ALIBABA CLOUD

Alibaba Cloud

Intelligent Speech Interaction
Short Sentence Recognition

Document Version: 20220414

Alibaba Cloud

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.

2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company or individual in any form or by any means without the prior written consent of Alibaba Cloud.

3. The content of this document may be changed because of product version upgrade, adjustment, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and an updated version of this document will be released through Alibaba Cloud-authorized channels from time to time. You should pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.

4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides this document based on the "status quo", "being defective", and "existing functions" of its products and services. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not take legal responsibility for any errors or lost profits incurred by any organization, company, or individual arising from download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, take responsibility for any indirect, consequential, punitive, contingent, special, or punitive damages, including lost profits arising from the use or trust in this document (even if Alibaba Cloud has been notified of the possibility of such a loss).

5. By law, all the contents in Alibaba Cloud documents, including but not limited to pictures, architecture design, page layout, and text description, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of this document shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates.

6. Please directly contact Alibaba Cloud for any errors of this document.

Document conventions

Style	Description	Example
Danger	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	**Danger:** Resetting will result in the loss of user configuration data.
Warning	A warning notice indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	**Warning:** Restarting will cause business interruption. About 10 minutes are required to restart an instance.
Notice	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	**Notice:** If the weight is set to 0, the server no longer receives new requests.
Note	A note indicates supplemental instructions, best practices, tips, and other content.	**Note:** You can use Ctrl + A to select all files.
>	Closing angle brackets are used to indicate a multi-level menu cascade.	Click **Settings**> **Network**> **Set network type**.
Bold	Bold formatting is used for buttons , menus, page names, and other UI elements.	Click **OK**.
`Courier font`	Courier font is used for commands	Run the `cd /d C:/window` command to enter the Windows system folder.
Italic	Italic formatting is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	This format is used for an optional value, where only one item can be selected.	`ipconfig [-all|-t]`
{} or {a|b}	This format is used for a required value, where only one item can be selected.	`switch {active|stand}`

Table of Contents

1.NUI SDK for mobile clients

1.1. Overview

The short sentence recognition service provides the Natural User Interaction (NUI) SDK for mobile clients to recognize speeches that last within 60 seconds in real time. The SDK applies to scenarios such as chat conversation, voice command control, voice search in applications, and speech input.

Description

Compared with common SDKs, the NUI SDK is smaller in size and supports more comprehensive status management. The NUI SDK provides comprehensive speech processing capabilities and can also serve as an atomic SDK, meeting diverse user requirements. In addition, the NUI SDK uses a unified API.

