

百度语音合成 REST API

帮助文档

v1.0



北京百度网讯科技有限公司
(版权所有, 翻版必究)

修改记录

版本号	修改内容简介	修改日期
1.0	第一版	2015-03-06

目录

目录.....	3
1 概念解释.....	4
2 简介.....	4
2.1 功能介绍.....	4
3 集成指南.....	4
3.1 注册百度开发者.....	4
3.2 创建应用.....	5
3.3 开通语音合成服务.....	6
3.4 获取 Access Token.....	7
4 API 请求方式基本说明	8
4.1 文本上传模式.....	8
4.2 下行接口定义.....	9
4.3 错误码解释.....	10
5 注意事项.....	10

1 概念解释

对本文中将提到的名词约定如下：

语音合成：也被称为文本转换技术（TTS），它是将计算机自己产生的、或外部输入的文字信息转变为可以听懂的、流利的口语输出的技术。

应用程序：在开发中调用了 API、具有语音合成功能的产品线产品。

开发者：想在应用程序中使用 API、正在阅读本文档的开发人员。

2 简介

百度语音合成通过 REST API 的方式给开发者提供一个通用的 HTTP 接口，基于该接口，开发者可以轻松获取语音合成能力，本文档描述了使用语音合成服务 REST API 的方法。

2.1 功能介绍

REST API 支持对一段文本的合成，其中对于文本格式以及参数有一定的要求。请按照文档中描述进行相应请求操作。

本文档适用使用 HTTP 接口的开发人员。

3 集成指南

本章将进行 Step-By-Step 的讲解，如何快速的获取开发者权限以及使用语音合成 REST API。

3.1 注册百度开发者

使用百度语音技术服务首先需要注册成为百度开发者，如果您在登陆百度帐号后进入“应用管理”出现如下提示，说明您的百度帐号还没有注册成为百度开发者。请按要求

填写开发者信息并完成百度开发者的注册。

* 类型:

☒ 个人 ☐ 公司

* 开发者来源:

开发者

* 开发者姓名:

* 开发者简介:

* Email地址:

发送验证码

* 邮箱验证码:

* 手机号:

发送验证码

* 手机验证码:

提交

3.2 创建应用

在首页登陆已经注册为开发者的百度账号，然后点击页面上方的“应用管理”进入应用管理页面。如果您之前没有创建过应用，则需要新创建一个应用。

- 1、点击右上角

创建新应用

填写应用名称并选择应用垂类信息，即可创建一个新的应用。



创建新应用

应用名称:

应用类别: 游戏娱乐 ▼

创建新应用 取消

2、创建成功后，新创建的应用会出现在应用列表最上方的位置。

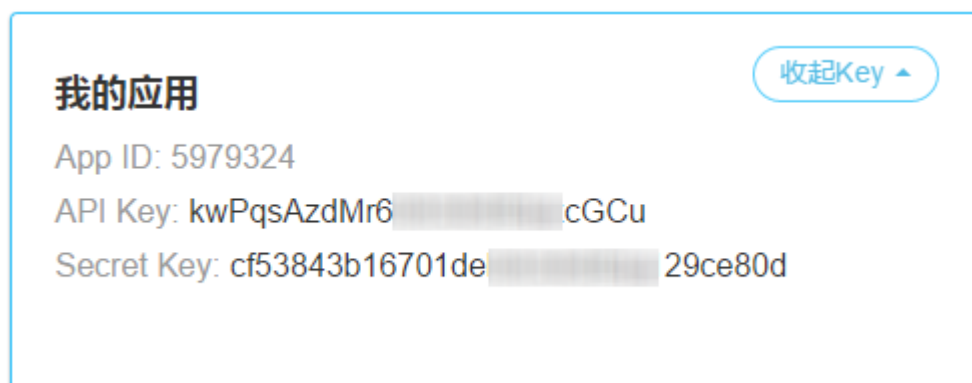


我 **我的应用** [查看Key ▼](#)

