声音。

> **注意**
>
> 字级别音素边界接口：语音合成服务在输出音频的同时，可输出每个汉字/英文单词在音频中的时间位置，即时间戳。该时间信息可用于驱动虚拟人口型、做视频配音字幕等。详情请参见语音合成时间戳功能介绍。

名称	voice参数值	类型	适用场景	支持语言	支持采样率(Hz)	支持字级别音素边界接口	备注
小云	Xiaoyun	标准女声	通用场景	中文及中英文混合场景	8K/16K	否	无
小刚	Xiaogang	标准男声	通用场景	中文及中英文混合场景	8K/16K	否	无
若兮	Ruoxi	温柔女声	通用场景	中文及中英文混合场景	8K/16K/24K	否	无
思琪	Siqi	温柔女声	通用场景	中文及中英文混合场景	8K/16K/24K	是	无
思佳	Sijia	标准女声	通用场景	中文及中英文混合场景	8K/16K/24K	否	无
思诚	Sicheng	标准男声	通用场景	中文及中英文混合场景	8K/16K/24K	是	无

名称	voice参数值	类型	适用场景	支持语言	支持采样率(Hz)	支持字级别音素边界接口	备注
艾琪	Aiqi	温柔女声	通用场景	中文及中英文混合场景	8K/16K	是	无
艾佳	Aijia	标准女声	通用场景	中文及中英文混合场景	8K/16K	是	无
艾诚	Aicheng	标准男声	通用场景	中文及中英文混合场景	8K/16K	是	无
艾达	Aida	标准男声	通用场景	中文及中英文混合场景	8K/16K	是	无
宁儿	Ninger	标准女声	通用场景	纯中文场景	8K/16K/24K	否	无
瑞琳	Ruilin	标准女声	通用场景	纯中文场景	8K/16K/24K	否	无
思悦	Siyue	温柔女声	客服场景	中文及中英文混合场景	8K/16K/24K	否	无
艾雅	Aiya	严厉女声	客服场景	中文及中英文混合场景	8K/16K	是	无
艾夏	Aixia	亲和女声	客服场景	中文及中英文混合场景	8K/16K	是	无
艾美	Aimei	甜美女声	客服场景	中文及中英文混合场景	8K/16K	是	无

名称	voice参数值	类型	适用场景	支持语言	支持采样率(Hz)	支持字级别音素边界接口	备注
艾雨	Aiyu	自然女声	客服场景	中文及中英文混合场景	8K/16K	是	无
艾悦	Aiyue	温柔女声	客服场景	中文及中英文混合场景	8K/16K	是	无
艾婧	Aijing	严厉女声	客服场景	中文及中英文混合场景	8K/16K	是	无
小美	Xiaomei	甜美女声	客服场景	中文及中英文混合场景	8K/16K/24K	否	无
艾娜	Aina	浙普女声	客服场景	纯中文场景	8K/16K	是	无
伊娜	Yina	浙普女声	客服场景	纯中文场景	8K/16K/24K	否	无
思婧	Sijing	严厉女声	客服场景	纯中文场景	8K/16K/24K	是	无
思彤	Sitong	儿童音	童声场景	纯中文场景	8K/16K/24K	否	无
小北	Xiaobei	萝莉女声	童声场景	纯中文场景	8K/16K/24K	是	无
艾彤	Aitong	儿童音	童声场景	纯中文场景	8K/16K	是	无

