



JSON Output Reference

for use with V-Spark

January 2022

Abstract

This document provides a detailed overview of the JSON output elements and schemata used in V-Spark and other Voci Automatic Speech Recognition products.

Legal Notice

The information contained in this document is the proprietary and confidential information of Voci Technologies or its licensor(s) (Voci). You may not disclose, provide or make available this document, or any information contained in this document, to any third party, without the prior written consent of Voci.

The information in this document is provided for use with Voci products. No license, express or implied, to any intellectual property associated with this document or such products is granted by this document.

All Voci Technologies products described in this document are protected by patents, trade secrets, copyrights, or other industrial property rights.

The Voci Technologies products described in this document may still be in development. The final form of each product and release date thereof is at the sole and absolute discretion of Voci. Your purchase, license and/or use of Voci products shall be subject to Voci's then current sales terms and conditions.

The following terms are trademarks of Voci Technologies in the United States and other countries:

- Voci
- V-Blaze
- V-Cloud
- V-Match
- V-Purify
- V-Spark

Other third-party disclaimers or notices may be set forth in Voci online or printed documentation.

All other product and service names, and trademarks, not owned by Voci are the property of their respective owners.

Table of Contents

Voci JSON Transcripts	4
Confidence Scores	4
Score Values and Calculation	5
Interpreting Confidence Scores	5
Confidence Scores and Clarity	5
Emotion and Text Analysis	5
Diarization Scores	6
Other Score Values	7
Top-level Elements	7
The textinfo Object	16
The utterances Array	19
The app_data Object	23
The scorecard Object	25
The client_data Object	27
V-Spark JSON Schema	31
JSON Output Example	64

Voci JSON Transcripts

All Voci products use the JSON file format to store transcript data derived from source audio. This data includes the text decoded from speech, along with metadata that describes audio attributes and the results of linguistic and emotional analysis performed by Voci products.



NOTE

When a data value is not defined (null), the data name will not appear as an element in JSON output. Therefore, not all of the elements listed in this section may be present in a specific JSON output file.

The data fields included in Voci JSON transcription output vary depending on the products, circumstances, and optional features that were used to generate the output. There are three categories of Voci JSON data:

1. Core ASR data generated by the ASR engine under most circumstances and whenever text is decoded from speech. For example, fields like the top-level **asr** and **model** elements will always be included in JSON output because they refer to ASR engine and language model attributes. Similarly, the **utterances** array is included if audio was successfully transcribed because it is generated any time speech is decoded from audio.
2. Conditional and parameter data generated only under certain conditions or when certain transcription parameters are specified. For example, the top-level **chaninfo** object is included only for stereo or diarized audio, and the top-level **emotion** field is included only when the transcription request includes the stream tag **emotion = true**.
Under some conditions, fields with identical names appear at different levels of the JSON data hierarchy. For example, the field **agentscore** is included at the top level only when processing undiarized mono audio with the stream tag **agentid = true**. When processing stereo audio with **agentid = true**, the **agentscore** field instead appears in the **chaninfo** objects for each channel of audio.
3. Data added to JSON transcripts when output is processed by V-Spark for analytics and application scoring. This data can be viewed in the V-Spark UI or in the top-level **app_data** and **client_data** objects and **last_modified** field generated for V-Spark JSON output.

Confidence Scores

Voci products use a confidence scoring system to represent the ASR engine's estimate of the probability that the correct words were selected during its transcription of speech to text and its classification of certain audio properties.

For example, when a word is decoded from speech, that word is assigned a numeric confidence score value. This confidence score is not an accuracy measure. It is a measure of how confident the ASR engine is that the output was transcribed correctly. In other words, a confidence score indicates the system's measure of probability that it has selected the word most likely to be correct out of all the words it believes a region of speech could represent. These confidence values can be used to filter output at various thresholds.

Score Values and Calculation

Confidence scores have a value between 0 and 1 for the text that is produced, and they are available at the word-event, utterance, and top level in JSON output. Confidence is calculated for each word event, then averaged up to the utterance level. The top-level confidence score is an average of utterance-level confidence scores.

Confidence scores tend to decrease as calls become more noisy, and as the overall speech signal moves further away from the speech used to train the Voci acoustic and language models. Noise, compression artifacts, and accents all contribute to lower confidence scores.

Interpreting Confidence Scores

A confidence score is only relevant to the language model used to transcribe the audio, and confidence should not be used to measure performance across multiple models. To measure performance across models, perform accuracy comparisons against reference transcript data for each of the models.

A confidence score on its own is not meaningful unless you have alternative results to compare against, either from the same recognition operation or from previous recognitions of the same input. Confidence values are relative and unique to each recognition engine. Confidence values returned by two different recognition engines cannot be meaningfully compared. A speech recognition engine may assign a low confidence score to spoken input for various reasons, including background interference, inarticulate speech, or unanticipated words or word sequences.

Confidence Scores and Clarity

The values for clarity scores that appear in JSON transcripts refer to both the confidence scores generated by the ASR engine and analytics scores as calculated by V-Spark. These different values are in different parts of the JSON depending on the product that generated them. Clarity values that appear in the V-Spark UI and in [the app_data object](#) are generated by V-Spark during the transcript analysis phase of call processing.

Emotion and Text Analysis

As of V-Blaze version 7.3 and V-Cloud version 1.6-2021-10.25, some emotion and text analysis is performed as part of ASR processing. The latest version of V-Spark does not account for the text analytics values generated by the ASR engine. As a result, transcripts generated by V-Spark systems using V-Blaze 7.3 or greater, or using V-Cloud, may contain redundant data depending on audio properties and transcription parameters. Although these redundant data values describe identical aspects of the transcribed text, they are calculated differently, and they can be distinguished by name or by JSON object hierarchy. These data fields include the following:

Table 1. ASR and V-Spark Analytics JSON Element Overlap

ASR Field	V-Spark Field	Description
chaninfo > emotion	app_data > agent_emotion and app_data > client_emotion	<p>Lower-level emotion values are more limited in scope than top-level emotion elements; a lower-level value may be derived from a single channel or utterance, but a top-level value is calculated using lower-level values.</p> <p>ASR emotion fields are listed in the chaninfo object that corresponds to the channel in which those emotions were detected. Those channels correspond to agent or client depending on the agentscore for the channel object.</p> <p>Emotion scores generated in V-Spark are directly assigned the agent or client classification; this designation may be configured for each channel, or it may have been detected using side classification.</p>
diascore	diarization	Each value indicates the system's certainty that it correctly diarized the audio. ASR and V-Spark diarization scores are calculated differently.
emotion	overall_emotion	<p>A top-level emotion value that represents a calculation based on lower-level values. Although emotion and overall_emotion describe the same metric, they are calculated differently by the ASR engine and V-Spark.</p> <p>The ASR engine and V-Spark use different value sets for emotion calculation results. Refer to the following topics for more information:</p> <ul style="list-style-type: none"> • Top-level Elements for the ASR engine's emotion field. • The app_data Object for V-Spark's overall_emotion field and its speaker-specific agent_emotion and client_emotion fields.
textinfo > overtalk	overtalk	Both the ASR engine and V-Spark calculate overtalk, but the overtalk object in the textinfo object generated by the ASR engine provides a more detailed breakdown of the underlying data.
textinfo > silence	silence	Both the ASR engine and V-Spark calculate silence as any audio segment without speech. This may include music or other noise. Values may vary between the two sources because they use different processes to detect speech.

Diarization Scores

Diarization score values displayed in the V-Spark UI and JSON transcripts function similarly to clarity and confidence scores, but diarization is a completely different metric from clarity and confidence. Diarization is the process by which multiple speakers in mono audio are separated onto separate channels, and a diarization score refers to the ASR engine's certainty that it separated speakers correctly. Diarization and its score values

apply only to mono audio. Since clarity and confidence measure the probability that the ASR engine selected the best text match, these values are always applicable when speech is decoded from text.

Other Score Values

Aside from individual words, other ASR metrics use score values that are similar to confidence scores in theory but vary in practice. Score values typically refer to a parent element, and they are usually labeled **score** in JSON output, but sometimes a score value is a weighted average of other scores in that transcription data.

For example, the **score** value in the **musicinfo** object at the **utterances** level indicates the ASR engine's level of confidence that it correctly classified the speech segment described by the **utterances** object as music. However, if the **musicinfo** object appears at the top level or in a **chaninfo** object, the score value is a time-weighted average of individual utterance scores.

Top-level Elements

The following table describes the top-level elements included in a JSON transcript.

Table 2. Top-level Elements

Element	Availability	Type	Description
emotion	V-Blaze version 7.2 and earlier	value	<p>Emotional intelligence consists of both acoustic and linguistic information. Events can be given the following values:</p> <ul style="list-style-type: none"> • Positive • Mostly Positive • Neutral • Mostly Negative • Negative <p>Emotion must be the same for all utterances to be included at the top level. Additional emotion scoring is available in The utterances Array.</p>

Element	Availability	Type	Description
emotion	V-Blaze version 7.3+	string	<p>Describes the emotion detected in decoded speech.</p> <p>Emotional intelligence consists of both acoustic and linguistic information. Events can be given the following values:</p> <ul style="list-style-type: none"> • Positive • Improving • Neutral • Negative • Worsening <p>As of V-Blaze version 7.3, the emotion field is always included at the top level, and the value describing detected emotion is more dynamic.</p> <p>The emotion detected toward the end of a call is compared to the emotion detected closer to the beginning. The emotion value describes what speaker emotion was, or how speaker emotion changed in transcribed audio.</p>
asr	V-Blaze version 6.1+	string	Version number of the automatic speech recognition server being used.
confidence	All	number	A measure of how confident the speech recognition system is in its transcription results. Results range between 0 and 1 with 1 being the most confident.
rawemotion	All	string	<p>Acoustic emotion values. Possible values in version 7.1+ include:</p> <ul style="list-style-type: none"> • ANGRY • NEUTRAL • HAPPY <p>Acoustic emotion values prior to version 7.1 include:</p> <ul style="list-style-type: none"> • NONANGRY • ANGRY
donedate	All	string	Date and time the file transcription was completed by the speech-to-text engine, meaning the last utterance finished.

Element	Availability	Type	Description
utterances	All	array	Each audio file is broken up into segments of speech called utterances. The utterances array contains the word transcripts and corresponding metadata organized by utterances.
license	All	string	Identification information for the license used.
audiosecs	V-Blaze version 6.1+	number	Duration of audio, in seconds, in the stream. As of V-Blaze 7.2, this element will not appear in the JSON output if there was a problem processing audio.
started	V-Blaze version 6.1+	string	Date and time the stream started. This is most useful for measuring real-time transcription.
streamtags	V-Blaze version 6.1+	object	A list of the parameters or other values specified by the user. This is useful for debugging and verification. It is also useful for tagging the output with user-level metadata (for example, tags that have meaning to the user for filtering or association). For example: <pre> "streamtags": { "emotion": "xa", "lid": true, "subst_rules": "<17 chars>", "gender": true, "rawemotion": "xa", "lidutt": true, "substinfo": true, "lidthreshold": 1.0, "subst": true, "scrubtext": true, "datahdr": "WAVE", "nsubs": "true" } </pre>
nchannels	All	number	Number of channels in the audio file unless diarization is set to true, in which a single (1) channel file is broken up into 2 based on speaker separation As of V-Blaze 7.2, this element will not appear in the JSON output if there was a problem processing audio.


