



**SPEECHMATICS**

Alignment service 3.3.0

## Table of Contents

- [Alignment](#)
  - [Supported Audio Formats](#)
  - [Data Retention](#)
  - [Supported Text Formats](#)
  - [Text Formatting](#)
    - [Input](#)
    - [Output](#)
  - [Supported endpoints](#)
  - [Egress IPs](#)
  - [Submitting Alignment Jobs](#)
  - [Checking Alignment Job Status](#)
  - [Poll for more than one job](#)
  - [Retrieving Alignment Job Files](#)
  - [Deleting Alignment Job](#)
  - [Fetching files from an online location](#)
  - [Callback Notifications](#)

# Alignment

Alignment allows the user to submit an audio file and a text file, and get back the speech timing information. This allows users to determine when exactly a given word was spoken in the context of the supplied audio file.

If you do not have access to use the alignment feature, and you would like to, please speak to your Account Manager.

The following documentation will show you how to request alignment, and how to retrieve an aligned file.

## Supported Audio Formats

The following audio formats are supported:

- aac
- amr
- flac
- m4a
- mp3
- mpg
- ogg
- wav

## Data Retention

Alignment correspond to Speechmatics' SaaS policy. All files are stored for seven days, after which point they are deleted. Files can be deleted earlier by explicitly requesting so. How to do so is documented below.

## Supported Text Formats

The input text file must be UTF-8 encoded plain text file. Characters outside this format will mean the job is rejected.

## Text Formatting

### Input

During the alignment process, Speechmatics tries to extract words from the text. Any string of characters separated by whitespace (space, tab, newline, etc.) is considered as a word. Any markup in the text file, with SGML-like tags with angled-brackets is considered as comments. For example, text within the comment delimiters ( `<!--` , `-->` ) or angle brackets ( `<` , `>` ) is ignored. Therefore, given this text:

```
Hello <markup> world <!-- comment > comment --> how are you?
```

The following words will be aligned with the provided audio file:

```
Hello world how are you?
```

### Output

The timing information (termed as *alignment files*) are available in two formats:

- Word Start and End ( `word_start_and_end` ): This is the default format:

```
<time=0.12>Hello<time=0.23> <markup> <time=0.34>world<time=0.45> <!-- comment >
comment -->
<time=0.56>how<time=0.67> <time=0.78>are<time=0.89> <time=0.90>you?<time=1.00>
```

- One per Line ( `one_per_line` ). This must be specified when you request the transcript via HTTP request.

```
[00:00:00.1] Hello <markup> world <!-- comment > comment --> how are you?
```

## Supported endpoints

Alignment uses the Speechmatics SaaS which supports the following endpoints for production use:

Region	Environment	Endpoints
EU	EU1	<b>eu.asr.api.speechmatics.com</b> <b>eu1.asr.api.speechmatics.com</b> <b>asr.api.speechmatics.com</b>
EU	EU2	<b>eu2.asr.api.speechmatics.com</b>
US	US1	<b>us.asr.api.speechmatics.com</b> <b>us1.asr.api.speechmatics.com</b>
US	US2	<b>us2.asr.api.speechmatics.com</b>

Authorization Tokens are replicated between all environments in the same region. Therefore, you can use any environment in a region that you are entitled to access.

All production environments are active and highly available. Multiple environments can be used to balance requests or provide a failover in the event of disruption to one environment.

Note that jobs are created in the environment corresponding to the endpoint used. You must use the same endpoint for all requests relating to a specific job.

If you attempt to use an endpoint for a region you are not contracted to use, that request will be unsuccessful. If you want to use a different region, please contact [sales@speechmatics.com](mailto:sales@speechmatics.com).

## Egress IPs

If you wish to receive an aligned transcript via notification, you must whitelist the relevant IPs below to ensure successful delivery:

Regions	IP Addresses
EU1	40.74.41.91, 52.236.157.154, 40.74.37.0, 20.73.209.153, 20.73.142.44
EU2	20.105.89.153, 20.105.89.173, 20.105.89.184, 20.105.89.98, 20.105.88.228
US1	52.149.21.32, 52.149.21.10, 52.137.102.83, 40.64.107.92, 40.64.107.99
US2	52.146.58.224, 52.146.58.159, 52.146.59.242, 52.146.59.213, 52.146.58.64

## Submitting Alignment Jobs

Creating an alignment job is similar in process to transcription job. An HTTP POST request must be made to `/v2/jobs` endpoint with following form fields:

- `config` : The job config for alignment
- `data_file` : The media file containing the speech. Can be passed in via `config` if the file is stored in an online location
- `text_file` : The text file containing the transcript. Can be passed in via `config` if the file is stored in an online location

If you do not provide all of the above the job will be rejected.

