info@readspeaker.com
www.readspeaker.com

ReadSpeaker® speechMaker Studio

ReadSpeaker speechMaker Studio (speechMaker Studio or speechMaker for short) is a web application that converts text input into audio files with a spoken version of the text using text-to-speech technology. For example, to create narration, audio descriptions, or talking subtitles.

Product Description

speechMaker Studio is a SaaS (Software as a Service) solution that is available to users who have been assigned a license and provided with login details.

To use speechMaker, the licensed user logs in with a username and password. Companies can buy a speechMaker license for five users. speechMaker administrators can invite other users.

To convert text into an audio track with speechMaker, the user types or copy/pastes the text into the tool, chooses the language/voice, the format for the audio output, the desired speaking rate, pitch, and volume, and simply exports the audio file to the local device. The user can preview (listen to) the audio before exporting.

The audio is generated using ReadSpeaker text-to-speech technology.

The customer has access to all languages/voices that are available by default in the product. If the customer has purchased custom voices from ReadSpeaker, these can also be made available in speechMaker.

Features

Dashboard

After logging into speechMaker Studio, the user is automatically redirected to the dashboard which is divided into sections to give the user an overview of the application, e.g. recent activity in the speechMaker account and system information.

The information displayed on the dashboard may vary depending on the user's role.

Content Library

Text content can be stored in speechMaker Studio's content library. A file containing text content is called a script. A script can contain one or more text blocks.

Scripts can be organized in folders, and folders and scripts can be organized in projects.

All users can access all content in the library of the customer account. Users with the role Viewer cannot make any changes. They can only view the content.

Text to Speech

speechMaker Studio can convert text to speech using state-of-the-art text to speech (TTS). For a complete list of available languages and voices, see the Voice Gallery section below. If you wish to use speechMaker with a language not currently listed, or even with your own custom voice, please contact ReadSpeaker, as other languages and voices can be made available on request.

Content Studio

Clicking on a script in the content library or on the dashboard will open it in the content studio.

Each script can contain one or more text blocks. Text blocks can have individual settings for:

- Reading voice
- Speaking rate (speed)
- Pitch
- Volume

The user can adjust the pauses between text blocks and the length of the silence at the beginning of the audio, to fit voice prompts into predefined time windows, for example.

To control the speech output at a more detailed level, the user can use the following tools:

- Pronunciation editor
- Intonation editor
- Pause tool

Adjustable Reading Voice

The reading voice can be selected from a menu and a different voice can be set for each text block. This means that a script can contain a dialogue between two or more speakers.

Adjustable Speaking Rate

The user can choose the speaking rate from slow to fast on a sliding scale to produce audio suitable for listeners who need a slower or faster speaking rate.

Adjustable Pitch

The user can choose the pitch from low to high on a sliding scale. A different pitch can be set for each text block.

Adjustable Volume

The user can choose the volume from low to high on a sliding scale. A different volume can be set for each text block.

Preview: Listen to Audio Before Exporting

The user can listen to the result of the text-to-speech conversion before actually exporting the audio content. You can preview the entire script or a specific text block.

Pronunciation Editor

The user can select a word and edit its pronunciation in the pronunciation editor. The pronunciation can be changed in one of two ways:

- Alternate spelling - Type the word with a different spelling to alter the way the TTS pronounces the word. E.g. iPhone could alternatively be spelled "eye phone".
- Phonetic transcription - Select phonetic symbols from a digital IPA keyboard to create a phonetic transcription.

Regardless of the method used to change the pronunciation, each edit can be applied to different predefined scopes:

- This Instance - The pronunciation is changed only once.
- This Project - The edited pronunciation will be used every time the word appears in the project.
- Always - The pronunciation will be used every time the word appears in any project in the customer account. This can be useful for company names or trademarks that always require a certain pronunciation.

Pronunciation edits can be previewed before they are applied to check they sound correct.

Intonation Editor

The intonation can be changed at the sentence level by adjusting the pitch and/or duration of each word in the intonation editor.

The intonation editor displays the selected sentence as a curve where each word is represented by a point that can be dragged/dropped to make the pitch higher or lower than normal. Each word can be assigned a duration value to make it shorter or longer than normal.

Note that the pitch and duration changes are relative to the way the TTS would normally read them. Therefore, the same edits can sound different for different reading voices, based on the characteristics of each voice.

Intonation edits can be previewed before they are applied to check they sound correct.

Pause Tool

Pauses can be inserted anywhere in the text. The user selects the desired length of the pause from a menu. The pauses are labelled and displayed in different colors.

Export Audio

The user can download audio as audio files to their device. This is referred to as exporting.

Audio can be exported at different levels:

- Text block - Export the audio for a single text block from within the content studio.
- Script - Export a single script from within the content studio or the content library.
- Project - Export a project from the content library.

Exporting an entire project will result in a zip file that mirrors the folder structure for that project in the content library. However, only folders that contain scripts will be included in the exported file.

Voice Gallery

All the voices that are available by default are listed in the voice gallery. The list can be made shorter by using the search filters. The voice list is sorted by language and then by voice name.

Each voice is presented with an avatar, a flag, and a text that the voice can read aloud so the user can preview the voice.

The user can mark a voice as a favourite which will make it show at the top of the voice menu in the content studio as well as in the voice gallery.

Any customer-specific voices created by ReadSpeaker can be added to the voice gallery on request.