Element	Availability	Type	Description
lidinfo	V-Blaze version 5.6+	object	<p>The <code>lidinfo</code> section is a global, top-level dictionary that contains the following fields:</p> <ul style="list-style-type: none">• <code>lang</code> – the three-letter language code specifying the language that was identified for the stream• <code>speech</code> – the number of seconds of automatically detected speech that were used to determine the language used in the stream• <code>langfinal</code> – (V-Blaze7.1+) Added when the language specified in LID is below threshold and not the default language.• <code>conf</code> – the confidence score of the language identification decision <p>For example:</p> <pre>"lidinfo": { "lang": "spa", "speech": 1.35, "langfinal": "eng", "conf": 0.81 }</pre>

Element	Availability	Type	Description
langinfo	V-Blaze version 7.1+	string	<p>Breakdown of language information that is added when there was more than one language detected. The dictionary contains several fields:</p> <ul style="list-style-type: none"> • <code>utts</code> - the number of utterances spoken for the language identified • <code>speech</code> – the number of seconds of automatically detected speech that were used to determine the language used in the stream • <code>conf</code> – the confidence score of the language identification decision • <code>time</code> - the number of seconds that the language was identified for the whole stream <p>For example:</p> <pre> "langinfo": { "spa": { "utts": 1, "speech": 17.46, "conf": 1.0, "time": 21.56 }, "eng": { "utts": 1, "speech": 1.35, "conf": 0.81, "time": 0.93 } } </pre>
ended	V-Blaze version 6.1+	string	<p>Date and time the stream ended. This is most useful for measuring real-time transcription.</p> <p>As of V-Blaze 7.2, this element will not appear in the JSON output if there was a problem processing audio.</p>
recvtz	All	array	<p>An array containing two values:</p> <ul style="list-style-type: none"> • time zone abbreviation of the time zone in which the ASR engine is running • offset in seconds from UTC for the time on the ASR engine

Element	Availability	Type	Description
scrubbed	All	boolean	If true then audio is purified so numbers are all redacted. If false, the data name does not appear in the JSON output.
sentiment	All	string	Linguistic sentiment value: <ul style="list-style-type: none"> • Positive • Mostly Positive • Neutral • Mostly Negative • Negative • Mixed (contains both Positive and Negative in the file)
sentiment_scores	All	array	Array of length 2. [0]=Positive phrase counts and [1]=Negative phrase counts in the file
source	All	string	The audio file name.
gender	All	string	The gender identified for the audio.
model	All	string containing model name if one model was specified; array of model names if multiple models were specified	Language model(s) specified for transcription. For example: <pre>"model" : "enl:callcenter"</pre> As of V-Blaze 7.2, this element will not appear in the JSON output if there was a problem processing audio.
rcvdate	All	string	Date and time the audio file was received by the ASR engine and placed in queue
requestid	All	string	The unique identifier for the request.

Element	Availability	Type	Description
nsubs	V-Blaze version 7.1+	number	The number of substitutions applied. This tag will not appear if no substitutions were applied. This value does not include <code>numtrans</code> substitutions.

Element	Availability	Type	Description
substinfo	V-Blaze version 7.1+	object	<p>Detail for substitutions that is included when <code>substinfo=true</code>.</p> <ul style="list-style-type: none">• nsubs (V-Blaze 7.1+) – The number of substitutions applied, including <code>numtrans</code> substitutions.• counts (V-Blaze 7.1+) – array that contains one nested array for each source of substitution data. Nested arrays show the source (string value) and count (number value) of substitutions that were performed on the audio, along with an object containing the transformation patterns used to perform the substitution.• numtrans (V-Blaze 7.1+) – array that contains details on individual <code>numtrans</code> substitutions. <p>For example, the object below shows 4 total substitutions from 2 sources using 3 patterns:</p> <pre>"substinfo": { "counts": [["subst_rules", 2, { "yeah => yes": 2 }], ["numtrans", 2, { "four => 4": 1, "last 4 or => last four or": 1 }]], "nsubs": 4 }</pre>

Element	Availability	Type	Description
warning	V-Blaze version 5.6.0-3+	string	This field describes a problem or issue that was encountered during transcription. A common example is substitutions errors.
app_data	V-Spark only	object	A JSON object that stores metadata and application scores generated by V-Spark.
client_data	V-Spark only	object	A JSON object that stores user-supplied call metadata associated with the audio file.
last_modified	V-Spark version 4.0.2-1+ only	string	<p>The date and time at which an update to the last_modified field was last triggered in the Elasticsearch record associated with a transcript. If the last_modified field is not present or has no date and time value, its return value is false.</p> <p>The following events trigger an update to the last_modified field:</p> <ul style="list-style-type: none"> • Creating a new transcript. • Updating transcript scores by reprocessing an application. • Deleting an application or application category associated with the transcript. • Unlinking an application from the transcript's folder, if that application has previously been used to score the transcript. • Updating transcript metadata using the API. <div style="background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p> NOTE</p> <p>The last_modified field was implemented in V-Spark version 4.0.2-1. As a result, last_modified is not included in static JSON transcripts generated by older versions. Dynamically generated JSON output, such as that downloaded from the Files View, includes a last_modified field even if its audio record is from an older version.</p> <p>To add the last_modified field to a transcript's Elasticsearch record that was generated in a version older than 4.0.2-1, an action that triggers an update to last_modified must occur.</p> </div>
textinfo	V-Blaze version 7.3+	object	Text metrics for the audio transcript, including the amount of transcribed audio that was silence or contained words, overtalk metrics, and the total number of words spoken. If the initial audio was stereo or diarized mono, textinfo also includes the total number of speaker turns.

Element	Availability	Type	Description
diascore	V-Blaze version 7.3+	number	Indicates the level of confidence the system has in its classification of agent and client for audio with two speakers on a single channel. Expressed as a range between 0 and 1, where 1 indicates the best speaker separation.
agentscore	V-Blaze version 7.3+	number	<p>Predicts whether the speaker is the agent or client. Expressed as a value between -1 and 1, where a negative value means the speaker is believed to be the client. A positive value corresponds to an agent. A value closer to -1 or 1 indicates the system is more confident in its prediction, where -1 and 1 indicate the most confidence.</p> <p>Appears at the top level only when processing mono audio that was not diarized and was submitted for transcription with the stream tag agentid=true.</p>
musicinfo	V-Blaze version 7.3+	object	Appears only for stereo audio in which music was detected when audio was submitted for transcription with the stream tag music=true or music=info .
chaninfo	V-Blaze version 7.3+	array	<p>Appears only for stereo or diarized audio. Contains one object for each audio channel. Each channel object may contain the elements in the following list depending on audio attributes and the stream tags specified with the request.</p> <ul style="list-style-type: none"> • emotion • textinfo • musicinfo • agentscore <p>Elements in the chaninfo array's channel objects contain the same information as the top-level elements described in this table.</p>

The textinfo Object

The **textinfo** object is included in a JSON transcript by default when any text is decoded from an audio file. To exclude the **textinfo** object, specify the stream tag **textinfo = false** when submitting audio for transcription. This element was first introduced with V-Blaze version 7.3.

The **textinfo** object includes the following elements:

Table 3. Elements in the `textinfo` object

Element	Type	Description
turns	number	The number of distinct speaker turns detected in the audio. Calculated for stereo or diarized mono audio only.
wordtime	array	An array with the following number values: <ul style="list-style-type: none"> The total audio time in seconds during which words were detected. Calculated by adding the durations of each utterance with words. The percentage of total audio time during which words were detected.
overtalk	object	Metrics for overtalk throughout the audio file. Calculated for multi-channel audio only.
count	array	Contains the following number elements: <ul style="list-style-type: none"> The number of overtalk occurrences. The percentage of total speaker turns on which overtalk occurred.
avgtime	number	The average duration of all overtalk occurrences.
time	array	Contains the following number elements: <ul style="list-style-type: none"> The total audio time in seconds with overtalk. Calculated by adding the durations of each utterance with overtalk. The percentage of total audio time during which overtalk occurred.
words	number	The total of number of words spoken in the transcribed audio file.
silence	array	Contains the following number elements: <ul style="list-style-type: none"> The total amount of audio in seconds with no sound. The percentage of total audio time with no sound.

The following JSON example shows a `textinfo` object generated from stereo audio:

```
"textinfo": {
  "turns": 229,
  "wordtime": [
    945.62,
    0.702
  ],
  "overtalk": {
    "count": [
      92,
      0.402
    ],
    "avgtime": 1.19,
    "time": [
      109.44,
      0.116
    ]
  },
  "words": 3652,
  "silence": [
    401.64,
    0.298
  ]
}
```

The utterances Array

The top-level **utterances** array is included in a JSON transcript when any text is decoded from audio. The **utterances** array is an array of objects, and it contains one object for each utterance.

An utterance is defined in this context as an uninterrupted chain of spoken language by a single speaker. An utterance ends with a period of silence that exceeds a threshold duration or that exceeds the maximum utterance duration threshold. Each object in the **utterances** array may contain the elements in the following table:

Table 4. Elements in utterances Array Objects

Element	Availability	Type	Description
emotion	All	string	Emotional intelligence consists of both acoustic and linguistic information. Events can be given the following values: <ul style="list-style-type: none"> • Positive • Mostly Positive • Neutral • Mostly Negative • Negative
confidence	All	number	A measure of how confident the speech recognition system is in its utterance transcription results. <ul style="list-style-type: none"> • Range between 0 and 1 • 1 is most confident
end	All	number	End time of the utterance in seconds
recvtz	All	array	An array containing two values: <ul style="list-style-type: none"> • time zone abbreviation of the time zone in which the ASR engine is running • offset in seconds from UTC for the time on the ASR engine
sentiment	All	string	Utterance-level linguistic sentiment value: <ul style="list-style-type: none"> • Positive • Mostly Positive • Neutral

Element	Availability	Type	Description
			<ul style="list-style-type: none"> • Mostly Negative • Negative • Mixed (contains both Positive and Negative in the file)
gender	All	string	Gender prediction of the speaker
rawemotion	All	string	Acoustic emotion values (version 7.1+): <ul style="list-style-type: none"> • ANGRY • NEUTRAL • HAPPY Acoustic emotion values (prior to version 7.1): <ul style="list-style-type: none"> • NONANGRY • ANGRY
lidinfo	V-Blaze version 7.1+ V-Cloud version 1.6+	array	The <code>lidinfo</code> section is a global, top-level dictionary that contains the following fields: <ul style="list-style-type: none"> • <code>lang</code> – the three-letter language code specifying the language that was identified for the stream • <code>speech</code> – the number of seconds of automatically detected speech that were used to determine the language used in the stream • <code>conf</code> – the confidence score of the language identification decision • <code>langfinal</code> - added when the language specified in LID is below threshold and not the default language For example: <pre>"lidinfo": { "lang": "spa", "speech": 17.46, "conf": 1.0 }</pre>
sentimentex	All	array	Contains sentiment information for each utterance

Element	Availability	Type	Description
			<ul style="list-style-type: none"> • [0][0] = Positive phrase counts • [0][1] = Negative phrase counts in utterance • [1][*] consist of an array of sentiment segments where [0] = '+' or '-' for Positive and Negative, and [1] is the position range of the phrase • [0] is beginning and [1] is end position
start	All	number	Start time of the utterance in seconds
doneDate	All	string	Date and time the utterance transcription was completed by the speech-to-text engine
recvDate	All	string	Date and time the utterance was received by the speech-to-text engine
events	All	array	<p>Contains information about individual words. Each element is a word object that contains the following values:</p> <ul style="list-style-type: none"> • confidence – number indicating the ASR engine's word-level transcription confidence level, expressed as a value between 0 and 1 where 1 is the most confident. • end – number indicating the end time of the word in seconds. • start – number indicating start time the word in seconds. • word – string indicating the normalized word. • wordex – string indicating the raw dictionary word. This value may not be present in each object in the events array, and it is often used to disambiguate different pronunciations that have the same spelling. <p>For example:</p> <pre>"events": [{ "confidence": 0.69, "end": 2.32, "start": 1.81, "word": "Stephanie" }, { "confidence": 0.76, "end": 2.74, "start": 2.32,</pre>

Element	Availability	Type	Description
			<pre> "word": "so" } </pre>
metadata	All	object	<p>Speaker and information of the utterance. Each object contains the following values:</p> <ul style="list-style-type: none"> • channel – number indicating the audio channel on which the utterance was recorded. • model – string indicating the model that decoded the utterance. • source – string indicating the audio file name. • nsubs (V-Blaze 7.1+) – a number indicating the count of substitutions applied for the utterance, not including <code>numtrans</code> counts. • uttid – number indicating the utterance segment. • substinfo (V-Blaze 7.1+) – object with detail about substitutions, included in an utterance object when audio is processed with the stream tag substinfo = true and substitutions were performed on the utterance. Includes the following data: <ul style="list-style-type: none"> • subs (V-Blaze 7.1+) – array that contains one nested array for each substitution performed. Nested arrays contain number elements describing the start and end time in seconds for the substituted speech, and an additional array of objects with string and number values describing the source and substitution performed. • nsubs (V-Blaze 7.1+) – number indicating the count of substitutions applied to the utterance, including <code>numtrans</code> counts. <p>For example:</p> <pre> "metadata": { "uttid": 3, "substinfo": { "subs": [[55.38, 55.74, [{ "source": "subst_rules", "end": 55.74, </pre>

Element	Availability	Type	Description
			<pre> "sub": "yeah => yes", "rule": "0", "start": 55.38 }]], "nsubs": 1 }, </pre>
music	V-Blaze version 7.3+	boolean	Appears only if audio was processed using the stream tag music = true . Has a value of true if music was detected in the utterance.
musicinfo	V-Blaze version 7.3+	object	<p>Appears only if audio was processed using the stream tag music = info. Contains the following number values:</p> <ul style="list-style-type: none"> • used is the total audio time seconds during which the utterance contains music. • score has a value in the range of -1 to 1, where a negative value means the utterance is not music. A positive value means the utterance is music. A value closer to -1 or 1 indicates the system is more confident in its classification, where -1 and 1 indicate the most confidence.

