

Microsoft Azure Cognitive Services- Text Translator API Technical Specification Document

©Pruthvi Kasula (#801254275)

University of North Carolina at Charlotte

TEXT TRANSLATOR API TECHNICAL DOCUMENT

Document Control Information

Document Information

Document Name	Text translator API technical document
Professor	Prof. Fabio Nolasco
Course	System Integration – ITIS 6177
Document Author	Pruthvi Kasula (801254275)
Document Version	1.0
Document Status	Final
Date Released	05/1/2022

Document Edit History

Version	Date	Additions/Modifications	Prepared/Revised by
1.0	05/1/2022	Initial draft document	Pruthvi Kasula

Document Review/Approval History

Date	Name	Organization/Title	Comments
	Prof. Fabio Nolasco	Professor	

Microsoft Azure Cognitive Services- Text Translator API

Text translation is a cloud-based REST API feature of the Translator service that uses neural machine translation technology to enable quick and accurate source-to-target text translation in real time across all supported languages.

Text translation features

The following methods are supported by the Text Translation feature:

- **Languages:** Returns a list of languages supported by Translate, Transliterate, and Dictionary Lookup operations.

API endpoint: GET [/api.v1.TextTranslator.com/AllLanguages/:scope](#)

- **Translate:** Renders single source-language text to multiple target-language texts with a single request.

API endpoint: POST [/api.v1.TextTranslator.com/Translate](#)

- **Transliterate:** Converts characters or letters of a source language to the corresponding characters or letters of a target language.

API endpoint: POST [/api.v1.TextTranslator.com/Transliterate](#)

- **Detect:** Returns the source code language code and a boolean variable denoting whether the detected language is supported for text translation and transliteration.

API endpoint: POST [/api.v1.TextTranslator.com/Detect](#)

- **Break Sentence:** Returns sentence length without translation using the breaksentence endpoint.

API endpoint: POST [/api.v1.TextTranslator.com/BreakSentence](#)

- **Alternate Translations:** Returns alternate translations for a word or phrase.

API endpoint: POST [/api.v1.TextTranslator.com/AlternateTranslations](#)

- **Swagger Documentation endpoint:** <http://137.184.210.75:3000/swaggerFinalProjDoc>

API End points

Languages:

```
GET /api.v1.TextTranslator.com/AllLanguages/:scope
```

Description:

- This REST endpoint is used to obtain a set of languages and codes currently supported by Translate, Transliterate, and Dictionary Lookup operations.
- scope=translation provides languages supported to translate text from one language to another language.
- scope=transliteration provides capabilities for converting text in one language from one script to another script.
- scope=dictionary provides language pairs for which Dictionary operations return data.
- Output is a list of languages and their corresponding language codes which can be used in its respective context.

GET Request URL:

<http://137.184.210.75:3000/api.v1.TextTranslator.com/AllLanguages/translation>

Parameters Required for GET request:

Parameter Name	Mandatory	Rule
scope	YES	String query parameter. Allowed values are translation / transliteration / dictionary

Response body:

- A successful response is a JSON array with a 200 Success status.
- The **scope** query parameter is used to define which group of languages is required by the user.

TEXT TRANSLATOR API TECHNICAL DOCUMENT

- scope: translation provides languages and their codes supported by the `/api.v1.TextTranslator.com/Translate`
- scope: transliteration provides capabilities of converting text in one language from one script to another script used by the `/api.v1.TextTranslator.com/Transliterate` api.
- scope: dictionary provides language pairs from the Dictionary and returns data to be implemented by `/api.v1.TextTranslator.com/AlternateTranslations` api.
- An unsuccessful request returns the error with a 400 Bad Request status.

Response Status Codes:

<u>Status Code</u>	<u>Description</u>
200	Success Operation!
400	Error!

