

语音交互服务

SDK 参考

文档版本 01
发布日期 2024-03-05



版权所有 © 华为云计算技术有限公司 2024。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为云计算技术有限公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为云计算技术有限公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为云计算技术有限公司

地址：贵州省贵安新区黔中大道交兴功路华为云数据中心 邮编：550029

网址：<https://www.huaweicloud.com/>

目录

1 SDK 简介	1
2 申请使用	2
3 AK/SK 认证	3
4 获取 SDK	4
5 准备环境	6
5.1 配置 Java 环境.....	6
5.2 配置 Python 环境.....	7
6 Java SDK	8
6.1 一句话识别 Http 接口.....	8
7 Python SDK	12
7.1 一句话识别 Http 接口.....	12
8 附录	16
8.1 示例音频.....	16
9 修订记录	17

1 SDK 简介

语音交互概述

语音交互服务（Speech Interaction Service，简称SIS）是一种人机交互方式，用户通过实时访问和调用API获取语音交互结果。

Java 接口与 API 对应关系

Java接口与API对应关系请参见[表 Java接口与API对应关系表](#)。

表 1-1 Java 接口与 API 对应关系表

Class	Method	API	功能名称
AsrCustomizationClient	AsrCustomShortResponse getAsrShortResponse(AsrCustomShortRequest request)	POST /v1/{project_id}/asr/short-audio	一句话识别

Python 接口与 API 对应关系

Python接口与API对应关系请参见[表 Python接口与API对应关系表](#)。

表 1-2 Python 接口与 API 对应关系表

Class	Method	API	功能名称
AsrCustomizationClient	get_short_response(request)	POST /v1/{project_id}/asr/short-audio	一句话识别

2 申请使用

语音交互服务的一句话识别当前已商用，无需申请公测。

QPS数量限制：3QPS。

3 AK/SK 认证

使用服务API需要进行认证，目前SDK仅支持AK/SK认证方式。

使用AK/SK方式，需要用户提供AK和SK。

1. 注册并登录华为云管理控制台。
2. 在控制台中，鼠标移动至右上角的用户名处，在下拉列表中单击“我的凭证”。
3. 单击“访问密钥”页签，在页签中，单击“新增访问密钥”。



4. 在“身份验证”对话框中，输入当前用户的登录密码，通过邮箱或者手机进行验证，输入对应的验证码。
5. 单击“确定”，下载认证帐号的AK/SK，请妥善保管AK/SK信息。

4 获取 SDK

下载 SDK 包

语音交互SDK软件包获取请参见[表 下载SDK包](#)。

示例音频参见[示例音频](#)。

表 4-1 下载 SDK 包

SDK语言	下载地址
Java	https://sis-open-data.obs.ap-southeast-3.myhuaweicloud.com/java/huaweicloud-java-sdk-sis-1.8.1.zip
Python	https://sis-open-data.obs.ap-southeast-3.myhuaweicloud.com/python/huaweicloud-python-sdk-sis-1.2.0.zip

Maven 引入 SDK

Java SDK也可以通过Maven引入，示例文件可参见[下载SDK包](#)，操作如下。

1. 在工程的pom.xml文件里添加依赖。

```
<dependency>
  <groupId>com.huawei.sis</groupId>
  <artifactId>huaweicloud-java-sdk-sis</artifactId>
  <version>1.8.1</version>
</dependency>
```

2. 在工程的pom.xml文件里添加仓库，这里以仓库ID为“sis-repo”为例。

仓库地址为：<https://mirrors.huaweicloud.com/repository/maven/huaweicloudsdk>。

```
<repositories>
  <repository>
    <id>sis-repo</id>
    <name>Sis Release Repository</name>
    <url>https://mirrors.huaweicloud.com/repository/maven/huaweicloudsdk</url>
  </repository>
</repositories>
```

```
<snapshots>
  <enabled>>false</enabled>
</snapshots>
</repository>
</repositories>
```

3. 在Maven的配置文件setting.xml中，添加镜像。

setting.xml文件一般存放在用户目录下，例如，Windows系统：“C:\Users\\${用户名}\.m2\setting.xml”，\${用户名}要替换成个人电脑的用户名。

```
<mirror>
  <id>huaweicloud</id>
  <name>Huaweicloud Repository</name>
  <url>https://mirrors.huaweicloud.com/repository/maven/</url>
  <mirrorOf>*,!sis-repo</mirrorOf>
</mirror>
```

