



# JSON Output Reference

---

for use with V-Spark

September 2021

## **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

JSON Output Format .....	4
Top-level Elements .....	4
The utterances Array .....	7
The app_data Object .....	10
The scorecard Object .....	13
The client_data Object .....	15
V-Spark JSON Schema .....	19
JSON Output Example .....	27

# JSON Output Format

All Voci products use the JSON file format to store the text information in an organized fashion to provide the most content that can be derived from the audio source. Some products may have additional JSON elements that are not included in the base JSON structure.



## 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.

## Top-level Elements

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



## NOTE

Metrics such as clarity, silence, and overtalk are only shown in V-Spark as they are calculated off of data in the V-Blaze JSON output file.

**Table 1. Top-level Elements**

Element	Availability	Type	Definition
emotion	All	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"> <li>• Positive</li> <li>• Mostly Positive</li> <li>• Neutral</li> <li>• Mostly Negative</li> <li>• Negative</li> </ul> <p>Emotion must be the same for all utterances to be included at the top level. Additional emotion scoring is available in <a href="#">The utterances Array</a></p>

Element	Availability	Type	Definition
confidence	All	value	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	value	Acoustic emotion values. Possible values in version 7.1+ include: <ul style="list-style-type: none"> <li>• ANGRY</li> <li>• NEUTRAL</li> <li>• HAPPY</li> </ul> Acoustic emotion values prior to version 7.1 include: <ul style="list-style-type: none"> <li>• NONANGRY</li> <li>• ANGRY</li> </ul>
doneDate	All	value	Date and time the file transcription was completed by the speech-to-text engine, meaning the last utterance finished.
<b>utterances</b>	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	value	Identification information for the license used.
nchannels	All	value	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.
scrubbed	All	value	If true then audio is purified so numbers are all redacted. If false, the data name does not appear in the JSON output.

Element	Availability	Type	Definition
sentiment	All	value	Linguistic sentiment value: <ul style="list-style-type: none"> <li>• Positive</li> <li>• Mostly Positive</li> <li>• Neutral</li> <li>• Mostly Negative</li> <li>• Negative</li> <li>• Mixed (contains both Positive and Negative in the file)</li> </ul>
sentiment_scores	All	array	Array of length 2. [0]=Positive phrase counts and [1]=Negative phrase counts in the file
source	All	value	The audio file name.
gender	All	value	The gender identified for the audio.
model	All	<b>string</b> containing model name if one model was specified;  <b>array</b> of model names if multiple models were specified	Language model(s) specified for transcription. For example:  <pre>"model": "en1:callcenter"</pre> As of V-Blaze 7.2, this element will not appear in the JSON output if there was a problem processing audio.
recvdate	All	value	Date and time the audio file was received by the ASR engine and placed in queue
requestid	All	value	The unique identifier for the request.
<a href="#">client_data</a>	V-Spark only	object	A JSON object that stores call metadata associated with the audio file.
<a href="#">app_data</a>	V-Spark only	object	A JSON object that stores metadata generated by V-Spark.

## The utterances Array

The top-level **utterances** element contains an array of speech segment information. Each utterances array consists of the following elements.



### NOTE

Metrics such as clarity, silence, and overtalk are only shown in V-Spark as they are calculated off of data in the V-Blaze JSON output file.

**Table 2. Data contained in the utterances Array**

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

Object	Availability	Type	Definition
			<ul style="list-style-type: none"> <li>• Mostly Negative</li> <li>• Negative</li> <li>• Mixed (contains both Positive and Negative in the file)</li> </ul>
gender	All	value	Gender prediction of the speaker
rawemotion	All	value	<p>Acoustic emotion values (version 7.1+):</p> <ul style="list-style-type: none"> <li>• ANGRY</li> <li>• NEUTRAL</li> <li>• HAPPY</li> </ul> <p>Acoustic emotion values (prior to version 7.1):</p> <ul style="list-style-type: none"> <li>• NONANGRY</li> <li>• ANGRY</li> </ul>
sentimentex	All	array	<p>Contains sentiment information for each utterance</p> <ul style="list-style-type: none"> <li>• [0][0] = Positive phrase counts</li> <li>• [0][1] = Negative phrase counts in utterance</li> <li>• [1][*] consist of an array of sentiment segments where [0] = '+' or '-' for Positive and Negative, and [1] is the position range of the phrase</li> <li>• [0] is beginning and [1] is end position</li> </ul>
start	All	value	Start time of the utterance in seconds
donedate	All	value	Date and time the utterance transcription was completed by the speech-to-text engine
recvdate	All	value	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"> <li>• <b>confidence</b>: word level transcription confidence value between 0 and 1</li> <li>• <b>end</b>: end time of the word in seconds</li> <li>• <b>start</b>: start time in seconds</li> <li>• <b>word</b>: normalized word.</li> </ul>