Example Requests:

- **Scope: translation**

Request URL
http://137.184.210.75:3000/api.v1.TextTranslator.com/AllLanguages/translation
REQUEST
GET API.V1.TEXTTRANSLATOR.COM/ALLLANGUAGES/TRANSLATION Content-type: application/json
RESPONSE
200 Content-Type: application/json <pre>{ "translation": { //... set of languages supported to translate text (scope=translation) }, }</pre>

Translate :

POST API.V1.TEXTTRANSLATOR.COM/TRANSLATE

Description:

- This REST endpoint is used to translate the provided input ‘text’ to the desired language which is provided as an input body parameter in ‘to’.
- /api.v1.TextTranslator.com/Translate will auto-detect the language of the text provided in the body and indicates the confidence score of the detected language.

Request URL:

<http://137.184.210.75:3000/api.v1.TextTranslator.com/Translate>

Request body:

- The body of the request is a JSON object.
- The JSON object has a string property named ‘**text**’, which represents the string to translate.
- The other string property which is an input param in the JSON request body is ‘**to**’ which specifies the language of the output text.
- The target language must be one of the supported language codes / included in the translation scope.

For example, use "to": "es" to translate the input text to Spanish.

- It is possible to translate to multiple languages simultaneously by passing an array to the ‘to’ input param.

For example, use "to": "it, de" to translate the same input text to Italian and German simultaneously.

Parameters Required for POST request:

Parameter Name	Mandatory	Rule
text	YES	String input. Cannot exceed 10,000 characters including spaces
to	YES	String input. Must be one of the language code supported in the Microsoft Azure Cognitive Services Text Translation document . OR refer http://137.184.210.75:3000/api.v1.TextTranslator.com/AllLanguages/translation

Example JSON Request Body:

JSON
<pre>{ "text": "hello", "to": "it, de" }</pre>

- The entire text included in the request cannot exceed 10,000 characters including spaces.

Response body:

- A successful response is a JSON array with a 200 Success status.
- A result object includes the following properties ^[1]:
 - detectedLanguage: An object that describes the detected language with the following properties:
 - language: A string representing the code or codes of the detected language.
 - score: A score is a float value that indicates the confidence in the result of the language detected. This score is between zero(low) and one(high). A lower score indicates lower confidence in the result of the detected language.

- translations: The JSON array represents the result of each target language translation specified through the 'to' request body parameter. Each element in the array includes the following properties:
 - text: A string output which gives the translation of the input text to the target language code. With the 'profanityMarker' enabled, profanities are handled by replacing with '***' sign in the response target language translation.
 - to: A string that represents the language code of the output target language.
- An unsuccessful request returns the error with a 400 Bad Request status.

Response Status Codes:

<u>Status Code</u>	<u>Description</u>
200	Success Operation!
400	Error!

Example Requests:

- Translate input text to single language

URL

<http://137.184.210.75:3000/api.v1.TextTranslator.com/Translate>

REQUEST

POST API.V1.TEXTTRANSLATOR.COM/TRANSLATE

Content-type: application/json

```
{
  "text": "Flower ",
  "to": "te"
}
```

RESPONSE

200

Content-Type: application/json

```
[
  {
    "detectedLanguage": {
      "language": "en",
      "score": 1
    },
    "translations": [
      {
        "text": "పూవు",
        "to": "te"
      }
    ]
  }
]
```

- Translate input text to multiple languages simultaneously.

URL

<http://137.184.210.75:3000/api.v1.TextTranslator.com/Translate>

REQUEST

POST API.V1.TEXTTRANSLATOR.COM/TRANSLATE

Content-type: application/json

```
{
  "text": "hello ",
  "to": "it,de"
}
```

RESPONSE

200

Content-Type: application/json

```
[
  {
    "detectedLanguage": {
      "language": "en",
      "score": 1
    },
    "translations": [
      {
        "text": "Ciao",
        "to": "it"
      },
      {
        "text": "Hallo",
        "to": "de"
      }
    ]
  }
]
```

