

# Speech Analytics Guide

The system offers speech transcription to allow searching in voice conversations.

- [Speech transcription](#)

# Speech transcription

## Overview

Speech transcription automatically creates a searchable, exportable full-text transcript of the audio portion of the recorded conversations. This allows users to identify, investigate and correct compliance and policy breaches. The system supports integration with multiple transcription engines and providers. Refer to the corresponding guides to understand the capabilities, language support, limitations, and license requirements for each.

- [Deploying transcription](#)
- [Configuring and running transcription](#)
- [Verint transcription](#)
- [Intelligent Voice transcription](#)
- [Speechmatics transcription](#)
- [Searching and viewing transcripts](#)
- [Exporting transcripts](#)

## Supported transcription providers

The following table provides a brief overview of the options:

	Verint Speech Transcription	Speechmatics	comitFS Voice Transcription Service	Intelligent Voice
<b>Status</b>	Available	Available	Available	Available
<b>Vendor</b>	Verint Systems <a href="https://www.verint.com/">https://www.verint.com/</a>	Speechmatics <a href="https://www.speechmatics.com/">https://www.speechmatics.com/</a>	comitFS <a href="https://www.comitfs.com/">https://www.comitfs.com/</a>	Intelligent Voice <a href="https://intelligentvoice.com/">https://intelligentvoice.com/</a>
<b>Technology</b>	LVCSR*	LVCSR*	LVCSR*	LVCSR*
<b>Languages</b>	Arabic-BH, Arabic-EG, Arabic (Gulf), Arabic (Levantine), Azeri, Bengali, Cantonese, Catalan, Czech, Danish, Dutch, English-AU, English-GB, English (Hong Kong), English-IN, English-MY, English-ZA, English-SG, English-US, French-CA, French-FR, Georgian, German, Hebrew, Hindi, Indonesian, Italian, Japanese, Kannada, Kazakh, Korean, Malayalam, Mandarin, Polish, Portuguese-BR, Portuguese-PT, Russian, Slovak, Swedish, Spanish-SA, Spanish-SP, Tagalog, Tamil, Telugu, Thai, Turkish, Ukrainian, Urdu	<a href="https://www.speechmatics.com/our-technology">https://www.speechmatics.com/our-technology</a>	North American English, British English, Latin American Spanish	<a href="https://intelligentvoice.com/languages/">https://intelligentvoice.com/languages/</a>
<b>On-premise /Cloud</b>	On-premise	On-premise and cloud	On-premise	On-premise and partner cloud
<b>Deployment</b>	Integrated into Verba Speech Server role	Integration via REST API	Integration via REST API	Integration via REST API

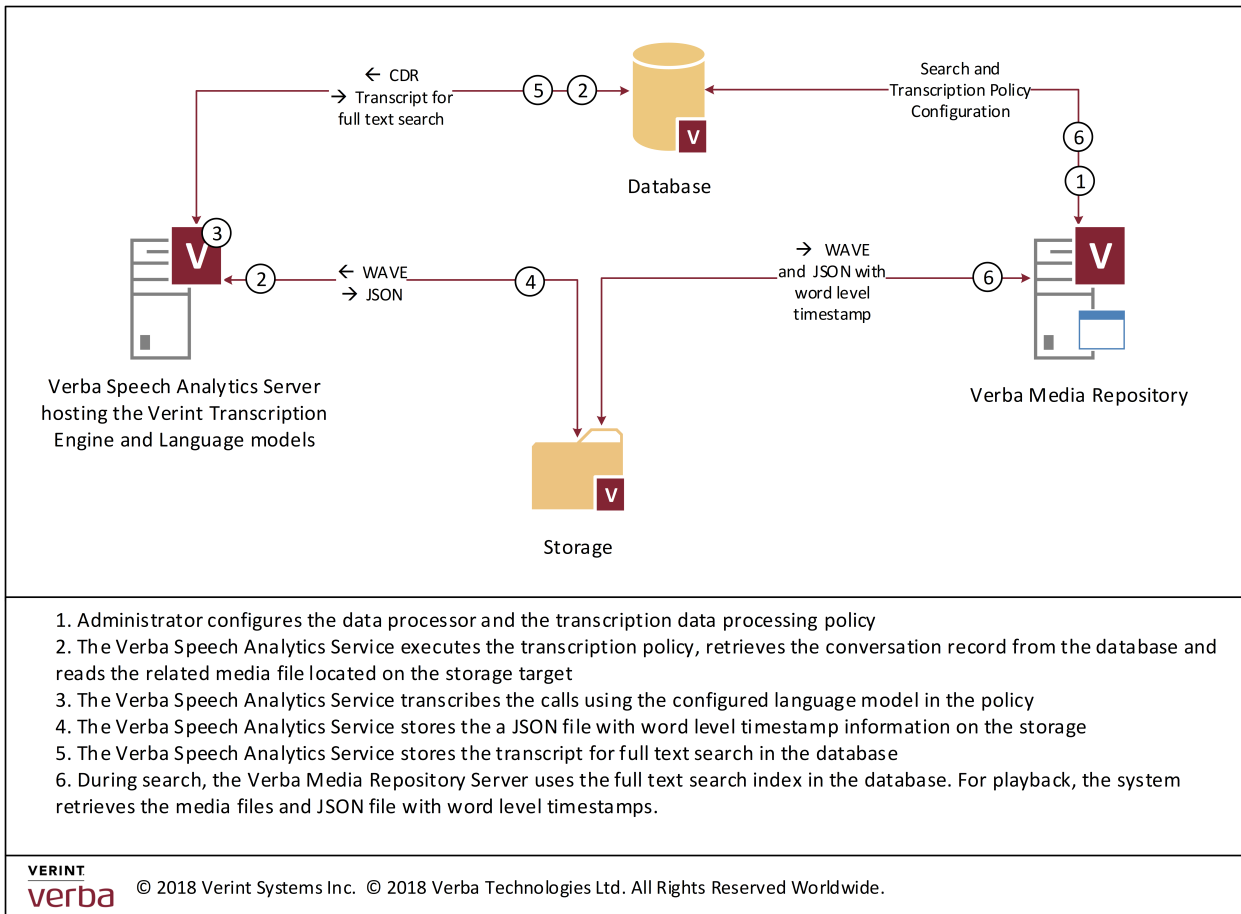
<b>Licensing</b>	Requires VFC Advanced Archive licenses to enable the integration and separate VFC Speech Transcription license to use the Verint Speech Transcription engine. For more information contact your Verint representative.	Requires VFC Advanced Archive licenses to enable the integration and separate Speechmatics license or subscription. For more information contact your Speechmatics representative.	Requires VFC Advanced Archive licenses to enable the integration and separate comitFS license or subscription. For more information contact your comitFS representative.	Requires VFC Advanced Archive licenses to enable the integration and separate Intelligent Voice license or subscription. For more information contact your Verint representative.
<b>Customization</b>	Available through professional service (phonetic boosting or language customization)	Contact Speechmatics for more information	Contact comitFS for more information	Contact Verint for more information
<b>Diarization / Speaker Recognition</b>	Yes	Yes	Yes	Yes
<b>Sentiment Score</b>	No	No	No	Yes
<b>Topics</b>	No	No	No	Yes
<b>Automatic Language Detection</b>	No	No	No	Yes