Featrures

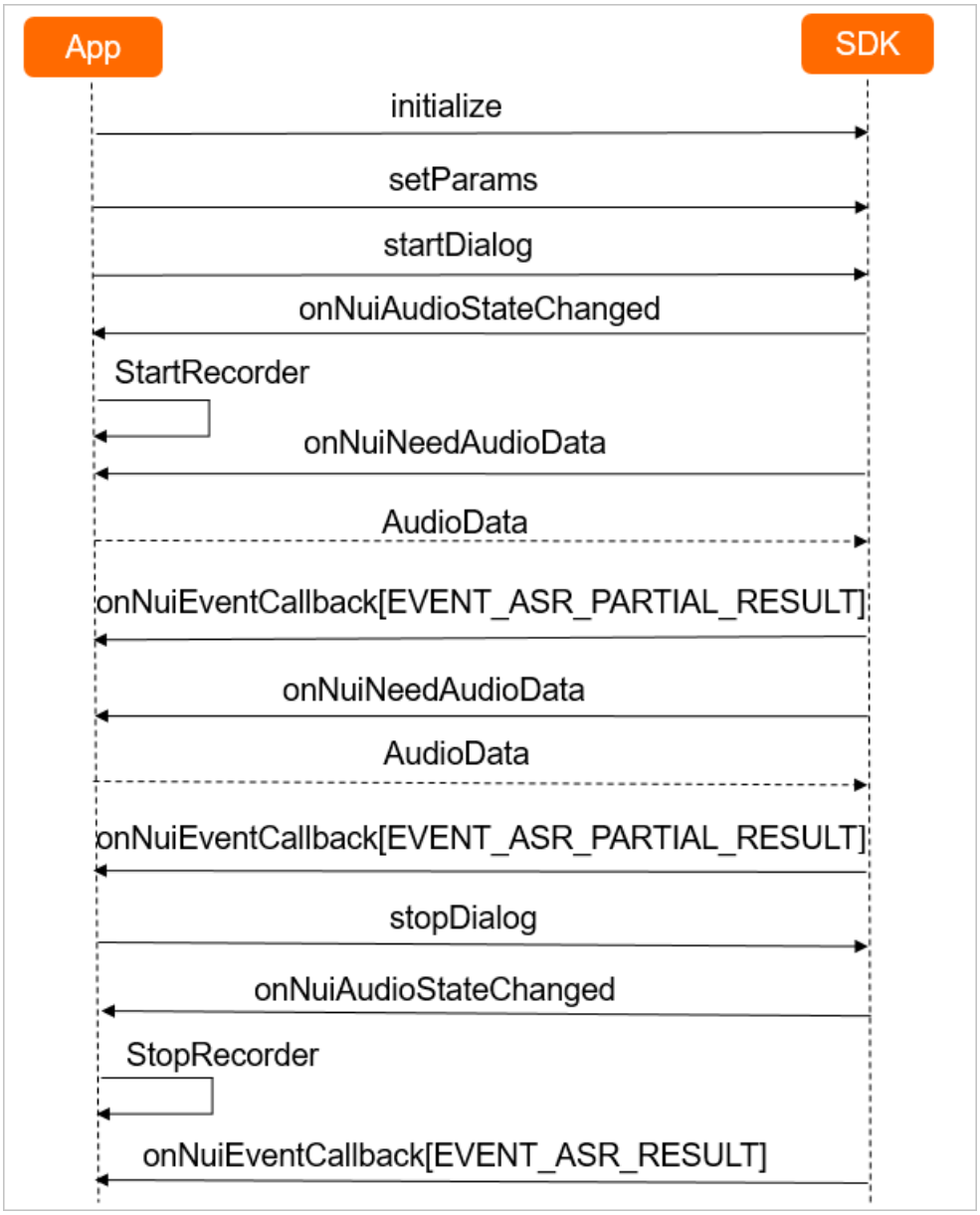
- Supports pulse-code modulation (PCM) encoded 16-bit mono audio files.
- Supports the audio sampling rates of 8,000 Hz and 16,000 Hz.
- Supports recognizing speeches that last within 60 seconds.
- Allows you to specify whether to return intermediate results, whether to add punctuation marks during post-processing, and whether to convert Chinese numerals to Arabic numerals.
- Allows you to select linguistic models to recognize speeches in different languages when you manage projects in the Intelligent Speech Interaction console. For more information, see Manage projects.

Endpoints

Access type	Description	URL
External access from the Internet	This endpoint allows you to access the short sentence recognition service from any host over the Internet. By default, the Internet access URL is built in the SDK.	wss://nls-gateway.ap-southeast-1.aliyuncs.com/ws/v1

Interaction process

The following figure shows the interaction process between the SDK and an Android or iOS client.



> **Note**
>
> The server adds the task_id parameter to the response header for all responses to indicate the ID of the recognition task. You can record the value of this parameter. If an error occurs, you can submit a ticket to report the task ID and error message.

1. Authenticate the client and initialize the SDK

To establish a WebSocket connection with the server, the client must use a token for authentication. For more information about how to obtain the token, see Obtain a token.

The following table describes the parameters used for authentication and initialization.

Parameter	Type	Required	Description
workspace	String	Yes	The working directory from which the SDK reads the configuration file.
app_key	String	Yes	The appkey of your project created in the Intelligent Speech Interaction console.
token	String	Yes	The token provided as the credential for you to use Intelligent Speech Interaction. Make sure that the token is valid. You can set the token when you initialize the SDK and update the token when you set the request parameters.
device_id	String	Yes	The unique identifier of the device, for example, the media access control (MAC) address, serial number, or pseudo unique ID of the device.
debug_path	String	No	The directory where audio files generated during the debugging are stored. If the save_log parameter is set to true when you initialize the SDK, intermediate results are stored in this directory.

Parameter	Type	Required	Description
save_wav	String	No	This parameter is valid if the save_log parameter is set to true when you initialize the SDK. This parameter specifies whether to store audio files generated during the debugging in the directory specified by the debug_path parameter. Make sure that the directory is writable.

2. Send a request to use the short sentence recognition service

You must set the request parameters for the client to send a service request. You can set the request parameters in the JSON format by using the setParams method in the SDK. The parameter configuration applies to all service requests. The following table describes the request parameters.

Parameter	Type	Required	Description
appkey	String	No	The appkey of your project created in the Intelligent Speech Interaction console. This parameter is generally set when you initialize the SDK.
token	String	No	The token provided as the credential for you to use Intelligent Speech Interaction. You can update the token as required by setting this parameter.