名称	voice参数值	类型	适用场景	支持语言	支持采样率(Hz)	支持字级别音素边界接口	备注
艾薇	Aiwei	萝莉女声	童声场景	纯中文场景	8K/16K	是	无
艾宝	Aibao	萝莉女声	童声场景	纯中文场景	8K/16K	是	无
Harry	Harry	英音男声	英文场景	英文场景	8K/16K	否	无
Abby	Abby	美音女声	英文场景	英文场景	8K/16K	否	无
Andy	Andy	美音男声	英文场景	英文场景	8K/16K	否	无
Eric	Eric	英音男声	英文场景	英文场景	8K/16K	否	无
Emily	Emily	英音女声	英文场景	英文场景	8K/16K	否	无
Luna	Luna	英音女声	英文场景	英文场景	8K/16K	否	无
Luca	Luca	英音男声	英文场景	英文场景	8K/16K	否	无
Wendy	Wendy	英音女声	英文场景	英文场景	8K/16K/24K	否	无
William	William	英音男声	英文场景	英文场景	8K/16K/24K	否	无
Olivia	Olivia	英音女声	英文场景	英文场景	8K/16K/24K	否	无
姗姗	Shanshan	粤语女声	方言场景	标准粤文（简体）及粤英文混合场景	8K/16K/24K	否	无

名称	voice参数值	类型	适用场景	支持语言	支持采样率(Hz)	支持字级别音素边界接口	备注
小玥	Xiaoyue	四川话女声	方言场景	中文及中英文混合场景	8K/16K	否	公测版
Lydia	Lydia	英中双语女声	英文场景	英文场景	8K/16K	否	公测版
艾硕	Aishuo	自然男声	客服场景	中文及中英文混合场景	8K/16K	是	公测版
青青	Qingqing	中国台湾话女声	方言场景	中文场景	8K/16K	否	公测版
翠姐	Cuijie	东北话女声	方言场景	中文场景	8K/16K	否	公测版
小泽	Xiaoze	湖南重口音男声	方言场景	中文场景	8K/16K	是	公测版

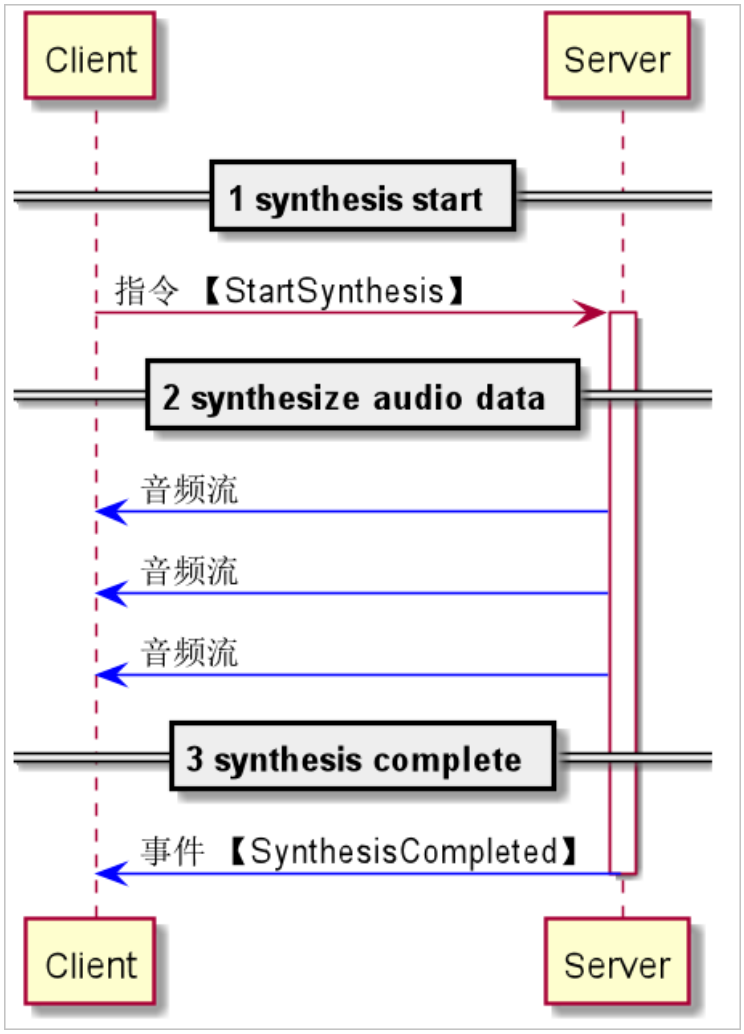
调用限制

- 输入文本必须采用 `UTF-8` 编码。
- 输入文本不能超过300个字符，超过300字符的内容会被截断。

服务地址

访问类型	说明	URL
外网访问	所有服务器均可使用外网访问URL（SDK中默认设置了外网访问URL）。	wss://nls-gateway.ap-southeast-1.aliyuncs.com/ws/v1

