

ReadSpeaker® speechMaker

ReadSpeaker speechMaker (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

To use speechMaker, the licensed user logs in to ReadSpeaker's web application service platform (WASP) with a username and password (and clicks on the speechMaker app icon if there are multiple apps in the account). Larger companies that wish to have multiple people use speechMaker for audio production can purchase a speechMaker package, i.e. a batch of speechMaker licenses which are distributed to users by the speechMaker administrator, who can add a user for each license, after which the user will receive login credentials.

To convert text into an audio track with speechMaker, the user types or copy/pastes the text into the tool, chooses the language/voice from the licensed voices in the account, the format for the audio output, the desired speaking rate, pitch, and volume, and simply downloads the audio file. The user can preview (listen to) the audio before downloading.

It is also possible to produce multiple audio files in one go using a script in Microsoft Excel.

The audio is generated using ReadSpeaker text-to-speech technology. The customer can purchase licenses for one or multiple languages and for one or multiple voices in each language.

speechMaker is a SAAS (Software As A Service) solution that is available to users who have been assigned a license and provided with login details.

Features

Text to speech

speechMaker can convert text to speech using state-of-the-art text to speech (TTS). speechMaker can produce audio files in 50+ languages. For a complete list of available languages go to [our languages web page](#).

Accessibility

speechMaker's full functionality is keyboard accessible.

Downloadable audio file

The user can download the produced audio files to her or his device.

Support for uploading audio files to Dropbox

The user can save the audio files produced directly to Dropbox.

Support for SSML

The input text can include SSML. With SSML, the user has more control over how the text is read by the TTS engine, specifically the prosody of the speech. SSML can be used to insert pauses/breaks of certain length, to insert phonetic transcriptions and to switch voices and languages in the same text. SSML support

is not activated by default. It can be activated by ReadSpeaker for customers wishing to use SSML.

Produce multiple files at once

If the user has many lines or sections of text that must be converted to multiple audio files (s)he can use batch processing. This is done via an Excel sheet where each row contains a field with the section of text and a field with the name of the file the corresponding speech should be saved in.

Adjustable speaking rate

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

Adjustable pitch

The user can choose the pitch from low to high on a sliding scale.

Adjustable volume

The user can choose the volume for the audio output from low to high on a sliding scale to produce files suitable for listeners with a need for adjusted volume.

Multiple languages and voices

Available if the customer has purchased a license for more than one language or for more than one voice. speechMaker supports many different languages and voices. There can be multiple voices for the same language. For many languages there are both female and male voices. For some languages there are also different regional accents, e.g. American, Australian, British, and Indian English. For each text-to-speech conversion a different language/voice can be used. The languages and voices licensed are specified in the order confirmation.

Pronunciation dictionaries

speechMaker can be customized to pronounce words or sequences of words in a specific way to increase the quality of the audio. The linguists at ReadSpeaker can correct mispronunciations in the dictionary for the voice used upon request. If a different pronunciation is site specific, it will be added to the customer's personalized dictionary.

View remaining balance

Shows the user how many credits are left, i.e. how much text, measured in characters excluding spaces, the user can still convert to speech. When the balance has been depleted, the user can purchase extra credits.

Administrator features

The users in a speechMaker package are managed by a speechMaker administrator, a user who has been assigned login credentials as well as administration rights.

Assign licenses to users

speechMaker administrators can create new users, i.e. assign speechMaker licenses to users and assign available credits to users. Users receive their login credentials for speechMaker by email automatically.

Access to online usage statistics

The administrator has access to usage statistics per selected period on the Usage tab within speechMaker. The tab shows the number of characters and words converted per user, per language/voice and in total for the selected period.

View remaining balance

The administrator can see how many credits (characters) are left in the account. When the balance has been depleted, the administrator should contact the ReadSpeaker Account Manager to order extra speechMaker credits.

Technical Specifications

Device independent

speechMaker 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 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 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

Modifiable text size limit

speechMaker has a default limit of 20,000 characters, excluding spaces, for text-to-speech conversion per audio file produced. This limit can be modified on a per customer basis (see options).

Formats for audio output file

The output file can be in the following formats:

- MP3 (24, 48, 128 Kbit/s)
- 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, U-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 is the ReadSpeaker speechMaker 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.

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 at a cost. Please contact your Account Manager for pricing information.

Larger maximum size for text input

The default maximum size text to be converted is 20,000 characters. This can be increased on a per-customer basis.

Extra languages or voices

Please contact your ReadSpeaker account manager if you would like to add more available languages for the licenses in your speechMaker package.

Requirements and Limitations

Internet connection

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

Maximum size for text input

The default maximum size for text to be converted into an audio file is 20,000 characters excluding spaces.

Not all audio formats supported for batch-processing

When multiple audio files are produced at once, the following audio formats are supported:

- MP3 (24, 48, 128 Kbit/s)
- OGG Vorbis (variable bitrate Q=0,2)
- WAV (16 bit/22.5 KHz)

Restriction to the use of the audio files

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

Support for Flash or HTML5

The user must use a web browser with support for Flash or HTML5 to be able to enjoy the full functionality provided by speechMaker.

Price Model

speechMaker is charged for with a set-up fee and a fixed monthly or annual service and maintenance fee. The audio produced using speechMaker is charged on usage by number of characters converted according to the pricing agreed in the Order Confirmation.

Intended Use

Customers may only use speechMaker 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.