



# 云知声智能语音交互平台

## Web Service SDK

### 开发指南

V2.0.01

北京云知声信息技术有限公司

Beijing Unisound Information Technology Co., Ltd.

# 重要声明

## 版权声明

版权所有 © 2013，北京云知声信息技术有限公司，保留所有权利。

## 商标声明

北京云知声信息技术有限公司的产品是北京云知声信息技术有限公司专有。在提及其他公司及其产品时将使用各自公司所拥有的商标，这种使用的目的仅限于引用。本文档可能涉及北京云知声信息技术有限公司的专利（或正在申请的专利）、商标、版权或其他知识产权，除非得到北京云知声信息技术有限公司的明确书面许可协议，本文档不授予使用这些专利（或正在申请的专利）、商标、版权或其他知识产权的任何许可协议。

## 不作保证声明

北京云知声信息技术有限公司不对此文档中的任何内容作任何明示或暗示的陈述或保证，而且不对特定目的的适销性及适用性或者任何间接、特殊或连带的损失承担任何责任。本手册内容若有变动，恕不另行通知。本手册例子中所用的公司、人名和数据若非特别声明，均属虚构。未得到北京云知声信息技术有限公司明确的书面许可，不得为任何目的、以任何形式或手段（电子的或机械的）复制或传播手册的任何部分。

## 保密声明

本文档（包括任何附件）包含的信息是保密信息。接收人了解其获得的本文档是保密的，除用于规定的目的外不得用于任何目的，也不得将本文档泄露给任何第三方。

本软件产品受最终用户许可协议（EULA）中所述条款和条件的约束，该协议位于产品文档和/或软件产品的联机文档中，使用本产品，表明您已阅读并接受了 EULA 的条款。

版权所有©北京云知声信息技术有限公司

Copyrights © Beijing Unisound Information Technology Co., Ltd.

## 目录

1. 概述.....	1
1.1. 目的.....	1
1.2. 范围.....	1
2. 使用说明.....	1
2.1. 开发说明.....	1
2.2. 开发前准备.....	1
2.3. 支持的平台.....	2
3. 环境搭建.....	2
4. HTTP 接口.....	3
4.1. 识别请求和响应.....	3
4.2. HTTP 请求.....	4
4.3. HTTP 响应.....	5
FAQ.....	7

# 1. 概述

云知声智能语音交互平台旨在使第三方应用便利的集成和使用语音理解(语音云和语义云)服务。本 SDK 开发包针对 WebService 方式调用进行定制优化,使开发者可以迅速开发基于通用语音识别、语义理解及语音合成需求的客户端软件。

本文档默认读者已经掌握 WebService 开发的相关知识。

## 1.1. 目的

本文档对 WebService SDK 接口定义进行说明。

文档读者为使用云知声 WebService SDK 进行开发的产品设计师、软件工程师。

## 1.2. 范围

本文档定义语音识别、语音合成的使用说明、体系结构、API 接口。

不包含核心引擎的性能定义,也不包含其它配套或附赠产品的使用说明。

# 2. 使用说明

## 2.1. 开发说明

本文属于入门级文档,旨在帮助开发者快速学习云知声 SDK 的使用,并应用到实际的开发工作中,开发者仅需关注文档中所提供的接口方法而不用了解具体实现。

## 2.2. 开发前准备

对于个人开发者使用云知声语音服务,请到“<http://dev.hivoice.cn>”注册成为云知声开发者,并创建应用,在“我的应用”中获取 AppKey。使用应用授权码可以帮助开发者监控语音识别服务的使用情况。

## 2.3. 支持的平台

所有支持HTTP 1.1协议的平台。

# 3. 环境搭建

### Chunked Transfer Coding

HTTP 客户端必须支持 HTTP 1.1 (RFC2616) 和 chunked transfer coding (<http://www.w3.org/Protocols/rfc2616/rfc2616-sec3.html#sec3.6.1>).

当语音音频流的总长度不能提前知道的时候，使用 chunked encoding 将音频流作为一系列片段（chunk）来传输，其中每个片段都有一个描述自身长度的标记。

关于 chunked encoding 更多信息请查看以下链接

- [four.livejournal.com/887211.html](http://four.livejournal.com/887211.html)
- [developers.sun.com/mobility/midp/questions/chunking/](http://developers.sun.com/mobility/midp/questions/chunking/)
- [en.wikipedia.org/wiki/Chunked\\_transfer\\_encoding](http://en.wikipedia.org/wiki/Chunked_transfer_encoding)

## 4. HTTP 接口

### 4.1. 识别请求和响应

下面是一个使用 HTTP POST 一段语音到识别服务器并得到响应的示例：

```
POST http://api.hivoice.cn/USCSservice/WebApi?appkey=appkey&userid=dev_user_id
&id=id HTTP/1.1

Transfer-Encoding: chunked

Content-Type: audio/x-wav;codec=pcm;bit=16;rate=16000

Accept: text/plain

Accept-Language: zh_CN

600
...
45
....
0
```

```
HTTP/1.1 200 OK

Date: Fri, 15 Mar 2013 12:50:35 GMT

Content-Type: text/plain;charset=utf-8

Content-Language: zh_CN

Content-Length: 11

x-usc-sessionid: 8eb77583036f7aa197bd6505b7d6420a
```

