Speech Interaction Service

SDK Reference

Issue 01

Date 2024-03-05

Huawei Cloud Computing Technologies Co., Ltd.

Address: Huawei Cloud Data Center Jiaoxinggong Road
Qianzhong Avenue
Gui'an New District
Gui Zhou 550029
People's Republic of China

Website: https://www.huaweicloud.com/intl/en-us/

Contents

1 SDK Introduction

SIS Overview

Speech Interaction Service (SIS) allows you to obtain speech interaction results by calling application programming interfaces (APIs) in real time.

Mappings Between Java Interfaces and APIs

Table Mappings between Java interfaces and APIs lists the mappings between the Java interfaces and SIS APIs.

Table 1-1 Mappings between Java interfaces and APIs

Class	Method	API	Function
AsrCustomizationClient	AsrCustomShortResponse getAsrShortResponse(AsrCustomShortRequest request)	POST /v1/{project_id}/asr/short-audio	Sentence Transcription

Mappings Between Python Interfaces and APIs

Table Mappings between Python interfaces and APIs lists the mappings between the Python interfaces and SIS APIs.

Table 1-2 Mappings between Python interfaces and APIs

Class	Method	API	Function
AsrCustomizationClient	get_short_response(request)	POST /v1/{project_id}/asr/short-audio	Sentence Transcription

2 Applying for SIS

The **Short Sentence Recognition** service is now commercially available. You do not need to apply for OBT.

Maximum QPS: 3

3 AK/SK-Based Authentication

Authentication is required to use service APIs. Currently, authentication using AK/SK is supported for SDKs.

If the AK/SK mode is used, you need to obtain the AK and SK as follows.

1. Register an account and log in to the HUAWEI CLOUD management console.
2. Hover the cursor on the username and select **My Credentials** from the drop-down list box.
3. Click **Access Keys**. On the corresponding tab, click **Create Access Key**.
4. Enter the login password of the current user and the verification code received through email or SMS.
5. Click **OK** to download the AK/SK. Keep the AK/SK properly.

4 Obtaining SDKs

Downloading SDK Software Packages

Table **Downloading SDK software packages** describes how to download the SIS SDK software packages.

To obtain audio examples, go to **Audio Examples**.

Table 4-1 Downloading SDK software packages

SDK Language	Download Link
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**

Importing SDKs Using Maven

The Java SDK can also be imported through Maven. For details about the sample file, see **Downloading SDK software packages**. The procedure is as follows:

1. Add the following dependencies to the **pom.xml** file:
   ```
   <dependency>
       <groupId>com.huawei.sis</groupId>
       <artifactId>huaweicloud-java-sdk-sis</artifactId>
       <version>1.8.1</version>
   </dependency>
   ```