Object	Availability	Type	Definition
			<ul style="list-style-type: none"> <li>• <b>wordex:</b> raw dictionary word. This value is often used to disambiguate different pronunciations that have the same spelling.</li> </ul> <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,     "word": "so"   } ]</pre>
metadata	All	object	<p>Speaker information of the utterance. Each object contains the following values:</p> <ul style="list-style-type: none"> <li>• <b>channel:</b> channel number</li> <li>• <b>model:</b> model that decoded the utterance</li> <li>• <b>source:</b> audio file name</li> <li>• <b>nsubs:</b> (V-Blaze 7.1+) count of substitutions applied for the utterance, not including <code>numtrans</code> counts.</li> <li>• <b>uttid:</b> utterance segment number</li> <li>• <b>substinfo:</b> (V-Blaze 7.1+) substitutions detail included when <code>substinfo=true</code> <ul style="list-style-type: none"> <li>• <b>nsubs:</b> (V-Blaze 7.1+) count of substitutions applied for the utterance, including <code>numtrans</code> counts.</li> </ul> </li> </ul> <p>For example:</p> <pre>"metadata": {   "uttid": 1,</pre>


Object	Availability	Type	Definition
			<pre> "substinfo": {   "subs": [     [       19.56,       20.14,       [         {           "source": "subst_rules",           "end": 20.14,           "sub": "persona =&gt; /Persona/",           "rule": "0",           "start": 19.56         }       ]     ]   ],   "nsubs": 1 }, "source": "spa-eng-sample.wav", "channel": 0, "model": "spa3:travel", "nsubs": 1 } "metadata": {   "source": "spa-eng-sample.wav",   "model": "eng1:callcenter",   "uttid": 0,   "channel": 0 } </pre>

## The app\_data Object

Another top-level JSON object is **app\_data**. Metadata in this object applies to the whole call, rather than to individual utterances.

**Table 3. Data contained in the app\_data Object**

Data Value	Definition
agent_channel	The channel the agent is on
agent_clarity	How clear the agent channel/speech is <ul style="list-style-type: none"> <li>• Range between 0 and 1</li> <li>• 1 is clearest.</li> </ul>
agent_emotion	Overall agent emotional intelligence consists of both acoustic and linguistic information and can have the following values: <ul style="list-style-type: none"> <li>• Positive</li> <li>• Improving</li> <li>• Negative</li> <li>• Worsening</li> </ul>
agent_gender	Agent gender prediction
client_clarity	How clear the client channel/speech is <ul style="list-style-type: none"> <li>• Range between 0 and 1</li> <li>• 1 is clearest</li> </ul>
client_emotion	Overall client emotional intelligence consists of both acoustic and linguistic information and can have the following values: <ul style="list-style-type: none"> <li>• Positive</li> <li>• Improving</li> <li>• Negative</li> <li>• Worsening</li> </ul>
client_gender	Client gender prediction
datetime	Transcript date and time. Recorded in UTC
diarization	This value provided in 2 speaker, 1 channel calls. <ul style="list-style-type: none"> <li>• Range between 0 and 1.</li> <li>• 1 is best speaker separation.</li> </ul>

Data Value	Definition
duration	Call duration
last_modified	<p>The date and time at which an update to the <b>last_modified</b> field was last triggered in the Elasticsearch record associated with a transcript. If the <b>last_modified</b> field is not present or has no date and time value, its return value is <b>false</b>.</p> <p>The following events trigger an update to the <b>last_modified</b> field:</p> <ul style="list-style-type: none"> <li>• Creating a new transcript.</li> <li>• Updating transcript scores by reprocessing an application.</li> <li>• Deleting an application or application category associated with the transcript.</li> <li>• Unlinking an application from the transcript's folder, if that application has previously been used to score the transcript.</li> <li>• Updating transcript metadata using the API.</li> </ul> <div style="background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p> <b>NOTE</b></p> <p>The <b>last_modified</b> field was implemented in V-Spark version 4.0.2-1. As a result, <b>last_modified</b> is not included in static JSON transcripts generated by older versions. Dynamically generated JSON output, such as that downloaded from the <b>Files View</b>, includes a <b>last_modified</b> field even if its audio record is from an older version.</p> <p>To add the <b>last_modified</b> 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 <b>last_modified</b> must occur.</p> </div>
overall_emotion	<p>Overall file emotional intelligence consists of both acoustic and linguistic information and can have the following values:</p> <ul style="list-style-type: none"> <li>• Positive</li> <li>• Improving</li> <li>• Negative</li> <li>• Worsening</li> </ul>

Data Value	Definition
overtalk	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.
<a href="#">scorecard</a>	Contains any application scores that have been calculated for the transcript
silence	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 is not decoded to word-events, they would be counted as silence.
tld	Transcript ID number
url	Location where the files can be viewed
words	Total number of words in a transcript

## The scorecard Object

Some JSON files include a **scorecard** object under **app\_data**, as shown in the following screenshot.

This object contains the Leaf, Hit/Miss, and Coverage scores for each category in each application. For more information about how these scores are calculated and when they are included in a JSON output file, refer to the [Application Development Guide](#) for V-Spark.