\* Large Vocabulary Continuous Speech Recognition

## Transcription process

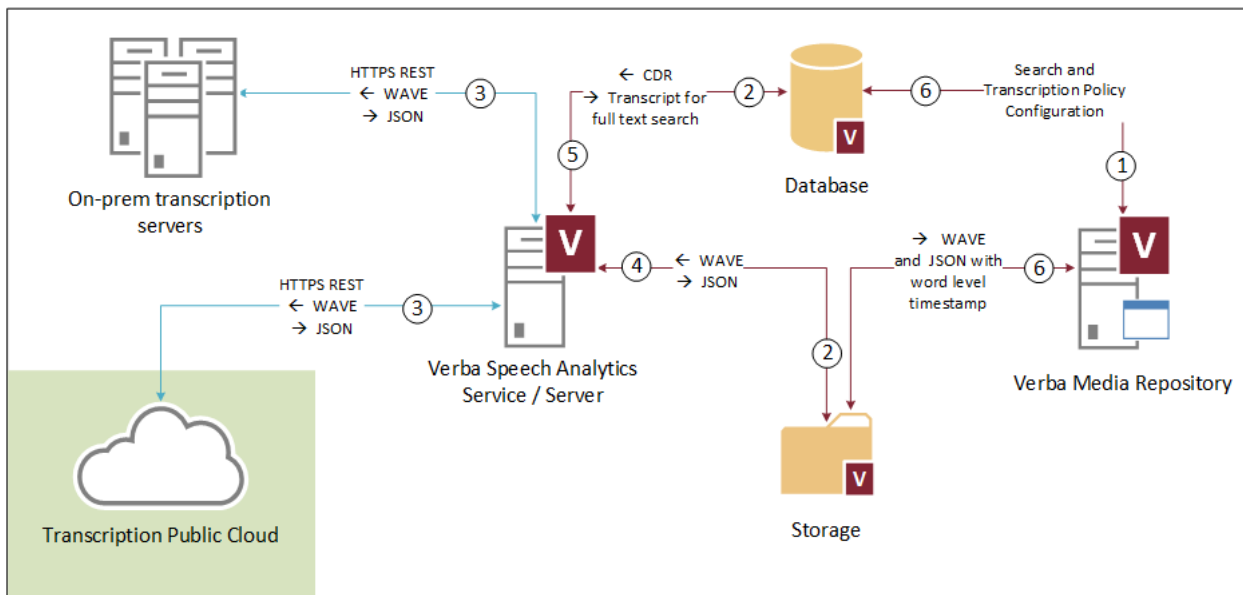
### Verint Speech Transcription

The following figure provides an overview of the transcription process when the Verint Speech Transcription engine is used. Note: in this case, the transcription process runs on the Verba Speech Servers directly.



### 3rd party Speech Transcription

The following figure provides an overview of the transcription process with a 3rd party transcription provider (cloud or on-premise). Note: in this case, the transcription process runs in the cloud or on 3rd party servers.



1. Administrator configures the data processor and the transcription data processing policy
2. The Verba Speech Analytics Service executes the transcription policy, retrieves the conversation record from the database and reads the related media file
3. The Verba Speech Analytics Service sends the media file to the transcription server or cloud. Once the transcription is complete, the transcription service returns the result in JSON
4. The Verba Speech Analytics Service stores the a JSON file with word level timestamp information on the storage
5. The Verba Speech Analytics Service stores the transcript for full text search in the database
6. During search, the Verba Media Repository Server uses the full text search index in the database. For playback, the system retrieves the media files and JSON file with word level timestamps.

# Deploying transcription

- [Server sizing](#)
  - [Server Roles](#)
  - [Verint Speech Transcription](#)
  - [3rd party Speech Transcription](#)
- [Storage requirement](#)
- [Installation](#)
  - [Installing the Verba Speech Analytics Server](#)
  - [Enabling the Speech Analytics Service](#)
- [Configuration](#)

## Server sizing

Server sizing can be different for various speech transcription providers. When using the Verint Speech Transcription, the Verba Speech Analytics Server runs the transcription process and requires additional resources. In the case of 3rd party speech transcription providers, the transcription process runs in the cloud or on on-premise 3rd party servers. The Verba Speech Analytics Service has the following jobs (with impact):

- Execute the configured Speech Transcription policies which query the database to create a list of call records for the transcription process (low)
- Download the audio files from the storage target (low)
- Optionally transcode the audio to the format supported by the transcription provider (medium)
- Run the transcription directly on the server in the case of Verint Speech Transcription (high) or in the cloud or on 3rd party server (no impact)
- Insert plain text transcription to the database (low)
- Upload transcript file to the storage target (low)

## Server Roles

The Verba Speech Analytics Service can be enabled on the following server roles:

- Speech Analytics Server
- Media Repository Server

Do not enable the service on any other server role. If the service is enabled on the Media Repository Server, the service has to be configured to limit the number of simultaneous transcription processes to 1 (Speech Analytics / Transcription / Task Processing Threads Count), this will use 1 core for transcription.

The service may be enabled on multiple servers, in this case, the system will automatically handle the load balancing between servers.

## Verint Speech Transcription

When the Verint Speech Transcription engine is used, stand-alone Verba Speech Analytics Servers have to be deployed. The speech transcription process cannot be running on other server roles. This is due to the extensive resource usage nature of the speech transcription service which might interfere with other processes on the servers.

This table outlines typical server sizing and recommended hardware and software configurations:

Server Role	Speech Analytics Server
Server Platform	Industry standard PC servers Physical or virtual

<b>CPU</b>	Intel Xeon 2.4 GHz or higher  Up to 16 CPU cores or vCPUs  2 cores/vCPUs must be "reserved" for the OS, only the rest can be used for application sizing  Numbers only applicable when Receive-side scaling (RSS) is enabled in the OS
	1 vCPU for every 200 hours of calls transcribed in 24 hours  Individual results may vary due to the talk time in recordings, the storage codec, and the language model used
<b>Memory</b>	4GB RAM + 2GB / language model
<b>Hard disk</b>	System partition (OS and applications): 80 GB or more  Media partition: not required, media files are only temporarily stored on the server  Always use redundant disks with RAID and have separate Media and System volumes
<b>Network</b>	All server clocks must be synchronized, typically either with the domain controller or time server  All servers must have the latest Time Zone configuration  Custom time zones are not supported
	1x Gigabit Ethernet
<b>Operating system</b>	Microsoft Windows Server 2012 R2 Microsoft Windows Server 2016 Microsoft Windows Server 2019  Standard Edition, latest service packs installed We support the English versions of Microsoft server software
<b>Virtualization</b>	VMware and Hyper-V are recommended, see <a href="#">Virtualization</a> .
<b>Antivirus</b>	Make sure your <a href="#">Antivirus software does not scan database, media and log folders</a> .
<b>Power</b>	Redundant Power Supply UPS recommended

### 3rd party Speech Transcription

In the case of 3rd party speech transcription providers, the Verba Speech Analytics Service initiates the transcription process, which runs in the cloud or on on-premise 3rd party servers. For that reason, the Verba Speech Analytics Service can be run on dedicated Verba Speech Analytics Servers or on servers with Media Repository / Application Server roles.