- Translate a text with missing or invalid request body parameters

URL
<u>http://137.184.210.75:3000/api.v1.TextTranslator.com/Translate</u>
REQUEST
POST API.V1.TEXTTRANSLATOR.COM/TRANSLATE Content-type: application/json { "text": "Good to see you", "to": "" }
RESPONSE
400 Content-Type: application/json "Request failed with status code 400: Missing parameter 'to'"

DETECT:

POST API.V1.TEXTTRANSLATOR.COM/DETECT

Description:

- This REST endpoint identifies the language code of a text that is passed as part of the request.
- This will auto-detect the language code of the text, provide a confidence score for the detected language, and list alternative languages with their confidence score.

Request URL:

<http://137.184.210.75:3000//api.v1.TextTranslator.com/Detect>

Request body:

- The body of the request is a JSON object.
- The JSON object has a string property names '**text**', which represents the string whose language code is to be detected.
- The language auto-detection works better with longer input text. ^[2]

Parameters Required for POST request:

Parameter Name	Mandatory	Rule
text	YES	String input. Cannot exceed 10,000 characters including spaces

Example JSON Request Body:

JSON
{ "text": "Good to see you" }

- The entire text included in the request cannot exceed 50,000 characters including spaces.

Response body:

- A successful response is a JSON array with a 200 Success status.
- The result object includes the following properties:
 - PrimaryLanguage: A string representing the code or output code of the detected language.
 - score: A score is a float value that indicates the confidence in the result of the language detected. This score is between zero(low) and one(high). A lower score indicates lower confidence in the result of the detected language.
 - SecondaryLanguage: This is a JSON array of other possible languages detected from the input 'text' body parameter. This will again have language and score properties identified.
- An unsuccessful request returns the error with a 400 Bad Request status.

Response Status Codes:

<u>Status Code</u>	<u>Description</u>
200	Success Operation!
400	Error!

Example Requests:

- Detect Language Code of the uniform input text

URL
http://137.184.210.75:3000/api.v1.TextTranslator.com/Detect
REQUEST
POST API.V1.TEXTTRANSLATOR.COM/DETECT Content-type: application/json <pre>{ "text": "Good to see you" }</pre>
RESPONSE
200 Content-Type: application/json <pre>[{ "PrimaryLanguage": "en", "score": 1, "SecondaryLanguage": [{ "language": "", "score": "" }] }]</pre>

- Detect Language Code of a mixed input text

URL
http://137.184.210.75:3000/api.v1.TextTranslator.com/Detect
REQUEST
POST API.V1.TEXTTRANSLATOR.COM/DETECT Content-type: application/json <pre>{ "text": "你好 Good to see you" }</pre>
RESPONSE
200 Content-Type: application/json <pre>[{ "PrimaryLanguage": "en", "score": 0.67, "SecondaryLanguage": [{ "language": "zh-Hans", "score": 0.33 }] }]</pre>

TRANSLITERATE:

POST API.V1.TEXTTRANSLATOR.COM/TRANSLITERATE

Description:

- This REST endpoint converts text in one language from one script to another script, that is, it gives a phonetic translation from of the input text from one language to another.
- /api.v1.TextTranslator.com/Transliterate supports the transliteration of input text from one script to another based on the rules in the document.

Request URL:

<http://137.184.210.75:3000/api.v1.TextTranslator.com/Transliterate>

Request body:

- The body of the request is a JSON object.
- The JSON object has a string property named '**text**', which represents the string which needs a phonetic translation, that is, needs transliteration.
- Another required string property is '**language**' which indicates the language code mentioned in the transliteration table .
- The request body takes two other string properties '**fromScript**' and '**toScript**' which indicates the name of the script supported by the transliteration table or supported scripts.

Parameters Required for POST request:

Parameter Name	Mandatory	Rule
text	YES	String input. Cannot exceed 10,000 characters including spaces
language	YES	String input. Must be one of the language code supported in the Microsoft Azure Cognitive Services Text Transliteration <u>document</u> or <u>supported codes</u> .