Figure 1. Example of scorecard Object data

The first level under scorecard contains the name of the application

The application name object contains scores for each top-level category, which also contains a **subcategories** object for each lower-level category

All leaf-level application categories contain a **score** value equivalent to that category's Count score

All top-level application categories contain a **score** value equivalent to that category's Coverage score

## The `client_data` Object

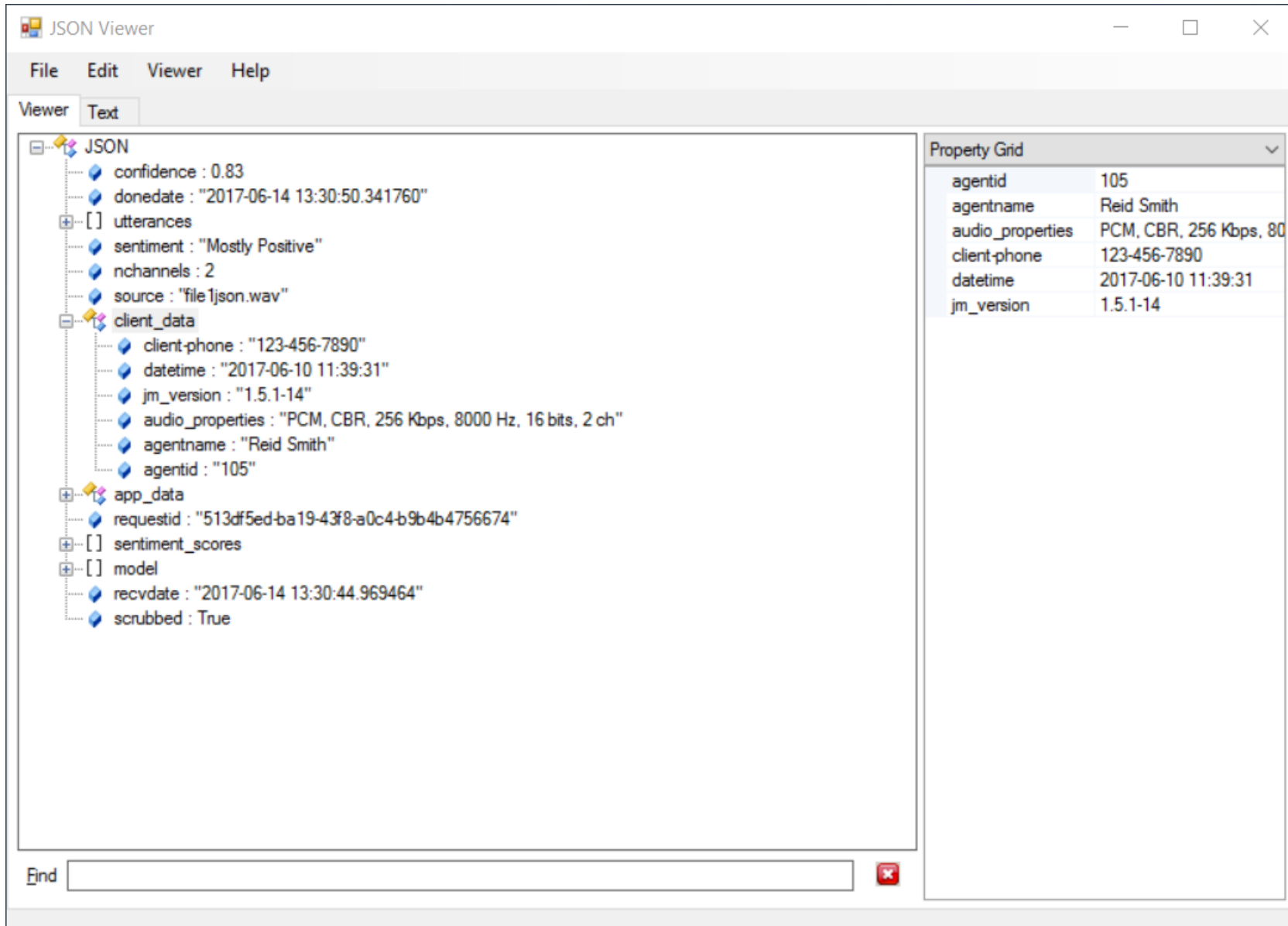
The `client_data` section appears if a file of custom metadata is submitted with an audio file for transcription. The `client_data` section contains one or more instances of name/value pairs that correspond to the additional metadata that was provided.

**NOTE**

This data is only available using V-Spark.

The following screenshot shows an example of the type of information contained in the `client_data` portion of a JSON file:

Figure 2. client\_data Object





The following is an example of the `client_data` section of a JSON transcript:

```
"client_data": {  
  "client-phone": "123-456-7890",  
  "datetime": "2020-01-10 11:39:31",  
  "agentname": "Reid Smith",  
  "agentid": "105"  
},  
S
```

Customer metadata can be provided in either JSON or XML format. The sample `<client_data>` section that was just shown was produced in response to uploading the following customer metadata in a JSON file along with the audio file that was being transcribed:

```
{  
  "metadata": {  
    "datetime": "2020-01-10 11:39:31",  
    "agentid": "105",  
    "agentname": "Reid Smith",  
    "client-phone": "123-456-7890"  
  }  
}
```

The XML equivalent of that customer metadata file would be the following:

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>  
<metadata>  
  <agentid>105</agentid>  
  <datetime>2020-01-10 11:39:31</datetime>  
  <agentname>Reid Smith</agentname>  
  <client-phone>123-456-7890</client-phone>  
</metadata>
```

**NOTE**

Prior to V-Spark version 3.2, the client data field contained a `metadata` section that provided the same sort of information. This information described the custom metadata that was provided with the audio file when it was submitted for transcription.

## V-Spark JSON Schema

