

NUANCE

The experience speaks for itself™

DATASHEET

Dragon NaturallySpeaking 10 SDK Server Edition

Dragon NaturallySpeaking 10 SDK Server Edition (DSS) provides a set of tools and interfaces to help software developers successfully build and integrate automated speech recognition into dictation and transcription workflow systems.

DSS is specifically designed for integration into solutions that require speech recognition to run in the background, without any interactive dictation or self-editing by users. Back-end speech recognition enables the batch processing of recorded speech from a range of devices, including digital recorders, PDAs, tablet PCs and telephone dictation systems.

Although DSS is not a transcription workflow in itself, it can be integrated within existing systems to automate transcription, increasing transcription productivity and reducing transcription costs. Instead of manually transcribing the original dictation, the transcriptionist simply edits the draft transcription, correcting any errors in the text.

DSS contains the following components:

- **A Dragon NaturallySpeaking runtime engine** that is configured for back-end transcription tasks. This engine does not require the presence of a sound card, enabling it to run on a server system.
- **IMPROVED Acoustic models.** The acoustic model, a statistical representation of the sound patterns that make up individual words, determines how each word sounds, given the acoustic-phonetic properties of the language, the speaker's voice, the audio channel (microphone, bandwidth, etc), and the acoustic environment (noise levels, room acoustics, etc). This enables the software to choose words that sound most similar to the spoken utterance. DSS comes with a collection of base acoustic models that have been trained on vast amounts of relevant speech recordings from many speakers. DSS supports different accents and dictation sources, including dictation recorded over the telephone.
- **IMPROVED Vocabulary.** DSS comes with a set of vocabularies, which include a word list with information about all the words a program can recognize and a language model that contains usage information about those words. DSS uses the vocabulary to recognize words correctly based not only on the sound of the words, but on their context. Legal vocabularies are now available for US English and German.
- **A set of ActiveX controls** that provide interfaces for integration into dictation/transcription workflow software. ActiveX provides flexible server software integration and is supported by many compiled and interpreted (script) languages such as C++, C#, Visual Basic, Python, Perl, etc.
- **Fully functional sample programs** that are designed for use in the transcription workflow.

TRANSCRIPTION WORKFLOW WITH DSS

Before back-end transcription can be implemented as part of a transcription workflow, a user and vocabulary must be created for each speaker, who should then be enrolled/trained within the system. To obtain optimal accuracy for each speaker/dictation-source combination, the application needs to adapt the factory-built acoustic models to the specific characteristics of the individual speaker's voice, accent and speaking style using speech recordings from the same dictation source (telephone, voice recorder, or microphone). DSS provides an "Enrollment from Transcriptions" capability (called EFEnroll, for eyes-free enrollment) that uses actual dictation and the associated transcriptions for each speaker to adapt an acoustic model. DSS also offers a Speaker Triage command-line utility that predicts the effectiveness of automatic speech recognition for an enrolled user using a set of recordings and corrected transcripts.

Programming Environment

The Dragon NaturallySpeaking SDK Server Edition (DSS) provides developers with the tools that needed to integrate Dragon's transcription capabilities into a transcription workflow application. Among these tools are the Dragon APIs used to access Dragon's ActiveX controls when programming in the DSS environment. Developers can program in any language that ActiveX controls support, but the Dragon APIs are documented for Visual Basic and C++.

Supported audio formats. DSS supports data formats from 8khz (for telephony) to 99khz in PCM and MS-ADPCM and IMA-ADPCM format. DSS supports the following audio file types in both mono and stereo:

- WAVE PCM
- MS ADPCM
- IMA ADPCM
- a-law
- mu-law
- VOX
- MP3
- WMA

System Requirements

- CPU: Intel® Pentium4® or later or AMD Athlon 64 1 GHz or later. (SSE2 instruction set required).
- Memory: 512 MB RAM (1 GB RAM for Windows Vista™)
- Free hard disk space: 1 GB (2 GB for localized non-English versions)
- L2 Cache: 512 KB
- Supported Operating Systems: Windows Server 2000, Windows Server 2003, Windows XP SP2 or higher - 32 bit, Windows 2000 SP4 or higher, Windows Vista™ or Windows Vista™ SP1, 32-bit
- DVD-ROM drive (required for installation)
- Microsoft® Internet Explorer 6 or higher (free download available at www.microsoft.com)
- Nuance-approved noise-canceling headset microphone (included, Creative® Labs Sound Blaster® 16 or equivalent sound card supporting 16-bit recording, speakers (required for playback of recorded speech and text to speech features)
- For Bluetooth wireless microphone support, please visit <http://support.nuance.com/compatibility/>
- A Web connection is required for activation

During the install process the software checks to make sure your system meets these minimum requirements. If you do not meet the requirements, the software will not be installed.

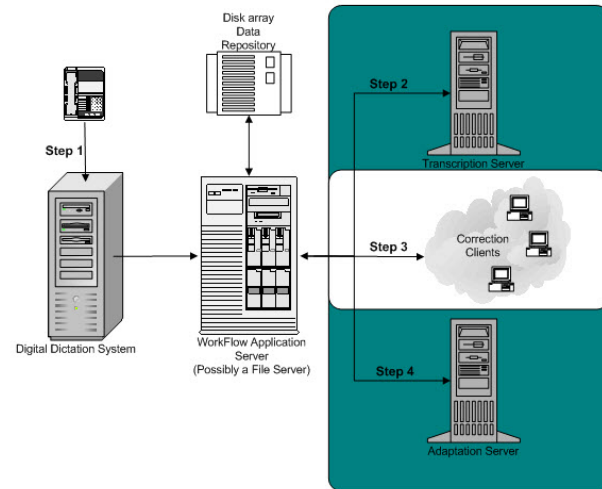
* This edition of Dragon does not support dictation directly into Electronic Medical Record (EMR) Systems. For EMR support, please use Dragon Medical Edition.

Recommended Specifications

- CPU: Intel® Pentium4® / 2.4 GHz (1.6 GHz dual core) or equivalent AMD processor. (SSE2 instruction set required).
- Memory: 1 GB RAM
- L2 Cache: 1 MB



©2009, Nuance Communications, Inc. All rights reserved. Nuance, the Nuance logo, Dragon, the Dragon logo, DragonBar, NaturallySpeaking, and RealSpeak are trademarks or registered trademarks of Nuance Communications, Inc. or its affiliates in the United States and/or other countries. All other trademarks are the property of their respective owners



- **Step One: Dictation.** A speaker dictates into the digital dictation system, creating a .WAV file. The transcription workflow component uses the metadata that is included with each .WAV file to determine the identity of the speaker and retrieve that speaker's user profile and the appropriate vocabulary. The workflow component passes the .WAV file, the user's profile, and the appropriate base speaker model and base topic model to the transcription component.
- **NEW Support for Audio Streaming.** To further streamline workflow, recognition can now be performed from a streaming file, eliminating the need to produce an audio file.
- **Step Two: Transcribing the Dictation.** The transcription component transcribes the recorded text to either a .TXT file or a .DRA file. If necessary, the transcription component will also create a concordance file for the transcribed text. The .DRA file includes speech and text data, as well as time stamps for each utterance, alternative word choices, and other information about the transcript.
- **Step Three: Correcting the Transcription.** A transcriptionist compares the transcribed document with the original .WAV file and fixes any errors. The transcription application's workflow component passes a correction client all of the files it needs to update a draft transcript, including user information, .WAV data, and an .IDX file that contains timestamps for each utterance. After the transcriptionists have corrected the draft transcript, the speaker verifies that the corrected document is accurate. The files, including the corrected transcript, are then passed back to the application's workflow component, where they are stored until it can be passed to the adaptation component.
- **Step Four: Adaptation.** After the transcriptionist has corrected the draft transcription and the speaker has reviewed and accepted the corrected transcript, the corrected transcript can be used to update the speaker's user profile. Updating the speaker's profile increases Dragon's accuracy for that speaker's subsequent transcriptions.