For server sizing, use the standard Media Repository / Application Server guidelines at [Server sizing and requirements](#).

## Storage requirement

**AVAILABLE IN 9.7.5 AND ABOVE**

Speech transcription is supported on any storage target, except EMC Centera and Hitachi Content Platform (which do not allow uploading individual files after setting the retention on the object). In the case of WORM storage, the system automatically applies the retention period on the uploaded transcription files to match the original retention of the conversation.

Transcription policies only run on conversations stored on storage targets, which means files must be uploaded to the storage first.

Before version 9.7.5 the transcription only works when the media is stored on SMB storage. If the media has to be stored on a WORM or non-SMB storage, the files should be hosted temporarily on SMB and moved to the final storage target after transcription.

## Installation

### Installing the Verba Speech Analytics Server

For installing the server, refer to [Installing a Verba Speech Analytics Server](#).

### Enabling the Speech Analytics Service

**Step 1** - On the Verba Web Interface, go to **System / Servers**. Select your server and click on the **Service Activation** tab.

**Step 2** - Activate the **Verba Speech Analytics Service** by clicking on the



icon.

**Step 3** - Click on the **Service Control** tab.

**Step 4** - Start the **Verba Speech Analytics Service** by clicking on the



icon.

## Configuration

Refer to [Configuring and running transcription](#) in order to configure speech transcription. You can find additional, integration specific information at:

- [Verint transcription](#)
- [Intelligent Voice transcription](#)
- [Speechmatics transcription](#)



# Configuring and running transcription

After deploying the servers and enabling the Speech Analytics Service, the system can be configured to transcribe voice calls. The configuration consists of the following steps:

- Creating a speech transcription data processor which enables the specific transcription integration in the system. The system supports multiple transcription engines. Multiple data processors can be created and enabled.
- Assigning transcription licenses to users to allow transcribing the calls of the users. Different transcription integrations may require different licenses.
- Creating speech transcription data management policies to define which calls will be transcribed by which processor and which ASR model.

## Creating a Data Processor

A data processor represents the connection with the transcription engine. In order to set up a processor, follow the steps below.

**Step 1** - In the Verba Web Interface go to **Data / Data Processors**

**Step 2** - In the top right corner click on the **Add new Data Processor** link

**Step 3** - Enter the name of the data processor

**Step 4** - Select **Speech Transcription** for **Type**

**Step 5** - Select the **Engine**

**Step 6** - Follow the guidelines specific to the transcription integration to configure the connection parameters:

- [Verint transcription](#)
- [Intelligent Voice transcription](#)
- [Speechmatics transcription](#)

**Step 7** - Click **Save** to save your data processor

## Assigning transcription user licenses

Transcription is licensed on a per-user basis. You can assign these licenses to roles (and the roles to users) by enabling one of the **Speech Analytics / Transcribe Conversations** checkboxes on the role configuration page. The transcription policy only selects calls that belong to users where there is transaction permission enabled, otherwise the call will not be transcribed. Certain integrations have specific license requirements or additional limitations. For more information on these, refer to the integration specific page.

## Creating a Transcription Policy

A Data Management policy configured with the Speech Transcription action is used for specifying the recorded conversations to be transcribed. Servers running the Verba Speech Analytics Service read the media file from the storage, then store the recorded conversations on their local disk temporarily while transcribing. After the transcript is created, the media files will be discarded, and the transcript files are relocated to the media's storage location. Each policy you define can be set up for different ASR models and a set of filters to determine which conversations should be transcribed. These filters are based on the metadata stored in the database for each conversation.

This is a good way to account for differences in languages used for conversations of different locations, groups, etc.

**Step 1** - Follow the generic policy creation steps described on the following page: [Data management policies](#)

**Step 2** - Select **Speech Transcription** as the action

**Step 3** - Select the **Data Processor**

**Step 4** - Select the **Language** for the transcription

**Step 5** - Select your filters to specify which conversations should be checked by this policy

**Step 6** - Click on **Save**. The policy will run periodically,

ID*	<input type="text" value="10"/>
Name*	<input type="text" value="Corp Transcription NY site"/>
Enabled*	<input type="text" value="Yes"/>
Priority*	<input type="text" value="30"/> <small>Higher priority policies are processed first when the 'older than' dates are equal.</small>
Action*	<input type="text" value="Speech Transcription"/>
Data Processor*	<input type="text" value="Transcription - Verint (Speech Transcription / Eliza) -"/>
Language	<input type="text" value="English (USA)"/>

## Running and monitoring the transcription process

The transcription actions are executed by the Verba Speech Analytics Service on the servers where this service is enabled.

The transcription process can be monitored on the Dashboard with a [Background Tasks Widget](#) or in the list of background tasks in the Verba Web Interface at **System / Background Tasks**.

The [Log files](#) for the service can be found at (your application path might be different):

```
C:\Program Files\Verba\log\speech-analytics.log
```

# Verint transcription

## Deploying and configuring Verint transcription

Verint transcription is using the Verint engine to transcribe voice calls. The engine is built into the Verba Speech Analytics Service which can be enabled on the Verba Speech Analytics Server role. The Speech Analytics Server requires special sizing when the Verint transcription engine is used, for more information see [Deploying transcription](#).

Once, the server(s) are deployed, the required Data Processor has to be created to enable the integration with the Verint transcription engine. Follow the steps described in [Configuring and running transcription](#) to create the processor and select the Verint engine. No additional settings are required for the data processor configuration.

After creating the data processor, you can follow the guidelines at [Configuring and running transcription](#) to configure one or more data management policies.

## License requirements

The Verint transcription engine requires specific licenses to be purchased and uploaded to the system. Please contact your Verint sales representative for more information.

After uploading the necessary licenses, the licenses have to be assigned to users through the role configuration. The following permission(s) are required for the Verint speech transcription:

- Transcribe Conversations (Built-in)

## Uploading new ASR models

Follow the instructions below to upload an ASR model (often referred to as language models) to the system.

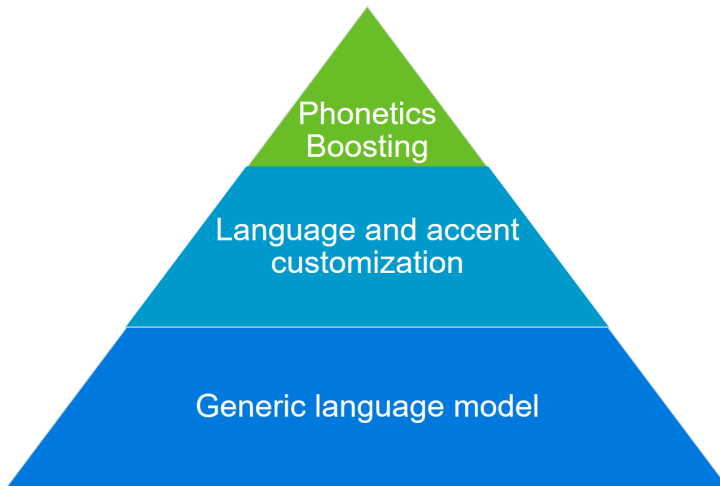
Copy the language model file, which is a .zip file, to `C:\Program Files\Verba\resources\transcription\eliza` to every server where the Verba Speech Analytics Service is enabled. For additional languages, only this step needs to be repeated. The service will detect the new language model within 30 minutes, or you can force the detection by restarting the service. After the service detects the new model, it will be listed on the speech transcription data management policy configuration page.