```
{
  "$schema": "http://json-schema.org/draft-03/schema#",
  "id": "Voci-JSON-Schema",
  "required": true,
  "type": "object",
  "properties": {
    "confidence": {
      "id": "confidence",
      "required": true,
      "type": "number"
    },
    "donedate": {
      "id": "donedate",
      "required": true,
      "type": "string"
    },
    "utterances": {
      "id": "utterances",
      "required": true,
      "type": "array",
      "items": {
        "id": "0",
        "required": true,
        "type": "object",
        "properties": {
          "emotion": {
            "id": "emotion",
            "required": true,
            "type": "string"
          },
          "confidence": {
            "id": "confidence",
            "required": true,

```

```
        "type": "number"
    },
    "end": {
        "id": "end",
        "required": true,
        "type": "number"
    },
    "sentiment": {
        "id": "sentiment",
        "required": true,
        "type": "string"
    },
    "gender": {
        "id": "gender",
        "required": true,
        "type": "string"
    },
    "sentimentex": {
        "id": "sentimentex",
        "required": true,
        "type": "array",
        "items": {
            "id": "0",
            "required": true,
            "type": "array",
            "items": {
                "id": "0",
                "required": true,
                "type": "number"
            }
        }
    },
    "start": {
        "id": "start",
        "required": true,
        "type": "number"
    },
    },
```

```
"donedate": {
  "id": "donedate",
  "required": true,
  "type": "string"
},
"recvdate": {
  "id": "recvdate",
  "required": true,
  "type": "string"
},
"events": {
  "id": "events",
  "required": true,
  "type": "array",
  "items": {
    "id": "0",
    "required": true,
    "type": "object",
    "properties": {
      "start": {
        "id": "start",
        "required": true,
        "type": "number"
      },
      "confidence": {
        "id": "confidence",
        "required": true,
        "type": "number"
      },
      "end": {
        "id": "end",
        "required": true,
        "type": "number"
      },
      "word": {
        "id": "word",
        "required": true,
```

```
        "type": "string"
      }
    }
  },
  "metadata": {
    "id": "metadata",
    "required": true,
    "type": "object",
    "properties": {
      "source": {
        "id": "source",
        "required": true,
        "type": "string"
      },
      "model": {
        "id": "model",
        "required": true,
        "type": "string"
      },
      "uttid": {
        "id": "uttid",
        "required": true,
        "type": "number"
      },
      "channel": {
        "id": "channel",
        "required": true,
        "type": "number"
      }
    }
  }
},
"sentiment": {
  "id": "sentiment",
```

```
    "required": true,
    "type": "string"
  },
  "nchannels": {
    "id": "nchannels",
    "required": true,
    "type": "number"
  },
  "source": {
    "id": "source",
    "required": true,
    "type": "string"
  },
  "client_data": {
    "id": "client_data",
    "required": true,
    "type": "object",
    "properties": {
      "datetime": {
        "id": "datetime",
        "required": true,
        "type": "string"
      }
    }
  },
  "app_data": {
    "id": "app_data",
    "required": true,
    "type": "object",
    "properties": {
      "silence": {
        "id": "silence",
        "required": true,
        "type": "string"
      },
      "client_gender": {
        "id": "client_gender",
```

```
    "required": true,  
    "type": "string"  
  },  
  "agent_clarity": {  
    "id": "agent_clarity",  
    "required": true,  
    "type": "null"  
  },  
  "agent_emotion": {  
    "id": "agent_emotion",  
    "required": true,  
    "type": "string"  
  },  
  "client_emotion": {  
    "id": "client_emotion",  
    "required": true,  
    "type": "string"  
  },  
  "overall_emotion": {  
    "id": "overall_emotion",  
    "required": true,  
    "type": "string"  
  },  
  "datetime": {  
    "id": "datetime",  
    "required": true,  
    "type": "string"  
  },  
  "scorecard": {  
    "id": "scorecard",  
    "required": true,  
    "type": "object"  
  },  
  "client_clarity": {  
    "id": "client_clarity",  
    "required": true,  
    "type": "string"  
  }
```



```
    },
    "overtalk": {
      "id": "overtalk",
      "required": true,
      "type": "null"
    },
    "url": {
      "id": "url",
      "required": true,
      "type": "string"
    },
    "words": {
      "id": "words",
      "required": true,
      "type": "number"
    },
    "client_channel": {
      "id": "client_channel",
      "required": true,
      "type": "number"
    },
    "tId": {
      "id": "tId",
      "required": true,
      "type": "number"
    },
    "agent_gender": {
      "id": "agent_gender",
      "required": true,
      "type": "null"
    },
    "duration": {
      "id": "duration",
      "required": true,
      "type": "string"
    }
  }
}
```

```
    },
    "requestid": {
      "id": "requestid",
      "required": true,
      "type": "string"
    },
    "sentiment_scores": {
      "id": "sentiment_scores",
      "required": true,
      "type": "array",
      "items": {
        "id": "0",
        "required": true,
        "type": "number"
      }
    },
    "model": {
      "id": "model",
      "required": true,
      "type": "array",
      "items": {
        "id": "0",
        "required": true,
        "type": "string"
      }
    },
    "recvdate": {
      "id": "recvdate",
      "required": true,
      "type": "string"
    }
  }
}
```

## JSON Output Example

The following output example shows a dual-language output (English and Spanish) using V-Blaze version 7.1.