2. Add a repository to the **pom.xml** file of the project, for example, **sis-repo**.

   The repository address is 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>
           <releases>
   ```

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

3. Add an image to the configuration file **setting.xml** on Maven.

   Generally, the **setting.xml** file is stored in the user directory. In the Windows OS, the directory is **C:\Users\${Username}\.m2\setting.xml**. Replace **${Username}** with the username of the PC.

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

   ☐ **NOTE**

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

   When configuring the image, to make the repository address configured in step **2** take effect, filter out the repository in the image settings so that the **sis-repo** repository in the image does not take effect.

5 Preparing Environment

5.1 Configuring the Java Environment

1. Configure the environment.

    Table Development environment describes the environment to be prepared for using SIS SDKs.

    📖 **NOTE**

    > Currently, Java SDKs cannot be used in the Android OS.

    Table 5-1 Development environment

Item	Description
Operating system (OS)	Windows OS. Windows 7 or later is recommended.
JDK	Basic configuration for the Java development environment. The 1.8 version is strongly recommended.
Eclipse	Download Eclipse, for example, **eclipse-jee-mars-R-win32-x86_64.zip**, from the **Eclipse official website**.

2. Import the SDK to Eclipse.

    a. Decompress Eclipse and open it. Then, **download the SDK**.

    b. Configure the JRE path correctly in **Window** > **Preferences** > **Java** > **installed JREs**.

    c. Create a project and choose **New > Folder** to create a folder named **lib** under the project. Copy the downloaded JAR file to the **lib** folder.

    d. Right-click the project and choose **Build Path > Configure Build Path > Libraries > Add JARs**. In the window that is displayed, select the JAR file that you just put in the **lib** folder. Click **OK**.
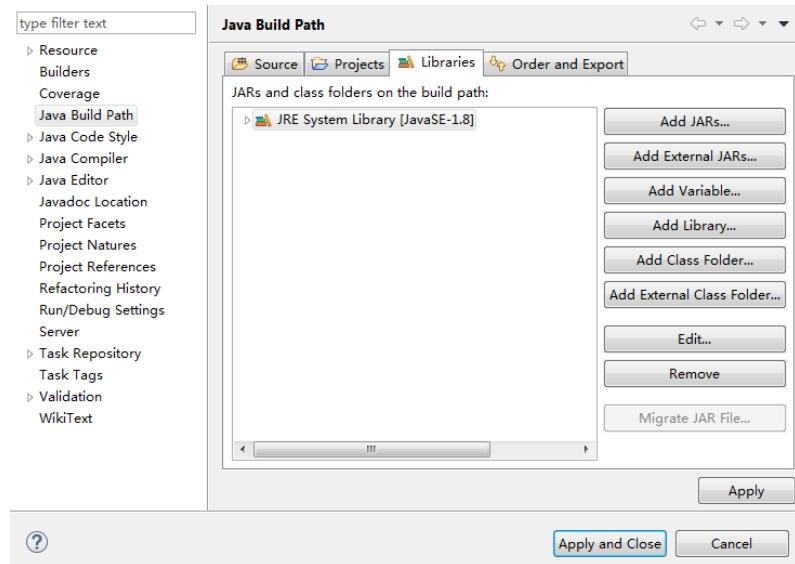
5.2 Configuring the Python Environment

Prerequisites

- Ensure that Python 3 has been installed. Currently, Python SDKs support only Python 3.

- Ensure that **setuptools** and **websocket-client** have been installed.

Procedure

1. **Download the SDK** and run the **pip-list** command to view the installation package. If **setuptools**, **requests**, and **websocket-client** have not been installed, run the following commands:
   ```
   pip install setuptools
   pip install requests
   ```

2. Go to the directory of the downloaded Python SDK and run the **python setup.py install** command in the directory where **setup.py** is located to install the SDK.

6 Java SDKs

6.1 HTTP API for Short Sentence Recognition

Prerequisites

- You have completed the configurations as instructed in **Configuring the Java Environment**.
- Ensure that the audio file to be recognized exists. If needed, you can obtain the sample audio file from the downloaded SDK package.

Initializing Client

Initialize AsrCustomizationClient. Its parameters include **AuthInfo** and **SisConfig**.

Table 6-1 AuthInfo

Name	Mandatory	Type	Description
ak	Yes	String	AK of the user. For details, see **AK/SK-Based Authentication**.
sk	Yes	String	SK of the user. For details, see **AK/SK-Based Authentication**.
region	Yes	String	Region, for example, **ap-southeast-3**. For details, see **Endpoints**.
projectId	Yes	String	Project ID corresponding to the region. For details about how to obtain a project ID, see **Obtaining a Project ID**.
endpoint	No	String	Endpoint. Generally, the default value is used.

Table 6-2 SisConfig

Name	Mandatory	Type	Description
connectionTimeout	No	Integer	Connection timeout interval (ms). The default value is **10000**ms.
readTimeout	No	Integer	Read timeout interval (ms). The default value is **10000**ms.

Request Parameters

The request class is **AsrCustomShortRequest**. For details, see **Table 6-3**.

Table 6-3 AsrCustomShortRequest

Name	Mandatory	Type	Description
data	Yes	String	Base64-encoded character string of the local audio file. The audio duration must be within 30 seconds.
audioFormat	Yes	String	Audio format. For details, see **Sentence Transcription** in *SIS API Reference*.
property	Yes	String	Property string in "language_sampling rate_model" format, for example, **english_8k_common**. For details, see **Sentence Transcription** in *SIS API Reference*.
addPunc	No	String	Whether to add punctuation marks to the recognition result. Possible values are **yes** and **no**. The default value is **no**.

Response Parameters

The response class is **AsrCustomShortResponse**. For details, see **Table 6-4**.

Table 6-4 AsrCustomShortResponse

Name	Mandatory	Type	Description
trace_id	Yes	String	Service internal token used to trace a specific process in logs. This parameter is not included when the API fails to be called. In some error cases, this field may not exist.
result	Yes	Object	Recognition result when the API is successfully called. In the case of an API calling failure, this parameter is not included. For details, see **Table 6-5**.

Table 6-5 Result

Name	Mandatory	Type	Description
text	Yes	String	Recognition result of a successful call.
score	Yes	Float	Confidence of a successful call. The value ranges from 0 to 1.

Sample Code

The following example is for reference only. Obtain and run the latest code in section **Obtaining SDKs**.