## Customizing the Verint ASR models

Please read the [Improving transcription accuracy for Verint Speech Transcription](#) for more information.

# Improving transcription accuracy for Verint Speech Transcription

This article is a guide on the possible ways of improving the recognition rate of the transcriptions when using the Verint Speech Transcription option. The speech transcription engine uses a language model that describes the spoken language acoustically and linguistically. There are two options for improving the generic language model, language and accent customization (LAC) and phonetic boosting.

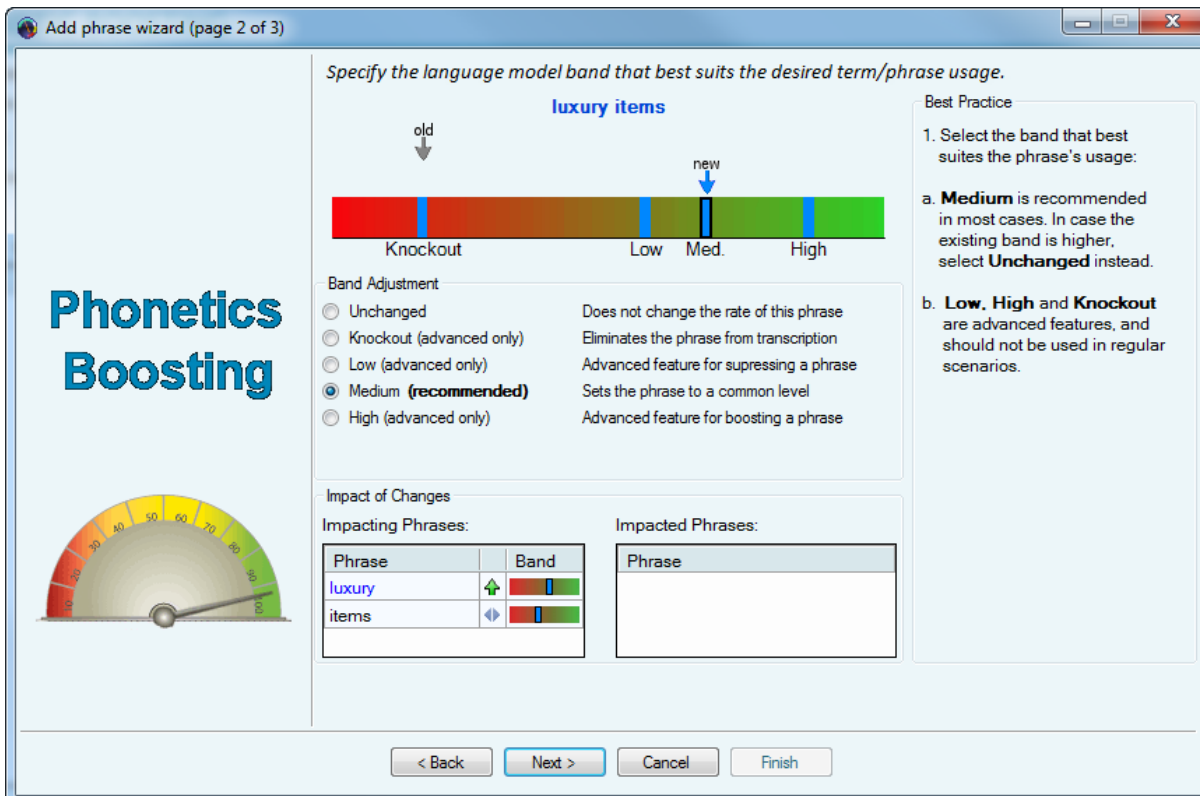


## Language and accent customization (LAC)

The language and accent customization is a professional service that focuses on the acoustic characteristics of the language model. It takes into consideration specific environmental and spoken elements, such as background noises, organization scripts or language patterns. For this process, it is necessary to collect audio.

Please contact us for further details.

## Phonetic boosting



Phonetic boosting is a fine tuning customization option for the linguistic characteristics. It helps to increase the recognition rate if the transcription system needs improvements with company-specific words. The following actions can be done:

- Expanding the language model vocabulary  
Add terms and phrases that are unique to your business or to an ongoing campaign to boost their recognition by the Speech Engine.
- Boosting recognition of existing terms and phrases  
The most common cause for poor recognition of a term or phrase is the fact that its probability is higher in your field. Changing the weight of an existing term or phrase increases its probability, changing its impact on the Speech Analytics results.
- Suppressing a competing phrase  
An incorrect transcription of a term or phrase causes it to compete with the “correct” transcription. With Phonetics Boosting, you can suppress the incorrect phrase by decreasing its band probability.
- Supplementing pronunciations  
If a term has several pronunciations, adding all the pronunciation variants boosts recognition

Please contact the support service for assistance in phonetic boosting.

# Intelligent Voice transcription

AVAILABLE IN 9.8 AND ABOVE

- [Overview](#)
- [License requirements](#)
- [Deploying and configuring Intelligent Voice transcription](#)
  - [Failover and load balancing](#)
  - [Failure scenarios](#)
  - [Performance requirements](#)
  - [Data processor](#)
  - [Transcription policy](#)
  - [Environment configuration](#)
- [Managing ASR models](#)
- [Detailed User Guides](#)
- [Limitations](#)

## Overview

The proliferation of voice capture across, Turrets, Unified Communication, Mobile and more, means that IT and Compliance users have more data to manage and less time to do so. To meet these challenges, Verint Financial Compliance's Compliance Application functionality has been enhanced with Intelligent Voice's speech capabilities to automate and simplify Conversation Search.

The screenshot displays the Verint Financial Compliance interface. At the top, there is a navigation bar with tabs for Conversations, Quality Management, Workflows, Communication Policies, Reports, Users, Data, and System. The main area is titled 'Conversations' and shows a search results table with columns for Start Date, Start Time, Duration, User, From, From Info, To, To Info, Language Switched, Languages, Sentiment Gradient, TCU Sentiment, and Topics. Below the table, there is a 'Conversation View' section with a waveform and a transcription view. The transcription view shows a dialogue between Speaker 1 and Speaker 2, with Speaker 1 mentioning 'Jane?' and 'realised Gaynor \$5000, and then again at 7500, on the games'.