```
{
  "emotion": "Neutral",
  "asr": "7.1.1-2",
  "done": "2020-09-30 11:44:37.653107",
  "nchannels": 1,
  "started": "2020-09-30 11:44:31.462813",
  "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"
  },
  "substinfo": {
    "counts": [
      [
        "subst_rules",
        1,
        {
          "persona => /Persona/": 1
        }
      ]
    ]
  },
  "nsubs": 1
},
```

```
"ended": "2020-09-30 11:44:37.654406",
"recvtz": [
  "EDT",
  -14400
],
"requestid": "8454-1601480671.450-172.27.104.121-233-184163",
"nsubs": 1,
"confidence": 0.84,
"utterances": [
  {
    "emotion": "Neutral",
    "confidence": 0.72,
    "end": 2.74,
    "recvtz": [
      "EDT",
      -14400
    ],
    "gender": "female",
    "rawemotion": "HAPPY",
    "lidinfo": {
      "lang": "spa",
      "speech": 1.35,
      "langfinal": "eng",
      "conf": 0.81
    },
    "start": 1.81,
    "done date": "2020-09-30 11:44:37.651913",
    "recvdate": "2020-09-30 11:44:31.578089",
    "events": [
      {
        "confidence": 0.69,
        "end": 2.32,
        "start": 1.81,
        "word": "Stephanie"
      },
      {
        "confidence": 0.76,
```

```
        "end": 2.74,  
        "start": 2.32,  
        "word": "so"  
      }  
    ],  
    "metadata": {  
      "source": "spa-eng-sample.wav",  
      "model": "en1:callcenter",  
      "uttid": 0,  
      "channel": 0  
    }  
  },  
  {  
    "emotion": "Neutral",  
    "confidence": 0.84,  
    "end": 26.12,  
    "recvtz": [  
      "EDT",  
      -14400  
    ],  
    "gender": "female",  
    "rawemotion": "HAPPY",  
    "lidinfo": {  
      "lang": "spa",  
      "speech": 17.46,  
      "conf": 1.0  
    },  
    "start": 4.56,  
    "done date": "2020-09-30 11:44:37.653107",  
    "recvdate": "2020-09-30 11:44:31.722011",  
    "events": [  
      {  
        "confidence": 0.9,  
        "end": 5.26,  
        "start": 4.56,  
        "word": "f??jate"  
      },  
    ]  
  }  
]
```

```
{
  "confidence": 0.9,
  "end": 5.55,
  "start": 5.46,
  "word": "que"
},
{
  "confidence": 0.88,
  "end": 5.93,
  "start": 5.55,
  "word": "estoy"
},
{
  "confidence": 0.73,
  "end": 6.44,
  "start": 5.93,
  "word": "anotando"
},
{
  "confidence": 0.9,
  "end": 6.82,
  "start": 6.44,
  "word": "Pero"
},
{
  "confidence": 0.75,
  "end": 7.0,
  "start": 6.82,
  "word": "El"
},
{
  "confidence": 0.83,
  "end": 7.49,
  "wordex": "me(2)",
  "start": 7.42,
  "word": "me"
},
}
```

```
{
  "confidence": 0.8,
  "end": 7.77,
  "start": 7.49,
  "word": "est??"
},
{
  "confidence": 0.9,
  "end": 8.07,
  "start": 7.77,
  "word": "llamando"
},
{
  "confidence": 0.71,
  "end": 8.13,
  "start": 8.07,
  "word": "Es"
},
{
  "confidence": 0.9,
  "end": 8.31,
  "start": 8.13,
  "word": "su"
},
{
  "confidence": 0.9,
  "end": 8.85,
  "start": 8.31,
  "word": "celular"
},
{
  "confidence": 0.6,
  "end": 9.03,
  "start": 8.85,
  "word": "por"
},
{
```

```
    "confidence": 0.74,  
    "end": 9.24,  
    "start": 9.03,  
    "word": "El"  
  },  
  {  
    "confidence": 0.9,  
    "end": 9.66,  
    "start": 9.24,  
    "word": "celular."  
  },  
  {  
    "confidence": 0.52,  
    "end": 9.78,  
    "start": 9.66,  
    "word": "No"  
  },  
  {  
    "confidence": 0.9,  
    "end": 10.02,  
    "start": 9.78,  
    "word": "tiene"  
  },  
  {  
    "confidence": 0.48,  
    "end": 10.6,  
    "start": 10.02,  
    "word": "carga"  
  },  
  {  
    "confidence": 0.9,  
    "end": 11.82,  
    "start": 11.52,  
    "word": "tiene"  
  },  
  {  
    "confidence": 0.9,
```