说明

```
<mirrorOf>*,!sis-repo</mirrorOf>
```

配置镜像时，为使步骤2配置的仓库地址生效。需要在镜像设置中过滤掉这个仓库，从而镜像中的这个ID为“sis-repo”的仓库不会生效。

5 准备环境

5.1 配置 Java 环境

配置环境

在使用语音交互SDK时，需要准备的环境请参见[表 开发环境](#)。

📖 说明

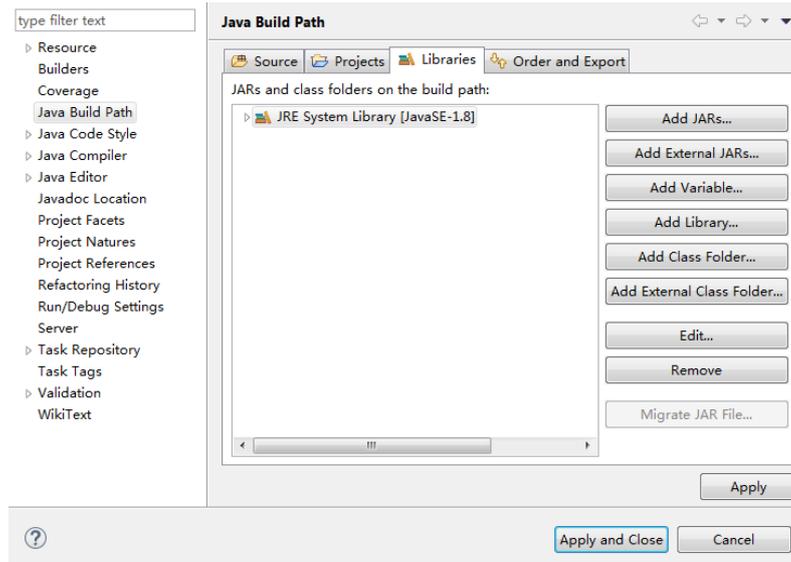
目前Java SDK不支持在android中使用。

表 5-1 开发环境

准备项	说明
操作系统	Windows系统，推荐Windows 7及以上版本。
JDK	Java开发环境的基本配置。版本要求：强烈推荐使用1.8版本。
Eclipse	在 Eclipse官网 下载对应平台的Eclipse版本，比如：eclipse-jee-mars-R-win32-x86_64.zip。

导入 SDK

1. Eclipse导入SDK。
 - a. 解压eclipse后，直接打开。同时[下载SDK](#)。
 - b. 选择“Window -> Preferences -> Java -> installed JREs”配置正确的JRE路径。
 - c. 新建工程，在工程下建立一个文件（New -> Folder），命名为lib。将下载的jar包拷贝至lib中。
 - d. 选中新建的工程，单击右键，下拉选择“Build Path -> Configure Build Path”，在“Java Build Path”对话框中，单击“Libraries”页签，选择“Add JARs”。在打开的窗口中，选择刚放进lib的jar包。单击“OK”，导入完成。



5.2 配置 Python 环境

前提条件

- 确保已安装Python3，目前Python SDK仅支持Python3。
- 确保已安装setuptools、requests。

操作步骤

1. **下载SDK**，通过**pip-list**命令查看安装包。若未安装，则执行以下命令：

```
pip install setuptools  
pip install requests
```
2. 进入下载的Python SDK目录，在setup.py所在层目录执行 `python setup.py install` 命令，完成SDK安装。

6 Java SDK

6.1 一句话识别 Http 接口

前提条件

- 确保已按照[配置Java环境](#)配置完毕。
- 确保已存在待识别的音频文件。如果需要请在下载的SDK压缩包中获取示例音频。

初始化 Client

初始化AsrCustomizationClient，其参数包括AuthInfo和SisConfig。

表 6-1 AuthInfo

参数名称	是否必选	参数类型	描述
ak	是	String	用户的ak，可参考 AK/SK认证 。
sk	是	String	用户的sk，可参考 AK/SK认证 。
region	是	String	区域，如ap-southeast-3，参考 终端节点 。
projectId	是	String	项目ID，同region一一对应，参考 获取项目ID 。
endpoint	否	String	终端节点，一般使用默认即可。