TEXT TRANSLATOR API TECHNICAL DOCUMENT

fromScript	YES	String input. Must be one of the from script included in the To/From Script pair supported in the Microsoft Azure Cognitive Services Text Transliteration document or supported fromScripts .
toScript	YES	String input. Must be one of the to script included in the To/From Script pair supported in the Microsoft Azure Cognitive Services Text Transliteration document or supported toScripts .

Example JSON Request Body:

JSON
<pre>{ "text": "สวัสดี", "language": "th", "fromScript": "Thai", "toScript": "Latn" }</pre>

- The text value of JSON object cannot exceed 1,000 characters including spaces.
- The entire text included in the request cannot exceed 5,000 characters including spaces.

Response body:

- A successful response is a JSON array with a 200 Success status.
- The result object includes the following properties :
 - text: The output string which is a phonetic translation or transliteration of the input text in the script passed in the 'toScript' property of the request body.
 - script: The script which is used to transliterate the input text to, that is, the value of the script mentioned in the 'toScript' property. This property is used to verify the script of the transliterated output text.

- An unsuccessful request returns the error with a 400 Bad Request status.

Response Status Codes:

<u>Status Code</u>	<u>Description</u>
200	Success Operation!
400	Error!

Example Requests:

- Transliterate input text to script according to the transliteration table.

URL
http://137.184.210.75:3000/api.v1.TextTranslator.com/Transliterate
REQUEST
POST API.V1.TEXTTRANSLATOR.COM/TRANSLITERATE Content-type: application/json <pre>{ "text": "สวัสดี", "language": "th", "fromScript": "Thai", "toScript": "Latn" }</pre>
RESPONSE
200 Content-Type: application/json <pre>[{ "text": "sawatdi", "script": "Latn" }]</pre>

- Transliterate input text to script not supported in transliteration table.

URL
http://137.184.210.75:3000/api.v1.TextTranslator.com/Transliterate
REQUEST
POST API.V1.TEXTTRANSLATOR.COM/TRANSLITERATE Content-type: application/json <pre>{ "text": "สวัสดี", "language": "th", "fromScript": "Thai", "toScript": "Hans" }</pre>
RESPONSE
400 Content-Type: application/json "Request failed with status code 400"

BREAK SENTENCE:

POST API.V1.TEXTTRANSLATOR.COM/BREAKSENTENCE

Description:

- This REST endpoint is used to identify the positioning of sentence boundaries for the input text provided.
- /api.v1.TextTranslator.com/BreakSentence auto-detects the language code of the provided input text and outputs each sentence's length from the given input text.

Request URL:

<http://137.184.210.75:3000/api.v1.TextTranslator.com/BreakSentence>

Request body:

- The body of the request is a JSON Object.
- The JSON object has a string property named '**text**', which represents the string whose value is used to compute the sentence boundaries.

Parameters Required for POST request:

Parameter Name	Mandatory	Rule
text	YES	String input. Cannot exceed 10,000 characters including spaces

Example JSON Request Body:

JSON
<pre>{ "text": "Hello, how are you? Hope you are doing great!" }</pre>

- The text value of JSON object cannot exceed 50,000 characters including spaces.
- The entire text included in the request cannot exceed 50,000 characters including spaces.

Response body:

- A successful response is a JSON array with a 200 Success status.
- A result object includes the following properties ^[3]:
 - SentenceLength: An array of representing the sentence number and length of each sentence present in the text element. Also, array length represents the number of sentences present in the given input text.
 - PrimaryLanguageDetected: A JSON object which describes the detected language with the following properties:
 - language: A string representing the code of the detected language.
 - score: A score is a float value that indicates the confidence in the result of the language detected. This score is between zero(low) and one(high). A lower score indicates lower confidence in the result of the detected language.
- When an input text consists of mixed languages, the `/api.v1.TextTranslator.com/BreakSentence` will detect the language that has a greater number of words and state it's confidence score.
- An unsuccessful request returns the error with a 400 Bad Request status.