The following functionality is available within Verint Financial Compliance:

- Transcription of audio for voice and video recordings
- Diarization: the ability to separate/identify speakers
- Punctuation and capitalization: the ability to identify the beginning and the end of sentences
- Conversation Search: Includes the ability to search within Transcribed Text, including keyword and phrase searches.
- Language Search: Includes the ability to search for Conversations in a specific (supported and configured language) and also conversations where a Language Switch took place during the conversation.
- Sentiment Search: Includes the ability to search for Sentiment Gradient swings, whereby sentiment changes (for example from positive to negative) during a conversation. In addition, users can search for high occurrences of negative or positive sentiment in a given conversation.
- Topic Search: Includes the ability to search for conversations that include IV derived Key Topics.

- Search Results: Search grid includes; Topics, Sentiment Information, Language Switches and Spoken Languages
- Conversation View: Includes a new tab for Analytics. This will present Topics, Sentiment Information, Language Switches and Spoken Languages to the user. In addition, the user can interact with the Topics and use this to navigate and jump to the point in the conversation relative to the topic.
- Support for multiple languages, see <https://intelligentvoice.com/languages/>
- Adaptation/customization of automatic speech recognition (ASR) models (Model adaptation is the process of taking an existing ASR model and adapting it to suit a specific use case, by incorporating new words and new patterns of speech. Speech recognition models reflect the patterns of speech in the training dataset they were built with. The general ASR models distributed by Intelligent Voice reflect the patterns of speech in the general population of the region. Improved results can be obtained by tailoring the model to reflect the speech in a given domain.)
- Integrated via REST APIs
- Intelligent Voice solution is deployed separately on-premise or in Verint partner cloud

## License requirements

The Intelligent Voice transcription engine requires specific licenses. Please contact your Verint sales representative for more information.

After uploading the necessary licenses, the licenses have to be assigned to users through the role configuration. The following permission (s) are required for the Intelligent Voice speech transcription:

Role Permission	Required License	Speaker Diarization	Export	Sentiment Analysis
Transcription (Profiling Speech)	Communications Profiling Speech	Yes	No	No
Transcription (Profiling Speech Advanced)	Communications Profiling Speech - Advanced	Yes	Yes	Yes
Transcription (Risk Profiling)	Communications Risk Profiling Speech	Yes	Yes	Yes

## Deploying and configuring Intelligent Voice transcription

The Intelligent Voice transcription is considered a 3rd party transcription engine, which requires using a separate on-premise or cloud based Intelligent Voice infrastructure to run the transcription service. The Verba Speech Analytics Service is connecting to the Intelligent Voice platform and sends audio files for transcription. For more information see [Deploying transcription](#).

### Failover and load balancing

Multiple servers can process the transcription policies simultaneously. If 1500 records have already been sent to the Intelligent Voice engine, then no new records will be selected for the policy (the records already selected will be sent though).

### Failure scenarios

- Database query or update problem: the Speech Analytics Service retries automatically
- Communication problem with the Intelligent Voice service: the Speech Analytics Service automatically retries
- Pending tasks can be monitored in the speech\_pending database table
- Getting stuck
  - The conversation is selected for sending by the service but it is never sent: Such entries are deleted from the speech\_pending table in one hour so the record will be selected for sending again
  - The conversation is selected for receiving the results by the service but it is never finished: Such entries are updated in the speech\_pending table pending so downloading will be retried by one of the Media Repositories again
  - Pending entries in the Verba database are cross-checked with the entries in the Intelligent Voice side in every 4 hours by the Speech Analytics Service

- Intelligent Voice engine is not properly configured or missing components: consult with IV about the installation and licenses.

## Performance requirements

Considering performance requirements, the most important factor is the storage and the network should be capable of reading and sending the expected amount of audio files to the IV cluster. As all analysis happens on the IV servers, CPU and memory are only used for lightweight database querying and parsing, and storing the results from the IV system.

## Data processor

Once, the Intelligent Voice platform is available, the required Data Processor has to be created to enable the integration with the Intelligent Voice transcription engine. Follow the steps described in [Configuring and running transcription](#) to create the processor and select the Intelligent Voice engine. The following table describes the settings available for an Intelligent Voice data processor:

Configuration item	Description
Name	Name of the data processor. This name will identify this processor across the system.
Type	Select Speech Transcription
Engine	Select Intelligent Voice
API Root URL	URL of the Intelligent Voice API (on-prem server or cloud)
API User	API user name, the group ID defined in Intelligent Voice
API Token	API token
Enable Speaker Diarization	Allows separating participants in conversations and producing a dialog like output

## Transcription policy

After creating the data processor, you can follow the guidelines at [Configuring and running transcription](#) to configure one or more data management policies.

When an Intelligent Voice Data Processor is selected, then the ASR Models must be defined:

Data Processor\* Intelligent Voice (Speech Transcription / Intelligent Voice) - API Root URL: <https://x1.intelligentvoice>

**Available ASR Models**

Intelligent Voice(#1)/4 pt-BR (IntelligentVoice\_pt-BR\_8kHz\_65000\_general\_V1\_ASRv5)

Intelligent Voice(#1)/5 nl-001 (IntelligentVoice\_nl-001\_16kHz\_170000\_general\_V1\_NASRv2)

Intelligent Voice(#1)/6 en-001 (IntelligentVoice\_en-001\_16kHz\_1024\_general\_V0\_NASRv3.1)

Intelligent Voice(#1)/7 pl (IntelligentVoice\_pl\_16kHz\_170000\_general\_V1\_NASRv2)

Intelligent Voice(#1)/8 es-001 (IntelligentVoice\_es-001\_16kHz\_512\_general\_V1\_NASRv3.1)

Intelligent Voice(#1)/9 de-DE (IntelligentVoice\_de-DE\_16kHz\_1024\_general\_V4\_NASRv3.1)

**ASR Models**

Add
Remove

Intelligent Voice(#1)/2 en-001 (IntelligentVoice\_en-001\_8kHz\_145000\_general\_V2.1\_ASRv6)

Intelligent Voice(#1)/3 en-001 (IntelligentVoice\_en-001\_16kHz\_200000\_general\_V1\_NASRv2)

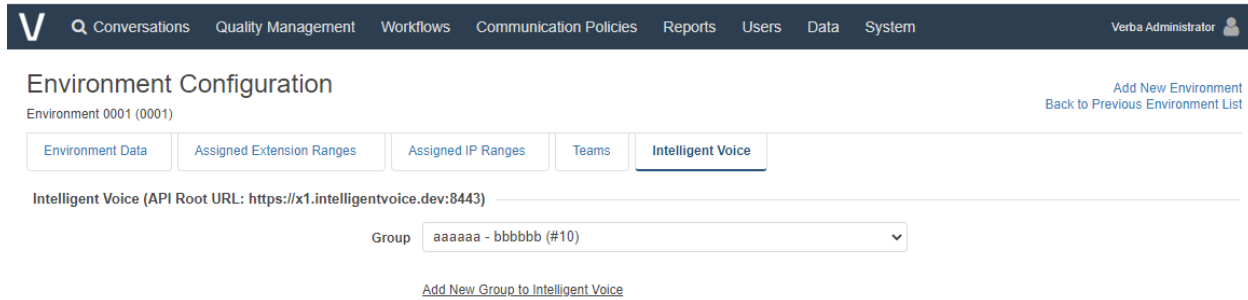
The maximum number of selected models is 4, because Intelligent Voice supports up to 3 languages for language detection, and the 4th item should be the special ASR Model that signals the language detection requirement.