The app_data Object

The top-level **app_data** object contains information calculated from the entire transcript, rather than from individual utterances. The **app_data** object also includes metadata describing the audio's basic attributes and the score results from any applications configured for the folder that processed the audio and transcript.

The **app_data** object includes the elements in the following table. Note that some agent- or client-specific fields may not appear for single-channel audio that is not diarized.

Table 5. Elements in the app_data object

Element	Type	Description
agent_channel	number	The audio channel with agent speech.
agent_clarity	string	How clear the speech on the agent channel is. Expressed as a range between 0 and 1, where 1 is clearest.

Element	Type	Description
agent_emotion	string	Indicates overall agent emotion, calculated using both acoustic and linguistic information, with one of the following values: <ul style="list-style-type: none">• Positive• Improving• Negative• Worsening
agent_gender	string	Gender prediction for the speaker classified as agent.
client_channel	number	The audio channel with client speech.
client_clarity	string	How clear the speech on the agent channel is. Expressed as a range between 0 and 1, where 1 is clearest.
client_emotion	string	Indicates overall client emotion, calculated using both acoustic and linguistic information, with one of the following values: <ul style="list-style-type: none">• Positive• Improving• Negative• Worsening
client_gender	string	Gender prediction for the speaker classified as client.
datetime	string	Transcript date and time expressed in Coordinated Universal Time (UTC).
diarization	number	Indicates the level of confidence the system has in its classification of agent and client for audio with two speakers on a single channel. Expressed as a range between 0 and 1, where 1 indicates the best speaker separation.
duration	string	The duration of the initial audio file, expressed in hours, minutes, and seconds using the format HH:MM:SS .

Element	Type	Description
overall_emotion	string	Indicates the audio file's overall emotion, calculated using both acoustic and linguistic information, with one of the following values: <ul style="list-style-type: none"> • Positive • Improving • Negative • Worsening
overtalk	string	Percentage of call when the agent talks over or interrupts the client. Equal to the number of turns where the agent initiated overtalk divided by the total number of agent turns.
scorecard	object	Contains any application scores that have been calculated for the transcript.
silence	string	Percentage of overall duration that is silence. Equal to all non-speech time, calculated as call duration minus the sum of the duration of each word. If music and noise are not decoded to word events, they are counted as silence.
tId	number	The unique transcriptID used to reference a particular transcript.
url	string	Location of the audio file associated with the transcript.
words	number	Total number of words in a transcript.

The scorecard Object

JSON transcripts include the **scorecard** object as part of the top-level **app_data** object when the V-Spark folder that processed the audio is configured to analyze transcripts using one or more applications.

V-Spark applications may have multiple levels of category, up to a maximum of 3 subcategories for each top-level category. The **scorecard** object uses the same nested structure. At its first level, the **scorecard** object contains one object for each application. Each application object contains a **score** element and a **subcategories** object. Each **subcategories** object contains lower-level subcategories and scores.

The following table lists elements of the **scorecard** object:

Table 6. Elements in the scorecard object

Element	Type	Description
<i>APPLICATION-NAME</i>	object	Represents an application, its score, and all of its subcategories in a subcategories object.
score	number	The score value for the application overall.
subcategories	object	Contains all subcategories below the first level.
score	number	The score value for subcategories.
subcategories	object	Contains all subcategories below the second level.
score	number	The score value for the lowest-level subcategories.
subcategories	object	An empty object, since this is the lowest level of application subcategory.

The following JSON example shows a **scorecard** object from a transcript processed by the application **WordsApp**. One category, named **Words**, has the maximum number of subcategories. The other category, named **MoreWords**, has no subcategories.

```
"scorecard": {
  "WordsApp": {
    "Words": {
      "subcategories": {
        "SubWords": {
          "subcategories": {
            "SubSubWords": {
              "subcategories": {
                "SubSubSubWords": {
                  "subcategories": {},
                  "score": 1
                },
                "SubSubSubWords2": {
                  "subcategories": {},
                  "score": 1
                }
              }
            },
            "score": 1
          },
          "score": 1
        },
        "score": 1
      },
      "MoreWords": {
        "subcategories": {},
        "score": 1
      }
    }
  }
},
```

The client_data Object

The top-level `client_data` object contains all custom metadata fields and values submitted with the initial ASR transcription request. The `client_data` object also includes the following fields created and populated by the system:

**IMPORTANT**

The `client_data` object is generated by V-Spark, and does not appear in V-Blaze or V-Cloud JSON transcripts.

Table 7. Elements in the `client_data` object created by V-Spark

Element	Type	Description
audio_duration	number	The initial audio file's duration in seconds.
audio_properties	string	The initial audio's file format, encoding type, bitrate, frequency, and number of channels. These properties may not be available for all audio. In that case, the comma-separated list of values will include blanks, as in the following example: <div data-bbox="976 613 1892 732" style="background-color: #2e3436; color: #eeeeec; padding: 5px; margin-top: 10px;"> <pre>{ ... "audio_properties": "Opus, , , 48000 Hz, , 2 ch", ... }</pre> </div>
jm_version	string	The version of Job Manager that processed the transcript. As of V-Spark 4.0, this field always has the value "None" because Job Manager version now matches the V-Spark version.
performance_stats	object	Stores the following metrics from ASR processing and transcript analysis:
job_latency	number	Represents the efficiency of the analysis job as a function of end time, start time, and audio duration as in the formula (EndTime - StartTime) / (AudioDuration) . In other words, a measure of how long it took to process the transcript in comparison to the length of the audio file.
job_end	string	The time at which the job ended.
asr_latency	number	Represents the efficiency of the ASR engine's transcription job as a function of end time, start time, and audio duration as in the formula (EndTime - StartTime) / (AudioDuration) . In other words, a measure of how long it took to transcribe the audio in comparison to the length of the audio file.
asr_end	string	The time at which ASR transcription ended.
job_start	string	The time at which the job began.
asr_start	string	The time at which ASR transcription began.

Custom metadata submitted with the initial audio file is always provided in JSON output exactly as submitted, but custom metadata is indexed and searchable from the UI only if the metadata field is configured for the audio's V-Spark folder when that audio is processed. Custom metadata that is submitted with the audio but not configured for the processing folder is passed through to the **client_data** object in JSON output, but that metadata is neither indexed nor searchable.

The following is an example of the **client_data** object in a JSON transcript:

```
"client_data": {
  "agentname": "Reid Smith",
  "client-phone": "123-456-7890",
  "branch-loc": "Raleigh",
  "account_priority": "Medium",
  "jm_version": "None",
  "audio_properties": "AAC, CBR, 175 kb/s, 48000 Hz, , 2 ch",
  "performance_stats": {
    "job_latency": 0.4851,
    "job_end": "2021-10-06_19:22:41 +0000",
    "asr_latency": 0.1541,
    "asr_end": "2021-10-06_19:22:40 +0000",
    "job_start": "2021-10-06_19:22:39 +0000",
    "asr_start": "2021-10-06_19:22:39 +0000"
  },
  "audio_duration": 4.1387
}
```

Custom metadata can be provided in either JSON or XML format. The preceding example **client_data** section was produced in response to uploading the following custom metadata JSON in a file zipped with a corresponding audio file:

```
{
  "metadata": {
    "agentname": "Reid Smith",
    "client-phone": "123-456-7890",
    "branch-loc": "Raleigh",
    "account_priority": "Medium"
  }
}
```

The following example shows the XML equivalent of the preceding JSON metadata file:

```
<metadata>  
  <agentname>Reid Smith</agentname>  
  <client-phone>123-456-7890</client-phone>  
  <branch-loc>Raleigh</branch-loc>  
  <account_priority>Medium</account_priority>  
</metadata>
```

V-Spark JSON Schema

```
{
  "$schema": "http://json-schema.org/draft-07/schema#",
  "$anchor": "V-Spark-JSON-Schema",
  "type": "object",
  "properties": {
    "client_data": {
      "title": "client_data",
      "type": "object",
      "properties": {
        "country": {
          "title": "country",
          "type": "string",
          "examples": [
            "USA"
          ]
        },
        "audio_duration": {
          "title": "audio_duration",
          "type": "number",
          "examples": [
            78.48
          ]
        },
        "audio_properties": {
          "title": "audio_properties",
          "type": "string",
          "examples": [
            "PCM, CBR, 256 kb/s, 8000 Hz, 16 bits, 2 ch"
          ]
        },
        "performance_stats": {
          "title": "performance_stats",
          "type": "object",