```
    "end": 11.95,  
    "start": 11.82,  
    "word": "que"  
  },  
  {  
    "confidence": 0.9,  
    "end": 12.41,  
    "start": 11.95,  
    "word": "llamar"  
  },  
  {  
    "confidence": 0.8,  
    "end": 12.55,  
    "start": 12.41,  
    "word": "por"  
  },  
  {  
    "confidence": 0.88,  
    "end": 12.68,  
    "start": 12.55,  
    "word": "la"  
  },  
  {  
    "confidence": 0.9,  
    "end": 13.24,  
    "start": 12.68,  
    "word": "Casa"  
  },  
  {  
    "confidence": 0.9,  
    "end": 14.18,  
    "start": 13.79,  
    "word": "n??mero"  
  },  
  {  
    "confidence": 0.9,  
    "end": 14.89,
```

```
    "start": 14.18,  
    "word": "Dos"  
  },  
  {  
    "confidence": 0.73,  
    "end": 15.64,  
    "start": 15.35,  
    "word": "cuando"  
  },  
  {  
    "confidence": 0.69,  
    "end": 16.14,  
    "start": 15.64,  
    "word": "tengo"  
  },  
  {  
    "confidence": 0.9,  
    "end": 16.48,  
    "start": 16.14,  
    "word": "chance."  
  },  
  {  
    "confidence": 0.9,  
    "end": 16.68,  
    "start": 16.48,  
    "word": "Si"  
  },  
  {  
    "confidence": 0.82,  
    "end": 16.81,  
    "start": 16.68,  
    "word": "t??"  
  },  
  {  
    "confidence": 0.68,  
    "end": 17.14,  
    "start": 16.81,
```

```
    "word": "puedas"
  },
  {
    "confidence": 0.9,
    "end": 17.71,
    "start": 17.14,
    "word": "llamar"
  },
  {
    "confidence": 0.72,
    "end": 18.51,
    "start": 18.15,
    "word": "trato"
  },
  {
    "confidence": 0.79,
    "end": 18.61,
    "wordex": "De(2)",
    "start": 18.51,
    "word": "De"
  },
  {
    "confidence": 0.6,
    "end": 19.28,
    "start": 18.61,
    "word": "comunicarme"
  },
  {
    "confidence": 0.9,
    "end": 19.45,
    "start": 19.28,
    "word": "con"
  },
  {
    "confidence": 0.9,
    "end": 19.56,
    "start": 19.45,
```

```
    "word": "la",
  },
  {
    "confidence": 1.0,
    "end": 20.14,
    "start": 19.56,
    "word": "Persona"
  },
  {
    "confidence": 0.9,
    "end": 20.29,
    "start": 20.14,
    "word": "Del"
  },
  {
    "confidence": 0.9,
    "end": 20.67,
    "start": 20.29,
    "word": "seguro"
  },
  {
    "confidence": 0.9,
    "end": 20.85,
    "start": 20.67,
    "word": "De"
  },
  {
    "confidence": 0.9,
    "end": 21.13,
    "start": 20.85,
    "word": "Alta"
  },
  {
    "confidence": 0.9,
    "end": 21.26,
    "start": 21.13,
    "word": "Es"
  }
}
```

```
},  
{  
  "confidence": 0.9,  
  "end": 21.42,  
  "start": 21.26,  
  "word": "que"  
},  
{  
  "confidence": 0.9,  
  "end": 21.83,  
  "start": 21.42,  
  "word": "no"  
},  
{  
  "confidence": 0.9,  
  "end": 22.03,  
  "start": 21.83,  
  "word": "no"  
},  
{  
  "confidence": 0.61,  
  "end": 22.12,  
  "start": 22.03,  
  "word": "Le"  
},  
{  
  "confidence": 0.52,  
  "end": 22.23,  
  "start": 22.12,  
  "word": "Han"  
},  
{  
  "confidence": 0.9,  
  "end": 22.85,  
  "start": 22.23,  
  "word": "respondido"  
},  
}
```

```
{
  "confidence": 0.9,
  "end": 23.03,
  "start": 22.85,
  "word": "ya"
},
{
  "confidence": 0.9,
  "end": 23.23,
  "start": 23.03,
  "word": "tiene"
},
{
  "confidence": 0.9,
  "end": 23.66,
  "start": 23.23,
  "word": "bastante"
},
{
  "confidence": 0.9,
  "end": 24.12,
  "start": 23.66,
  "word": "tiempo"
},
{
  "confidence": 0.9,
  "end": 25.78,
  "start": 25.43,
  "word": "cualquier"
},
{
  "confidence": 0.9,
  "end": 26.12,
  "start": 25.78,
  "word": "cosa."
}
],
```

```
    "metadata": {
      "uttid": 1,
      "substinfo": {
        "subs": [
          [
            19.56,
            20.14,
            [
              {
                "source": "subst_rules",
                "end": 20.14,
                "sub": "persona => /Persona/",
                "rule": "0",
                "start": 19.56
              }
            ]
          ]
        ],
        "nsubs": 1
      },
      "source": "spa-eng-sample.wav",
      "channel": 0,
      "model": "spa3:travel",
      "nsubs": 1
    }
  ],
  "license": "ASRUse-voci_doctest-12345",
  "source": "spa-eng-sample.wav",
  "gender": "female",
  "rawemotion": "HAPPY",
  "lidinfo": {
    "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
      }
    },
    "lang": "spa",
    "speech": 18.81,
    "conf": 0.99
  },
  "audiosecs": 30.52,
  "model": "en1:callcenter",
  "recvdate": "2020-09-30 11:44:31.578089"
}
{
  "emotion": "Neutral",
  "asr": "7.1.1-2",
  "confidence": 0.83,
  "gender": "female",
  "doneat": "2020-09-18 14:16:56.760942",
  "utterances": [
    {
      "emotion": "Neutral",
      "confidence": 0.72,
      "end": 2.74,
      "sentiment": "Neutral",
      "recvtz": [
        "EDT",
        -14400
      ],
      "gender": "female",
      "rawemotion": "HAPPY",
      "lidinfo": {
        "lang": "spa",

```