Response Status Codes:

<u>Status Code</u>	<u>Description</u>
200	Success Operation
400	Error!

Example Requests:

- **Sentence Break with uniform input text**

URL
<u>http://137.184.210.75:3000/api.v1.TextTranslator.com/BreakSentence</u>
REQUEST
POST API.V1.TEXTTRANSLATOR.COM/BREAKSENTENCE Content-type: application/json <pre>{ "text": "Hello, how are you? Hope you are doing great! Have a good time" }</pre>
RESPONSE
200 Content-Type: application/json <pre>[{ "SentenceLength": ["Sentence 1 : 20", "Sentence 2 : 26", "Sentence 3 : 16"], "PrimaryLanguageDetected": { "language": "en", "score": 1 } }]</pre>

- Sentence Break with mixed input text

URL
http://137.184.210.75:3000/api.v1.TextTranslator.com/BreakSentence
REQUEST
POST API.V1.TEXTTRANSLATOR.COM/BREAKSENTENCE Content-type: application/json <pre>{ "text": "Hello, how are you? Hope you are doing great! Хорошо тебе провести время" }</pre>
RESPONSE
200 Content-Type: application/json <pre>[{ "SentenceLength": ["Sentence 1 : 20", "Sentence 2 : 26", "Sentence 3 : 26"], "PrimaryLanguageDetected": { "language": "en", "score": 0.69 } }]</pre>

ALTERNATE TRANSLATIONS:

POST API.V1.TEXTTRANSLATOR.COM/ALTERNATETRANSLATIONS

Description:

- This REST endpoint provides alternate translations for a word or certain idiomatic phrases in the input language provided.
- /api.v1.TextTranslator.com/AlternateTranslations responses with a list of alternate translations for the provided input text, in the input language provided and also lists the part-of-speech of the listed text in the translated language along with a list of back-translations in the original language of the text to provide context.

Request URL:

<http://137.184.210.75:3000/api.v1.TextTranslator.com/AlternateTranslations>

Request body:

- The body of the request is a JSON object.
- The JSON object has a string property names '**text**', which represents the sting which is to be translated.
- The '**from**' string property specifies the language code of the input text. It must be one of the language codes included in the dictionary scope.
- The '**to**' string property specifies the language code of the target output text i.e the language to which the input text is to be translated. It must be one of It must be one of the language codes included in the dictionary scope.

TEXT TRANSLATOR API TECHNICAL DOCUMENT

Parameters Required for POST request:

Parameter Name	Mandatory	Rule
text	YES	String input. Cannot exceed 100 characters including spaces.
from	YES	String input. Must be one of the language code supported in the Microsoft Azure Cognitive Services Text Translation document or languages supported for alternative translations
to	YES	String input. Must be one of the language code supported in the Microsoft Azure Cognitive Services Text Translation document or languages supported for alternative translations

Example JSON Request Body:

JSON
<pre>{ "text": "house", "from": "en", "to": "es" }</pre>

- The text value of JSON object cannot exceed 100 characters including spaces.

Response body:

- A successful response is a JSON array with a 200 Success status.
- If the term in the input text is not defined in the dictionary, the response is 200 (OK) but the result array will be empty.
- The result object includes the following properties :

- normalizedTarget: A string with the normalized form of the output text translated in the provided input language.
- displayTarget: A string which is a better representation of the normalizedTarget term for best end-user display.
- postTag: This property indicates the part-of-speech of the input text in the translated language. The following Tags represent the given part-of-speech

Tag name	Description
ADJ	Adjectives
ADV	Adverbs
CONJ	Conjunctions
DET	Determiners
MODAL	Verbs
NOUN	Nouns
PREP	Prepositions
PRON	Pronouns
VERB	Verbs
OTHER	Other

- confidence: A float value that indicates the confidence in the result of the translated output text. This score is between zero(low) and one(high). A lower score indicates lower confidence in the translation pair.
- prefixWord: A string value, which is a gendered determination of nouns, in the provided language for translation if it has any gendered determinations. If there is no prefix, the output for this property is an empty string.