Credits for Production

speechMaker Studio requires credits to export audio. New accounts typically include a trial period with a limited number of credits for testing purposes. Administrators can purchase credit packages through speechMaker. These packages contain a specific credit amount valid for a defined period. Each purchase requires confirmation from a ReadSpeaker account manager before the credits can be applied to the account.

One credit is charged per character when an audio file is exported. The characters are counted excluding whitespaces and including any changes dictated by pronunciation dictionary entries.

View Remaining Balance

The number of remaining credits, i.e. how much text, measured in characters excluding spaces, the users can still convert to speech, is displayed in the top bar of the speechMaker Studio interface. The remaining balance is visible to users of all roles.

When the balance has been depleted, speechMaker administrators can purchase extra credits.

Users and Roles

The users in a speechMaker Studio account can have different roles:

- Owner - This role is assigned to the first user created for the customer account. The owner cannot be deleted as customer accounts must have an owner. The owner has the same permissions as an administrator.
- Admin - Administrators can edit all content in the library, invite users, and order speechMaker credits.
- Editor - Editors can edit all content in the library.

- Viewer - Viewers can view all content. They cannot make any changes. Viewers can preview audio in the studio.

Technical Specifications

Device Independent

speechMaker Studio is a device-independent and cross-platform solution. It should work on all major platforms and in all major browsers.

Fully Server-Based, No Software Downloads or Installation

speechMaker Studio users need not download or install any software for speechMaker to work on their computers and/or mobile devices. speechMaker is cloud-based and hosted by ReadSpeaker.

Product Hosting

speechMaker Studio is hosted by ReadSpeaker in a high-security environment with a high level of redundancy and failover, and with the following properties:

- double redundant enterprise load balancing and failover
- 100% Gigabit network (internally and to the Internet)
- IBM server hardware with virtualization engine
- redundant power supply
- highly scalable and high availability design
- environmentally friendly
- Storage Area Network (SAN) storage over Gigabit iScsi
- UPS and diesel engine power backup
- 24/7 video and audio surveillance
- Level 3 security

- Argonite-based automatic fire-fighting system
- automatic server monitoring every 3 seconds with alarm system
- automatic daily backup

Supported Input Formats

speechMaker Studio accepts plain text input. The text must be entered in speechMaker. This can be done by writing or by copying a text from a file and pasting it into the editor.

Text Size Limit

speechMaker Studio has a default limit of 5,000 characters, excluding spaces, for text-to-speech conversion per audio file produced.

Supported Audio Formats

The output file can be in the following formats:

- MP3 (24, 48, and 128 Kbits/s, 22.1 KHz and 44.1 KHz; 192 Kbit/s, 44.1 KHz)
- Ogg Vorbis (variable bitrate Q=0,2)
- PCM (8 bit/16 KHz, 16 bit/8 KHz, 16 bit/16 KHz, 16 bit/44.1 KHz)
- IMA-ADPCM (4 bit/8 KHz)
- WAV (U-law 16 bit/8 KHz, A-law 16 bit/8 KHz)
- WAV (8 bit/16 KHz, 16 bit/8 KHz, 16 bit/16 KHz, 16 bit/44.1 KHz, 16 bit/48 KHz)

The Order Confirmation will specify which formats are available for the account.

Product Components

SaaS Application

The core of speechMaker Studio is the ReadSpeaker speechMaker Studio SaaS (Software as a Service) application, hosted by ReadSpeaker, and accessible to all users with login credentials for a speechMaker license.

Login Credentials

The only “product component” for the user is the login credentials, which provide access to speechMaker Studio.

User Support and Documentation

Users are provided with a link to a user manual. The ReadSpeaker support team can be contacted for further support and questions. See https://www.readspeaker.com/support/ for opening hours and contact details.

Options

Options are features that can be added to speechMaker Studio at a cost. Please contact your Account Manager for pricing information.

Larger Maximum Size for Text Input

The default limit of 5,000 characters for text-to-speech conversion per audio file produced can be increased on a per-customer basis.

Requirements and Limitations

Internet Connection

The user must have an Internet connection to be able to use speechMaker Studio.

Maximum Size for Text Input

The maximum size for text-to-speech conversion per audio file produced is 20,000 characters excluding spaces. Note: the default limit is 5,000 characters. This can be increased on a per-customer basis, see under Options above.

Voice Types

ReadSpeaker solutions use voices of different types. All the voices available by default in speechMaker Studio have the voice type Neural Premium. Voices of other voice types can be added to the customer account, for example a custom voice. It is, however, not possible to use voices of different voice types in the same script.

Restriction to the Use of the Audio Files

The customer may use the audio files produced with speechMaker Studio only within the context set out in the order confirmation.

Support for HTML5 and CSS3

The user must use a web browser with support for HTML5 and CSS3 to enjoy the full functionality provided by speechMaker Studio.

Price Model

speechMaker Studio is charged for with a set-up fee and a fixed monthly or annual service and maintenance fee. The audio produced is charged based on the number of characters converted, in accordance with the pricing specified in the Order Confirmation.

Intended Use

Customers may only use speechMaker Studio to convert text segments into audio files in accordance with this product specification.
Any other use of speechMaker is not considered intended use.

Disclaimer

Features listed in this document are guaranteed only if the requirements listed in this product specification are met. Specifications and features as described in this product specification can be changed by the manufacturer without prior notice.