```

```
"properties": {
  "asr_end": {
    "title": "asr_end",
    "type": "string",
    "examples": [
      "2021-11-12_16:04:42 +0000"
    ]
  },
  "asr_start": {
    "title": "asr_start",
    "type": "string",
    "examples": [
      "2021-11-12_16:04:36 +0000"
    ]
  },
  "job_start": {
    "title": "job_start",
    "type": "string",
    "examples": [
      "2021-11-12_16:04:36 +0000"
    ]
  },
  "asr_latency": {
    "title": "asr_latency",
    "type": "number",
    "examples": [
      0.0717
    ]
  },
  "job_latency": {
    "title": "job_latency",
    "type": "number",
    "examples": [
      0.0767
    ]
  },
  "job_end": {
```



```
        "title": "job_end",
        "type": "string",
        "examples": [
            "2021-11-12_16:04:42 +0000"
        ]
    },
    "additionalProperties": true,
    "required": [
        "asr_end",
        "asr_start",
        "job_start",
        "asr_latency",
        "job_latency",
        "job_end"
    ]
},
"region": {
    "title": "region",
    "type": "string",
    "examples": [
        "SELA"
    ]
},
"jm_version": {
    "title": "jm_version",
    "type": "string",
    "examples": [
        "None"
    ]
},
"brand": {
    "title": "brand",
    "type": "string",
    "examples": [
        "ACME"
    ]
}
```

```
    },
    "additionalProperties": true,
    "required": [
      "audio_duration",
      "audio_properties",
      "performance_stats",
      "jm_version",
    ]
  },
  "sentiment": {
    "title": "sentiment",
    "type": "string",
    "examples": [
      "Negative"
    ]
  },
  "asr": {
    "title": "asr",
    "type": "string",
    "examples": [
      "7.3.0-1"
    ]
  },
  "source": {
    "title": "source",
    "type": "string",
    "examples": [
      "audio.wav"
    ]
  },
  "doneDate": {
    "title": "doneDate",
    "type": "string",
    "examples": [
      "2021-11-12 11:04:44.367598"
    ]
  }
}
```

```
},
  "streamtags": {
    "title": "streamtags",
    "type": "object",
    "properties": {
      "agentid": {
        "title": "agentid",
        "type": "boolean",
        "examples": [
          true
        ]
      },
      "gender": {
        "title": "gender",
        "type": "boolean",
        "examples": [
          true
        ]
      },
      "rawemotion": {
        "title": "rawemotion",
        "type": "boolean",
        "examples": [
          true
        ]
      },
      "nsubs": {
        "title": "nsubs",
        "type": "string",
        "examples": [
          "true"
        ]
      },
      "lid": {
        "title": "lid",
        "type": "boolean",
        "examples": [
```

```
    true
  ]
},
"scrubtext": {
  "title": "scrubtext",
  "type": "boolean",
  "examples": [
    true
  ]
},
"datahdr": {
  "title": "datahdr",
  "type": "string",
  "examples": [
    "WAVE"
  ]
},
"subst_rules": {
  "title": "subst_rules",
  "type": "string",
  "examples": [
    "<10 chars>"
  ]
},
"billing": {
  "title": "billing",
  "type": "string",
  "examples": [
    "col-orgA-folder0"
  ]
},
"token": {
  "title": "token",
  "type": "string",
  "examples": [
    ""
  ]
}
```

```
    },
    "emotion": {
      "title": "emotion",
      "type": "string",
      "examples": [
        "xa"
      ]
    },
    "music": {
      "title": "music",
      "type": "string",
      "examples": [
        "info"
      ]
    },
    "substinfo": {
      "title": "substinfo",
      "type": "boolean",
      "examples": [
        true
      ]
    },
    "model": {
      "title": "model",
      "type": "array",
      "items": {
        "type": "string",
        "examples": [
          "en:callcenter"
        ]
      }
    }
  },
  "additionalProperties": true,
  "required": [
    "billing",
    "token",
```

```
    "model"
  ]
},
"sentiment_scores": {
  "title": "sentiment_scores",
  "type": "array",
  "items": {
    "type": "number",
    "examples": [
      0
    ]
  }
},
"substinfo": {
  "title": "substinfo",
  "type": "object",
  "properties": {
    "nsubs": {
      "title": "nsubs",
      "type": "number",
      "examples": [
        4
      ]
    }
  }
},
"counts": {
  "title": "counts",
  "type": "array",
  "items": {
    "type": "array",
    "items": {
      "anyOf": [
        {
          "type": "number",
          "examples": [
            2
          ]
        }
      ]
    }
  }
},
```

```
{
  "type": "object",
  "properties": {
    "yeah => yes": {
      "title": "yeah => yes",
      "type": "number",
      "examples": [
        2
      ]
    },
    "last 4 or => last four or": {
      "title": "last 4 or => last four or",
      "type": "number",
      "examples": [
        1
      ]
    },
    "four => 4": {
      "title": "four => 4",
      "type": "number",
      "examples": [
        1
      ]
    }
  },
  "additionalProperties": true
},
{
  "type": "string",
  "examples": [
    "subst_rules"
  ]
}
]
```

```
    },
    "additionalProperties": true,
    "required": [
      "nsubs",
      "counts"
    ]
  },
  "model": {
    "title": "model",
    "type": "array",
    "items": {
      "type": "string",
      "examples": [
        "en:callcenter"
      ]
    }
  }
},
"app_data": {
  "title": "app_data",
  "type": "object",
  "properties": {
    "agent_clarity": {
      "title": "agent_clarity",
      "type": "string",
      "examples": [
        "0.845"
      ]
    }
  },
  "client_channel": {
    "title": "client_channel",
    "type": "number",
    "examples": [
      1
    ]
  },
  "agent_emotion": {
    "title": "agent_emotion",
```



```
    "type": "string",
    "examples": [
      "Negative"
    ]
  },
  "client_clarity": {
    "title": "client_clarity",
    "type": "string",
    "examples": [
      "0.824"
    ]
  },
  "words": {
    "title": "words",
    "type": "number",
    "examples": [
      241
    ]
  },
  "agent_gender": {
    "title": "agent_gender",
    "type": "string",
    "examples": [
      "female"
    ]
  },
  "agent_channel": {
    "title": "agent_channel",
    "type": "number",
    "examples": [
      0
    ]
  },
  "client_emotion": {
    "title": "client_emotion",
    "type": "string",
    "examples": [
```

```
    "Positive"
  ]
},
"overall_emotion": {
  "title": "overall_emotion",
  "type": "string",
  "examples": [
    "Negative"
  ]
},
"scorecard": {
  "title": "scorecard",
  "type": "object",
  "additionalProperties": true
},
"url": {
  "title": "url",
  "type": "string",
  "examples": [
    "http://ip-10-253-150-159.us-east-2.compute.internal:3000/fileDetails/col-orgA/folder0/
JP71FH0BG4pTu3xoE2Wz/1/audio.wav"
  ]
},
"tId": {
  "title": "tId",
  "type": "number",
  "examples": [
    1
  ]
},
"duration": {
  "title": "duration",
  "type": "string",
  "examples": [
    "0:01:18"
  ]
},
}
```

```
"datetime": {
  "title": "datetime",
  "type": "string",
  "examples": [
    "2021-11-12 16:04:39 UTC"
  ]
},
"overtalk": {
  "title": "overtalk",
  "type": "string",
  "examples": [
    "0.143"
  ]
},
"silence": {
  "title": "silence",
  "type": "string",
  "examples": [
    "0.342"
  ]
},
"client_gender": {
  "title": "client_gender",
  "type": "string",
  "examples": [
    "male"
  ]
},
},
"additionalProperties": true,
"required": [
  "agent_clarity",
  "client_channel",
  "agent_emotion",
  "client_clarity",
  "words",
  "agent_gender",
```

```
    "agent_channel",
    "client_emotion",
    "overall_emotion",
    "scorecard",
    "url",
    "tId",
    "duration",
    "datetime",
    "overtalk",
    "silence",
    "client_gender"
  ]
},
"audiosecs": {
  "title": "audiosecs",
  "type": "number",
  "examples": [
    78.48
  ]
},
"last_modified": {
  "title": "last_modified",
  "type": "string",
  "examples": [
    "2021-11-12 16:04:56"
  ]
},
"chaninfo": {
  "title": "chaninfo",
  "type": "array",
  "items": {
    "type": "object",
    "properties": {
      "emotion": {
        "title": "emotion",
        "type": "string",
        "examples": [
```

```
    "Improving"
  ]
},
"agentscore": {
  "title": "agentscore",
  "type": "number",
  "examples": [
    0.38
  ]
},
"textinfo": {
  "title": "textinfo",
  "type": "object",
  "properties": {
    "overtalk": {
      "title": "overtalk",
      "type": "object",
      "properties": {
        "avgtime": {
          "title": "avgtime",
          "type": "number",
          "examples": [
            0.48
          ]
        }
      }
    }
  }
},
"count": {
  "title": "count",
  "type": "array",
  "items": {
    "type": "number",
    "examples": [
      1
    ]
  }
},
"time": {
  "title": "time",
```

```
        "type": "array",
        "items": {
          "type": "number",
          "examples": [
            0.48
          ]
        }
      },
      "additionalProperties": true,
      "required": [
        "avgtime",
        "count",
        "time"
      ]
    },
    "wordtime": {
      "title": "wordtime",
      "type": "array",
      "items": {
        "type": "number",
        "examples": [
          38.04
        ]
      }
    },
    "words": {
      "title": "words",
      "type": "number",
      "examples": [
        175
      ]
    },
    "silence": {
      "title": "silence",
      "type": "array",
      "items": {
```

```
        "type": "number",
        "examples": [
          40.44
        ]
      },
      "turns": {
        "title": "turns",
        "type": "number",
        "examples": [
          7
        ]
      },
      "additionalProperties": true,
      "required": [
        "overtalk",
        "wordtime",
        "words",
        "silence",
        "turns"
      ]
    },
    "additionalProperties": true,
    "required": [
      "emotion",
      "agentscore",
      "textinfo"
    ]
  },
  "utterances": {
    "title": "utterances",
    "type": "array",
    "items": {
      "type": "object",
```

```
"properties": {
  "sentiment": {
    "title": "sentiment",
    "type": "string",
    "examples": [
      "Negative"
    ]
  },
  "metadata": {
    "title": "metadata",
    "type": "object",
    "properties": {
      "substinfo": {
        "title": "substinfo",
        "type": "object",
        "properties": {
          "subs": {
            "title": "subs",
            "type": "array",
            "items": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "number",
                    "examples": [
                      3.83
                    ]
                  }
                ]
              },
              "type": "array",
              "items": {
                "type": "object",
                "properties": {
                  "sub": {
                    "title": "sub",
                    "type": "string",

```



```
    "examples": [
      "yeah => yes"
    ]
  },
  "start": {
    "title": "start",
    "type": "number",
    "examples": [
      3.83
    ]
  },
  "rule": {
    "title": "rule",
    "type": "string",
    "examples": [
      "0"
    ]
  },
  "end": {
    "title": "end",
    "type": "number",
    "examples": [
      4.13
    ]
  },
  "source": {
    "title": "source",
    "type": "string",
    "examples": [
      "subst_rules"
    ]
  }
},
"additionalProperties": true,
"required": [
  "sub",
  "start",
```

```
        "rule",
        "end",
        "source"
    ]
    }
  ]
},
"nsubs": {
  "title": "nsubs",
  "type": "number",
  "examples": [
    3
  ]
},
"additionalProperties": true,
"required": [
  "subs",
  "nsubs"
]
},
"nsubs": {
  "title": "nsubs",
  "type": "number",
  "examples": [
    1
  ]
},
"uttid": {
  "title": "uttid",
  "type": "number",
  "examples": [
    0
  ]
}
```

```
    },  
    "channel": {  
      "title": "channel",  
      "type": "number",  
      "examples": [  
        0  
      ]  
    },  
    "model": {  
      "title": "model",  
      "type": "string",  
      "examples": [  
        "en1:callcenter"  
      ]  
    },  
    "source": {  
      "title": "source",  
      "type": "string",  
      "examples": [  
        "audio.wav"  
      ]  
    }  
  },  
  "additionalProperties": true,  
  "required": [  
    "uttid",  
    "channel",  
    "model",  
    "source"  
  ]  
},  
"gender": {  
  "title": "gender",  
  "type": "string",  
  "examples": [  
    "female"  
  ]  
}
```

```
    },  
    "rawemotion": {  
      "title": "rawemotion",  
      "type": "string",  
      "examples": [  
        "ANGRY"  
      ]  
    },  
    "confidence": {  
      "title": "confidence",  
      "type": "number",  
      "examples": [  
        0.86  
      ]  
    },  
    "start": {  
      "title": "start",  
      "type": "number",  
      "examples": [  
        0.41  
      ]  
    },  
    "recvtz": {  
      "title": "recvtz",  
      "type": "array",  
      "items": {  
        "anyOf": [  
          {  
            "type": "string",  
            "examples": [  
              "EST"  
            ]  
          },  
          {  
            "type": "number",  
            "examples": [  
              -18000  
            ]  
          }  
        ]  
      }  
    }  
  ]  
}
```

```
    ]
  }
]
},
"doneDate": {
  "title": "doneDate",
  "type": "string",
  "examples": [
    "2021-11-12 11:04:44.325363"
  ]
},
"recvDate": {
  "title": "recvDate",
  "type": "string",
  "examples": [
    "2021-11-12 11:04:39.223285"
  ]
},
"emotion": {
  "title": "emotion",
  "type": "string",
  "examples": [
    "Negative"
  ]
},
"sentimentex": {
  "title": "sentimentex",
  "type": "array",
  "items": {
    "type": "array",
    "items": {
      "anyOf": [
        {
          "type": "number",
          "examples": [
            0
          ]
        }
      ]
    }
  ]
}
```

```
    ],
  },
  {
    "type": "array",
    "items": {
      "anyOf": [
        {
          "type": "string",
          "examples": [
            "_"
          ]
        },
        {
          "type": "array",
          "items": {
            "type": "number",
            "examples": [
              7
            ]
          }
        }
      ],
    },
    {
      "type": "number",
      "examples": [
        1
      ]
    }
  ]
}
],
}
},
"musicinfo": {
  "title": "musicinfo",
  "type": "object",
```

```
    "properties": {
      "score": {
        "title": "score",
        "type": "number",
        "examples": [
          -0.8
        ]
      },
      "used": {
        "title": "used",
        "type": "number",
        "examples": [
          12.67
        ]
      }
    },
    "additionalProperties": true,
    "required": [
      "score",
      "used"
    ]
  },
  "end": {
    "title": "end",
    "type": "number",
    "examples": [
      20.21
    ]
  },
  "events": {
    "title": "events",
    "type": "array",
    "items": {
      "type": "object",
      "properties": {
        "confidence": {
          "title": "confidence",
```

```
    "type": "number",
    "examples": [
      0.86
    ]
  },
  "start": {
    "title": "start",
    "type": "number",
    "examples": [
      0.41
    ]
  },
  "end": {
    "title": "end",
    "type": "number",
    "examples": [
      0.68
    ]
  },
  "word": {
    "title": "word",
    "type": "string",
    "examples": [
      "So"
    ]
  },
  "wordex": {
    "title": "wordex",
    "type": "string",
    "examples": [
      "for(4)"
    ]
  }
},
"additionalProperties": true,
"required": [
  "confidence",
```



```
        "start",
        "end",
        "word"
    ]
    }
  },
  "additionalProperties": true,
  "required": [
    "metadata",
    "confidence",
    "start",
    "recvtz",
    "donedate",
    "recvdate",
    "emotion",
    "end",
    "events"
  ]
}
},
"nsubs": {
  "title": "nsubs",
  "type": "number",
  "examples": [
    2
  ]
},
"confidence": {
  "title": "confidence",
  "type": "number",
  "examples": [
    0.84
  ]
},
"recvtz": {
  "title": "recvtz",
```

```
"type": "array",
"items": {
  "anyOf": [
    {
      "type": "string",
      "examples": [
        "EST"
      ]
    },
    {
      "type": "number",
      "examples": [
        -18000
      ]
    }
  ]
},
"started": {
  "title": "started",
  "type": "string",
  "examples": [
    "2021-11-12 11:04:39.050853"
  ]
},
"textinfo": {
  "title": "textinfo",
  "type": "object",
  "properties": {
    "overtalk": {
      "title": "overtalk",
      "type": "object",
      "properties": {
        "avgtime": {
          "title": "avgtime",
          "type": "number",
          "examples": [
```

```
    0.87
  ],
  "count": {
    "title": "count",
    "type": "array",
    "items": {
      "type": "number",
      "examples": [
        2
      ]
    }
  },
  "time": {
    "title": "time",
    "type": "array",
    "items": {
      "type": "number",
      "examples": [
        1.74
      ]
    }
  }
},
"additionalProperties": true,
"required": [
  "avgtime",
  "count",
  "time"
],
"wordtime": {
  "title": "wordtime",
  "type": "array",
  "items": {
    "type": "number",
    "examples": [
```

```
    51.36
  ]
}
},
"words": {
  "title": "words",
  "type": "number",
  "examples": [
    241
  ]
},
"silence": {
  "title": "silence",
  "type": "array",
  "items": {
    "type": "number",
    "examples": [
      27.12
    ]
  }
},
"turns": {
  "title": "turns",
  "type": "number",
  "examples": [
    13
  ]
},
"additionalProperties": true,
"required": [
  "overtalk",
  "wordtime",
  "words",
  "silence",
  "turns"
]
```

```
},
"recvdate": {
  "title": "recvdate",
  "type": "string",
  "examples": [
    "2021-11-12 11:04:39.223285"
  ]
},
"license": {
  "title": "license",
  "type": "string",
  "examples": [
    "string"
  ]
},
"emotion": {
  "title": "emotion",
  "type": "string",
  "examples": [
    "Improving"
  ]
},
"lidinfo": {
  "title": "lidinfo",
  "type": "object",
  "properties": {
    "speech": {
      "title": "speech",
      "type": "number",
      "examples": [
        12.67
      ]
    }
  }
},
"conf": {
  "title": "conf",
  "type": "number",
  "examples": [
```

```
    1
    ]
  },
  "lang": {
    "title": "lang",
    "type": "string",
    "examples": [
      "eng"
    ]
  },
  "additionalProperties": true,
  "required": [
    "speech",
    "conf",
    "lang"
  ]
},
"nchannels": {
  "title": "nchannels",
  "type": "number",
  "examples": [
    2
  ]
},
"requestid": {
  "title": "requestid",
  "type": "string",
  "examples": [
    "6efe9038-962f-4cdb-8ac4-7876809fae94"
  ]
},
"ended": {
  "title": "ended",
  "type": "string",
  "examples": [
    "2021-11-12 11:04:44.377419"
  ]
}
```

```
    ]
  },
  "additionalProperties": true,
  "required": [
    "client_data",
    "sentiment",
    "asr",
    "source",
    "donedate",
    "streamtags",
    "sentiment_scores",
    "model",
    "app_data",
    "audiosecs",
    "last_modified",
    "utterances",
    "nsubs",
    "confidence",
    "recvtz",
    "started",
    "textinfo",
    "recvdate",
    "license",
    "emotion",
    "nchannels",
    "requestid",
    "ended"
  ]
}
```