表 6-2 SisConfig

参数名称	是否必选	参数类型	描述
connecti onTimeo ut	否	Integer	连接超时，默认10000，单位ms。
readTim eout	否	Integer	读取超时，默认10000，单位ms。

请求参数

请求类为AsrCustomShortRequest，详见[表6-3](#)。

表 6-3 AsrCustomShortRequest

参数名称	是否必选	参数类型	描述
data	是	String	本地音频文件经过Base64编码后的字符串，音频文件时长不超过30s。
audioForm at	是	String	音频格式，具体信息请参见《API参考》中 一句话识别 章节。
property	是	String	属性字符串，语言_采样率_模型，如 english_8k_common。具体信息请参见《API参考》中 一句话识别 章节。
addPunc	否	String	表示是否在识别结果中添加标点，取值为yes、no，默认no。

响应参数

响应类为AsrCustomShortResponse，详见[表6-4](#)。

表 6-4 AsrCustomShortResponse

参数名	是否必选	参数类型	说明
trace_id	是	String	服务内部的令牌，可用于在日志中追溯具体流程，调用失败无此字段。在某些错误情况下可能没有此令牌字符串。
result	是	Object	调用成功表示识别结果，调用失败时无此字段。请参考 表6-5 。

表 6-5 Result

参数名	是否必选	参数类型	说明
text	是	String	调用成功表示识别出的内容。
score	是	Float	调用成功表示识别出的置信度，取值范围：0~1。

代码示例

如下示例仅供参考，最新代码请前往[获取SDK](#)章节获取并运行。

```
import com.huawei.sis.bean.SisConfig;
import com.huawei.sis.bean.SisConstant;
import com.huawei.sis.bean.request.AsrCustomShortRequest;
import com.huawei.sis.bean.response.AsrCustomShortResponse;
import com.huawei.sis.bean.AuthInfo;
import com.huawei.sis.client.AsrCustomizationClient;
import com.huawei.sis.exception.SisException;
import com.huawei.sis.util.IOUtils;
import java.util.List;
import com.huawei.sis.util.JsonUtils;

/**
 * 一句话识别
 *
 * Copyright 2021 Huawei Technologies Co.,Ltd.
 */
public class AsrCustomizationDemo {
    private static final int SLEEP_TIME = 500;
    private static final int MAX_POLLING_NUMS = 1000;

    // 认证用的ak和sk硬编码到代码中或者明文存储都有很大的安全风险，建议在配置文件或者环境变量中密文存放，使用时解密，确保安全；
    // 本示例以ak和sk保存在环境变量中来实现身份验证为例，运行本示例前请先在本地环境中设置环境变量HUAWEICLOUD_SDK_AK和HUAWEICLOUD_SDK_SK。
    private String ak = System.getenv("HUAWEICLOUD_SDK_AK");
    private String sk = System.getenv("HUAWEICLOUD_SDK_SK");

    private String region = ""; // 区域，如ap-southeast-3
    private String projectId = ""; // 项目id。登录管理控制台，鼠标移动到右上角的用户名上，在下拉列表中选择我的凭证，在项目列表中查看项目id。多项目时，展开“所属区域”，从“项目ID”列获取子项目ID。
    // 一句话识别参数
    private String path = ""; // 音频文件路径，如D:/test.wav等，sdk会将音频文件转化为base64编码
    private String pathAudioFormat = ""; // 文件格式，如wav等
    private String pathProperty = "chinese_16k_general"; // 属性字符串，language_sampleRate_domain, 16k模型推荐使用chinese_16k_general

    /**
     * 设置一句话识别参数，所有参数均有默认值，不配置也可使用
     *
     * @param request 一句话识别请求
     */
    private void setShortParameter(AsrCustomShortRequest request) {

        // 设置是否添加标点，默认是no
        request.setAddPunc("yes");
    }

    /**
     * 定义config，所有参数可选，设置超时时间等。
     */
}
```

```
*
* @return SisConfig
*/
private SisConfig getConfig() {
    SisConfig config = new SisConfig();
    // 设置连接超时, 默认10000ms
    config.setConnectionTimeout(SisConstant.DEFAULT_CONNECTION_TIMEOUT);
    // 设置读取超时, 默认10000ms
    config.setReadTimeout(SisConstant.DEFAULT_READ_TIMEOUT);
    // 设置代理, 一定要确保代理可用才启动此设置。代理初始化也可用不加密的代理, new
    ProxyHostInfo(host, port);
    // ProxyHostInfo proxy = new ProxyHostInfo(host, port, username, password);
    // config.setProxy(proxy);
    return config;
}

/**
 * 一句话识别demo
 */
private void shortDemo() {
    try {
        // 1. 初始化AsrCustomizationClient
        // 定义authInfo, 根据ak, sk, region, projectId
        AuthInfo authInfo = new AuthInfo(ak, sk, region, projectId);
        // 设置config, 主要与超时有关
        SisConfig config = getConfig();
        // 根据authInfo和config, 构造AsrCustomizationClient
        AsrCustomizationClient asr = new AsrCustomizationClient(authInfo, config);

        // 2. 配置请求
        String data = IOUtils.getEncodeDataByPath(path);
        AsrCustomShortRequest request = new AsrCustomShortRequest(data, pathAudioFormat, pathProperty);
        // 设置请求参数, 所有参数均为可选
        setShortParameter(request);

        // 3. 发送请求, 获取响应
        AsrCustomShortResponse response = asr.getAsrShortResponse(request);
        // 打印结果
        System.out.println(JsonUtils.obj2Str(response, true));
    } catch (SisException e) {
        e.printStackTrace();
        System.out.println("error_code:" + e.getErrorCode() + "\nerror_msg:" + e.getErrorMsg());
    }
}

public static void main(String[] args) {
    AsrCustomizationDemo demo = new AsrCustomizationDemo();
    demo.shortDemo();
}
}
```

