

# Installation Guide

Windows version

## System Requirements

The installer checks most of the requirements listed below, except the OS version. If your computer does not meet the minimum requirements, installation will not be possible.

### Minimum Requirements

- [x86-64](#) processor running 64-bit version of Windows 10 or Windows 11
- The processor should support [AVX2](#) instruction set extension
- Graphics card and GPU driver supporting at least Direct3D 11.0 feature level 11.0
- Approximately 36 MB disk space for the core software
- Additional disk space for speech-to-text models:
  - Voxtral Mini Realtime: 3.4 GB, 4.6 GB, or 8.5 GB depending on the compression
  - Whisper Turbo BCML 580 MB, Whisper Large BCML 1 GB

An internet connection is **not required** for day-to-day use. This software is not a subscription service.<sup>1</sup>

### Recommended

- Graphics card supporting Direct3D 12, shader model 6.4, and fp16 shader instructions.
- GPU with a theoretical performance of at least 3 teraflops (FP16 or FP32)

A discrete graphics card is recommended, but not required. Speech to text works very well on modern integrated GPUs such as the [Radeon 780M](#). Performance is acceptable on older integrated graphics like [Vega 7](#), which is a mobile iGPU introduced in 2021<sup>2</sup>.

The software is unlikely to perform well on the integrated graphics found in high-performance desktop CPUs such as the Ryzen 7 7700, as these provide too few GPU cores.

## Installation

After downloading and running Speech2TextSetup.exe, Windows will usually ask for permission to make changes to your device, with a prompt that reads “Do you want to allow this app to make changes to your device?”

---

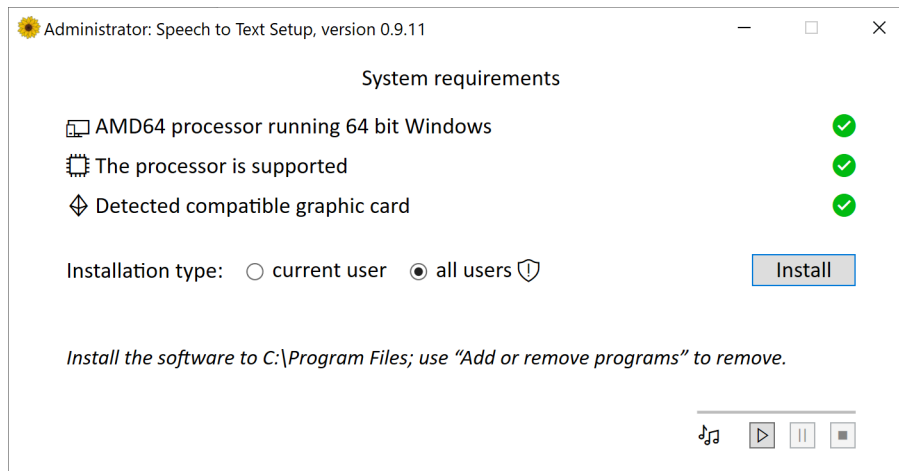
<sup>1</sup> You still need internet access (possibly on another computer) to activate the software, and download speech-to-text models.

<sup>2</sup> Performance on integrated graphics cards depends heavily on memory bandwidth. My old laptop with Vega 7 has dual channel DDR4-3200 memory, which is optimal for the Ryzen 5 5600U chip.

If you decline this prompt, the installer will only allow installation into %LOCALAPPDATA%

If you can approve the prompt (that is, you have administrator permissions), I recommend installing system-wide to C:\Program Files. This is mainly for security reasons: non-elevated processes cannot modify binaries installed under Program Files

The setup program consists of a single screen shown in the screenshot below.



If you choose to install for current user only, the program will be installed into "%LOCALAPPDATA%\Speech to Text\Binaries" and start menu shortcuts will be created in "%APPDATA%\Microsoft\Windows\Start Menu\Programs\Speech To Text"

If you choose to install for all users, the program will be installed into "%ProgramFiles%\Speech to Text" start menu shortcuts in "%ProgramData%\Microsoft\Windows\Start Menu\Programs\Speech to Text"

Once installation is complete, consider reading UserGuide.pdf document. It is located in the installation folder, and there is also a Start Menu shortcut that opens it directly.

## Compatibility Notes

The software is only compatible with 64-bit OS versions. Technically, it is fully functional on Windows 8.1. The software works on Windows 7 SP1, but there it requires audio sample rate to be exactly 16 kHz. To run the software on Windows 7 or 8, please install "Segoe MDL2 Assets" font from the "Segoe fonts v1710.zip" archive [available from Microsoft](#). These fonts are distributed under a custom licence and I'm afraid I do not have the right to redistribute them myself.

I have tested speech to text software in a VMware Workstation Pro virtual machine. That virtualisation software [is now free](#), by the way. With the default settings, the software fails to load models with an error message complaining about device interface or feature level not being supported. However, it runs successfully after switching the GPU backend from the default Direct3D 12 to the older, slower, but more compatible Direct3D 11.