应用类别: 游戏娱乐
创建时间: 2015年05月20日

+ 开通服务

3、点击“查看 Key”可以查看当前应用的 AppID、API Key、Secret Key 信息。



我的应用 [收起Key ▲](#)

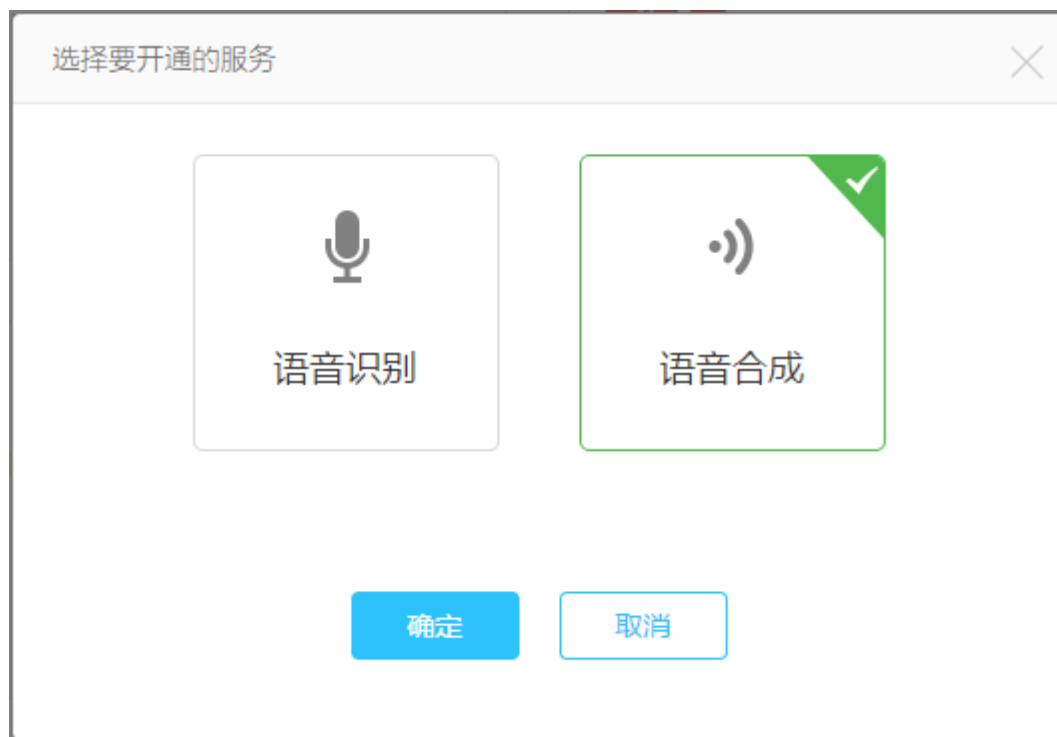
App ID: 5979324
API Key: kwPqsAzdMr6 [REDACTED] cGCu
Secret Key: cf53843b16701de [REDACTED] 29ce80d

3.3 开通语音合成服务

新创建或从未开通语音合成服务的应用，则需要开通语音合成服务之后才可以使用。



1、点击应用卡片上的按钮，选择“**语音合成**”进行开通。



2、“**语音合成**”服务开通成功后即可获得 **200000 次/日** 的在线调用次数配额。

使用语音合成 REST API 需要获取 Access Token。Access Token 是用户身份验证和授权的凭证。详细信息请参考百度开放云平台帮助文档的 OAuth 介绍部分。

语音合成采用的是 Client Credentials 授权方式，即采用应用公钥、密钥获取 Access Token，适用于任何带 server 类型应用，通过此授权方式获取 Access Token 仅可访问平台授权类的接口，详见 Client Credentials 授权说明。

3.4 获取 Access Token

根据百度 OAuth 的 Client Credentials 授权方式获取 Access Token。详见百度 OAuth 授权“[Client Credentials 授权](#)”部分。

使用 Client Credentials 获取 Access Token 需要应用在其服务端发送请求（推荐用 POST 方法）到百度 OAuth2.0 授权服务的“<https://openapi.baidu.com/oauth/2.0/token>”地址上，并带上以下参数：