Parameter	Type	Required	Description
service_type	Int	Yes	The type of speech service to be requested. Set this parameter to 0, which indicates the short sentence recognition service.
direct_ip	String	No	The IP address that is resolved from the Domain Name System (DNS) domain name. The client completes the resolution and uses the obtained IP address to access the service.
nls_config	JsonObject	No	The service parameters.

The following table describes the parameters in the nls_config parameter.

Parameter	Type	Required	Description
sr_format	String	No	The audio encoding format. The short sentence recognition service supports the Opus and PCM formats. Default value: OPUS. Note: This parameter must be set to PCM if the sample_rate parameter is set to 8000.
sample_rate	Integer	No	The audio sampling rate. Unit: Hz. Default value: 16000. After you set this parameter, you must specify a model or scene that is applicable to the audio sampling rate for your project in the Intelligent Speech Interaction console.

Parameter	Type	Required	Description
enable_intermediate_result	Boolean	No	Specifies whether to return intermediate results. Default value: False.
enable_punctuation_prediction	Boolean	No	Specifies whether to add punctuation marks during post-processing. Default value: False.
enable_inverse_text_normalization	Boolean	No	Specifies whether to enable inverse text normalization (ITN) during post-processing. Valid values: true and false. Default value: false. If you set this parameter to true, Chinese numerals are converted to Arabic numerals. Note: ITN is not implemented on words.
customization_id	String	No	The ID of the custom speech training model.
vocabulary_id	String	No	The vocabulary ID of custom extensive hotwords.
enable_voice_detection	Boolean	No	Specifies whether to enable voice detection. Default value: False.

Parameter	Type	Required	Description
max_start_silence	Integer	No	The maximum duration of start silence. Unit: milliseconds. If the actual duration of start silence exceeds the value of this parameter, the server sends a TaskFailed event to end the recognition task. This parameter takes effect only when the enable_voice_detection parameter is set to true.

Parameter	Type	Required	Description
max_end_silence	Integer	No	The maximum duration of end silence. Unit: milliseconds. Valid values: 200 to 2000. If the actual duration of end silence exceeds the value of this parameter, the server sends a RecognitionCompleted message to complete the recognition task. Then, the subsequent speech is no longer processed. This parameter takes effect only when the enable_voice_detection parameter is set to true.

3. Send audio data from the client

The client cyclically sends audio data to the server and continuously receives recognition results from the server.

- If the enable_intermediate_result parameter is set to true, the SDK reports multiple EVENT_ASR_PARTIAL_RESULT events by calling the onNuiEventCallback method to return intermediate results of a sentence. For example, the server returns the following response:

```
{
    "header": {
        "namespace": "SpeechRecognizer",
        "name": "RecognitionResultChanged",
        "status": 20000000,
        "message_id": "e06d2b5d50ca40d5a50d4215c7c8****",
        "task_id": "4c3502c7a5ce4ac3bdc488749ce4****",
        "status_text": "Gateway:SUCCESS:Success."
    },
    "payload": {
        "result": "Weather in Beijing"
    }
}
```

The following table describes the parameters in the header object.

Parameter	Type	Description
namespace	String	The namespace of the message.
name	String	The name of the message. The RecognitionResultChanged message indicates that an intermediate result is obtained.
status	Integer	The HTTP status code. It indicates whether the request is successful. For more information, see the "Error codes" section of this topic.
message_id	String	The ID of the message, which is automatically generated by the SDK.
task_id	String	The GUID of the task. Record the value of this parameter to facilitate troubleshooting.
status_text	String	The status message.

The following table describes the parameters in the payload object.

Parameter	Type	Description
result	String	The intermediate result of the recognition task.

> **Note**
>
> The latest intermediate result may be different from the final result. Use the result returned in the EVENT_ASR_RESULT event as the final result.

- If the enable_intermediate_result parameter is set to false, the server does not return any messages in this step.

4. Complete the recognition task

The client sends a request for stopping short sentence recognition to the server. The server returns the final recognition result. For example, the server returns the following response:

```
{
    "header": {
        "namespace": "SpeechRecognizer",
        "name": "RecognitionCompleted",
        "status": 20000000,
        "message_id": "10490c992aef44eaa4246614838f****",
        "task_id": "4c3502c7a5ce4ac3bdc488749ce4****",
        "status_text": "Gateway:SUCCESS:Success."
    },
    "payload": {
        "result": "Weather in Beijing. "
    }
}
```

The following table describes the parameters in the header object.

Parameter	Type	Description
namespace	String	The namespace of the message.
name	String	The name of the message. The RecognitionCompleted message indicates that the recognition task is completed.

