

ReadSpeaker® speechTime

ReadSpeaker speechTime (speechTime for short) is a web application that converts timed text manuscripts in SRT format (.srt files) into audio files using text-to-speech technology. For example to create narration, audio descriptions, or talking subtitles from timed text scripts.

Product Description

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

To convert an .srt file containing timed text manuscript into an audio track with speechTime, the user needs to create a job, give it a name, choose a format for the audio output, choose a language and voice, add a description to the job, upload the .srt file(s) and start the job. Once the conversion is complete, the user can preview (listen to) the audio and download the audio file, which can then be added to a movie to give it an audio description.

As input, speechTime uses an .srt standard subtitling file, where there is time code information and text. This is the type of file used for subtitling. It can be used for things like audio descriptions, i.e. a spoken description of what is happening in a movie, so that someone that can't see well, e.g. due to a visual impairment, can still understand the action.

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.

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

Features

Text to speech

speechTime can convert timed text manuscripts to speech using state-of-the-art text to speech (TTS).

Accessibility

speechTime's full functionality is keyboard accessible.

Downloadable audio file

The user can download an audio file in the chosen format with the audio version of the timed text manuscript.

Adaptive speaking rate

The timed text manuscript contains numbered text fragments, each with a start time and an end time. If the play time of the audio generated from a text fragment at default speaking rate will be longer than the time window specified,

i.e. the period between start time and end time, speechTime will adapt the speaking rate until the output does fit into the time window. This is very important since sometimes the script author does not know exactly how long it will take to read a certain text out loud.

Multiple languages and voices

Available if the customer has purchased a license for more than one language or for more than one voice. speechTime supports many different languages and voices. For some languages there are multiple voices, such as male and female and American, Australian, British, and Indian English. For each job a different language/voice can be used. The languages and voices licensed are specified in the order confirmation.

Pronunciation dictionaries

speechTime 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 edited in the customer's personalized dictionary.

View remaining balance

Shows the user how much speech time is left for that user, i.e. how many seconds of speech that user can still generate. When the balance has been depleted, the user can ask the administrator for extra balance in speechTime.

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 speech time 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 speechTime. The tab shows (in seconds) both speech time (actually generated speech) and play time (movie play time) generated per user, per language/voice and in total for the selected period.

View remaining balance and order extra balance

The administrator can see how much speech time (for actually generated speech) is still left in the account (in seconds) and per user and redistribute the remaining balance between the users. When the balance has been depleted, the administrator should contact the ReadSpeaker Account Manager to order extra speechTime balance.

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

SRT format for timed text input file

The input file with timed text must be in SRT format (SubRip subtitle). The file must have an .srt extension. The timing in the input file must be such that the durations set for each text fragment is sufficiently long for the audio output of that text fragment.

Formats for audio output file

The output file can be in the following formats:

- MP3 (128 Kbit)
- MP3 (32 Kbit)
- MP3 (64 Kbit)
- OGG Vorbis
- WAV (Mono 16bit, 22 KHz)

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

Product Components

SAAS application

The core of speechTime is the ReadSpeaker speechTime SAAS (Software As A Service) application, hosted by ReadSpeaker, and accessible to all users with login credentials for a speechTime license.

Login credentials

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

User support and documentation

Users are provided with a link to an instructional video and user documentation in the form of a Quick Start Guides for users and for administrators.

Options

Larger maximum file size for input file

The default maximum size for .srt files to be converted into audio is 10 MB.

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 speechTime.

SRT format

The input file with timed text must be in SRT format (SubRip subtitle). The file must have an .srt extension. The timing in the input file must be such that the durations set for each text fragment is sufficiently long for the audio output of that text fragment.

Maximum file size for input file

The default maximum size for .srt files to be converted into audio is 10 MB.

Support for HTML5

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

Price Model

speechTime is charged for with a set-up fee and a fixed monthly or annual service and maintenance fee. The audio produced using speechTime is charged on usage either by speech time (actually generated speech) or by play time (movie play time) according to the Order Confirmation.

Intended Use

Customers may only use speechTime to convert "timed text" in an .srt file (i.e. in SRT format) into an audio track and in accordance with this product specification. Any other use of speechTime 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.