JSON Output Example

The following output example shows JSON-formatted transcription output from a V-Spark 4.2.0-1 system configured to use V-Blaze 7.3.0-1 for ASR transcription.


```
{
  "emotion": "Improving",
  "done_date": "2021-11-12 11:04:44.367598",
  "recvtz": [
    "EST",
    -18000
  ],
  "ended": "2021-11-12 11:04:44.377419",
  "client_data": {
    "audio_duration": 78.48,
    "country": "USA",
    "region": "SELA",
    "jm_version": "None",
    "audio_properties": "PCM, CBR, 256 kb/s, 8000 Hz, 16 bits, 2 ch",
    "performance_stats": {
      "job_latency": 0.0767,
      "job_end": "2021-11-12_16:04:42 +0000",
      "asr_latency": 0.0717,
      "asr_end": "2021-11-12_16:04:42 +0000",
      "job_start": "2021-11-12_16:04:36 +0000",
      "asr_start": "2021-11-12_16:04:36 +0000"
    },
    "brand": "ACME"
  },
  "app_data": {
    "silence": "0.342",
    "agent_channel": 0,
    "agent_clarity": "0.845",
    "agent_emotion": "Negative",
    "client_emotion": "Positive",
    "overall_emotion": "Negative",
    "client_gender": "male",
    "datetime": "2021-11-12 16:04:39 UTC",
    "scorecard": {},
    "client_clarity": "0.824",
    "overtalk": "0.143",
    "url": "http://ip-10-253-150-159.us-east-2.compute.internal:3000/fileDetails/col-orgA/folder0/"
  }
}
```

```
JP71FH0BG4pTu3xoE2Wz/1/audio.wav",
  "words": 241,
  "client_channel": 1,
  "tId": 1,
  "agent_gender": "female",
  "duration": "0:01:18"
},
"textinfo": {
  "turns": 13,
  "wordtime": [
    51.36,
    0.654
  ],
  "overtalk": {
    "count": [
      2,
      0.154
    ],
    "avgtime": 0.87,
    "time": [
      1.74,
      0.034
    ]
  },
  "silence": [
    27.12,
    0.346
  ],
  "words": 241
},
"confidence": 0.84,
"sentiment": "Negative",
"source": "audio.wav",
"lidinfo": {
  "lang": "eng",
  "speech": 12.67,
  "conf": 1.0
}
```

```
},
"audiosecs": 78.48,
"recvdate": "2021-11-12 11:04:39.223285",
"asr": "7.3.0-1",
"nchannels": 2,
"started": "2021-11-12 11:04:39.050853",
"streamtags": {
  "emotion": "xa",
  "lid": true,
  "subst_rules": "<10 chars>",
  "billing": "col-orgA-folder0",
  "gender": true,
  "rawemotion": true,
  "substinfo": true,
  "token": "",
  "music": "info",
  "scrubtext": true,
  "datahdr": "WAVE",
  "model": [
    "eng1:callcenter"
  ],
  "agentid": true,
  "nsubs": "true"
},
"substinfo": {
  "counts": [
    [
      "subst_rules",
      2,
      {
        "yeah => yes": 2
      }
    ],
    [
      "numtrans",
      2,
      {
```

```
    "four => 4": 1,  
    "last 4 or => last four or": 1  
  }  
],  
  "nsubs": 4  
},  
"chaninfo": [  
  {  
    "emotion": "Improving",  
    "textinfo": {  
      "turns": 7,  
      "wordtime": [  
        38.04,  
        0.485  
      ],  
      "overtalk": {  
        "count": [  
          1,  
          0.143  
        ],  
        "avgtime": 0.48,  
        "time": [  
          0.48,  
          0.013  
        ]  
      },  
      "silence": [  
        40.44,  
        0.515  
      ],  
      "words": 175  
    },  
    "agentscore": 0.38  
  },  
  {  
    "emotion": "Neutral",
```

```
    "textinfo": {
      "turns": 6,
      "wordtime": [
        15.06,
        0.192
      ],
      "overtalk": {
        "count": [
          1,
          0.167
        ],
        "avgtime": 1.26,
        "time": [
          1.26,
          0.084
        ]
      },
      "silence": [
        63.42,
        0.808
      ],
      "words": 66
    },
    "agentscore": -0.0
  }
],
"requestid": "6efe9038-962f-4cdb-8ac4-7876809fae94",
"sentiment_scores": [
  0,
  2
],
"nsubs": 2,
"utterances": [
  {
    "emotion": "Negative",
    "confidence": 0.86,
    "rawemotion": "ANGRY",
```

```
"doneDate": "2021-11-12 11:04:44.325363",
"sentiment": "Negative",
"recvtz": [
  "EST",
  -18000
],
"gender": "female",
"sentimentex": [
  [
    0,
    2
  ],
  [
    [
      "-",
      1,
      [
        7,
        59
      ]
    ]
  ]
],
"start": 0.41,
"musicinfo": {
  "score": -0.8,
  "used": 12.67
},
"end": 20.21,
"recvdate": "2021-11-12 11:04:39.223285",
"events": [
  {
    "start": 0.41,
    "confidence": 0.86,
    "end": 0.68,
    "word": "So"
  }
],
```

```
{
  "start": 0.68,
  "confidence": 0.95,
  "end": 0.95,
  "word": "thank"
},
{
  "start": 0.95,
  "confidence": 0.98,
  "end": 1.01,
  "word": "you"
},
{
  "start": 1.01,
  "confidence": 0.97,
  "end": 1.19,
  "word": "so"
},
{
  "start": 1.19,
  "confidence": 0.98,
  "end": 1.37,
  "word": "much"
},
{
  "wordex": "for(4)",
  "confidence": 0.62,
  "end": 1.49,
  "word": "for",
  "start": 1.37
},
{
  "start": 1.49,
  "confidence": 0.81,
  "end": 1.85,
  "word": "holding"
},
}
```

```
{
  "start": 1.85,
  "confidence": 0.7,
  "end": 1.91,
  "word": "I"
},
{
  "start": 1.91,
  "confidence": 0.93,
  "end": 2.48,
  "word": "apologize"
},
{
  "wordex": "for(4)",
  "confidence": 0.95,
  "end": 2.66,
  "word": "for",
  "start": 2.48
},
{
  "start": 2.66,
  "confidence": 0.98,
  "end": 2.78,
  "word": "the"
},
{
  "start": 2.78,
  "confidence": 0.99,
  "end": 3.2,
  "word": "delay,"
},
{
  "start": 3.83,
  "confidence": 1.0,
  "end": 4.13,
  "word": "yes."
},
}
```



```
{
  "start": 4.13,
  "confidence": 0.98,
  "end": 4.31,
  "word": "The"
},
{
  "start": 4.31,
  "confidence": 0.98,
  "end": 4.67,
  "word": "only"
},
{
  "start": 4.7,
  "confidence": 0.95,
  "end": 4.91,
  "word": "I"
},
{
  "start": 4.91,
  "confidence": 0.82,
  "end": 5.12,
  "word": "will"
},
{
  "start": 5.12,
  "confidence": 0.94,
  "end": 5.33,
  "word": "be"
},
{
  "start": 5.33,
  "confidence": 0.79,
  "end": 5.6,
  "word": "at"
},
{