你好，很高兴认识你！

## 4.2. HTTP 请求

### 请求字符串 (Request query string)

请求字符串是 URL 的一部分，包含了传给 HTTP 接口的数据。为了获得本服务，需要提供以下所有 3 个参数，用 <key>=<value> 的键值对形式，且之间用 “&” 符号分隔：

- **userid**: 开发者网站(dev.hivoice.cn)上注册的用户 ID;
- **appkey**: 开发者网站上申请应用后获得的 appkey;
- **id**: 标记请求来源的标识，如用户所用设备的序列号(SN)，IMEI，MAC 地址等

注意：**id** 请注意填写，**仅用于标识不同用户**，以便利用用户口音自适应 (adaptation) 技术、个性化用户数据方案来提高不同用户的识别性能。此处可填写：客户设备的 IMEI、MAC 地址、用户微信号等。

### 请求头 (Request headers)

**Content-Type** 上传的语音数据流的格式，目前支持：

audio/x-wav;codec=pcm;bit=16;rate=16000      16Khz 16bit PCM

**Accept** 识别结果返回的格式，目前支持：

text/plain      文本格式

**Accept-Language** 语音数据流记录的语种及识别结果的语种：

zh\_CN      中文普通话

**Accept-Topic** 指定 ASR 识别的引擎。支持：

general      通用

movietv      影视

song      音乐

poi      POI

medical      医药

**Transfer-Encoding**

如不知道语音流的总长度则必须使用 *Transfer-Encoding: chunked*

**Content-Length**

如果事先知道语音输入流的长度，可以通过 **Content-Length** 来设置。

当指定 **Content-Length** 时，禁止设置 **Transfer-Encoding** 属性。

### 请求消息体 (Request message body)

HTTP POST 消息承载了该次请求的语音输入流。只有当消息头定义了“Transfer-Encoding: chunked”时，message-body 和 entity-body 才不一致（请参见 [tools.ietf.org/html/rfc2616#section-3.6.1](https://tools.ietf.org/html/rfc2616#section-3.6.1)）。

发送语音输入流应该首选“Transfer-Encoding: chunked”方式，同时每个分段（chunk）最好在 260ms 左右（约 xxx bytes）；

如果没有使用“Transfer-Encoding: chunked”，则整个消息体包含完整语音输入流，并且其长度等于 Content-Length。

注意：单次语音识别最大时长为 60 秒。

## 4.3. HTTP 响应

### 响应头 (Response headers)

Content-Type	识别结果返回的格式	
	text/plain	文本格式
Content-Language	语音数据流记录的语种	
	zh_CN	

### 响应消息体 (Response message body)

语音识别的结果，格式为 Content-Type 中指定的格式。

### 会话 ID (x-usc-sessionid)

标示该次识别会话在服务器端的标记，可以用来查询和确认调用记录。

### 响应状态码 (Response HTTP status codes)

- 2xx Success



2xx 类别的状态码表明客户端的请求已经被接收、并成功处理。

- **200 OK**  
标准的返回状态，识别请求处理成功。
- **204 No Content**  
服务成功地响应识别请求，但没有返回任何内容。

#### ■ 4xx Client Error

4xx 类的状态码表明客户端可能存在错误。

- **400 Bad Request**  
由于语法错误，无法处理客户端的请求。
- **401 Unauthorized**  
appkey 验证不通过
- **403 Forbidden**  
客户端提交的请求是合法的，但服务器拒绝响应。
- **404 Not Found**  
找不到识别服务器，请稍后再试。
- **405 Method Not Allowed**  
使用不支持的 HTTP 请求方式。如客户端使用了 GET 方法而不是 POST。
- **406 Not Acceptable**  
服务器不支持 Http Headers 中的 Accept 属性。
- **408 Request Timeout**  
请求超时。
- **410 Gone**  
Request 请求的资源已经永久失效。
- **413 Request Entity Too Large**  
请求的内容太长，超出服务器的限制。
- **414 Request-URI Too Long**  
URI 太长，服务器不能处理。
- **415 Unsupported Media Type**  
不支持的音频格式。

#### ■ 5xx Server Error

5xx 类别的错误表明服务器不能处理合法的请求。

- **500 Internal Server Error**  
通用的错误信息。该错误不属于以下状态码之一。
- **501 Not Implemented**  
服务器不能识别请求的方法，或者该服务的功能不能处理该请求。
- **503 Service Unavailable**

服务不可用（有可能是过载、宕机、正在维护）。一般情况下，这只是一个临时的状态。

- **504 Gateway Timeout**  
这是一个代理服务器，没有及时得到上游服务器的返回。
- **505 HTTP Version Not Supported**  
不支持客户端请求所使用的 HTTP 协议版本。

## FAQ

### 1. 如何联系我们?

如果应用开发过程中遇到问题，可随时与我们联系，联系方式如下：

QQ 语音云交流群：253369015

联系电话：(+8610-) 62369899-664      传真：(+8610-) 82601009

请发邮件至：[support@yunzhisheng.cn](mailto:support@yunzhisheng.cn)