Only the Transcription (Risk Profiling) license supports language detection, in the case of Transcription (Profiling Speech) and Transcription (Profiling Speech Advanced), the Users' Roles define which ASR Model will be available for the given user. In that case, the Data Management Policy must be set up to include the Users' ASR Model, otherwise, the Users' calls will not be processed by the policy.

## Environment configuration

The Intelligent Voice Group ID should be set up for each tenant before using the integration:



The screenshot shows the Verba Administrator interface. At the top is a navigation bar with a search icon and menu items: Conversations, Quality Management, Workflows, Communication Policies, Reports, Users, Data, System. The user is logged in as Verba Administrator. The main content area is titled "Environment Configuration" for "Environment 0001 (0001)". There are two links: "Add New Environment" and "Back to Previous Environment List". Below the title are five tabs: Environment Data, Assigned Extension Ranges, Assigned IP Ranges, Teams, and Intelligent Voice. The Intelligent Voice tab is active. It shows the "Intelligent Voice (API Root URL: https://x1.intelligentvoice.dev:8443)" and a "Group" dropdown menu with the value "aaaaaa - bbbbbb (#10)". There is a link "Add New Group to Intelligent Voice" below the dropdown.

## Managing ASR models

[ASR model management](#)

## Detailed User Guides

[Using analytics search with Intelligent Voice](#)

[Using transcription and analytics in player](#)

## Limitations

The Intelligent Voice integration has the following limitations:

- The single conversation can only be transcribed once. Once it is transcribed there is no way to transcribe it again.
- Automatic language detection for multi-language calls supports up to 3 languages for a call and requires the configuration of each model upfront. It is not able to detect a language that is not configured in the transcription policy.
- Records migrated from WFO cannot be transcribed

# Using analytics search with Intelligent Voice

## Overview

Organizations use Verint Financial Compliance's Analytics Search capabilities to reduce the time effort required to find potential compliance issues within Conversations. This guide will describe how to:

- Search within Transcripts
- Search by Language
- Search by Sentiment
- Search by Topics
- Understand the available columns within the Conversation Search results grid

## Search within transcripts

This functionality empowers the user to search within transcribed recordings.

The search fields support '\*' for a **starts-with or ends-with match**, or "" for an **exact match**.

Please note, if IV is configured to transcribe multiple languages, then the user can type phrases in the supported language set.

Search Field Name	Description
All of these phrases	Will search across transcribed conversations for transcripts that contain all of the listed phrases in one transcript.
Any of these phrases	Will search across transcribed conversations for transcripts that contain any of the listed phrases in one transcript.
None of these phrases	Will search across transcribed conversations for transcripts that contain none of the listed phrases in one transcript.

▼ Text Search and Analytics (\*) ?

Search in

Instant Messaging

Transcript

SMS

Marker

All of these phrases

compliance x

Any of these phrases

Enter phrase...

None of these phrases

Enter phrase...

[Create labeling rule from these](#)

## Language search

This functionality empowers the user to search for use of a specific language within a recorded and transcribed conversation. The functionality also empowers the user to search for conversations where the recorded user, switched from one language to another.



To use Language Search, the environment must be configured to transcribe the target language. [Please refer to the ASR model management](#) for more details.


Search Field Name	Description
Language switched during conversation	Will search across transcribed conversations where a user has switched to a second language during the conversation.
Language Spoken	Will search across transcribed conversations for transcripts that contain use of a recognized language. Please note, multiple instances of the language spoken search field can be added.

**Languages**

Language switched during conversation

**Language Spoken**

 --Choose-- 



## Search within sentiment


This functionality empowers the user to search by sentiment within transcribed recordings.



Prior to explaining the search capabilities, it is essential to understand the granular sentiment analysis capabilities of the system. Sentiment analysis is conducted at a fine-grained level, each diarized, speaker separated transcribed text is analyzed and given a sentiment score, within the system this is referred to as a TCU instance. As sentiment is analyzed and stored at such a granular level, richer sentiment analytics and search capabilities are enabled.

Search Field Name	Description
Sentiment Gradient	An additional filter is available to refine the sentiment shift. This enables the user to define a search filter (more than or less than) shift by Numeric Value, where the numeric value (-99 to +99) represents the sentiment change
Sentiment Gradient	Enables the user to search for conversations where the sentiment has shifted in different directions. For example: <ul style="list-style-type: none"> <li>• Conversations where there has been a significant sentiment shift from negative to positive sentiment.</li> <li>• Conversations where there has been a significant sentiment shift from positive to negative sentiment.</li> <li>• Conversations where there has been a sentiment shift from negative to positive sentiment.</li> <li>• Conversations where there has been a sentiment shift from positive tonegativesentiment.</li> </ul>
TCU Instances	Enables the user to search for conversations where there has been a large volume of negative or positive TCU sentiments in a given conversation. The user is able to refine the search to focus on 'Very Negative', or 'Negative', or 'Positive', or 'Verint Positive' instances of sentiment TCU. The user can set the numeric range.


**Sentiment**

**Sentiment Gradient**

--Select-- 

>   

**TCU Instances**

Negative 

Between  -

## Search within topics

This functionality empowers the user to search by Topic.

Prior to explaining the search capabilities, it is essential to understand Topics. Topics are generated for each conversation, topics represent linguistically significant text within a given conversation. The system automatically identifies key Topics spoken within a conversation, providing a quick summary of the audio file content.

Search Field Name	Description
Topics	Enables the user to search for conversations that contain specific Topics.

Topics

Account ×

Options ×

Position ×





Reset Search
Cannot find a conversation?
🔍 Search

## Search results

The Conversation Search includes columns for the enriched speech analytics insights.

The following columns can be added to the search results.

Search Result Column	Description
Language Switched	Highlights if a language switch was detected during the conversation.
Languages	Lists the detected languages spoken during the conversation.
Topics	Lists key topics spoken within a conversation, providing a quick summary of the audio file content.
Sentiment Gradient	<p>Provides a graphical display of the sentiment shift during the conversation. The angle and direction of the line reflects the sentiment direction and gradient.</p> <p>The sentiment gradient maps the changing sentiment (tone or mood) across the entirety of an interaction by calculating the change in the sentiment between the beginning and the end of the conversation for each speaker. Note that a positive sentiment gradient does not necessarily mean that the outcome of the conversation was positive, it only reflects the direction of the change in the sentiment across the call.</p> <p>The Verba system displays the sentiment gradient on the search result screen as follows:</p>

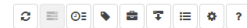
- The gradient scores are provided separately for each identified speaker by the Intelligent Voice solution. The Verba system only displays the best and the worst sentiment gradients for the conversation. Note that this means if there are more than 2 speakers in the conversation, the system will not display all sentiment gradients.
- The sentiment gradient is presented on a scale of 4 categories and displayed as a small line chart:
  - Sentiment has changed very negatively 
  - Sentiment has changed negatively 
  - Sentiment has changed positively 
  - Sentiment has changed very positively 

**TCU Sentiment**

Lists the granular count of TCU sentiment within the conversation.