7 Python SDK

7.1 一句话识别 Http 接口

前提条件

- 确保已按照[配置Python环境](#)配置完毕，Python SDK仅支持Python3。
- 确保已存在待识别的音频文件。如果需要请在下载的SDK压缩包中获取示例音频。

初始化 Client

初始化AsrCustomizationClient详见[表 AsrCustomizationClient初始化参数](#)。

表 7-1 AsrCustomizationClient 初始化参数

参数名称	是否必选	参数类型	描述
ak	是	String	用户的ak，可参考 AK/SK认证 。
sk	是	String	用户的sk，可参考 AK/SK认证 。
region	是	String	区域，如ap-southeast-3，参考 终端节点 。
project_id	是	String	项目ID，同region一一对应，参考 获取项目ID 。
service_endpoint	否	String	终端节点，一般使用默认即可。
sis_config	否	Object	详见 表7-2 。

表 7-2 SisConfig

参数名称	是否必选	参数类型	描述
connect_timeout	否	Integer	连接超时，默认10，单位s。
read_timeout	否	Integer	读取超时，默认10，单位s。
proxy	否	List	[host, port] 或 [host, port, username, password]。

请求参数

请求类为AsrCustomShortRequest，详见[表7-3](#)。

表 7-3 AsrCustomShortRequest

参数名称	是否必选	参数类型	描述
data	是	String	本地音频文件经过Base64编码后的字符串，音频文件时长不超过30s。
audio_format	是	String	音频格式，具体信息请参见《API参考》中 一句话识别 章节。
model_property	是	String	属性字符串，语言_采样率_模型，如 english_8k_common。具体信息请参见《API参考》中 一句话识别 章节。
add_punc	否	String	表示是否在识别结果中添加标点，取值为yes、no，默认no。