Parameter	Type	Description
status	Integer	The HTTP status code. It indicates whether the request is successful. For more information, see the "Error codes" section of this topic.
message_id	String	The ID of the message, which is automatically generated by the SDK.
task_id	String	The GUID of the task. Record the value of this parameter to facilitate troubleshooting.
status_text	String	The status message.

The following table describes the parameters in the payload object.

Parameter	Type	Description
result	String	The final recognition result.

Error codes

For more information about the error codes that the short sentence recognition service may return, see Error codes.

1.2. Error codes

This topic describes the error codes and error messages that may be returned during the use of Intelligent Speech Interaction to facilitate troubleshooting.

Description

An error code may be returned by the SDK or the server.

- Error codes returned by the SDK:

  These error codes may be generated during the use of the Intelligent Speech Interaction SDK. The SDK returns an error code in a callback.

- Error codes returned by the server:

  If a DEFAULT_NLS_ERROR or HTTP_SERVER_ERROR event is returned, an error occurred during the service use. The status parameter in the header object of the event displays the error code.

> **Note**
>
> The error codes and error messages described in this topic apply to the short sentence recognition and real-time speech recognition services.

Successful requests

Error code	Error message	Description
0	SUCCESS	The task is successful.

Configuration or Parameter errors

Error code	Error message	Description
240999	DEFAULT_ERROR	The error message returned because a default internal error occurred.
240001	NUI_CONFIG_INVALID	The error message returned because the configuration file is invalid.
240002	ILLEGAL_PARAM	The error message returned because a specified parameter is invalid.
240003	ILLEGAL_INIT_PARAM	The error message returned because an initialization parameter is invalid.
240004	NECESSARY_PARAM_LACK	The error message returned because a required parameter is missing.
240005	NULL_PARAM_ERROR	The error message returned because a parameter is left empty.
240006	NULL_LISTENER_ERROR	The error message returned because the listener callback is not specified.
240007	NULL_DIALOG_ERROR	The error message returned because no valid dialog instance is specified. This is generally an internal status error.

Error code	Error message	Description
240008	NULL_ENGINE_ERROR	The error message returned because no valid engine instance is specified. Check whether the engine initialization is successful.
240009	ILLEGAL_DATA	The error message returned because the URL of the audio data is invalid or the size of the audio data exceeds the upper limit.

Status errors related to the SDK

Error code	Error message	Description
240010	ILLEGAL_REENTRANT	The error message returned because you used the SDK after you exited the client.
240011	SDK_NOT_INIT	The error message returned because the SDK is not properly initialized.
240012	SDK_ALREADY_INIT	The error message returned because the SDK is repeatedly initialized.
240013	DIALOG_INVALID_STATE	The error message returned because the internal dialog instance is in an abnormal state.
240014	STATE_INVALID	The error message returned because the SDK is in an abnormal state.
240015	ILLEGAL_FUNC_CALL	The error message returned because the SDK is not used in a valid mode.

System errors

Error codes	Error message	Description
240020	MEM_ALLOC_ERROR	The error message returned because memory resources failed to be allocated.

Error codes	Error message	Description
240021	FILE_ACCESS_FAIL	The error message returned because the file failed to be accessed.
240022	CREATE_DIR_ERROR	The error message returned because the storage directory failed to be created.

Internal SDK call errors

Error code	Error message	Description
240030	CREATE_NUI_ERROR	The error message returned because the engine failed to be created.
240031	TEXT_DIALOG_START_FAIL	The error message returned because the text comprehension task failed to be started.
240032	TEXT_CANCEL_START_FAIL	The error message returned because the text comprehension task failed to be canceled.
240033	WUW_DUPLICATE	The error message returned because you specified repeated dynamic wake-up words.

Client engine errors

Error code	Error message	Description
240040	CEI_INIT_FAIL	The error message returned because the client engine failed to be initialized.
240041	CEI_SET_PARAM_FAIL	The error message returned because an engine parameter failed to be set.
240042	CEI_COMPILE_GRAMMER_FAIL	The error message returned because the code syntax failed to be compiled.

Error code	Error message	Description
240043	CEI_STOP_FAIL	The error message returned because the client engine failed to stop the recognition task.
240044	CEI_CANCEL_FAIL	The error message returned because the client engine failed to cancel the recognition task.
240045	CEI_UNLOAD_KWS_FAIL	The error message returned because the client engine failed to cancel the specified wake-up words.
240046	GET_WUW_ERROR	The error message returned because the client engine failed to obtain the specified wake-up words.