The job config must state that the job type is alignment and the language of the audio and text.

```
{
  "type": "alignment",
  "alignment_config": {
    "language": "en"
  }
}
```

The corresponding curl request looks like so:

```
curl https://asr.api.speechmatics.com/v2/jobs/ \
-X POST \
-H "Authorization: Bearer <TOKEN>" \
-F data_file=@/tmp/file.mp3 \
-F text_file=@/tmp/speech.txt \
-F config='{ "type": "alignment", "alignment_config": { "language": "en" } }'
```

A successful request will return a HTTP 201 response, and will contain a unique 10-digit alphanumeric Job ID, which will be returned as `id` in the HTTP response.

```
HTTP/2 201
date: Mon, 11 Oct 2021 16:45:44 GMT
content-type: application/vnd.speechmatics.v2+json
content-length: 20
strict-transport-security: max-age=15724800; includeSubDomains
request-id: 802b2603d62d23b5bb113836ec0a8d21

{"id": "r0btay8pxr"}
```

## Checking Alignment Job Status

You can retrieve the status of an alignment job by making a HTTP GET request that includes the Job ID in the request endpoint.

An example is below:

```
curl https://asr.api.speechmatics.com/v2/jobs/<JOB_ID> \
-X GET \
-H "Authorization: Bearer <TOKEN>"
```

An example response is below:

```
{
  "job": {
    "config": {
      "alignment_config": {
        "language": "en"
      }
    }
  }
}
```

```

    },
    "type": "alignment"
  },
  "created_at": "2021-09-24T10:51:13.641Z",
  "data_name": "63f662ce-4b82-4471-b0e0-380abb83f666.m4a",
  "duration": 281,
  "id": "g0sjrmiqng",
  "status": "done"
}
}

```

The status will be one of the following:

- `Done` : The file is ready to be retrieved
- `Running` : The file is still being processed and not yet ready
- `Rejected` : The job has failed

## Poll for more than one job

If you have submitted multiple jobs, you can retrieve a list of the 100 most recent jobs submitted in the past 7 days by making a GET request without a Job ID. If a job has been deleted it will not be included in the list.

An example is below:

```

curl https://asr.api.speechmatics.com/v2/jobs/ \
-X GET \
-H "Authorization: Bearer <TOKEN>"

```

## Retrieving Alignment Job Files

An aligned file can be retrieved from the `/v2/jobs/<JOB_ID>/alignment` endpoint. By default, the 'Word Start and End' alignment format is returned. This can be overridden with the query parameter `tags` in the HTTP GET request as illustrated below:

```

curl https://asr.api.speechmatics.com/v2/jobs/<JOB_ID>/alignment?tags=one_per_line \
-X GET \
-H "Authorization: Bearer <TOKEN>"

```

Use the following endpoints to retrieve the inputs files used for an alignment job:

- `/v2/jobs/<JOB_ID>/text` : to get the text file submitted
- `/v2/jobs/<JOB_ID>/data` : to get the audio file submitted

## Deleting Alignment Job

If you want to delete a submitted job you can do so via sending a HTTP DELETE request specifying the Job ID. All files, including aligned files, will be deleted from the Speechmatics SaaS.

```

curl https://asr.api.speechmatics.com/v2/jobs/<JOB_ID> \
-X DELETE \
-H "Authorization: Bearer <TOKEN>"

```

The response will show a `status` of `deleted` as shown below:

```

{
  "job": {
    "config": {
      "alignment_config": {
        "language": "en"
      },
      "type": "alignment"
    },
    "created_at": "2021-09-24T10:51:13.641Z",
    "data_name": "63f662ce-4b82-4471-b0e0-380abb83f666.m4a",
    "duration": 281,
    "id": "g0sjrmiqng",
    "status": "deleted"
  }
}

```

## Fetching files from an online location

Speechmatics supports retrieving files from an online location. If you store your digital media and transcripts in cloud storage (for example AWS S3 or Azure Blob Storage) you can also submit a job by providing the URL of the audio file or transcript.

To retrieve files from an online location, you must specify the location for the media and/or transcript in the configuration of your request. You can locally upload a media file and retrieve a text file from an online location (or vice versa):

```

{
  "type": "alignment",
  "fetch_data": {"url": "$MY_AUDIO_URL"},
  "fetch_text": {"url": "$MY_TRANSCRIPT"},
  "alignment_config": { "language": "en" }
}

```

You should not use `fetch_data` or `fetch_text` with locally uploaded files simultaneously, as this will cause the job to fail.

## Callback Notifications

Alignment jobs can also be used with callback notifications by including the `notification_config` section in the job config when submitting the job. Please ensure you have whitelisted Speechmatics' egress IPs to allow notifications.

```

{
  "type": "alignment",
  "alignment_config": {
    "language": "en"
  },
  "notification_config": [
    {
      "contents": [
        "alignment"
      ],
      "url": "https://lorem.ipsum/"
    }
  ],
}

```

```
{
  "contents": [
    "alignment.one_per_line", "text"
  ],
  "method": "post",
  "url": "https://dolor.sit.amet/"
}
]
```

The following outputs are supported:

- `alignment`, `alignment.one_per_line`, `alignment.word_start_and_end` : the Aligned transcript
- `text` : the non-aligned transcript submitted as part of the job request
- `data` : the media file submitted as part of the job request
- `jobinfo` : the summary information about the job, to support identification and tracking