- grant_type: 必须参数，固定为 “client_credentials”；
- client_id: 必须参数，应用的 API Key；
- client_secret: 必须参数，应用的 Secret Key；

例如：

```
1. https://openapi.baidu.com/oauth/2.0/token?
2. grant_type=client_credentials&
3. client_id=Va5yQRH1*****LT0vuXV4&
4. client_secret= 0rDSjzQ20XUj5i*****PQSzr5pVw2&
```

响应数据包如下所示，其中“access_token”字段即为请求 REST API 所需的令牌，默认情况下，Access Token 有效期为一个月，开发者需要对 Access Token 的有效性进行判断，如果 Access Token 过期可以重新获取。

例如：

```
1. HTTP/1.1 200 OK
2. Content-Type: application/json
3. Cache-Control: no-store
4.
5. {
6.     "access_token": "1.a6b7dbd428f731035f771b8d*****.
86400.1292922000-2346678-124328",
7.     "expires_in": 86400,
8.     "refresh_token": "2.385d55f8615fd9edb7c4b*****.
604800.1293440400-2346678-124328",
9.     "scope": "public",
10.    "session_key": "ANXxSNjwQDugf86150nqeik*****Cd1Lx
n",
11.    "session_secret": "248APxvxjCZ0VEC*****aK4oZExMB
",
12. }
```

4 API 请求方式基本说明

- 语音合成接口支持 POST 和 GET 两种方式
- 正式地址：<http://tsn.baidu.com/text2audio>

4.1 文本上传模式

➤ 上传参数

参数	必需	描述
tex	必填	合成的文本，使用 UTF-8 编码，请注意文本长度必须小于 1024 字节
lan	必填	语言选择,填写 zh

tok	必填	开放平台获取到的开发者 access_token
ctp	必填	客户端类型选择，web 端填写 1
cuid	必填	用户唯一标识，用来区分用户，web 端参考填写机器 mac 地址或 imei 码，长度为 60 以内
spd	选填	语速，取值 0-9，默认为 5
pit	选填	音调，取值 0-9，默认为 5
vol	选填	音量，取值 0-9，默认为 5
per	选填	发音人选择，取值 0-1；0 为女声，1 为男声，默认为女声

● GET 调用方式

将所有参数都填写到 URL 地址中，可以通过浏览器播放合成的语音结果。

```
1. http://tsn.baidu.com/text2audio?tex=***&lan=zh&cuid=***&ctp=1&tok=***
```

● POST 调用方式

将文本以及其他参数写入到 body 里面，利用表单的方式将参数传递到服务端。调用地址为 <http://tsn.baidu.com/text2audio> 所有的参数都在 body 中。body 里面的数据为：

```
1. tex=***&cuid=***&ctp=1&tok=***
```

注意：

不论是参数放到 URL 地址中，还是利用表单的方式放入到 http body 中，参数都需要经过两次 URLENCODE 操作。如果是直接在浏览器地址栏输入则不需要。

4.2 下行接口定义

如果合成成功，下行数据为二进制语音文件，具体 header 信息 Content-Type: audio/mp3；如果合成出现错误，则会返回 json 结果，具体 header 信息为：Content-Type: application/json。其中 sn 数据主要用于 DEBUG 追查问题，如果出现问题，可以提供 sn 帮助确认问题。

错误实例为：

```
{"err_no":500,"err_msg":"notsupport.", "sn":"abcdefgh", "idx":1}
```

4.3 错误码解释

错误码	含义
500	不支持输入
501	输入参数不正确
502	token 验证失败
503	合成后端错误

5 注意事项

请严格按照文档里描述的参数进行开发。请注意以下几个问题：

1. 合成文本长度必须小于 1024 字节，如果本文长度较长，可以采用多次请求的方式。切忌不可文本长度超过限制。
2. 语音合成 rest api 初次申请默认请求数配额 20 万次/天，，如果默认配额不能满足需求，请申请提高配额。
3. 必填字段中，严格按照文档描述中内容填写。