Errors related to audio data

Error code	Error message	Description
240050	SELECT_RECORDER_ERROR	The error message returned because the recording device is not properly selected.
240051	UPDATE_AUDIO_ERROR	The error message returned because the audio data failed to be pushed to the server. The general cause is that the size of the audio data exceeds the upper limit.
240052	MIC_ERROR	The error message returned because the microphone has not captured any audio data for 2 consecutive seconds.

Errors related to request timeout

Error code	Error message	Description

Error code	Error message	Description
240080	ENGINE_INIT_TIMEOUT	The error message returned because the request to initialize the engine timed out.
240081	SET_PARAM_TIMEOUT	The error message returned because the request to set parameters timed out.
240082	SET_WUW_TIMEOUT	The error message returned because the request to set wake-up words timed out.
240083	SELECT_RECORDER_TIMEOUT	The error message returned because the request to select the recording device timed out.
240084	STOP_TIMEOUT	The error message returned because the request to terminate the dialog timed out.
240085	ASR_ENGINE_STOP_TIMEOUT	The error message returned because the request to disable the engine timed out.
240086	UNLOAD_DYNAMIC_WUW_TIMEOUT	The error message returned because the request to cancel the dynamic wake-up words timed out.
240087	ADD_DYNAMIC_WUW_TIMEOUT	The error message returned because the request to add the dynamic wake-up words timed out.
240100	WAIT_TIMEOUT	The error message returned because the engine request timed out.
240101	HANDLE_API_TIMEOUT	The error message returned because the API request timed out.

Network errors

Error code	Error message	Description
240060	CREATE_DA_REQUEST_ERROR	The error message returned because the dialog assistant failed to be created.

Error code	Error message	Description
240061	START_DA_REQUEST_ERROR	The error message returned because the dialog assistant failed to be started.
240062	DEFAULT_NLS_ERROR	The error message returned because an error occurred on the server. Note: This error also generates an error code that is returned by the server. For more information, see the Error codes returned by the server table.
240063	SSL_ERROR	The error message returned because the Secure Sockets Layer (SSL) certificate failed to be created.
240064	SSL_CONNECT_FAILED	The error message returned because the SSL connection failed.
240065	HTTP_CONNECT_FAILED	The error message returned because the HTTP connection failed.
240066	DNS_FAILED	The error message returned because the Domain Name System (DNS) resolution failed.
240067	CONNECT_FAILED	The error message returned because the socket connection failed.
240068	SERVER_NOT_ACCESS	The error message returned because the server cannot be accessed.
240069	SOCKET_CLOSED	The error message returned because the socket is closed.
240070	AUTH_FAILED	The error message returned because the authentication failed.
240071	HTTPDNS_FAILED	The error message returned because the connection between the server and the client by using the specified IP address failed.

Errors related to network timeout

Error code	Error message	Description
240090	UPDATE_CONTEXT_TIMEOUT	The error message returned because the request to update the client timed out.
240091	CONNECTION_TIMEOUT	The error message returned because the network connection timed out.
240092	PARTIAL_ASR_TIMEOUT	The error message returned because the request to obtain the intermediate recognition result timed out.
240093	ASR_TIMEOUT	The error message returned because the request to obtain the final recognition result timed out.
240094	DIALOG_TIMEOUT	The error message returned because the request to obtain the dialog processing result timed out.
240095	WWV_TIMEOUT	The error message returned because the request to obtain the wake-up result of the server timed out.

Error codes returned by the server

If the client receives an EVENT_ASR_ERROR event, and the error code and error message returned by the SDK are 240062 and DEFAULT_NLS_ERROR respectively, the status parameter in the header object of the event displays the error code that is returned by the server.

Error code	Cause	Solution
40000001	The authentication failed.	Check whether the token expires or is invalid.
40000002	The message is invalid.	Check whether the message that is sent meets the requirement.
403	The token expires or is invalid.	1. Check whether the token expires.2. Check whether the token is valid.
40000004	The idle connection timed out.	Check whether no data has been sent to the server for 10 consecutive seconds.

