ReadSpeaker

pioneering voice technology

info@readspeaker.com
www.readspeaker.com

ReadSpeaker® speechServer

ReadSpeaker speechServer allows you to make text-to-speech technology available in a client application through a network. The server makes it possible for multiple client applications to simultaneously request for speech synthesis service.

Product Description

ReadSpeaker speechServer provides a complete text-to-speech (TTS) engine with voice specific files for each licensed voice and a Server API, along with the tools you need depending on your specific server requirements and enables flawless dynamic multithreaded speech sessions.

It offers high quality TTS with natural sounding human-like voices. It's perfect for IVR systems, virtual assistants, alert or notification systems, educational web apps, and more.

speechServer allows developers to integrate ReadSpeaker TTS voices in client-server architecture and run intelligent voice services efficiently, interacting with speechServer through a TCP/IP communication protocol, a command line tool, or REST API. speechServer allows you to simultaneously execute as many TTS conversions as you need.

Features

Leading synthetic speech for superior customer experience

ReadSpeaker's text-to-speech voices are extremely accurate, clear and natural, designed to deliver the highest quality sound and exceptional performance every time and are continually optimized.

90+ voices in 30+ languages available

New voices are being developed all the time. Presently, ReadSpeaker has TTS voices in 30+ languages and 90+ voices available. For a complete list of available languages go to our languages web page.

Modifiable speaking rate, pitch, and volume (DSP)

The speaking rate, pitch, and volume can be configured as desired.

Voice/language switch

Switching to another language or to another speaking voice in the same language during the conversion from text to speech in response to SSML code in the input text is supported.

User dictionary, IPA supported

The voice specific files included in speechServer for each licensed language include a dictionary file, the user dictionary, in which the customer can customize the pronunciation of words or sequences of words in a specific way to increase the quality of the reading. The user can create multiple user dictionaries per language if required. The user dictionaries accept IPA input.

Audio clip insertion supported

speechServer allows for (links to) audio clips (files) to be inserted in the text input so that the audio from the clip will be inserted in the correct place within the synthesized speech in the output audio streamed back to the client application.

Admin tool

speechServer comes with a powerful administration tool which makes it easy to operate and monitor and includes:

- Service status monitoring
- Service start & stop
- Usage statistics
- Log viewer
- Configurations setup

Mark information creation for event notification

speechServer allows for mark information to be created in addition to the audio output. Mark information allows creating events when a specific location in a text is read. This can be used to support running a device in sync with the audio output or to offer synchronized highlighting of the text during reading.

Technical Specifications

Supported Operating Systems

- Windows : Server 2008, 2012, 2016
- Linux :
  - CentOS 5.x / 6.x / 7.x
    Ubuntu 16.x
  - RHEL 5.x / 6.x / 7.x

Supported architecture

Intel 32 and 64 bit (x86/x86_64)

CPU

1 GHz or higher

RAM

4 GB or more is recommended

Voice footprint

64 MB ~ 2.2 GB per voice created with unit selection synthesis (USS)

5 MB ~ 15 MB per voice created with hidden Markov model (HMM) or deep neural network (DNN) technology

Runtime memory

- 100 MB for basic load
- 35 MB per voice
- 2 MB ~ 90 MB per channel

Development languages

- C / C++
- Java
- C# / VB .NET (Windows)

Supported input formats

- Plain text
- SSML

Supported character encoding for text input

- UTF8 (all voices)
- Multibyte code set (certain voices only)

Supported audio formats

- 16-bit linear PCM
- 16-bit linear PCM Wave
- 8-bit A-law PCM
- 8-bit A-law PCM Wave
- 8-bit µ-law PCM
- 8-bit µ-law PCM Wave
- 8-bit unsigned linear PCM Wave
- 4-bit Dialogic ADPCM
- MP3, requires downloading and installing the LAME package
- Ogg Vorbis
- Windows ASF

TCP/IP protocol

speechServer supports the use of TCP/IP protocol for communication between the server application and the TTS engine/speechServer.

Command line interface

speechServer allows the server application to communicate with the TTS engine/speechServer through a command line tool.

REST API

speechServer allows the server application to communicate with the TTS engine/speechServer through a REST API. REST API requires that the command

line application is reachable via a CGI enabled web server application such as Apache.

Voice/language switch

Switching to another language or to another speaking voice in the same language can be realized through SSML commands.

Simple implementation

ReadSpeaker speechServer is easy to install using the speechServer installer. Any ReadSpeaker voice can be licensed, installed and used in speechServer at any time.

Installation and implementation support

ReadSpeaker speechServer includes installation and implementation support by our Support Team, who will assist the customer wherever necessary in understanding and carrying out the installation of the speechServer following the instructions provided as well as the implementation necessary on the client application to communicate with speechServer.

Configuration settings

The use of a configuration file is supported in which the user can set the default value for certain parameters, including:

- speaking rate
- pitch
- volume
- the default user dictionary to use

License file required

The use of ReadSpeaker speechServer is governed by a required separate license agreement. This license agreement determines which voices are licensed and the context within which the product may be used, usually the name and a description of the server application that will be speech-enabled. It also specifies how many concurrent TTS ports are licensed, the synthesis speed rate for text-to-speech conversion, the number of servers licensed, and which method of communication with the speech server will be used (command line tool, REST API or TCP/IP communication protocol (Linux only)). A license file is a part of the product and technically enforces elements of the license agreement. For other elements, the agreement is based on trust that the customer will not exceed the specified use.

Product Components

ReadSpeaker speechServer

TCP/IP network-based speech synthesis service program

Client SDK

SDK (C, Java, .NET, COM) available on a client software

ReadSpeaker TTS Engine

The ReadSpeaker TTS Engine includes the voice database and a pronunciation dictionary for each licensed language/voice. The voices database(s) contain audio fragments which are used by the TTS Engine for voice synthesis. The TTS engine is called upon by the speechServer to convert text to voice.

User dictionary tool

Allows the user to create one or more user dictionaries for each licensed language to customize the pronunciation The user dictionary files are used by the TTS Engine for the voice synthesis.

Administration tool

User documentation

Includes the APIs needed and the installation and implementation instructions.

Requirements and Limitations

Scope of use

ReadSpeaker speechServer may only be used within the context and with the restrictions, such as number of ports, described in the license agreement.

Admin tool

The admin tool will only function on Linux machines with Apache web server and xinetd super-server daemon installed and on Windows machines with IIS web server.

Intended Use

Customers may only use ReadSpeaker speechServer in accordance with this product specification and the separate required license agreement.
Any other use of ReadSpeaker speechServer is not considered intended use.

Disclaimer

Features listed in this document are guaranteed only if ReadSpeaker speechServer has been implemented according to our implementation instructions. Specifications and features as described in this product specification can be changed by the manufacturer without prior notice.