```
    "speech": 1.35,
    "langfinal": "eng",
    "conf": 0.81
  },
  "sentimentex": [
    [
      0,
      0
    ],
    []
  ],
  "start": 1.81,
  "done": "2020-09-18 14:16:56.759691",
  "recvd": "2020-09-18 14:16:35.728058",
  "events": [
    {
      "confidence": 0.69,
      "end": 2.32,
      "start": 1.81,
      "word": "Stephanie"
    },
    {
      "confidence": 0.76,
      "end": 2.74,
      "start": 2.32,
      "word": "so"
    }
  ],
  "metadata": {
    "source": "spa-eng-sample.wav",
    "model": "en1:callcenter",
    "uttid": 0,
    "channel": 0
  }
},
{
  "emotion": "Neutral",
```

```
"confidence": 0.83,
"end": 26.12,
"sentiment": "Neutral",
"recvtz": [
  "EDT",
  -14400
],
"gender": "female",
"rawemotion": "HAPPY",
"lidinfo": {
  "lang": "spa",
  "speech": 17.46,
  "conf": 1.0
},
"sentimentex": [
  [
    0,
    0
  ],
  []
],
"start": 4.56,
"done": "2020-09-18 14:16:56.760942",
"recvdate": "2020-09-18 14:16:35.864053",
"events": [
  {
    "confidence": 0.9,
    "end": 5.26,
    "start": 4.56,
    "word": "f??jate"
  },
  {
    "confidence": 0.9,
    "end": 5.55,
    "start": 5.46,
    "word": "que"
  }
],
```

```
{
  "confidence": 0.88,
  "end": 5.93,
  "start": 5.55,
  "word": "estoy"
},
{
  "confidence": 0.73,
  "end": 6.44,
  "start": 5.93,
  "word": "anotando"
},
{
  "confidence": 0.9,
  "end": 6.82,
  "start": 6.44,
  "word": "Pero"
},
{
  "confidence": 0.75,
  "end": 7.0,
  "start": 6.82,
  "word": "El"
},
{
  "confidence": 0.83,
  "end": 7.49,
  "wordex": "me(2)",
  "start": 7.42,
  "word": "me"
},
{
  "confidence": 0.8,
  "end": 7.77,
  "start": 7.49,
  "word": "est??"
},
}
```

```
{
  "confidence": 0.9,
  "end": 8.07,
  "start": 7.77,
  "word": "llamando"
},
{
  "confidence": 0.71,
  "end": 8.13,
  "start": 8.07,
  "word": "Es"
},
{
  "confidence": 0.9,
  "end": 8.31,
  "start": 8.13,
  "word": "su"
},
{
  "confidence": 0.9,
  "end": 8.85,
  "start": 8.31,
  "word": "celular"
},
{
  "confidence": 0.6,
  "end": 9.03,
  "start": 8.85,
  "word": "por"
},
{
  "confidence": 0.74,
  "end": 9.24,
  "start": 9.03,
  "word": "El"
},
{
```

```
    "confidence": 0.9,  
    "end": 9.66,  
    "start": 9.24,  
    "word": "celular."  
  },  
  {  
    "confidence": 0.52,  
    "end": 9.78,  
    "start": 9.66,  
    "word": "No"  
  },  
  {  
    "confidence": 0.9,  
    "end": 10.02,  
    "start": 9.78,  
    "word": "tiene"  
  },  
  {  
    "confidence": 0.48,  
    "end": 10.6,  
    "start": 10.02,  
    "word": "carga"  
  },  
  {  
    "confidence": 0.9,  
    "end": 11.82,  
    "start": 11.52,  
    "word": "tiene"  
  },  
  {  
    "confidence": 0.9,  
    "end": 11.95,  
    "start": 11.82,  
    "word": "que"  
  },  
  {  
    "confidence": 0.9,
```

```
    "end": 12.41,  
    "start": 11.95,  
    "word": "llamar"  
  },  
  {  
    "confidence": 0.8,  
    "end": 12.55,  
    "start": 12.41,  
    "word": "por"  
  },  
  {  
    "confidence": 0.88,  
    "end": 12.68,  
    "start": 12.55,  
    "word": "la"  
  },  
  {  
    "confidence": 0.9,  
    "end": 13.24,  
    "start": 12.68,  
    "word": "Casa"  
  },  
  {  
    "confidence": 0.9,  
    "end": 14.18,  
    "start": 13.79,  
    "word": "n?mero"  
  },  
  {  
    "confidence": 0.9,  
    "end": 14.89,  
    "start": 14.18,  
    "word": "Dos"  
  },  
  {  
    "confidence": 0.73,  
    "end": 15.64,
```