Error code	Cause	Solution
40000005	The number of requests exceeds the upper limit.	Check whether the number of concurrent connections or the queries per second (QPS) exceeds the upper limit. If the number of concurrent connections exceeds the upper limit, we recommend that you upgrade Intelligent Speech Interaction from the trial edition to Commercial Edition. If you have upgraded the service to Commercial Edition, we recommend that you purchase more resources for higher concurrency.
40000000	A client error occurred. This is the default client error code.	Check the error message or submit a ticket.
41010120	The client timed out.	The client has not sent audio data for 10 or more consecutive seconds.
50000000	A server error occurred. This is the default server error code.	If the error code is occasionally returned, ignore it. If the error code is returned multiple times, submit a ticket.
50000001	An internal call error occurred.	If the error code is occasionally returned, ignore it. If the error code is returned multiple times, submit a ticket.
52010001	An internal call error occurred.	If the error code is occasionally returned, ignore it. If the error code is returned multiple times, submit a ticket.
40010001	The method is not supported.	If you use the SDK, submit a ticket.
40010002	The instruction is not supported.	If you use the SDK, submit a ticket.
40010003	The instruction is invalid.	If you use the SDK, submit a ticket.
40010004	The client is disconnected.	Check whether the client is disconnected before the server completes the requested task.

Error code	Cause	Solution
40010005	The task is in an abnormal state.	Check whether the instruction is supported in the current task status.
40020105	The specified appkey is invalid.	Resolve the route to check whether the application exists.
40020106	The specified appkey and token do not match.	Check whether the appkey of the application is valid and belongs to the same Alibaba Cloud account as the token.
40020503	Resource Access Management (RAM) user authentication fails.	Use your Alibaba Cloud account to authorize the RAM user to access the POP API.
41040201	The client has not sent data for 10 consecutive seconds.	Check the network connection or whether no business data needs to be sent.
41040202	The client sends data at a high transmission rate and consumes all resources of the server.	Check whether the client sends data at an appropriate transmission rate, for example, at the real-time factor of 1:1.
41040203	The client sends audio data in an invalid audio coding format.	Convert the audio coding format of audio data to a format supported by the SDK.
41040204	The client calls methods in an invalid order.	Check whether the client calls the relevant method to send a request before it calls other methods.
41040205	The specified MAXSILENCE_PARAM parameter is invalid.	Check whether the value of the MAXSILENCE_PARAM parameter is in the range of 200 to 2000.
41050008	The specified sampling rate does not match that of the selected model.	Check whether the audio sampling rate specified for the service call matches the audio sampling rate of the automatic speech recognition (ASR) model that is bound to the application in the console.
51040101	An internal error occurred on the server.	Resolve the error based on the error message.
51040102	Reserved.	N/A

Error code	Cause	Solution
51040103	The real-time speech recognition service is unavailable.	Check whether the number of real-time speech recognition tasks exceeds the upper limit.
51040104	The request for real-time speech recognition timed out.	Check the logs of the real-time speech recognition service.
51040105	The real-time speech recognition service failed to be called.	Check whether the real-time speech recognition service is enabled and whether the port works properly.
51040106	The load balancing of the real-time speech recognition service failed and the client failed to obtain the IP address of the real-time speech recognition service.	Check whether the real-time speech recognition server in the configured virtual private cloud (VPC) works properly.

2.SDK for Java

The short sentence recognition service provides an SDK for Java. This topic describes how to download and install the SDK. This topic also provides sample code for you to use the SDK.

Precautions

- Before you use the SDK, make sure that you understand how the SDK works. For more information, see Overview.
- The nls-sdk-short-asr SDK is renamed as nls-sdk-recognizer since version 2.1.0. If you use the nls-sdk-short-asr SDK and need to upgrade the SDK, you must delete it and add callbacks as prompted.

Download and installation

Download the latest version of the SDK from the Maven repository and the nls-sdk-java-demo package.

Add the following dependency:

```
<dependency>
    <groupId>com.alibaba.nls</groupId>
    <artifactId>nls-sdk-recognizer</artifactId>
    <version>2.1.6</version>
</dependency>
```

Decompress the .zip demo package. Run the `mvn package` command from the pom directory. An executable JAR package named nls-example-recognizer-2.0.0-jar-with-dependencies.jar is generated in the target directory. Copy the JAR package to the destination server. You can use the JAR package for quick service validation and stress testing.

Service validation:

Run the following command and set parameters as prompted.

Then, the logs/nls.log file is generated in the directory where the command is run.

```
java -cp nls-example-recognizer-2.0.0-jar-with-dependencies.jar com.alibaba.nls.client.SpeechRecognizerDemo
```

Stress testing:

Run the following command and set parameters as prompted.

Set the service URL to `wss://nls-gateway.ap-southeast-1.aliyuncs.com/ws/v1`. Provide .pcm audio files with a sampling rate of 16,000 Hz. Set the maximum number of concurrent calls based on your purchased resources.

```
java -jar nls-example-recognizer-2.0.0-jar-with-dependencies.jar
```

> **Note**
>
> You are charged if you make more than two concurrent calls to perform stress testing.