The numbers with the colored background represent the number of TCUs in the respective sentiment category for the recorded conversation.

**Conversations**



3 items found, displaying all items. Results per page 20

	Start Date	Start Time	Duration	Topics	User	From	To	Language Switched	Languages	Sentiment Gradient	TCU Sentiment
	Nov 12, 2022	1:14:57 PM	00:01:01	Records, Addition, Options, Account, Reverse, ... 29 more	Béla	1197	1082	No	en-001		<span style="background-color: red; color: white;">8</span> <span style="background-color: orange; color: white;">0</span> <span style="background-color: yellow; color: black;">15</span> <span style="background-color: green; color: white;">1</span> <span style="background-color: blue; color: white;">15</span>
	Nov 12, 2022	11:19:00 AM	00:00:21	Account, Clients, Auction, Eventually, Nineteen, ... 8 more	Béla	1121	1131	No	en-001		<span style="background-color: red; color: white;">9</span> <span style="background-color: orange; color: white;">0</span> <span style="background-color: yellow; color: black;">9</span> <span style="background-color: green; color: white;">1</span> <span style="background-color: blue; color: white;">3</span>
	Nov 12, 2022	10:00:17 AM	00:00:23	Account, History, Position, Context, Failure, ... 15 more	Béla	1179	1200	No	en-001		<span style="background-color: red; color: white;">4</span> <span style="background-color: orange; color: white;">1</span> <span style="background-color: yellow; color: black;">9</span> <span style="background-color: green; color: white;">0</span> <span style="background-color: blue; color: white;">4</span>

3 items found, displaying all items. Results per page 20

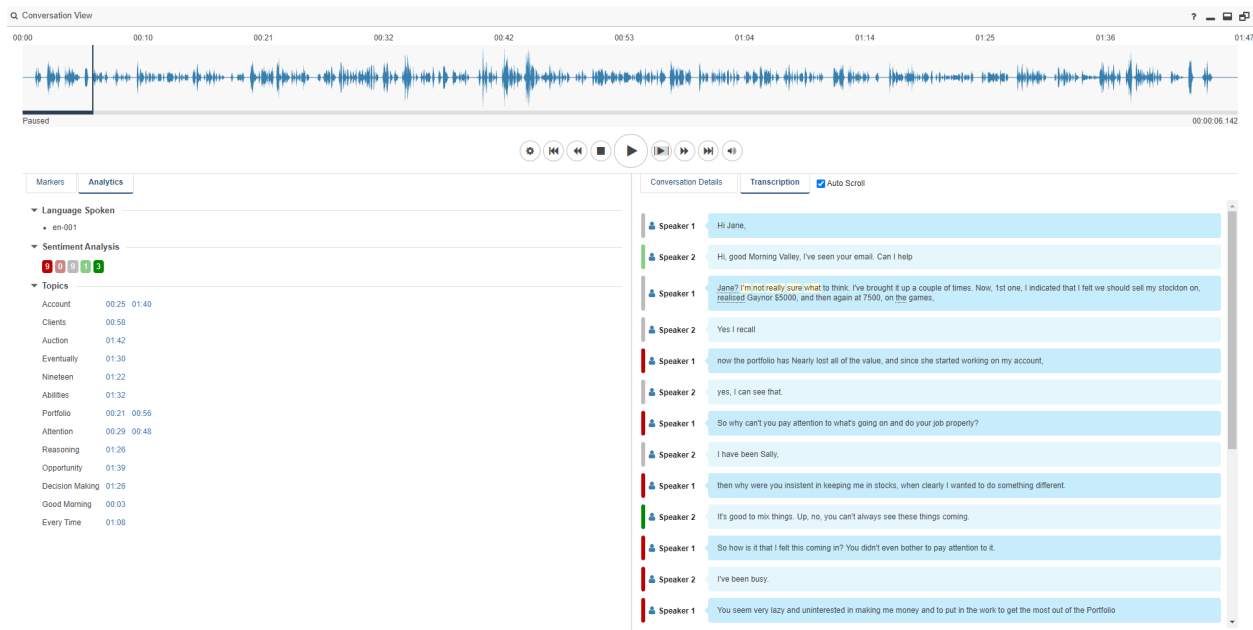
# Using transcription and analytics in player

## Overview

Organizations use Verint Financial Compliance's Analytics Search capabilities to reduce the time and effort required to find potential compliance issues within Conversations. This guide will describe how to interact with the:

- Speech Analytics Tab
- Speech Transcription Tab

Additional information on the media player guide can be found here: [Using the media player](#)



## Speech analytics tab

A new Speech Analytics Tab is available for the in focus conversation this includes the following information

Expandable Information	Description
Language Spoken	Lists the detected languages spoken during the conversation.
Sentiment Analysis	Lists the granular count of TCU sentiment within the conversation.  The numbers with the colored background represent the number of TCUs in the respective sentiment category for the recorded conversation.
Topics	Lists key topics spoken within a conversation, providing a quick summary of the audio file content.  This also includes a time stamp, highlighting at which point during the topic was detected. The timestamp is a hyperlink and will move the player to the appropriate point in the Wav File.

# Transcription tab

The transcription tab provides the following functionality

Feature	Description
Auto Scroll	The transcription will scroll, in synch with the audio and audio player.
Speaker Diarization	Identifies the number of speakers in an audio file and segments the transcript according to speaker identity. Speaker 1, Speaker 2 etc...
TCU Sentiment	Each diarized segment of transcribed text is given a color coding to infer the sentiment score for the transcribed segment.
Confidence	Words with a dashed underline are words with a low confidence score.
Sentiment	<p>The color bars on the left present the sentiment score of the TCU on a scale of 5 categories:</p> <ul style="list-style-type: none"><li>• Red: very negative sentiment</li><li>• Light red: negative sentiment</li><li>• Gray: neutral sentiment</li><li>• Light green: positive sentiment</li><li>• Green: very positive sentiment</li></ul> <p>By moving the cursor over the color bar, the exact sentiment score value is displayed.</p>

# ASR model management

Intelligent Voice (IV) provides the following features for ASR models:

- IV supports automatic language detection and the transcription of multi-language calls. When transcribing a call, language detection requires the definition of up to 3 languages (plus the language detection model).
- IV allows easy customization of ASR models (the lexical model), enabling customers to create new models (the process is called model adaptation, for more information, see <https://support.intelligentvoice.com/hc/en-us/articles/360044447274-Model-adaptation>). Both partners and customers can adopt models. In a multi-tenant configuration, models adapted for specific tenants are considered private and not accessible to any other tenant. Models adapted by partners can be potentially shared with multiple customers.
- IV license articles limit the number of languages used to transcribe calls for individual users. A basic license only allows to use of a single language per user (different users can have different languages configured), other licenses allow the usage of multiple languages per user.

## ASR model administration

The ASR Models are administered on the Intelligent Voice portal. VFC synchronizes the data between the two systems, and the Reference Environment Administrators can enable the ASR Models to be used by specific tenants.

Tenant Administrators are only allowed to list the ASR Models, they cannot change anything.