```
    "start": 15.35,  
    "word": "cuando"  
  },  
  {  
    "confidence": 0.69,  
    "end": 16.14,  
    "start": 15.64,  
    "word": "tengo"  
  },  
  {  
    "confidence": 0.9,  
    "end": 16.48,  
    "start": 16.14,  
    "word": "chance."  
  },  
  {  
    "confidence": 0.9,  
    "end": 16.68,  
    "start": 16.48,  
    "word": "Si"  
  },  
  {  
    "confidence": 0.82,  
    "end": 16.81,  
    "start": 16.68,  
    "word": "t??"  
  },  
  {  
    "confidence": 0.68,  
    "end": 17.14,  
    "start": 16.81,  
    "word": "puedas"  
  },  
  {  
    "confidence": 0.9,  
    "end": 17.71,  
    "start": 17.14,
```

```
    "word": "llamar",
  },
  {
    "confidence": 0.72,
    "end": 18.51,
    "start": 18.15,
    "word": "trato"
  },
  {
    "confidence": 0.79,
    "end": 18.61,
    "wordex": "De(2)",
    "start": 18.51,
    "word": "De"
  },
  {
    "confidence": 0.6,
    "end": 19.28,
    "start": 18.61,
    "word": "comunicarme"
  },
  {
    "confidence": 0.9,
    "end": 19.45,
    "start": 19.28,
    "word": "con"
  },
  {
    "confidence": 0.9,
    "end": 19.56,
    "start": 19.45,
    "word": "la"
  },
  {
    "confidence": 0.9,
    "end": 20.14,
    "start": 19.56,
```



```
    "word": "persona"  
  },  
  {  
    "confidence": 0.9,  
    "end": 20.29,  
    "start": 20.14,  
    "word": "Del "  
  },  
  {  
    "confidence": 0.9,  
    "end": 20.67,  
    "start": 20.29,  
    "word": "seguro"  
  },  
  {  
    "confidence": 0.9,  
    "end": 20.85,  
    "start": 20.67,  
    "word": "De "  
  },  
  {  
    "confidence": 0.9,  
    "end": 21.13,  
    "start": 20.85,  
    "word": "Alta"  
  },  
  {  
    "confidence": 0.9,  
    "end": 21.26,  
    "start": 21.13,  
    "word": "Es "  
  },  
  {  
    "confidence": 0.9,  
    "end": 21.42,  
    "start": 21.26,  
    "word": "que "
```

```
},  
{  
  "confidence": 0.9,  
  "end": 21.83,  
  "start": 21.42,  
  "word": "no"  
},  
{  
  "confidence": 0.9,  
  "end": 22.03,  
  "start": 21.83,  
  "word": "no"  
},  
{  
  "confidence": 0.61,  
  "end": 22.12,  
  "start": 22.03,  
  "word": "Le"  
},  
{  
  "confidence": 0.52,  
  "end": 22.23,  
  "start": 22.12,  
  "word": "Han"  
},  
{  
  "confidence": 0.9,  
  "end": 22.85,  
  "start": 22.23,  
  "word": "respondido"  
},  
{  
  "confidence": 0.9,  
  "end": 23.03,  
  "start": 22.85,  
  "word": "ya"  
},  
}
```

```
{
  "confidence": 0.9,
  "end": 23.23,
  "start": 23.03,
  "word": "tiene"
},
{
  "confidence": 0.9,
  "end": 23.66,
  "start": 23.23,
  "word": "bastante"
},
{
  "confidence": 0.9,
  "end": 24.12,
  "start": 23.66,
  "word": "tiempo"
},
{
  "confidence": 0.9,
  "end": 25.78,
  "start": 25.43,
  "word": "cualquier"
},
{
  "confidence": 0.9,
  "end": 26.12,
  "start": 25.78,
  "word": "cosa."
}
],
"metadata": {
  "source": "spa-eng-sample.wav",
  "model": "spa3:travel",
  "uttid": 1,
  "channel": 0
}
```

```
    },
  ],
  "license": "test_license_docs123",
  "audiosecs": 30.52,
  "started": "2020-09-18 14:16:35.623669",
  "streamtags": {
    "emotion": "xa",
    "lid": "spa3:travel",
    "gender": true,
    "rawemotion": true,
    "subst": true,
    "substinfo": true,
    "lidthreshold": 1.0,
    "lidutt": true,
    "scrubtext": true,
    "datahdr": "WAVE",
    "nsubs": "true"
  },
  "rawemotion": "HAPPY",
  "nchannels": 1,
  "lidinfo": {
    "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
      }
    },
    "lang": "spa",
    "speech": 18.81,
```

```
    "conf": 0.99
  },
  "sentiment": "Neutral",
  "ended": "2020-09-18 14:16:56.762195",
  "recvtz": [
    "EDT",
    -14400
  ],
  "source": "spa-eng-sample.wav",
  "sentiment_scores": [
    0,
    0
  ],
  "model": "en1:callcenter",
  "recvdate": "2020-09-18 14:16:35.728058",
  "requestid": "exampleid1235-78938576-20294894725"
}
```