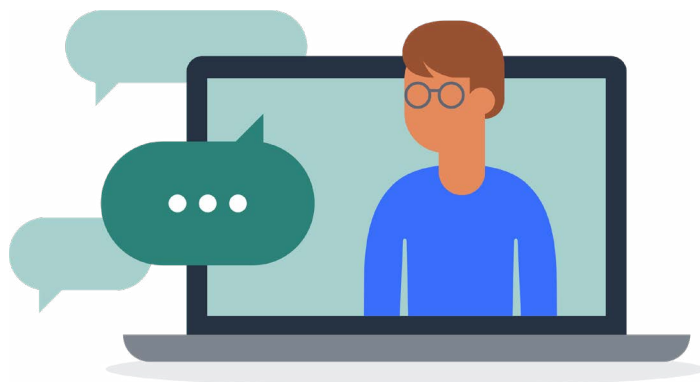


Batch ASR

Product Sheet

Transcribe pre-recorded media files at your convenience with Speechmatics' most powerful, inclusive, and accurate engine ever.



With a selection of deployment options, you can schedule your transcription at a time that suits you.

We've made it simple for you to choose what's right for you. Our on-premises ASR ensures data remains exactly where you want it to. While we also offer deployments in both Public Cloud – which is hosted by Speechmatics – or privately in your own Cloud environment which is hosted by you.

	Docker Container	Virtual Appliance	Cloud Offering
Overview	The software comprises proprietary code, language models, and open-source software. The proprietary code is restricted to the core automatic speech recognition engine.		
Deployment Method	<p>Transcription is provided after an audio file is passed to the Speechmatics engine within the container.</p> <p>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.</p>	<p>Provides transcription of pre-recorded audio in a standalone environment provided directly by a virtual machine.</p> <p>The virtual appliance contains a clean-up policy which automatically deletes the transcripts after a configurable period of time. Default 24-hour period.</p>	<p>Transcription is provided after the audio/video file is submitted to the Speechmatics engine which is deployed within a hosted environment, managed and operated by Speechmatics.</p> <p>The Cloud offering accelerates time to market, while reducing operational complexities and cost.</p>
Description	A fully-inclusive, lightweight, standalone software package. A Docker installation for Linux required.	A preconfigured virtual machine that's simple and easy to configure.	A low-risk, high-reward approach to highly accurate transcription. Deployed within a public cloud environment. Fully managed and operated by Speechmatics.
Languages	You can find all supported languages for both Virtual Appliances and Containers at Speechmatics.com .		
Operating Points	<p>Standard Operating Point:</p> <ul style="list-style-type: none"> Requires Intel Broadwell class architecture minimum. <p>Enhanced Operating Point:</p> <ul style="list-style-type: none"> Requires Intel Cascade Lake class architecture minimum. Recommended using hardware that supports AVX512_VNNI flag to improve transcription processing speed. 		
Supported Operating Environment	<p>Minimum specification: Intel® Xeon® CPU E5-2630 v4 (Sandy Bridge) 2.20GHz (or equivalent). Linux Docker runtime host must be Advanced Vector Extension (AVX) compatible.</p> <p>We support AVX2 compatible hardware to take advantage of the latest performance improvements.</p>	<p>Minimum specification: Intel® Xeon® CPU E5-2630 v4 (Sandy Bridge) 2.20GHz (or equivalent). We support AVX2 compatible hardware to take advantage of the latest performance improvements.</p> <p>Hypervisor: Oracle VirtualBox, VMWare ESXi 6.5 and onward, VMWare Workstation, Amazon Web Services EC2.</p>	Our production environment is hosted in both Western Europe and Western USA. Our trial environment is hosted in Western Europe.
Compute Requirement	<p>An individual Docker image is required for each transcription language.</p> <p>Each running container requires: 1 vCPU, 2-5GB RAM, 100MB hard disk space.</p>	<p>Base config – 2 vCPU, 8GB RAM this will process approximately 2 hours of audio per hour.</p> <p>Additional resources: 1 vCPU, up to 5GB RAM for every additional worker.</p>	N/A
Management Interface	Standard Docker, Kubernetes or other orchestration tools.	HTTPS REST API to manage the appliance including license config, log collection, scale config. For additional information see documentation.	N/A
Speech interface	Console/STDIO	REST API	REST API
Input File Formats	wav, mp3, aac, ogg, mpeg, amr, m4a, mp4, flac – no additional formats are supported.		