```

```
    "start": 6.86,  
    "confidence": 0.93,  
    "end": 7.07,  
    "word": "we"  
  },  
  {  
    "wordex": "can(2)",  
    "confidence": 0.84,  
    "end": 7.22,  
    "word": "can",  
    "start": 7.07  
  },  
  {  
    "start": 7.22,  
    "confidence": 0.96,  
    "end": 7.4,  
    "word": "look"  
  },  
  {  
    "start": 7.4,  
    "confidence": 0.9,  
    "end": 7.52,  
    "word": "at"  
  },  
  {  
    "start": 7.52,  
    "confidence": 0.95,  
    "end": 7.61,  
    "word": "up"  
  },  
  {  
    "start": 8.0,  
    "confidence": 0.82,  
    "end": 8.27,  
    "word": "would"  
  },  
  {  
    "start": 8.27,  
    "confidence": 0.97,  
    "end": 8.4,  
    "word": "and"  
  }  
]
```

```
    "start": 8.27,  
    "confidence": 0.92,  
    "end": 8.45,  
    "word": "be"  
  },  
  {  
    "wordex": "your(2)",  
    "confidence": 0.8,  
    "end": 9.14,  
    "word": "your",  
    "start": 8.6  
  },  
  {  
    "start": 9.2,  
    "confidence": 0.83,  
    "end": 9.62,  
    "word": "anti,"  
  },  
  {  
    "wordex": "is(3)",  
    "confidence": 0.55,  
    "end": 9.74,  
    "word": "is",  
    "start": 9.62  
  },  
  {  
    "wordex": "has(2)",  
    "confidence": 0.66,  
    "end": 9.95,  
    "word": "has",  
    "start": 9.74  
  },  
  {  
    "start": 9.95,  
    "confidence": 0.8,  
    "end": 10.04,  
    "word": "a"
```

```
},  
{  
  "start": 10.04,  
  "confidence": 0.78,  
  "end": 10.37,  
  "word": "cutie"  
},  
{  
  "start": 10.37,  
  "confidence": 0.98,  
  "end": 10.7,  
  "word": "number,"  
},  
{  
  "start": 10.7,  
  "confidence": 0.97,  
  "end": 10.88,  
  "word": "but"  
},  
{  
  "start": 10.88,  
  "confidence": 0.98,  
  "end": 10.94,  
  "word": "I"  
},  
{  
  "start": 10.94,  
  "confidence": 0.97,  
  "end": 11.21,  
  "word": "can't"  
},  
{  
  "start": 11.21,  
  "confidence": 0.94,  
  "end": 11.36,  
  "word": "look"  
},  
},
```

```
{
  "start": 11.36,
  "confidence": 0.65,
  "end": 11.42,
  "word": "it"
},
{
  "start": 11.42,
  "confidence": 0.97,
  "end": 11.51,
  "word": "up"
},
{
  "wordex": "with(2)",
  "confidence": 0.84,
  "end": 11.63,
  "word": "with",
  "start": 11.51
},
{
  "start": 11.63,
  "confidence": 0.98,
  "end": 11.72,
  "word": "the"
},
{
  "wordex": "last(3)",
  "confidence": 0.98,
  "end": 12.05,
  "word": "last",
  "start": 11.72
},
{
  "start": 12.05,
  "confidence": 0.78,
  "end": 12.23,
  "word": "four"
}
```

```
},
{
  "wordex": "or(2)",
  "confidence": 0.52,
  "end": 12.29,
  "word": "or",
  "start": 12.23
},
{
  "start": 12.29,
  "confidence": 0.96,
  "end": 12.65,
  "word": "anything"
},
{
  "start": 12.65,
  "confidence": 0.94,
  "end": 12.83,
  "word": "like"
},
{
  "start": 12.83,
  "confidence": 0.6,
  "end": 13.01,
  "word": "that"
},
{
  "wordex": "cause(2)",
  "confidence": 0.58,
  "end": 13.22,
  "word": "cause",
  "start": 13.01
},
{
  "start": 13.22,
  "confidence": 0.78,
  "end": 13.43,
```

```
    "word": "this"
  },
  {
    "start": 13.43,
    "confidence": 0.91,
    "end": 13.58,
    "word": "is"
  },
  {
    "start": 13.58,
    "confidence": 0.75,
    "end": 13.7,
    "word": "a"
  },
  {
    "start": 13.91,
    "confidence": 0.77,
    "end": 14.36,
    "word": "closed"
  },
  {
    "start": 14.36,
    "confidence": 0.98,
    "end": 14.72,
    "word": "account"
  },
  {
    "start": 14.72,
    "confidence": 0.57,
    "end": 14.9,
    "word": "or"
  },
  {
    "start": 14.9,
    "confidence": 0.87,
    "end": 15.2,
    "word": "maybe"
  }
}
```

```
},
{
  "start": 15.2,
  "confidence": 0.81,
  "end": 15.32,
  "word": "if"
},
{
  "start": 15.35,
  "confidence": 0.79,
  "end": 15.5,
  "word": "it"
},
{
  "start": 15.5,
  "confidence": 0.52,
  "end": 15.74,
  "word": "isn't"
},
{
  "start": 15.74,
  "confidence": 0.71,
  "end": 16.07,
  "word": "written"
},
{
  "start": 16.07,
  "confidence": 0.67,
  "end": 16.25,
  "word": "out"
},
{
  "start": 16.25,
  "confidence": 0.91,
  "end": 16.31,
  "word": "the"
},
},
```



```
{
  "start": 16.31,
  "confidence": 0.97,
  "end": 16.67,
  "word": "account"
},
{
  "start": 17.06,
  "confidence": 0.71,
  "end": 17.21,
  "word": "the"
},
{
  "start": 17.21,
  "confidence": 0.88,
  "end": 17.42,
  "word": "only"
},
{
  "start": 17.42,
  "confidence": 0.89,
  "end": 17.63,
  "word": "way"
},
{
  "start": 17.63,
  "confidence": 0.7,
  "end": 17.81,
  "word": "it's"
},
{
  "wordex": "gonna(3)",
  "confidence": 0.91,
  "end": 18.02,
  "word": "gonna",
  "start": 17.81
},
}
```

```
{
  "start": 18.02,
  "confidence": 0.97,
  "end": 18.17,
  "word": "be"
},
{
  "start": 18.17,
  "confidence": 0.88,
  "end": 18.53,
  "word": "pulled"
},
{
  "start": 18.53,
  "confidence": 0.93,
  "end": 18.68,
  "word": "up"
},
{
  "start": 18.68,
  "confidence": 0.73,
  "end": 18.77,
  "word": "is"
},
{
  "start": 18.77,
  "confidence": 0.82,
  "end": 18.92,
  "word": "through"
},
{
  "start": 18.92,
  "confidence": 0.84,
  "end": 19.01,
  "word": "the"
},
{

```

```
    "start": 19.01,  
    "confidence": 0.85,  
    "end": 19.52,  
    "word": "collections"  
  },  
  {  
    "start": 19.52,  
    "confidence": 0.89,  
    "end": 20.21,  
    "word": "department."  
  }  
],  
"metadata": {  
  "uttid": 0,  
  "substinfo": {  
    "subs": [  
      [  
        3.83,  
        4.13,  
        [  
          {  
            "source": "subst_rules",  
            "end": 4.13,  
            "sub": "yeah => yes",  
            "rule": "0",  
            "start": 3.83  
          }  
        ]  
      ]  
    ],  
    [  
      11.72,  
      12.29,  
      [  
        {  
          "source": "numtrans",  
          "end": 12.29,  
          "sub": "last 4 or => last four or",
```

```
        "rule": "43",
        "start": 11.72
      },
      {
        "source": "numtrans",
        "end": 12.23,
        "sub": "four => 4",
        "rule": "0",
        "start": 12.05
      }
    ]
  ],
  "nsubs": 3
},
"source": "audio.wav",
"channel": 0,
"model": "en1:callcenter",
"nsubs": 1
}
},
{
  "emotion": "Neutral",
  "confidence": 0.0,
  "end": 8.2,
  "recvtz": [
    "EST",
    -18000
  ],
  "gender": "male",
  "rawemotion": "NEUTRAL",
  "start": 6.8,
  "musicinfo": {
    "score": 0.24,
    "used": 0.26
  },
  "done": "2021-11-12 11:04:44.325824",