## Finding and listing ASR models

The synchronized ASR models can be listed under **Data / Data Management / ASR Models**. The list provides the following information about the available ASR models:

Name	Description	Sample Value
Full Name	The name of the ASR model defined in the IV system.	IntelligentVoice_en-001_8kHz_94000_general_V5.1_ASRv6
Data Processor	The name of the data processor defined in the Verba system which was used to synchronize in the ASR model.	-
Model ID	The ID of the ASR model in the IV system.	1
Created By	The name of the creator of the ASR model as defined in the IV system.	IntelligentVoice
Language Code	The language code associated with the ASR model as defined in the IV system.	en-001
Lexicon Size	The size of the lexicon in the ASR model.	94000
Description	The description for the ASR model as defined in the IV system.	general
Version	The version of the ASR model as defined in the IV system.	V5.1_ASRv6

## Associating ASR models with environments

By default, the ASR models are only available in the reference environment in a multi tenant configuration. In a single tenant system, the ASR models are automatically available. In order to associate an ASR model with one or more tenants, follow the steps below:



**Step 1** - Navigate to **Data / Data Management / ASR Models** and select the ASR model.

**Step 2** - Under **Choose Environments**, select the environments/tenants and click on the **>>** button to select the environment(s) to be associated with the ASR model. If you want to make the ASR model available for all existing and future environments, click the **Visible for All Tenants** checkbox.

**Step 3** - Press the **Save** button to save the new configuration.

## Activating and deactivating ASR models

If you decide to not use an ASR model temporarily, you can deactivate the ASR model. When deactivating an ASR model, the system checks if the ASR model is currently configured for any of the transcription policies, and it doesn't allow deactivating until it is configured for at least one policy. Deactivated ASR models are also automatically deactivated in the IV system. You can activate deactivated ASR models again. In order to activate or deactivate an ASR model, follow the steps below:

**Step 1** - Navigate to **Data / Data Management / ASR Models** and select the ASR model.

**Step 2** - Press the **Deactivate** or **Activate** button to deactivate the ASR model.

## Deleting ASR models

If you decide to not use an ASR model anymore (e.g. you have a new adopted model to replace the old one), you can delete the ASR model permanently. When deleting an ASR model, the system checks if the ASR model is currently configured for any of the transcription policies, and it doesn't allow deleting until it is configured for at least one policy. Deleted ASR models are also automatically activated in the IV system. In order to delete an ASR model, follow the steps below:

**Step 1** - Navigate to **Data / Data Management / ASR Models** and select the ASR model.

**Step 2** - Press the **Delete** button to permanently delete the ASR model.

# Speechmatics transcription

## Deploying and configuring Speechmatics transcription

The Speechmatics transcription is considered a 3rd party transcription engine, which requires using a separate on-premise or cloud based Speechmatics infrastructure to run the transcription service. The Verba Speech Analytics Service is connecting to the Speechmatics platform and sends audio files for transcription. For more information see [Deploying transcription](#).

The Speechmatics integration requires the Speechmatics API 1.0.

Once, the Speechmatics platform is available, the required Data Processor has to be created to enable the integration with the Speechmatics transcription engine. Follow the steps described in [Configuring and running transcription](#) to create the processor and select the Speechmatics engine. The following table describes the settings available for a Speechmatics data processor:

Configuration item	Description
Name	Name of the data processor. This name will identify this processor across the system.
Type	Select Speech Transcription
Engine	Select Speechmatics
API Root URL	URL of the Speechmatics API (on-prem server or cloud)
API User	Speechmatics API user name
API Token	Speechmatics API token
Enable Speaker Diarization	Allows separating participants in conversations and producing a dialog like output

After creating the data processor, you can follow the guidelines at [Configuring and running transcription](#) to configure one or more data management policies.

## License requirements

The Speechmatics transcription engine does not require specific licenses, only the Advanced Archive license is required to enable the integration. Please contact your Verint sales representative for more information.

Although there is no specific license requirement, the feature requires a license to be assigned to users through the role configuration. The following permission(s) are required for the Speechmatics speech transcription:

- Transcribe Conversations (3rd Party)

## Uploading new ASR models

Please contact Speechmatics for more information.

## Customizing the Speechmatics ASR models

Please contact Speechmatics for more information.



# Searching and viewing transcripts

The screenshot displays the Verba Web Interface. At the top, there are navigation tabs: Conversations, Quality Management, Workflows, Reports, Users, Data, and System. The user is logged in as Verba Administrator. The main area is titled 'Conversations' and shows a search results table with 2 items found. The search criteria are 'ticket'. The table columns are Start Date, Start Time, Duration, From, From Info, To, To Info, and Direction. The first result is from Jul 31, 2018, at 1:53:16 PM, with a duration of 00:45:00. The transcript text is: '...about she's well have a call i got a i've got a call ...well i've actually no really but we have some tickets'. The second result is from Jul 31, 2018, at 1:53:11 PM, with a duration of 00:45:00. The transcript text is: 'calling at&t for quality assurance purposes your call ...trouble press or say one this is at&t is trouble ticket'. Below the table is a 'Conversation View' section with a timeline from 00:00 to 05:51. The audio waveform is paused at 00:01:14.694. Below the waveform are playback controls (play, stop, previous, next, volume) and a 'Markers' section with two markers. A 'Transcription' window is open, showing the transcript text with speaker labels: Speaker2: number give me gave me what number did you; Speaker1: yeah i was he said; Speaker2: let me see two hundred five eight seven seven zero three zero; Speaker1: okay repeat that again i'm sorry.

## Full-text search

In the Verba Web Interface under **Conversations**, you can search for words or phrases mentioned in the transcribed conversations. The search results will be highlighted in the transcript and in a small window of context will be shown in the results as well. You can find more in the [Searching in IM conversations and voice transcripts](#) article.

## Playback

During the playback, the portion of the transcript corresponding to the currently played part of the media will be automatically highlighted. Clicking on a word on the transcription will jump the playback to the corresponding part of the media, and start the playback.

## Automatic labeling rules

You can use the [Automatic labeling](#) to add labels based on transcripts. You can create an automatic labeling rule based on your current search by clicking on the **'Create labeling rule from these'** link. The text search query will be copied as filtering criteria to a newly created Label Rule. The filtering criteria for searching in IM conversations, SMS and voice transcripts is **Text Search**.

# Exporting transcripts

The system allows exporting transcripts in the following ways:

- Running a report
- Exporting
- Using the API

## Report

Following the [Creating report](#) guide, you can create a **Users Speech Transcript Details** report containing the transcripts from the filtered conversations.

## Export

You can use the [Export](#) feature to obtain the media files and the transcripts. The transcript file is a compressed file with **VTR** extension, containing a plain text file. The text file contains the following fields for each recognized word:

Field Name	Value	Example
word	the recognized word	hello
confidence	the confidence level	0.79000000
speaker	participant recognised	speaker1
time	time elapsed in seconds	189.35000000
duration	length of the word in seconds	0.50000000

## API

It is possible to integrate an external application to retrieve a transcription file using the [HTTP Business API](#). The [Get Call Information](#) API call can return either a plain text file or the transcription format, depending on the call.