```
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;


/**
 * Short Sentence Recognition
 *
 * Copyright 2021 Huawei Technologies Co.,Ltd.
 */
public class AsrCustomizationDemo {
  private static final int SLEEP_TIME = 500;
  private static final int MAX_POLLING_NUMS = 1000;

  // Hard-coded or plaintext AK and SK are risky. For security purposes, encrypt your AK and SK and store
them in the configuration file or environment variables.
```

```
// In this example, the AK and SK are stored in environment variables for identity authentication. Before
running this example, configure environment variables HUAWEICLOUD_SDK_AK and
HUAWEICLOUD_SDK_SK.
 private String ak = System.getenv("HUAWEICLOUD_SDK_AK");
 private String sk = System.getenv("HUAWEICLOUD_SDK_SK");


 private String region = "";    // Region, such as ap-southeast-3
 private String projectId = ""; // Project ID corresponding to the region. Log in to the management console,
move the cursor over your username in the upper right corner, and choose My Credentials from the drop-
down list. On the displayed page, view the project ID. If there are multiple projects, expand Region and
obtain the sub-project ID from the Project ID column.
*
 // Sentence Transcription parameters
 private String path = "";             // Audio file path, for example, D:/test.wav. The SDK converts the audio
file into a Base64-encoded file.
 private String pathAudioFormat = "";  // File format, for example, WAV.
 private String pathProperty = "chinese_16k_general";     // Attribute string, language_sampleRate_domain.
chinese_16k_general is recommended for 16 kHZ models.


 /**
  * Configure Sentence Transcription parameters. All parameters have default values. You can use them
without configuring their values.
  *
  * @param request Sentence Transcription request
  */
 private void setShortParameter(AsrCustomShortRequest request) {

   // Set whether to add punctuation marks. The default value is no.
   request.setAddPunc("yes");
 }

 /**
  * Define config. All parameters are optional. Set the timeout intervals.
  *
  * @return SisConfig
  */
 private SisConfig getConfig() {
   SisConfig config = new SisConfig();
   // Set the connection timeout interval. The default value is 10000 ms.
   config.setConnectionTimeout(SisConstant.DEFAULT_CONNECTION_TIMEOUT);
  // Set the read timeout interval. The default value is 10000 ms.
   config.setReadTimeout(SisConstant.DEFAULT_READ_TIMEOUT);
   // Set the proxy. Ensure that the proxy is available before setting it. The unencrypted proxy can also be
used for proxy initialization, for example, new ProxyHostInfo(host, port).
   // ProxyHostInfo proxy = new ProxyHostInfo(host, port, username, password);
   // config.setProxy(proxy);
   return config;
 }


 /**
  * Sentence Transcription demo
  */
 private void shortDemo() {
   try {
     // 1. Initialize the AsrCustomizationClient.
     // Define authInfo based on the ak, sk, region, and projectId parameters.
     AuthInfo authInfo = new AuthInfo(ak, sk, region, projectId);
     // Set config, which is related to timeout settings.
     SisConfig config = getConfig();
     // Construct the AsrCustomizationClient based on authInfo and config.
     AsrCustomizationClient asr = new AsrCustomizationClient(authInfo, config);

     // 2. Configure the request.
     String data = IOUtils.getEncodeDataByPath(path);
     AsrCustomShortRequest request = new AsrCustomShortRequest(data, pathAudioFormat, pathProperty);
     // Set request parameters. All parameters are optional.
```

```java
        setShortParameter(request);

        // 3. Send the request and obtain a response.
        AsrCustomShortResponse response = asr.getAsrShortResponse(request);
        // Print the result.
        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 SDKs

7.1 HTTP API for Short Sentence Recognition

Prerequisites

- You have completed the configurations as instructed in **Configuring the Python Environment**. Python SDKs support only Python 3.
- Ensure that the audio file to be recognized exists. If needed, you can obtain the sample audio file from the downloaded SDK package.

Initializing Client

For details about AsrCustomizationClient initialization, see **Table 7-1**.

Table 7-1 AsrCustomizationClient initialization parameters

Name	Mandatory	Type	Description
ak	Yes	String	AK of the user. For details, see **AK/SK-Based Authentication**.
sk	Yes	String	SK of the user. For details, see **AK/SK-Based Authentication**.
region	Yes	String	Region, for example, **ap-southeast-3**. For details, see **Endpoints**.
project_id	Yes	String	Project ID corresponding to the region. For details about how to obtain a project ID, see **Obtaining a Project ID**.
service_endpoint	No	String	Endpoint. You can use the default value.
sis_config	No	Object	For details, see **Table 7-2**.