```

```
"recvdate": "2021-11-12 11:04:39.156901",
"events": [],
"metadata": {
  "source": "audio.wav",
  "model": "en1:callcenter",
  "uttid": 0,
  "channel": 1
}
},
{
  "emotion": "Neutral",
  "confidence": 0.87,
  "rawemotion": "HAPPY",
  "done": "2021-11-12 11:04:44.326578",
  "sentiment": "Neutral",
  "recvtz": [
    "EST",
    -18000
  ],
  "gender": "male",
  "sentimentex": [
    [
      0,
      0
    ],
    []
  ],
  "start": 21.91,
  "musicinfo": {
    "score": -0.65,
    "used": 2.44
  },
  "end": 25.78,
  "recvdate": "2021-11-12 11:04:39.227284",
  "events": [
    {
      "start": 21.91,
```

```
    "confidence": 0.98,  
    "end": 22.57,  
    "word": "Okay,"  
  },  
  {  
    "start": 23.17,  
    "confidence": 0.68,  
    "end": 23.44,  
    "word": "he"  
  },  
  {  
    "start": 23.44,  
    "confidence": 0.55,  
    "end": 23.65,  
    "word": "say,"  
  },  
  {  
    "start": 23.65,  
    "confidence": 0.84,  
    "end": 23.74,  
    "word": "I"  
  },  
  {  
    "start": 23.74,  
    "confidence": 0.87,  
    "end": 24.01,  
    "word": "will"  
  },  
  {  
    "start": 24.01,  
    "confidence": 0.97,  
    "end": 24.16,  
    "word": "be"  
  },  
  {  
    "start": 24.16,  
    "confidence": 0.92,
```

```
    "end": 24.31,  
    "word": "able"  
  },  
  {  
    "wordex": "to(3)",  
    "confidence": 0.94,  
    "end": 24.4,  
    "word": "to",  
    "start": 24.31  
  },  
  {  
    "start": 24.4,  
    "confidence": 0.98,  
    "end": 24.58,  
    "word": "take"  
  },  
  {  
    "start": 24.58,  
    "confidence": 0.98,  
    "end": 24.64,  
    "word": "a"  
  },  
  {  
    "start": 24.64,  
    "confidence": 0.98,  
    "end": 24.82,  
    "word": "look"  
  },  
  {  
    "start": 24.82,  
    "confidence": 0.89,  
    "end": 24.94,  
    "word": "at"  
  },  
  {  
    "start": 24.94,  
    "confidence": 0.99,  
  }  
}
```

```
    "end": 25.03,
    "word": "it"
  },
  {
    "wordex": "with(2)",
    "confidence": 0.96,
    "end": 25.15,
    "word": "with",
    "start": 25.03
  },
  {
    "start": 25.15,
    "confidence": 0.71,
    "end": 25.3,
    "word": "my"
  },
  {
    "start": 25.3,
    "confidence": 0.82,
    "end": 25.78,
    "word": "social."
  }
],
"metadata": {
  "source": "audio.wav",
  "model": "enl:callcenter",
  "uttid": 1,
  "channel": 1
},
"emotion": "Neutral",
"confidence": 0.83,
"rawemotion": "ANGRY",
"doneDate": "2021-11-12 11:04:44.333551",
"sentiment": "Neutral",
"recvtz": [
```



```
    "EST",
    -18000
  ],
  "gender": "female",
  "sentimentex": [
    [
      0,
      0
    ],
    []
  ],
  "start": 26.69,
  "musicinfo": {
    "score": -0.78,
    "used": 1.58
  },
  "end": 29.72,
  "recvdate": "2021-11-12 11:04:39.325637",
  "events": [
    {
      "start": 26.69,
      "confidence": 0.64,
      "end": 27.38,
      "word": "I'm"
    },
    {
      "start": 27.38,
      "confidence": 0.83,
      "end": 27.71,
      "word": "making"
    },
    {
      "start": 27.71,
      "confidence": 0.87,
      "end": 27.95,
      "word": "pull"
    }
  ],
```

```
{
  "start": 27.95,
  "confidence": 0.98,
  "end": 28.04,
  "word": "it"
},
{
  "start": 28.04,
  "confidence": 0.95,
  "end": 28.13,
  "word": "up"
},
{
  "wordex": "with(2)",
  "confidence": 0.96,
  "end": 28.22,
  "word": "with",
  "start": 28.13
},
{
  "start": 28.22,
  "confidence": 0.88,
  "end": 28.31,
  "word": "the"
},
{
  "start": 28.31,
  "confidence": 0.79,
  "end": 28.7,
  "word": "social"
},
{
  "start": 28.7,
  "confidence": 0.98,
  "end": 28.76,
  "word": "I"
},
}
```

```
{
  "wordex": "can(2)",
  "confidence": 0.98,
  "end": 28.94,
  "word": "can",
  "start": 28.76
},
{
  "start": 28.94,
  "confidence": 0.97,
  "end": 29.15,
  "word": "try"
},
{
  "start": 29.15,
  "confidence": 0.81,
  "end": 29.27,
  "word": "the"
},
{
  "start": 29.27,
  "confidence": 0.9,
  "end": 29.72,
  "word": "social."
}
],
"metadata": {
  "source": "audio.wav",
  "model": "en1:callcenter",
  "uttid": 1,
  "channel": 0
}
},
{
  "emotion": "Neutral",
  "confidence": 0.74,
  "rawemotion": "ANGRY",
```

```
"doneDate": "2021-11-12 11:04:44.334116",
"sentiment": "Neutral",
"recvtz": [
  "EST",
  -18000
],
"gender": "male",
"sentimentex": [
  [
    0,
    0
  ],
  []
],
"start": 32.11,
"musicinfo": {
  "score": -0.49,
  "used": 2.56
},
"end": 34.75,
"recvdate": "2021-11-12 11:04:39.366937",
"events": [
  {
    "start": 32.11,
    "confidence": 0.98,
    "end": 32.77,
    "word": "Okay, "
  },
  {
    "start": 33.46,
    "confidence": 0.55,
    "end": 33.64,
    "word": "I'm"
  },
  {
    "wordex": "gonna(3)",
    "confidence": 0.59,
```

```
    "end": 33.76,  
    "word": "gonna",  
    "start": 33.64  
  },  
  {  
    "start": 33.76,  
    "confidence": 0.56,  
    "end": 33.85,  
    "word": "say"  
  },  
  {  
    "start": 33.85,  
    "confidence": 0.76,  
    "end": 34.06,  
    "word": "number"  
  },  
  {  
    "start": 34.06,  
    "confidence": 0.49,  
    "end": 34.18,  
    "word": "I'm"  
  },  
  {  
    "wordex": "gonna(3)",  
    "confidence": 0.63,  
    "end": 34.3,  
    "word": "gonna",  
    "start": 34.18  
  },  
  {  
    "start": 34.3,  
    "confidence": 0.68,  
    "end": 34.54,  
    "word": "check"  
  },  
  {  
    "start": 34.54,
```

```
    "confidence": 0.54,
    "end": 34.75,
    "word": "Mark."
  }
],
"metadata": {
  "source": "audio.wav",
  "model": "en1:callcenter",
  "uttid": 2,
  "channel": 1
}
},
{
  "emotion": "Neutral",
  "confidence": 0.84,
  "rawemotion": "ANGRY",
  "doneat": "2021-11-12 11:04:44.345489",
  "sentiment": "Neutral",
  "recvtz": [
    "EST",
    -18000
  ],
  "gender": "female",
  "sentimentex": [
    [
      0,
      0
    ],
    []
  ],
  "start": 33.22,
  "musicinfo": {
    "score": -0.76,
    "used": 8.79
  },
  "end": 49.03,
  "recvdate": "2021-11-12 11:04:39.463435",
```

```
"events": [  
  {  
    "start": 33.22,  
    "confidence": 0.94,  
    "end": 33.4,  
    "word": "So"  
  },  
  {  
    "start": 33.4,  
    "confidence": 0.85,  
    "end": 33.67,  
    "word": "once"  
  },  
  {  
    "start": 33.67,  
    "confidence": 0.98,  
    "end": 33.79,  
    "word": "I"  
  },  
  {  
    "start": 33.79,  
    "confidence": 0.9,  
    "end": 34.0,  
    "word": "get"  
  },  
  {  
    "start": 34.0,  
    "confidence": 0.95,  
    "end": 34.09,  
    "word": "it"  
  },  
  {  
    "start": 34.12,  
    "confidence": 0.89,  
    "end": 34.27,  
    "word": "if"  
  },  
]
```

```
{
  "start": 34.27,
  "confidence": 0.99,
  "end": 34.39,
  "word": "it"
},
{
  "start": 34.39,
  "confidence": 0.94,
  "end": 35.05,
  "word": "doesn't"
},
{
  "start": 35.05,
  "confidence": 0.94,
  "end": 35.32,
  "word": "but"
},
{
  "start": 35.32,
  "confidence": 0.65,
  "end": 35.38,
  "word": "if"
},
{
  "start": 36.49,
  "confidence": 0.97,
  "end": 36.73,
  "word": "it"
},
{
  "start": 37.36,
  "confidence": 0.81,
  "end": 37.51,
  "word": "if"
},
{
```



```
    "start": 37.51,  
    "confidence": 0.79,  
    "end": 37.66,  
    "word": "up"  
  },  
  {  
    "wordex": "with(2)",  
    "confidence": 0.77,  
    "end": 37.81,  
    "word": "with",  
    "start": 37.66  
  },  
  {  
    "wordex": "you(2)",  
    "confidence": 0.76,  
    "end": 37.93,  
    "word": "you",  
    "start": 37.81  
  },  
  {  
    "start": 37.93,  
    "confidence": 0.74,  
    "end": 38.14,  
    "word": "so"  
  },  
  {  
    "start": 38.14,  
    "confidence": 0.79,  
    "end": 38.23,  
    "word": "if"  
  },  
  {  
    "start": 38.23,  
    "confidence": 0.87,  
    "end": 38.35,  
    "word": "you"  
  },  
}
```

```
{
  "wordex": "have(3)",
  "confidence": 0.68,
  "end": 38.56,
  "word": "have",
  "start": 38.35
},
{
  "start": 38.56,
  "confidence": 0.9,
  "end": 38.59,
  "word": "the"
},
{
  "start": 38.59,
  "confidence": 0.82,
  "end": 38.92,
  "word": "account"
},
{
  "start": 38.95,
  "confidence": 0.78,
  "end": 39.1,
  "word": "it"
},
{
  "wordex": "is(3)",
  "confidence": 0.69,
  "end": 39.25,
  "word": "is",
  "start": 39.1
},
{
  "start": 39.25,
  "confidence": 0.93,
  "end": 39.31,
  "word": "a"
}
```

```
},
{
  "start": 39.31,
  "confidence": 0.94,
  "end": 39.52,
  "word": "written"
},
{
  "start": 39.52,
  "confidence": 0.64,
  "end": 39.61,
  "word": "out."
},
{
  "start": 39.61,
  "confidence": 0.77,
  "end": 39.91,
  "word": "Okay,"
},
{
  "wordex": "probably(2)",
  "confidence": 0.67,
  "end": 40.21,
  "word": "probably",
  "start": 39.91
},
{
  "start": 40.21,
  "confidence": 0.81,
  "end": 40.33,
  "word": "not"
},
{
  "wordex": "gonna(3)",
  "confidence": 0.93,
  "end": 40.51,
  "word": "gonna",