	Docker Container	Virtual Appliance			Cloud Offering
Transcription Formats	16 kHz (Broadcast) and 8 kHz (Telephony) acoustic models built in, with automatic selection based on file sample rate.				
Output Format	JSON – Transcripts provided in JSON include metadata such as timing and confidence scores, speaker/channel labels plus more on a per word basis. TXT – TXT output does not include timing information and confidence scores, SRT.				
Resource Requirements	An individual Docker image is required for each transcription language. Each running container requires: 1 vCPU, 2-5GB RAM, 1GB hard disk space.	Pack	Supported Languages	Appliance Size	Supports files up to 1GB in size or 2 hours in length.
		Nano	1	40GB	
		Mini	3	40GB	
		Midi	8	60GB	
		Maxi	15	80GB	
		Plus	19	80GB	
Performance	Transcript can be provided in 2x real-time for files >5 minutes. Parallelization has the potential to have faster turnaround times. 1vCPU per concurrent transcription job. Multiple containers can be executed on the same Docker engine at the same time or across multiple Docker engines to enable large scale operations.	Transcript can be provided in 2x real-time for files >5 minutes for Standard Operating Point. 1vCPU per concurrent transcription job. Additional vCPUs can be added to enable multiple streams to be transcribed at the same time.			Transcript can be provided in 2x real-time for files >5 minutes. The real-time factor is calculated from when the file starts being processed to its completion.
		The real-time factor is calculated from the start of the file being processed to its completion.			
		Using our standard model, a transcript can be provided in 2x real-time for files >5 minutes. Our enhanced model may take longer.			
Data Retention Period	Data is not retained in the Docker Container.	Once the transcription is complete the transcript will be retained by the virtual appliance for 24 hours. The media file is removed immediately after the transcription has completed.			7 days
	After this period, any media and transcription data will be purged unless already done so by the customer using the API.				
Limitations	Can operate on input file sizes of up to 2 hours recorded length or 4GB in size, whichever is reached first.	Can operate on input file sizes up to 2 hours recorded length or 1GB in size, whichever is reached first on POSTS within the API.			
Connectivity Requirements	Can operate within own security boundary allowing you to keep control of your data.	HTTPS port 443 needed to enable access to the Cloud offering. Some IPs may need to be whitelisted for callback notifications.			
Admin	No ongoing maintenance needed for the containers. All administration is provided by direct use of Docker commands.	Full APIs for management and monitoring. Monitoring of the appliance can be done via a web GUI or APIs. Flexible licensing model. Local administration available via APIs.			
Support	Latest release (N) and previous release of the Speechmatics product (N-1).				

Additional Features

Features	Description
Confidence Scores	Visualize the confidence of every word in the transcript.
Speaker Diarization	Detect and label different speakers within the same channel.
Channel Diarization	Detect and label different speakers on up to six streams or channels.
All Major Files Formats Supported	Support for major audio and video formats, with automatic sample rate detection to get you started quickly.
Advanced Punctuation	Use an extensive set of supported punctuation marks to optimize the speed and ease of transcription.
Custom Dictionary and Sounds Feature	Add context-specific words to the dictionary to enhance your transcription accuracy.
Speaker Change	Easily identify a change of speaker within your transcript and improve its readability.

Ready to try Speechmatics?

Sign up for your [free trial](#) and we'll guide you through the implementation of our API. We pride ourselves on offering the best support for your business needs. If you have any questions, just ask.

Contact Us

For any other questions or comments, call or send us an email. Our office is open between 9am-5pm.

Email

hello@speechmatics.com

Phone

+44 (0)1223 794 497

UK office

Cambridge Science Park
Milton Road
Cambridge CB4 0WD
United Kingdom

