

UD8 INTERFACES NATURALES.

Servicios de voz y texto



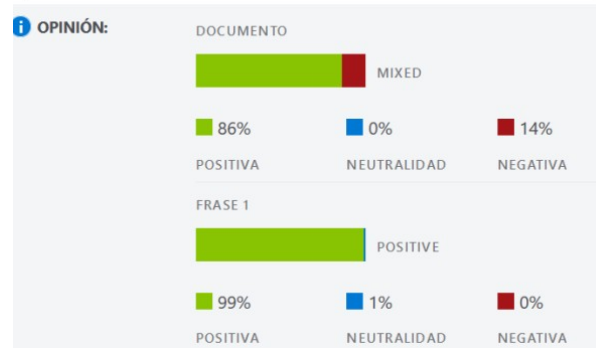
Ciclo: Desarrollo de Aplicaciones Multiplataforma
Módulo: Desarrollo de interfaces. Curso 2021-22



Funcionalidades de Text Analytic

- Text Analytics es el servicio de análisis de texto de Azure Cognitive Services. La última versión estable de este servicio es la 3.1. Esta versión del servicio nos ofrece cinco funcionalidades relacionadas con el análisis de texto:
- Análisis de opinión (Sentiment):** Esta funcionalidad nos permite saber si el sentimiento que se extrae de un texto es positivo, negativo o neutro.

We went to Contoso Steakhouse located at midtown NYC last week for a dinner party, and we adore the spot! They provide marvelous food and they have a great menu. The chief cook happens to be the owner (I think his name is John Doe) and he is super nice, coming out of the kitchen and greeted us all. We enjoyed very much dining in the place! The Sirloin steak I ordered was tender and juicy, and the place was impeccably clean. You can even pre-order from their online menu at www.contososteakhouse.com, call 312-555-0176 or send email to order@contososteakhouse.com! The only complaint I have is the food didn't come fast enough. Overall I highly recommend it!



- **Extracción de frases clave (Key phrases):** Con esta operación podremos extraer de un texto las palabras o grupos de palabras clave, que tienen mayor relevancia en el texto.

We went to Contoso Steakhouse located at midtown NYC last week for a dinner party, and we adore the spot! They provide marvelous food and they have a great menu. The chief cook happens to be the owner (I think his name is John Doe) and he is super nice, coming out of the kitchen and greeted us all. We enjoyed very much dining in the place! The Sirloin steak I ordered was tender and juicy, and the place was impeccably clean. You can even pre-order from their online menu at www.contososteakhouse.com, call 312-555-0176 or send email to order@contososteakhouse.com! The only complaint I have is the food didn't come fast enough. Overall I highly recommend it!



FRASES CLAVE:

place, online menu, great menu, marvelous food, midtown NYC, week, dinner party, Contoso Steakhouse, pre-order, John Doe, Sirloin steak, chief cook, owner, kitchen, spot, dining, complaint, email

- **Detección de idioma (Detect Language):** Esta funcionalidad permite determinar de forma automática el idioma en el que se ha escrito un texto, reconociendo multitud de idiomas y dialectos.

We went to Contoso Steakhouse located at midtown NYC last week for a dinner party, and we adore the spot! They provide marvelous food and they have a great menu. The chief cook happens to be the owner (I think his name is John Doe) and he is super nice, coming out of the kitchen and greeted us all. We enjoyed very much dining in the place! The Sirloin steak I ordered was tender and juicy, and the place was impeccably clean. You can even pre-order from their online menu at www.contososteakhouse.com, call 312-555-0176 or send email to order@contososteakhouse.com! The only complaint I have is the food didn't come fast enough. Overall I highly recommend it!



IDIOMAS:

English (confiabilidad: 100 %)

- **Reconocimiento de entidades (Named Entity Recognition):** El reconocimiento de entidades nos permite detectar en un texto la presencia de información que se puede incluir en alguna de las categorías predeterminadas del servicio (como por ejemplo un lugar, un nombre de persona o un número de teléfono).

We went to Contoso Steakhouse located at midtown NYC last week for a dinner party, and we adore the spot! They provide marvelous food and they have a great menu. The chief cook happens to be the owner (I think his name is John Doe) and he is super nice, coming out of the kitchen and greeted us all. We enjoyed very much dining in the place! The Sirloin steak I ordered was tender and juicy, and the place was impeccably clean. You can even pre-order from their online menu at www.contososteakhouse.com, call 312-555-0176 or send email to order@contososteakhouse.com! The only complaint I have is the food didn't come fast enough. Overall I highly recommend it!

**i ENTIDADES CON
NOMBRE:**

Contoso [Organization]
Steakhouse [Location]
NYC [Location-GPE]
last week [DateTime-DateRange]
dinner party [Event]
chief cook [PersonType]
owner [PersonType]
John Doe [Person]
kitchen [Location-Structural]
Sirloin steak [Product]
www.contososteakhouse.com [URL]
312-555-0176 [Phone Number]

- **Entidades enlazadas** (Linked entities from a well known knowledge base): Con esta funcionalidad podemos obtener enlaces a una fuente de información fiable (fundamentalmente Wikipedia) para entidades reconocidas en el texto.