Key objects

- NlsClient: the speech processing client. You can use this client to process short sentence recognition, real-time speech recognition, and speech synthesis tasks. This object is thread-safe. You can globally create one NlsClient object.
- SpeechRecognizer: the short sentence recognition object. You can use this object to set request parameters, send a request, and send audio data. This object is not thread-safe.
- SpeechRecognizerListener: the recognition result listener, which listens to recognition results. This object is not thread-safe.

For more information, see Java API overview.

> **Notice**
>
> Notes on SDK calls:
>
> - Based on Netty, the creation of an NlsClient object consumes time and resources, but the created NlsClient object can be reused. We recommend that you create and disable an NlsClient object based on the lifecycle of your project.
> - The SpeechRecognizer object cannot be reused. You must create a SpeechRecognizer object for each recognition task. For example, to process N audio files, you must create N SpeechRecognizer objects to complete N recognition tasks.
> - One SpeechRecognizerListener object corresponds to one SpeechRecognizer object. You cannot use one SpeechRecognizerListener object for multiple SpeechRecognizer objects. Otherwise, you may fail to distinguish recognition tasks.
> - The SDK for Java depends on Netty. If your application is dependent on Netty, make sure that the version of Netty is 4.1.17.Final or later.

Sample code

> **Note**
>
> - Download the nls-sample-16k.wav file.
>
> The demo uses an audio file with a sampling rate of 16,000 Hz. To obtain an accurate recognition result, set the model to **universal model** for the project to which the appkey is bound in the Intelligent Speech Interaction console. In actual use, you can select a model based on the audio sampling rate. For more information about model setting, see Manage projects.

```
import java.io.File;
import java.io.FileInputStream;
import java.io.IOException;
import com.alibaba.nls.client.protocol.InputFormatEnum;
import com.alibaba.nls.client.protocol.NlsClient;
import com.alibaba.nls.client.protocol.SampleRateEnum;
import com.alibaba.nls.client.protocol.asr.SpeechRecognizer;
import com.alibaba.nls.client.protocol.asr.SpeechRecognizerListener;
import com.alibaba.nls.client.protocol.asr.SpeechRecognizerResponse;
import org.slf4j.Logger;
import org.slf4j.LoggerFactory;
```

```
/**
 * The sample code demonstrates how to perform the following operations:
 *      Call the API of the short sentence recognition service.
 *      Dynamically obtain a token.
 *      Use local files to simulate the sending of real-time streams.
 *      Calculate time consumed for recognition.
 */
public class SpeechRecognizerDemo {
    private static final Logger logger = LoggerFactory.getLogger(SpeechRecognizerDemo.class
);
    private String appKey;
    NlsClient client;
    public SpeechRecognizerDemo(String appKey, String id, String secret, String url) {
        this.appKey = appKey;
        // Globally create an NlsClient object. The default endpoint is the Internet access
URL of the short sentence recognition service.
        // Obtain a token. You must obtain another token before the current token expires.
You can call the accessToken.getExpireTime() method to query the expiration time of a token
.
        AccessToken accessToken = new AccessToken(id, secret);
        try {
            accessToken.apply();
            System.out.println("get token: " + accessToken.getToken() + ", expire time: " +
accessToken.getExpireTime());
            if(url.isEmpty()) {
                client = new NlsClient(accessToken.getToken());
            }else {
                client = new NlsClient(url, accessToken.getToken());
            }
        } catch (IOException e) {
            e.printStackTrace();
        }
    }
    private static SpeechRecognizerListener getRecognizerListener(int myOrder, String userP
aram) {
        SpeechRecognizerListener listener = new SpeechRecognizerListener() {
            // Return intermediate results. This message is returned only if the setEnableI
ntermediateResult parameter is set to true.
            @Override
            public void onRecognitionResultChanged(SpeechRecognizerResponse response) {
                // getName means to obtain the name of the event. getStatus means to obtain
the status code. getRecognizedText means to obtain the recognized text.
                System.out.println("name: " + response.getName() + ", status: " + response.
getStatus() + ", result: " + response.getRecognizedText());
            }
            // Recognition is completed.
            @Override
            public void onRecognitionCompleted(SpeechRecognizerResponse response) {
                // getName means to obtain the name of the event. getStatus means to obtain
the status code. getRecognizedText means to obtain the recognized text.
                System.out.println("name: " + response.getName() + ", status: " + response.
getStatus() + ", result: " + response.getRecognizedText());
            }
            @Override
```

```
            public void onStarted(SpeechRecognizerResponse response) {
