Speech Interaction Service

Overview

Issue   01
Date   2024-04-03

Huawei Technologies Co., Ltd.

Address:	Huawei Industrial Base Bantian, Longgang Shenzhen 518129 People's Republic of China
Website:	https://www.huawei.com
Email:	support@huawei.com

Security Declaration

Vulnerability

Huawei's regulations on product vulnerability management are subject to the *Vul. Response Process.* For details about this process, visit the following web page:
https://www.huawei.com/en/psirt/vul-response-process
For vulnerability information, enterprise customers can visit the following web page:
https://securitybulletin.huawei.com/enterprise/en/security-advisory

Contents

1 What Is SIS?

Speech Interaction Service (SIS) allows you to obtain speech interaction results by calling application programming interfaces (APIs) in real time. For example, you can use Automatic Speech Recognition (ASR) to convert speech recordings to editable text, and use Text To Speech (TTS) to convert text into lifelike voices. SIS is applicable to scenarios such as voice customer service inspection, conference records, voice SMS messages, audio books, and telephone follow-ups.

Prerequisites

You must have programming capabilities and be familiar with the Java, Python, and iOS programming languages.

SIS provides APIs for you to convert speech into editable text and returns the recognition result in JSON format. You need to encode the recognition result and save it to a service system or save it in TXT or Excel format.

Using SIS for the First Time

If you are a first-time user, the following information will help you get familiar with SIS:

- Functions

  Functions describes SIS functions, including Real-time ASR, Short Sentence Recognition, TTS.

- Getting Started

  SIS provides services through open APIs. You can learn how to use SIS by referring to the **Speech Interaction Service Getting Started**.

- Using SIS

  If you are a development engineer familiar with code compilation and want to directly call SIS APIs, see the **Speech Interaction Service API Reference** or **Speech Interaction Service SDK Reference**.

- From Beginners to Experts

  You can learn how to use SIS by referring to **Progressive Knowledge**.

2 Functions

Short Sentence Recognition

Short Sentence Recognition converts audio recordings within 30s to text. Specifically, Short Sentence Recognition converts binary audio data to corresponding text. English is supported.

Highlights

- High Recognition Rate

  Utilizes the deep learning technology to optimize speech recognition for domain-specific scenarios, enabling an industry-leading recognition rate.

- Cutting-Edge Technologies

  Combines mature speech recognition algorithms currently in active use in the industry with the latest research to empower enterprises with unique competitive advantages.

- Customizable Models

  Increases accuracy by using speech recognition models designed for the specific requirements of the vertical industry you operate in for other specific scenarios.

3 Application Scenarios

- Voice Customer Service Inspection

  Recognizes the speech of the customer service personnel and customer, converts the speech into text, and checks whether it contains any violation, sensitive word, or phone number through text retrieval.

- Voice Message

  Converts voice messages you send or receive into text to deliver higher reading efficiency and interaction experience.

- Gaming and Entertainment

  Converts voice chats into text messages, improving reading efficiency and user experience.

4 Billing

Billing Item

- Sentence Transcription is billed based on the number of API calls. Failed API calls are not billed.

Billing Mode

- **Pay-per-use**

  SIS adopts tiered pricing based on the number of API calls. The sentence transcription function of SIS is not put into commercial use. For details about how to subscribe to the service and obtain pricing details, **Contact Sales**.

5 Related Services

Figure 5-1 shows the relationships between SIS and other services.

Figure 5-1 Relationships between SIS and other services
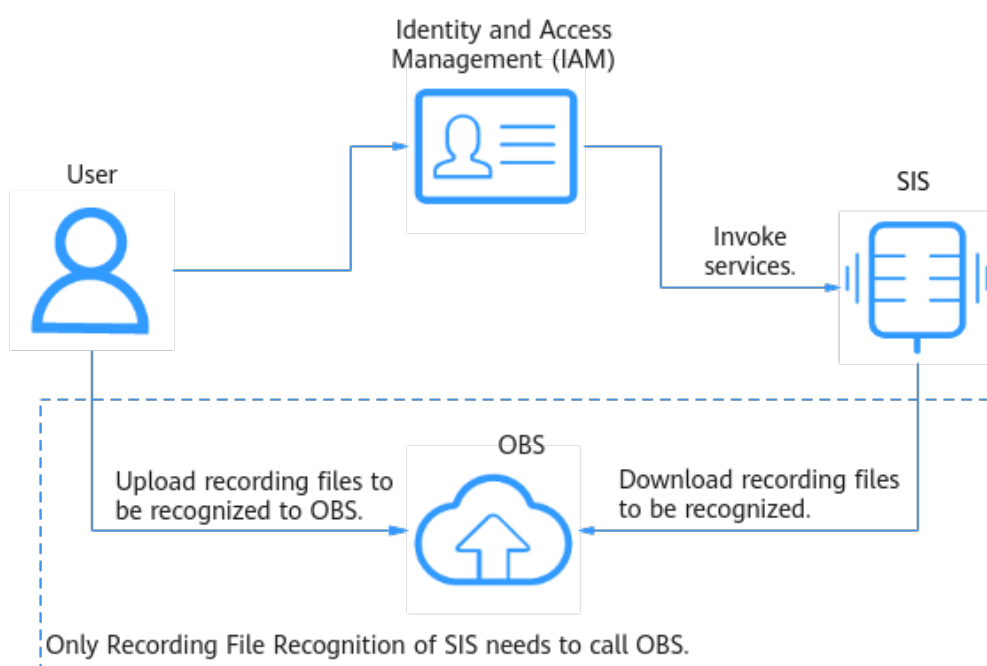
Identity and Access Management (IAM)

IAM provides user authentication and authorization for SIS.

6 Restrictions and Limitations

6.1 Short Sentence Recognition

- Short Sentence Recognition is available in the AP-Singapore3 region.
- Formats pcm8k16bit, pcm16k16bit, and wav are supported.
- The audio duration cannot exceed 30s.
- Only recognition of speech in English is supported.