```

```
    "start": 40.33
  },
  {
    "start": 40.51,
    "confidence": 0.97,
    "end": 40.72,
    "word": "come"
  },
  {
    "start": 40.72,
    "confidence": 0.97,
    "end": 40.87,
    "word": "up"
  },
  {
    "wordex": "for(4)",
    "confidence": 0.94,
    "end": 41.02,
    "word": "for",
    "start": 40.87
  },
  {
    "start": 41.02,
    "confidence": 0.95,
    "end": 41.2,
    "word": "me"
  },
  {
    "start": 41.2,
    "confidence": 0.82,
    "end": 41.65,
    "word": "anyway,"
  },
  {
    "start": 41.65,
    "confidence": 0.97,
    "end": 41.8,
```

```
    "word": "but"
  },
  {
    "start": 41.8,
    "confidence": 0.99,
    "end": 41.86,
    "word": "I"
  },
  {
    "wordex": "can(2)",
    "confidence": 0.97,
    "end": 42.07,
    "word": "can",
    "start": 41.86
  },
  {
    "start": 42.07,
    "confidence": 0.97,
    "end": 42.4,
    "word": "try"
  },
  {
    "start": 42.4,
    "confidence": 0.98,
    "end": 42.52,
    "word": "it"
  },
  {
    "start": 42.85,
    "confidence": 0.81,
    "end": 43.09,
    "word": "would"
  },
  {
    "start": 43.09,
    "confidence": 0.98,
    "end": 43.18,
```

```
    "word": "it"
  },
  {
    "start": 43.18,
    "confidence": 0.94,
    "end": 43.3,
    "word": "be"
  },
  {
    "start": 43.3,
    "confidence": 0.79,
    "end": 43.45,
    "word": "the"
  },
  {
    "start": 43.45,
    "confidence": 0.67,
    "end": 43.66,
    "word": "reason"
  },
  {
    "start": 43.66,
    "confidence": 0.53,
    "end": 43.81,
    "word": "our"
  },
  {
    "start": 43.81,
    "confidence": 0.76,
    "end": 44.11,
    "word": "accounts"
  },
  {
    "start": 44.11,
    "confidence": 0.74,
    "end": 44.26,
    "word": "I'm"
```

```
},
{
  "start": 44.26,
  "confidence": 0.93,
  "end": 44.44,
  "word": "still"
},
{
  "wordex": "gonna(2)",
  "confidence": 0.91,
  "end": 44.62,
  "word": "gonna",
  "start": 44.44
},
{
  "wordex": "have(2)",
  "confidence": 0.98,
  "end": 44.71,
  "word": "have",
  "start": 44.62
},
{
  "wordex": "to(2)",
  "confidence": 0.93,
  "end": 44.8,
  "word": "to",
  "start": 44.71
},
{
  "wordex": "get(2)",
  "confidence": 0.97,
  "end": 44.92,
  "word": "get",
  "start": 44.8
},
{
  "start": 44.92,
```

```
    "confidence": 0.98,  
    "end": 45.01,  
    "word": "you"  
  },  
  {  
    "start": 45.01,  
    "confidence": 0.98,  
    "end": 45.16,  
    "word": "over"  
  },  
  {  
    "wordex": "to(3)",  
    "confidence": 0.72,  
    "end": 45.22,  
    "word": "to",  
    "start": 45.16  
  },  
  {  
    "start": 45.22,  
    "confidence": 0.57,  
    "end": 45.37,  
    "word": "that"  
  },  
  {  
    "start": 45.37,  
    "confidence": 0.87,  
    "end": 45.85,  
    "word": "department"  
  },  
  {  
    "wordex": "because(3)",  
    "confidence": 0.97,  
    "end": 46.24,  
    "word": "because",  
    "start": 45.85  
  },  
  {
```



```
    "start": 46.72,  
    "confidence": 0.89,  
    "end": 47.05,  
    "word": "this"  
  },  
  {  
    "start": 47.2,  
    "confidence": 0.56,  
    "end": 47.41,  
    "word": "out"  
  },  
  {  
    "start": 47.41,  
    "confidence": 0.94,  
    "end": 47.47,  
    "word": "of"  
  },  
  {  
    "start": 47.47,  
    "confidence": 0.53,  
    "end": 47.77,  
    "word": "mine"  
  },  
  {  
    "wordex": "is(3)",  
    "confidence": 0.55,  
    "end": 47.98,  
    "word": "is",  
    "start": 47.77  
  },  
  {  
    "start": 48.01,  
    "confidence": 0.83,  
    "end": 48.31,  
    "word": "not"  
  },  
  {
```

```
    "start": 48.31,
    "confidence": 0.89,
    "end": 48.49,
    "word": "my"
  },
  {
    "start": 48.49,
    "confidence": 0.9,
    "end": 49.03,
    "word": "area."
  }
],
"metadata": {
  "source": "audio.wav",
  "model": "enl:callcenter",
  "uttid": 2,
  "channel": 0
}
},
{
  "emotion": "Neutral",
  "confidence": 0.88,
  "rawemotion": "ANGRY",
  "done": "2021-11-12 11:04:44.346109",
  "sentiment": "Neutral",
  "recvtz": [
    "EST",
    -18000
  ],
  "gender": "female",
  "sentimentex": [
    [
      0,
      0
    ]
  ],
  []
],
```

```
"start": 50.45,
"musicinfo": {
  "score": -0.62,
  "used": 1.81
},
"end": 52.73,
"recvdate": "2021-11-12 11:04:39.630183",
"events": [
  {
    "start": 50.45,
    "confidence": 0.98,
    "end": 50.99,
    "word": "Okay."
  },
  {
    "start": 51.14,
    "confidence": 0.81,
    "end": 51.29,
    "word": "Now"
  },
  {
    "start": 51.29,
    "confidence": 0.92,
    "end": 51.44,
    "word": "do"
  },
  {
    "start": 51.44,
    "confidence": 0.98,
    "end": 51.62,
    "word": "you"
  },
  {
    "start": 51.62,
    "confidence": 0.92,
    "end": 51.89,
    "word": "guys"
  }
]
```

```
    },
    {
      "start": 51.89,
      "confidence": 0.84,
      "end": 52.1,
      "word": "have"
    },
    {
      "start": 52.1,
      "confidence": 0.69,
      "end": 52.37,
      "word": "caller"
    },
    {
      "start": 52.37,
      "confidence": 0.8,
      "end": 52.73,
      "word": "ID."
    }
  ],
  "metadata": {
    "source": "audio.wav",
    "model": "enl:callcenter",
    "uttid": 3,
    "channel": 1
  }
},
{
  "emotion": "Neutral",
  "confidence": 0.74,
  "rawemotion": "ANGRY",
  "done": "2021-11-12 11:04:44.356009",
  "sentiment": "Neutral",
  "recvtz": [
    "EST",
    -18000
  ]
},
]
```

```
"gender": "female",
"sentimentex": [
  [
    0,
    0
  ],
  []
],
"start": 54.27,
"musicinfo": {
  "score": -0.59,
  "used": 2.98
},
"end": 58.38,
"recvdate": "2021-11-12 11:04:39.773581",
"events": [
  {
    "start": 54.27,
    "confidence": 0.13,
    "end": 55.05,
    "word": "Um,"
  },
  {
    "start": 55.38,
    "confidence": 1.0,
    "end": 55.74,
    "word": "yes,"
  },
  {
    "wordex": "your(2)",
    "confidence": 0.68,
    "end": 55.98,
    "word": "your",
    "start": 55.74
  },
  {
    "start": 55.98,
```

```
    "confidence": 0.69,  
    "end": 56.25,  
    "word": "number"  
  },  
  {  
    "start": 56.25,  
    "confidence": 0.78,  
    "end": 56.46,  
    "word": "came"  
  },  
  {  
    "start": 56.46,  
    "confidence": 0.97,  
    "end": 56.58,  
    "word": "up"  
  },  
  {  
    "wordex": "on(2)",  
    "confidence": 0.97,  
    "end": 56.73,  
    "word": "on",  
    "start": 56.58  
  },  
  {  
    "start": 56.73,  
    "confidence": 0.95,  
    "end": 56.94,  
    "word": "my"  
  },  
  {  
    "start": 56.94,  
    "confidence": 0.95,  
    "end": 57.27,  
    "word": "phone."  
  },  
  {  
    "start": 57.27,
```

```
    "confidence": 0.98,
    "end": 57.33,
    "word": "I"
  },
  {
    "wordex": "can(2)",
    "confidence": 0.98,
    "end": 57.48,
    "word": "can",
    "start": 57.33
  },
  {
    "start": 57.48,
    "confidence": 0.97,
    "end": 57.66,
    "word": "see"
  },
  {
    "start": 57.66,
    "confidence": 0.94,
    "end": 57.75,
    "word": "it"
  },
  {
    "wordex": "on(2)",
    "confidence": 0.94,
    "end": 57.87,
    "word": "on",
    "start": 57.75
  },
  {
    "start": 57.87,
    "confidence": 0.96,
    "end": 58.02,
    "word": "my"
  },
  {
```

```
    "start": 58.02,  
    "confidence": 0.93,  
    "end": 58.38,  
    "word": "phone."  
  }  
],  
"metadata": {  
  "uttid": 3,  
  "substinfo": {  
    "subs": [  
      [  
        55.38,  
        55.74,  
        [  
          {  
            "source": "subst_rules",  
            "end": 55.74,  
            "sub": "yeah => yes",  
            "rule": "0",  
            "start": 55.38  
          }  
        ]  
      ]  
    ]  
  },  
  "nsubs": 1  
},  
"source": "audio.wav",  
"channel": 0,  
"model": "en1:callcenter",  
"nsubs": 1  
}  
},  
{  
  "emotion": "Neutral",  
  "confidence": 0.79,  
  "rawemotion": "HAPPY",  
  "doneDate": "2021-11-12 11:04:44.357116",
```



```
"sentiment": "Neutral",
"recvtz": [
  "EST",
  -18000
],
"gender": "male",
"sentimentex": [
  [
    0,
    0
  ],
  []
],
"start": 59.74,
"musicinfo": {
  "score": -0.56,
  "used": 5.34
},
"end": 66.34,
"recvdate": "2021-11-12 11:04:40.471104",
"events": [
  {
    "start": 59.74,
    "confidence": 0.76,
    "end": 59.86,
    "word": "There"
  },
  {
    "wordex": "was(2)",
    "confidence": 0.95,
    "end": 60.04,
    "word": "was",
    "start": 59.86
  },
  {
    "start": 60.04,
    "confidence": 0.94,
```

```
    "end": 60.31,  
    "word": "more"  
  },  
  {  
    "wordex": "than(2)",  
    "confidence": 0.98,  
    "end": 60.46,  
    "word": "than",  
    "start": 60.31  
  },  
  {  
    "start": 60.46,  
    "confidence": 0.67,  
    "end": 60.67,  
    "word": "say"  
  },  
  {  
    "start": 60.67,  
    "confidence": 0.57,  
    "end": 60.91,  
    "word": "to"  
  },  
  {  
    "start": 60.91,  
    "confidence": 0.89,  
    "end": 61.3,  
    "word": "assume"  
  },  
  {  
    "wordex": "to(2)",  
    "confidence": 0.85,  
    "end": 61.39,  
    "word": "to",  
    "start": 61.3  
  },  
  {  
    "start": 61.39,
```

```
    "confidence": 0.92,  
    "end": 62.02,  
    "word": "collections"  
  },  
  {  
    "start": 62.02,  
    "confidence": 0.76,  
    "end": 62.32,  
    "word": "would"  
  },  
  {  
    "start": 62.32,  
    "confidence": 0.9,  
    "end": 62.68,  
    "word": "to"  
  },  
  {  
    "start": 63.34,  
    "confidence": 0.1,  
    "end": 63.82,  
    "word": "um,"  
  },  
  {  
    "start": 63.88,  
    "confidence": 0.98,  
    "end": 64.0,  
    "word": "I"  
  },  
  {  
    "wordex": "was(2)",  
    "confidence": 0.96,  
    "end": 64.18,  
    "word": "was",  
    "start": 64.0  
  },  
  {  
    "wordex": "gonna(3)",
```

```
    "confidence": 0.72,  
    "end": 64.39,  
    "word": "gonna",  
    "start": 64.18  
  },  
  {  
    "start": 64.39,  
    "confidence": 0.64,  
    "end": 64.57,  
    "word": "write"  
  },  
  {  
    "wordex": "them(2)",  
    "confidence": 0.62,  
    "end": 64.66,  
    "word": "them",  
    "start": 64.57  
  },  
  {  
    "start": 64.66,  
    "confidence": 0.78,  
    "end": 64.72,  
    "word": "a"  
  },  
  {  
    "start": 64.72,  
    "confidence": 0.97,  
    "end": 65.02,  
    "word": "letter"  
  },  
  {  
    "start": 65.02,  
    "confidence": 0.9,  
    "end": 65.14,  
    "word": "I'll"  
  },  
  {  
  }
```

```
    "start": 65.14,  
    "confidence": 0.96,  
    "end": 65.32,  
    "word": "see"  
  },  
  {  
    "start": 65.32,  
    "confidence": 0.99,  
    "end": 65.47,  
    "word": "how"  
  },  
  {  
    "start": 65.47,  
    "confidence": 0.95,  
    "end": 65.68,  
    "word": "that"  
  },  
  {  
    "start": 65.68,  
    "confidence": 0.76,  
    "end": 66.34,  
    "word": "works."  
  }  
],  
  "metadata": {  
    "source": "audio.wav",  
    "model": "en1:callcenter",  
    "uttid": 4,  
    "channel": 1  
  }  
},  
{  
  "emotion": "Neutral",  
  "confidence": 0.98,  
  "rawemotion": "HAPPY",  
  "doneat": "2021-11-12 11:04:44.366438",  
  "sentiment": "Neutral",
```

```
"recvtz": [
  "EST",
  -18000
],
"gender": "female",
"sentimentex": [
  [
    0,
    0
  ],
  []
],
"start": 66.89,
"musicinfo": {
  "score": -0.64,
  "used": 0.58
},
"end": 67.4,
"recvdate": "2021-11-12 11:04:40.571069",
"events": [
  {
    "start": 66.89,
    "confidence": 0.98,
    "end": 67.4,
    "word": "Okay."
  }
],
"metadata": {
  "source": "audio.wav",
  "model": "eng1:callcenter",
  "uttid": 4,
  "channel": 0
}
},
{
  "emotion": "Neutral",
  "confidence": 0.89,
```

```
"rawemotion": "ANGRY",
"donedate": "2021-11-12 11:04:44.367031",
"sentiment": "Neutral",
"recvtz": [
  "EST",
  -18000
],
"gender": "male",
"sentimentex": [
  [
    0,
    0
  ],
  []
],
"start": 68.97,
"musicinfo": {
  "score": -0.68,
  "used": 2.11
},
"end": 70.95,
"recvdate": "2021-11-12 11:04:40.798243",
"events": [
  {
    "start": 68.97,
    "confidence": 0.67,
    "end": 69.09,
    "word": "Do"
  },
  {
    "start": 69.09,
    "confidence": 0.79,
    "end": 69.39,
    "word": "this?"
  },
  {
    "start": 69.54,
```

```
    "confidence": 0.97,  
    "end": 69.9,  
    "word": "Okay."  
  },  
  {  
    "start": 69.9,  
    "confidence": 0.9,  
    "end": 70.14,  
    "word": "Thank"  
  },  
  {  
    "start": 70.14,  
    "confidence": 0.98,  
    "end": 70.23,  
    "word": "you"  
  },  
  {  
    "wordex": "for(4)",  
    "confidence": 0.92,  
    "end": 70.35,  
    "word": "for",  
    "start": 70.23  
  },  
  {  
    "wordex": "your(2)",  
    "confidence": 0.98,  
    "end": 70.47,  
    "word": "your",  
    "start": 70.35  
  },  
  {  
    "start": 70.47,  
    "confidence": 0.98,  
    "end": 70.68,  
    "word": "help"  
  },  
  {
```



```
    "start": 70.68,  
    "confidence": 0.86,  
    "end": 70.95,  
    "word": "maam."  
  }  
],  
"metadata": {  
  "source": "audio.wav",  
  "model": "en1:callcenter",  
  "uttid": 5,  
  "channel": 1  
}  
},  
{  
  "emotion": "Neutral",  
  "confidence": 0.98,  
  "rawemotion": "HAPPY",  
  "doneDate": "2021-11-12 11:04:44.367364",  
  "sentiment": "Neutral",  
  "recvtz": [  
    "EST",  
    -18000  
  ],  
  "gender": "female",  
  "sentimentex": [  
    [  
      0,  
      0  
    ],  
    []  
  ],  
  "start": 70.34,  
  "musicinfo": {  
    "score": -0.64,  
    "used": 1.17  
  },  
  "end": 71.6,
```

```
"recvdate": "2021-11-12 11:04:40.851078",
"events": [
  {
    "start": 70.34,
    "confidence": 0.97,
    "end": 70.82,
    "word": "Okay,"
  },
  {
    "start": 71.03,
    "confidence": 0.98,
    "end": 71.42,
    "word": "thank"
  },
  {
    "start": 71.42,
    "confidence": 0.98,
    "end": 71.6,
    "word": "you."
  }
],
"metadata": {
  "source": "audio.wav",
  "model": "enl:callcenter",
  "uttid": 5,
  "channel": 0
},
{
  "emotion": "Neutral",
  "confidence": 0.0,
  "end": 76.67,
  "recvtz": [
    "EST",
    -18000
  ],
  "gender": "male",
```

```
    "rawemotion": "NEUTRAL",
    "start": 72.05,
    "musicinfo": {
      "score": -0.17,
      "used": 1.58
    },
    "done date": "2021-11-12 11:04:44.367598",
    "recv date": "2021-11-12 11:04:40.965757",
    "events": [],
    "metadata": {
      "source": "audio.wav",
      "model": "en1:callcenter",
      "uttid": 6,
      "channel": 1
    }
  ],
  "license": "string",
  "model": [
    "en1:callcenter"
  ],
  "last_modified": "2021-11-12 16:04:56"
}
```