Voice Assets

This chapter describes how to create, configure, and manage the voice assets used for automatic outbound call tasks.

- Recording File Management
- Text-to-Speech (TTS) Voice Generation

Recording File Management

Overview

The **Recording File Management** module is used to upload and manage pre-recorded audio messages (such as notifications and reminders) for automatic outbound call tasks. Once uploaded, recording files are added to the voice asset library and can be used by outbound call tasks.

Accessing the Feature

After logging in to the Call Center platform, navigate to **Voice** > **Voice Assets** > **Recording File Management**.

Feature Description

Upload a Recording File

1. On the **Recording File Management** page, click **Add**.
2. In the dialog box, enter the following information:

- **Voice Name:** Required. It is recommended to include the business scenario and version in the name (e.g., **March Payment Reminder_V2**). The name must be between 1 and 20 characters.
- **Audio File:** Click **Select File** to upload a local **.mp3** or **.wav** file.
- **Description:** Optional. Used to describe the applicable scenario or validity period. Up to 100 characters.
- Click **Confirm** to complete the upload.

Note: The uploaded audio file must meet the following requirements:

- **Audio Format:** MP3 (recommended) or WAV
- **Sample Rate:** 8000 Hz
- **Channels:** Mono
- **Duration:** No longer than 60 seconds
- **File Size:** No larger than 1 MB

Files that do not meet these requirements cannot be uploaded.

Edit Recording File Information

1. On the **Recording File Management** page, locate the target recording file and click **Edit** in the **Actions** column.

2. In the dialog box, you can modify the following information:

- **Voice Name**
- **Description**

Click **Confirm** to save the changes.

Note: Uploaded audio files cannot be replaced. Only the metadata (Voice Name and Description) can be modified.

Preview a Recording File

1. On the **Recording File Management** page, locate the target recording file and click **Play** in the **Actions** column.
2. The system plays the recording in the embedded audio player.

You can use the player controls to pause playback, adjust the playback speed, or download the recording.

Tip: It is recommended to preview the recording after uploading it to verify the audio quality, playback speed, and content before using it for outbound call tasks.

Delete a Recording File

1. On the **Recording File Management** page, locate the target recording file and click **Delete** in the **Actions** column.
2. In the confirmation dialog box, click **Confirm**.

If the recording file is deleted successfully, it is removed from the list.

Important: A recording file can be deleted only if it is not referenced by any automatic outbound call or predictive outbound call task. If the recording file is in use, the system displays the following message: **"This recording file is currently in use and cannot be deleted."**

Text-to-Speech (TTS) Voice Generation

This page provides detailed instructions on generating voice assets using Text-to-Speech (TTS) technology, including text configuration, variable insertion, and pronunciation rules for special formats.

Overview

The **Text-to-Speech (TTS) Voice Generation** module converts text into speech and generates audio files in real time, making it ideal for dynamic voice scenarios such as bill amount notifications and verification code announcements. Users can configure both fixed text and parameter variables, preview the generated audio before saving, and add the synthesized audio to the **Voice Assets** library for use in outbound call tasks.

Accessing the Feature

After logging in to the Call Center platform, navigate to **Voice** > **Voice Assets** > **Generate Voice**.

Feature Description

Create a TTS Voice

1. On the **Generate Voice** page, click **Add**.
2. In the dialog box, configure the following information:

- **Voice Name:** Required. It is recommended to include the business scenario and version in the name (e.g., **Double 11 Promotion Notification_Male Voice_V2**). The name must be between 1 and 20 characters.
- **Text Type:** Required. Select one of the following modes:
  - **Static Text:** Enter fixed text that remains unchanged during playback.
  - **Variable Text:** Enter template text containing variables, which are dynamically replaced during outbound calls.
- **Language:** Required. Select the target language from the drop-down list (e.g., **Chinese (China)**, **English (United States)**, or **Japanese (Japan)**). The system loads the corresponding speech engine based on your selection.
- **Text Editor:** Enter the text to be converted into speech.
  - **Insert Variable:** Click **Insert** on the toolbar to insert predefined variables (e.g., **{name}** or **{amount}**) for dynamic replacement.

    - **Insert Pause:** Click the **Pause** icon on the toolbar to insert pauses of a specified duration (e.g., **0.5 s** or **1 s**) between sentences for more natural playback.
    - **Reading Mode:** For the selected variable or text, choose one of the following reading rules:
        - **Default:** Reads the text naturally based on the context.
        - **Number:** Reads a sequence of digits as a complete number, suitable for phone numbers or card numbers.
        - **Read Characters Individually:** Reads each letter or digit separately (e.g., **ABC123** is read as **A, B, C, 1, 2, 3**), suitable for verification codes or serial numbers.
        - **Date:** Automatically recognizes and reads dates in the appropriate format.
        - **Currency:** Automatically converts and reads monetary amounts using the appropriate currency format.
- **Preview (Recommended):**
    - Before saving, click **Preview** below the editor.
    - The system generates and plays the audio immediately using the current text, sample variable values, pauses, and reading rules.
    - You can adjust the configuration and preview the audio repeatedly until you are satisfied.

After confirming that all settings are correct, click **Generate and Save**.

Note: The generated content must meet the following requirements:

- **Text Length:** Up to 500 characters, including variable placeholders.
- **Variable Names:** Variable names must exactly match the column headers in the outbound call data file; otherwise, the values cannot be replaced.
- **Special Characters:** Complex mathematical expressions and unsupported special symbols or emojis are not supported and may be filtered or cause an error.
- **Audio Duration:** The generated audio must not exceed 60 seconds.

View and Play a Generated TTS Voice

1. On the **Generate Voice** page, locate the target entry and click **Play Voice** in the **Actions** column.
2. The system plays the generated audio file associated with the selected entry.

You can use the player controls to pause playback, adjust the volume, or download the audio file.

Delete a TTS Voice

1. On the **Generate Voice** page, locate the target entry and click **Delete** in the **Actions** column.
2. In the confirmation dialog box, click **Confirm**.

If the TTS voice is deleted successfully, the entry is removed from the list.

Important: A TTS voice can be deleted only if it is not referenced by any automatic outbound call or predictive outbound call task. If the TTS voice is in use, the system displays the following message: **"This voice asset is currently in use and cannot be deleted."**