交互流程



说明

- 上图不包含RESTful API的交互流程，关于RESTful API的交互流程图请参见RESTful API。
- 服务端的响应除了音频流之外，都会在返回信息的header包含本次识别任务的task_id参数，请记录该值。如果出现错误，请将task_id和错误信息提交到工单。

1. 鉴权

客户端在与服务端建立WebSocket连接时，使用Token进行鉴权。关于Token获取请参见获取Token。

2. 开始合成

客户端发起语音合成请求，在请求消息中进行参数设置，各参数通过SDK中SpeechSynthesizer对象的set方法设置，含义如下。

参数	类型	是否必需	说明
appkey	String	是	管控台创建的项目appkey。

参数	类型	是否必需	说明
text	String	是	待合成文本，文本内容必须采用 `UTF-8` 编码，长度不超过300个字符（英文字母之间需要添加空格）。
voice	String	否	发音人，默认是 `xiaoyun` 。
format	String	否	音频编码格式，默认值：PCM。支持PCM/WAV/MP3格式。
sample_rate	Integer	否	音频采样率，默认值：16000。
volume	Integer	否	音量，取值范围：0～100。默认值：50。
speech_rate	Integer	否	语速，取值范围：-500～500，默认值：0。
pitch_rate	Integer	否	语调，取值范围：-500～500，默认值：0。

3. 接收合成数据

服务端返回合成的语音二进制数据，SDK接收并处理二进制数据。

4. 结束合成

语音合成完毕，服务端发送合成完毕事件通知，举例如下。

```
{
    "header": {
        "message_id": "05450bf69c53413f8d88aed1ee60****",
        "task_id": "640bc797bb684bd6960185651307****",
        "namespace": "SpeechSynthesizer",
        "name": "SynthesisCompleted",
        "status": 20000000,
        "status_message": "GATEWAY|SUCCESS|Success."
    }
}
```

说明

文档示例将合成的音频保存在文件中，如果您需要播放音频且对实时性要求较高，建议使用流式播放，即边接收语音数据边播放，减少延时。

服务状态码

服务的每一次响应都包含status字段，即服务状态码，各状态码含义如下。

通用错误：

错误码	原因	解决办法
40000001	身份认证失败	检查使用的令牌是否正确，是否过期。
40000002	无效的消息	检查发送的消息是否符合要求。
403	令牌过期或无效的参数	首先检查使用的令牌是否过期，然后检查参数值设置是否合理。
40000004	空闲超时	确认是否长时间（10秒）没有发送数据到服务端。
40000005	请求数量过多	检查是否超过了并发连接数或者每秒钟请求数。如果超过并发数，建议从免费版升级到商用版，或者商用版扩容并发资源。
40000000	默认的客户端错误码	查看错误消息或提交工单。
50000000	默认的服务端错误	如果偶现可以忽略，重复出现请提交工单。

错误码	原因	解决办法
50000001	内部调用错误	如果偶现可以忽略，重复出现请提交工单。

网关错误：

错误码	原因	解决办法
40010001	不支持的接口	使用了不支持的接口，如果使用SDK请提交工单。
40010002	不支持的指令	使用了不支持的指令，如果使用SDK请提交工单。
40010003	无效的指令	指令格式错误，如果使用SDK请提交工单。
40010004	客户端提前断开连接	检查是否在请求正常完成之前关闭了连接。
40010005	任务状态错误	发送了当前任务状态不能处理的指令。

配置错误：

错误码	原因	解决办法
40020105	应用不存在	检查应用appkey是否正确，是否与令牌归属同一个账号。

TTS（Text to Speech）错误：

错误码	原因	解决办法
41020001	参数错误	检查是否传递了正确的参数。
51020001	TTS服务端错误	如果偶现可以忽略，重复出现请提交工单。