



SPEECHMATICS

Batch Containerised ASR

Our Batch Containerised ASR provides transcription of pre-recorded audio within a Docker container.

The transcription is provided after an audio file is passed to the Speechmatics engine within the container. The container technology provides very simple and really fast installation, management and use with a customer's deployment. The container does not store any audio or transcripts making it easy to use within secure environments and to maintain any audio and transcripts within the customer's own security boundaries.

Transcription:

- Supports pre-recorded audio
- Transcripts provided in JSON or TXT format, containing:
 - Timing and confidence scores for each transcribed word
 - Sentence boundary information
- Diarisation is available in 2 modes:
 - Speaker diarisation provides the ability to detect and label speakers within the same channel
 - Multi-channel diarisation provides speaker labelling based on up to 6 audio channels or streams
- Custom dictionary allows up to 1,000 additional words to be added to the dictionary on a per-transcription basis
 - Allows users to quickly add context-specific words, for example company names, place names or unusual names.
 - Custom Dictionary Sounds is an extension that allows alternate spellings or pronunciations to be used

Languages supported:

- Core languages
 - Global English (en), German (de), Spanish (es), French (fr), Dutch (nl), Portuguese (pt), Japanese (ja), Korean (ko), Italian (it), Swedish (sv), Danish (da), Polish (pl), Catalan (ca), Hindi (hi), Russian (ru)
- Research languages
 - Bulgarian (bg), Czech (cs), Greek (el), Finnish (fi), Croatian (hr), Hungarian (hu), Lithuanian (lt), Latvian (lv), Romanian (ro), Slovakian (sk), Slovenian (sl)

Programmatic use:

- Console/STDIO is used to directly interact with the

container to provide a transcription

- Docker layering allows simple integration into customer deployments

Hardware needs:

- The container needs to run in a supported Docker environment:
 - 1 vCPU per concurrent transcription
 - 2-5GB RAM for each executing container*

*At least 100 MB free disk space must be available for operating overheads and, if streaming in audio file space, for a copy of the audio file being transcribed. Memory requirements will depend on audio quality

Storage:

- 2GB storage footprint per language container
- No audio or transcripts stored within the container

Performance:

- 1vCPU per concurrent transcript
- Transcript can be provided in better than 2x real time
- Multiple containers can be executed on the same Docker engine at the same time or across multiple Docker engines to enable large scale operations

Connectivity requirements:

- Can operate within your own security boundary allowing you to keep control of your own data

Administration:

- No ongoing maintenance needed for the containers
- All administration is provided by direct use of Docker commands
- Flexible licensing model

Contact us for more information on features and pricing:

info@speechmatics.com | +44 (0)1223 794 497 | www.speechmatics.com