响应参数

Python SDK响应结果为Json格式，[表7-4](#)。

表 7-4 响应结果

参数名称	是否必选	参数类型	描述
result	是	Object	详见 表7-5 。
trace_id	是	String	用于后台日志问题追溯。

表 7-5 Result

参数名称	是否必选	参数类型	描述
text	是	String	识别结果。
score	是	Float	识别结果置信度评分。

代码示例

如下示例仅供参考，最新代码请前往[获取SDK](#)章节获取并运行。

```
# -*- coding: utf-8 -*-
from huaweicloud_sis.client.asr_client import AsrCustomizationClient
from huaweicloud_sis.bean.asr_request import AsrCustomShortRequest
from huaweicloud_sis.exception.exceptions import ClientException
from huaweicloud_sis.exception.exceptions import ServerException
from huaweicloud_sis.utils import io_utils
from huaweicloud_sis.bean.sis_config import SisConfig
import json
import os
# 鉴权参数
# 认证用的ak和sk硬编码到代码中或者明文存储都有很大的安全风险，建议在配置文件或者环境变量中密文存
# 本示例以ak和sk保存在环境变量中来实现身份验证为例，运行本示例前请先在本地环境中设置环境变量
HUAWEICLOUD_SIS_AK/HUAWEICLOUD_SIS_SK。
ak = os.getenv("HUAWEICLOUD_SIS_AK")
assert ak is not None, "Please add ak in your develop environment"
sk = os.getenv("HUAWEICLOUD_SIS_SK")
assert sk is not None, "Please add sk in your develop environment"
project_id = "" # project id 同region——对应
region = "" # region, 如ap-southeast-3
"""
    todo 请正确填写音频格式和模型属性字符串
    1. 音频格式一定要相匹配。
        例如wav音频，格式是wav。具体参考api文档。
        例如音频是pcm格式，并且采样率为8k，则格式填写pcm8k16bit。
        如果返回audio_format is invalid 说明该文件格式不支持。具体支持哪些音频格式，需要参考一些api文
    档。
    2. 音频采样率要与属性字符串的采样率要匹配。
        例如格式选择pcm16k16bit，属性字符串却选择chinese_8k_common，则会返回'audio_format' is not
    match model
        例如wav本身是16k采样率，属性选择chinese_8k_common，同样会返回'audio_format' is not match
    model
    """
# 一句话识别参数，以音频文件的base64编码传入，1min以内音频
path = "" # 文件位置，需要具体到文件，如D:/test.wav
path_audio_format = "" # 音频格式，如wav等，详见api文档
path_property = 'chinese_16k_general' # language_sampleRate_domain, 如chinese_16k_general，详见api文
档
def sasr_example():
    """ 一句话识别示例 """
    # step1 初始化客户端
    config = SisConfig()
    config.set_connect_timeout(10) # 设置连接超时
    config.set_read_timeout(10) # 设置读取超时
    # 设置代理，使用代理前一定要确保代理可用。代理格式可为[host, port] 或 [host, port, username,
password]
    # config.set_proxy(proxy)
    asr_client = AsrCustomizationClient(ak, sk, region, project_id, sis_config=config)
    # step2 构造请求
    data = io_utils.encode_file(path)
    asr_request = AsrCustomShortRequest(path_audio_format, path_property, data)
    # 所有参数均可不设置，使用默认值
```

```
# 设置是否添加标点, yes or no, 默认no
asr_request.set_add_punc('yes')
# 设置是否将语音中数字转写为阿拉伯数字, yes or no, 默认yes
asr_request.set_digit_norm('yes')
# 设置是否添加热词表id, 没有则不填
# asr_request.set_vocabulary_id(None)
# 设置是否需要word_info, yes or no, 默认no
asr_request.set_need_word_info('no')
# step3 发送请求, 返回结果,返回结果为json格式
result = asr_client.get_short_response(asr_request)
print(json.dumps(result, indent=2, ensure_ascii=False))
if __name__ == '__main__':
    try:
        sasr_example()
    except ClientException as e:
        print(e)
    except ServerException as e:
        print(e)
```

8 附录

8.1 示例音频

测试音频如[表 示例音频](#)所示，音频文件标题表示采样率。

表 8-1 示例音频

音频格式	下载链接
8k pcm	https://sis-open-data.obs.ap-southeast-3.myhuaweicloud.com/audio/8k.pcm
8k wav	https://sis-open-data.obs.ap-southeast-3.myhuaweicloud.com/audio/8k.wav
16k pcm	https://sis-open-data.obs.ap-southeast-3.myhuaweicloud.com/audio/16k.pcm
16k wav	https://sis-open-data.obs.ap-southeast-3.myhuaweicloud.com/audio/16k.wav

9 修订记录

发布日期	修订说明
2020-04-21	第一次正式发布。