                System.out.println("myOrder: " + myOrder + "; myParam: " + userParam + "; t
ask_id: " + response.getTaskId());
            }
            @Override
            public void onFail(SpeechRecognizerResponse response) {
                // task_id is the unique identifier that indicates the interaction between
the caller and the server. If an error occurs, you can submit a ticket and provide the task
ID to Alibaba Cloud to facilitate troubleshooting.
                System.out.println("task_id: " + response.getTaskId() + ", status: " + resp
onse.getStatus() + ", status_text: " + response.getStatusText());
            }
        };
        return listener;
    }
    // Calculate the equivalent voice length based on the binary data size.
    // Set the sampling rate to 8,000 Hz or 16,000 Hz.
    public static int getSleepDelta(int dataSize, int sampleRate) {
        // Set the sampling size to 16-bit.
        int sampleBytes = 16;
        // Use a single sound channel.
        int soundChannel = 1;
        return (dataSize * 10 * 8000) / (160 * sampleRate);
    }
    public void process(String filepath, int sampleRate) {
        SpeechRecognizer recognizer = null;
        try {
            // Pass the user-defined parameters.
            String myParam = "user-param";
            int myOrder = 1234;
            SpeechRecognizerListener listener = getRecognizerListener(myOrder, myParam);
            recognizer = new SpeechRecognizer(client, listener);
            recognizer.setAppKey(appKey);
            // Set the audio encoding format. For the .opus file, name the file as InputFor
matEnum.OPUS.
            recognizer.setFormat(InputFormatEnum.PCM);
            // Set the audio sampling rate.
            if(sampleRate == 16000) {
                recognizer.setSampleRate(SampleRateEnum.SAMPLE_RATE_16K);
            } else if(sampleRate == 8000) {
                recognizer.setSampleRate(SampleRateEnum.SAMPLE_RATE_8K);
            }
            // Specify whether to return intermediate results.
            recognizer.setEnableIntermediateResult(true);
            // Serialize preceding parameter settings to the JSON format. Then, send the JS
ON file to the server for confirmation.
            long now = System.currentTimeMillis();
            recognizer.start();
            logger.info("ASR start latency : " + (System.currentTimeMillis() - now) + " ms"
);
            File file = new File(filepath);
            FileInputStream fis = new FileInputStream(file);
            byte[] b = new byte[3200];
            int len;
```

```
            while ((len = fis.read(b)) > 0) {
                logger.info("send data pack length: " + len);
                recognizer.send(b, len);
                // In this example, local files are read to simulate real-time speech data
streams. You must set the sleep duration because files are fast read.
                // To recognize real-time speech, you do not need to set the sleep duration
. If the audio sampling rate is 8,000 Hz, you must set the second parameter to 8000.
                int deltaSleep = getSleepDelta(len, sampleRate);
                Thread.sleep(deltaSleep);
            }
            // Notify the server that audio data has been sent. Wait until the server compl
etes processing.
            now = System.currentTimeMillis();
            // Calculate the latency. The time when a response is returned after the stop m
ethod is called is regarded as the time when the final recognition result is returned.
            logger.info("ASR wait for complete");
            recognizer.stop();
            logger.info("ASR stop latency : " + (System.currentTimeMillis() - now) + " ms")
;
            fis.close();
        } catch (Exception e) {
            System.err.println(e.getMessage());
        } finally {
            // Close the connection.
            if (null != recognizer) {
                recognizer.close();
            }
        }
    }
    public void shutdown() {
        client.shutdown();
    }
    public static void main(String[] args) throws Exception {
        String appKey = null; // Enter the appkey.
        String id = null; // Enter the AccessKey ID.
        String secret = null; // Enter the AccessKey secret.
        String url = ""; // Default value: wss://nls-gateway.ap-southeast-1.aliyuncs.com/ws
/v1
        if (args.length == 3) {
            appKey    = args[0];
            id        = args[1];
            secret    = args[2];
        } else if (args.length == 4) {
            appKey    = args[0];
            id        = args[1];
            secret    = args[2];
            url       = args[3];
        } else {
            System.err.println("run error, need params(url is optional): " + "<app-key> <Ac
cessKeyId> <AccessKeySecret> [url]");
            System.exit(-1);
        }
        SpeechRecognizerDemo demo = new SpeechRecognizerDemo(appKey, id, secret, url);
        // In this example, local files are used to simulate the sending of real-time strea
```

```
ms.
        demo.process("./nls-sample-16k.wav", 16000);
        //demo.process("./nls-sample.opus", 16000);
        demo.shutdown();
    }
}
```