- backTranslations: This output a list of backtranslations of the text in context to the normazliedTarget value in the original language of the input text. Each element of the backTranslations list is an object that has the following properties:
 - normalizedText: A string with the normalized form of the source input text term which is backtranslated in the original language of the input text.
 - displayText: A string which is a better representation of the normalized Text term for best end-user display.
 - numExamples: An integer value that represents the number of examples that are available, for the given translation pair of the input text and the target translated text, in the training model.
 - frequencyCount: An integer value which represents the frequency of the given translation pair of the input text and the output translated text. This count can be used to sort the back-translations to get the most frequent pairs.
- An unsuccessful request returns the error with a 400 Bad Request status.

Response Status Codes:

<u>Status Code</u>	<u>Description</u>
200	Success Operation!
400	Error!

Example Requests:

- **Alternate Translation for input text listed in the dictionary with available alternate translations in the target language**

URL

<http://137.184.210.75:3000/api.v1.TextTranslator.com/AlternateTranslations>

REQUEST

POST API/ALT_TRANSALTE
Content-type: application/json

```
{
  "text": "shark",
  "from": "en",
  "to": "es"
}
```

RESPONSE

200

Content-Type: application/json

```
[
  {
    "normalizedTarget": "tiburón",    "displayTarget": "tiburón",
    "posTag": "NOUN",
    "confidence": 1,    "prefixWord": "",
    "backTranslations": [
      {
        "normalizedText": "shark",
        "displayText": "shark",
        "numExamples": 15,
        "frequencyCount": 1848
      },
      {
        "normalizedText": "jaws",
        "displayText": "jaws",
        "numExamples": 5,
        "frequencyCount": 46
      },
      {
        "normalizedText": "tiburon",
        "displayText": "Tiburon",
        "numExamples": 0,
        "frequencyCount": 16
      }
    ]
  }
]
```

- Alternate Translation for input text not listed in the dictionary

URL
<u>http://137.184.210.75:3000/api.v1.TextTranslator.com/AlternateTranslations</u>
REQUEST
POST API.V1.TEXTTRANSLATOR.COM/TRANSLATE Content-type: application/json { "text": "It's a wonderful day!", "from": "en", "to": "es" }
RESPONSE
200 Content-Type: application/json []

- Alternate Translations with missing Parameter in the request body

URL
<u>http://137.184.210.75:3000/api.v1.TextTranslator.com/AlternateTranslations</u>
REQUEST
POST API.V1.TEXTTRANSLATOR.COM/TRANSLATE Content-type: application/json { "text": "It's a wonderful day!", "from": "", "to": "es" }
RESPONSE
400 Content-Type: application/json "Request failed with status code 400: Missing parameter 'from'"

References :

- <https://docs.microsoft.com/en-us/azure/cognitive-services/translator/reference/v3-0-languages>
- <https://docs.microsoft.com/en-us/azure/cognitive-services/translator/reference/v3-0-translate>
- <https://docs.microsoft.com/en-us/azure/cognitive-services/translator/reference/v3-0-detect>
- <https://docs.microsoft.com/en-us/azure/cognitive-services/translator/reference/v3-0-transliterate>
- <https://docs.microsoft.com/en-us/azure/cognitive-services/translator/reference/v3-0-break-sentence>
- <https://docs.microsoft.com/en-us/azure/cognitive-services/translator/reference/v3-0-dictionary-lookup>
- <https://docs.microsoft.com/en-us/azure/cognitive-services/translator/translator-info-overview>
- <https://docs.microsoft.com/en-us/azure/cognitive-services/translator/quickstart-translator?tabs=nodejs>