Table 7-2 SisConfig

Name	Mandatory	Type	Description
connect_timeout	No	Integer	Connection timeout interval (s). The default value is **10**s.
read_timeout	No	Integer	Read timeout interval (s). The default value is **10**s.
proxy	No	List	[host, port] or [host, port, username, password].

Request Parameters

The request class is **AsrCustomShortRequest**. For details, see **Table 7-3**.

Table 7-3 AsrCustomShortRequest

Name	Mandatory	Type	Description
data	Yes	String	Base64-encoded character string of the local audio file. The audio duration must be within 30 seconds.
audio_format	Yes	String	Audio format. For details, see **Sentence Transcription** in *SIS API Reference*.
model_property	Yes	String	Property string in "language_sampling rate_model" format, for example, **english_8k_common**. For details, see **Short Sentence Recognition** in *SIS API Reference*.
add_punc	No	String	Whether to add punctuation marks to the recognition result. Possible values are **yes** and **no**. The default value is **no**.

Response Parameters

The responses of Python SDKs are in JSON format. For details, see **Table 7-4**.

Table 7-4 Response result

Name	Mandatory	Type	Description
result	Yes	Object	For details, see **Table 7-5**.
trace_id	Yes	String	ID used for fault tracing in background logs.

Table 7-5 Result

Name	Mandatory	Type	Description
text	Yes	String	Recognition result
score	Yes	Float	Recognition result confidence

Sample Code

The following example is for reference only. Obtain and run the latest code in section **Obtaining SDKs**.

```
# -*- 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
# Authentication parameters
# Hard-coded or plaintext AK and SK are risky. For security purposes, encrypt your AK and SK and store
them in the configuration file or environment variables.
# In this example, the AK and SK are stored in environment variables. Before running this example, set
environment variables HUAWEICLOUD_SIS_AK or 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 = ""     # Correspond to the region.
region = ''         # region, for example, ap-southeast-3
"""
   todo Enter the correct audio format and model property character string.
   1. The audio formats must match.
       For example, for a WAV audio file, set the format to WAV. For details, see the API document.
       If the audio format is PCM and the sampling rate is 8 kHz, enter pcm8k16bit.
       If "audio_format is invalid" is returned, the file format is not supported. For details about the
supported audio formats, see the API Reference.
   2. The audio sampling rate must match the sampling rate of the attribute character string.
       If the format is pcm16k16bit but the attribute character string is set to chinese_8k_common,
"'audio_format' is not match model" is returned.
       If the sampling rate of the WAV file is 16 kHz but the attribute character string is set to
chinese_8k_common, "'audio_format' is not match model" is returned.
"""
# Sentence Transcription parameters, which are passed through the Base64 code of the audio file. The
audio duration must be within 1 minute.
path = ''                     # File location, which must be specific to a file, for example, D:/test.wav.
path_audio_format = ''        # Audio format, for example, WAV. For details, see the API Reference.
path_property = 'chinese_16k_general'   # language_sampleRate_domain, for example, chinese_16k_general.
For details, see the API Reference.
def sasr_example():
    """ Sentence Transcription demo """
    # Step 1 Initialize the client.
    config = SisConfig()
    config.set_connect_timeout(10)  # Set the connection timeout interval.
    config.set_read_timeout(10)  # Set the read timeout interval.
    # Set the proxy. Ensure that the proxy is available before using it. The proxy format can be [host, port] or
[host, port, username, password].
    # config.set_proxy(proxy)
    asr_client = AsrCustomizationClient(ak, sk, region, project_id, sis_config=config)
    # Step 2 Construct a request.
    data = io_utils.encode_file(path)
```

```
    asr_request = AsrCustomShortRequest(path_audio_format, path_property, data)
    # You can use the default values for all parameters.
    # Set whether to add punctuation marks. Possible values are yes and no. The default value is no.
    asr_request.set_add_punc('yes')
    # Set whether to convert numbers in the speech into Arabic numerals. Possible values are yes and no.
The default value is yes.
    asr_request.set_digit_norm('yes')
    # Set whether to add the hot word table ID. If no hot word table ID is added, leave this parameter blank.
    # asr_request.set_vocabulary_id(None)
    # Set whether word_info is required. The value can be yes or no. The default value is no.
    asr_request.set_need_word_info('no')
    # Step 3 Send the request and the result is returned in JSON format.
    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 Appendix

8.1 Audio Examples

Table Audio examples lists the test audio recordings. The title of each audio recording file contains the sampling rate.

Table 8-1 Audio examples

Audio Format	Download Link
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 Change History

Release Date	Description
2020-04-21	This issue is the first official release.