We went to Contoso Steakhouse located at midtown NYC last week for a dinner party, and we adore the spot! They provide marvelous food and they have a great menu. The chief cook happens to be the owner (I think his name is John Doe) and he is super nice, coming out of the kitchen and greeted us all. We enjoyed very much dining in the place! The Sirloin steak I ordered was tender and juicy, and the place was impeccably clean. You can even pre-order from their online menu at www.contososteakhouse.com, call 312-555-0176 or send email to order@contososteakhouse.com! The only complaint I have is the food didn't come fast enough. Overall I highly recommend it!

**i ENTIDADES
/INCLADAS:**

We went to Contoso [Steakhouse](#) located at midtown [NYC](#) last week for a dinner party, and we adore the spot! They provide marvelous food and they have a great menu. The chief cook happens to be the owner (I think his name is [John Doe](#)) and he is super nice, coming out of the kitchen and greeted us all. We enjoyed very much dining in the place! The [Sirloin steak](#) I ordered was tender and juicy, and the place was impeccably clean. You can even pre-order from their online menu at www.contososteakhouse.com, call 312-555-0176 or send email to order@contososteakhouse.com! The only complaint I have is the food didn't come fast enough. Overall I highly recommend it!

Características del servicio

- Estas son algunas de las características y limitaciones más importantes de Text Analytics:
 - No todas las funcionalidades comentadas están disponibles en todos los idiomas (el castellano está soportado por todas las funcionalidades).
 - Cada uno de los documentos que se envían en una solicitud no puede exceder de 5.120 caracteres. Si el texto a analizar es más grande, se debe dividir en trozos.
 - El número máximo de documentos que se pueden incluir en una solicitud depende de la funcionalidad.
 - En cualquier caso, el texto total incluido en una solicitud no puede ser mayor de 1 MB.

Análisis de opinión (Sentiment)

- El servicio de análisis de opinion (Sentiment) permite obtener el tipo de opinión presente en un texto (positiva, negativa o neutra). La petición al servicio REST será de tipo POST, y se configurará de la siguiente forma:
 - URL: `https://{endpoint}/text/analytics/v3.0/sentiment` (el endpoint lo obtendremos en el portal de Azure).
 - Cabeceras:
 - Content-Type: `application/json` o `text/json`.
 - Ocp-Apim-Subscription-Key: clave para acceder a la API, que obtendremos del portal de Azure.

– Cuerpo:

- Los documentos a analizar en JSON, por ejemplo:

```
{
  "documents": [
    {
      "language": "en",
      "id": "1",
      "text": "Hello world. This is some input text that I love."
    },
    {
      "language": "en",
      "id": "2",
      "text": "It's incredibly sunny outside! I'm so happy."
    }
  ]
}
```

- La respuesta se recibirá en formato JSON, y contendrá la siguiente información para cada uno de los documentos:
 - **sentiment**: sentimiento general del documento (positive, negative, neutral o mixed).
 - **confidenceScores**: confianza asociada a cada uno de los sentimientos (entre 0 y 1).
 - **sentences**: detalle del sentimiento para cada oración del documento. Se incluye además la posición de la oración en el documento.

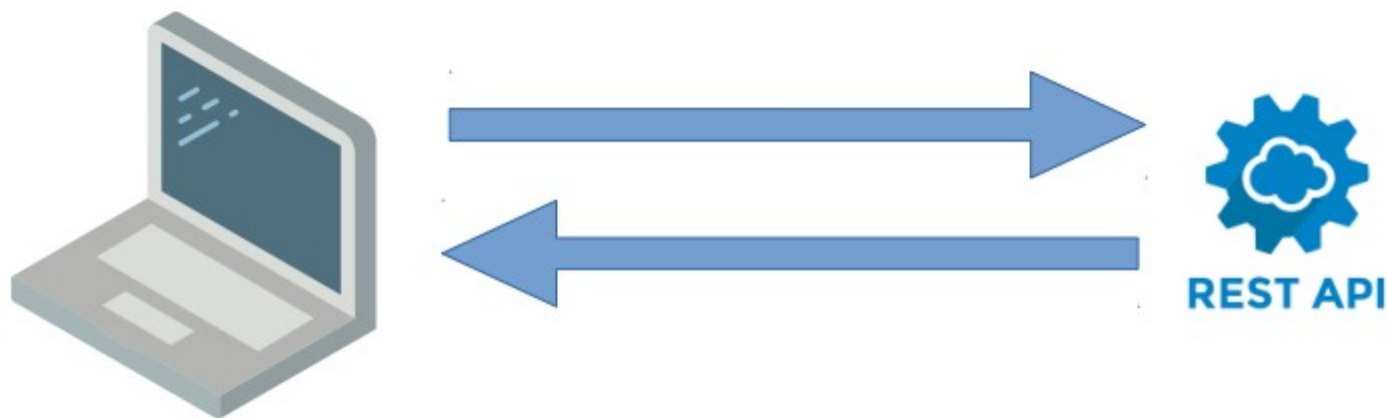
```
{
  "documents": [
    {
      "id": "1",
      "sentiment": "positive",
      "confidenceScores": {
        "positive": 0.98,
        "neutral": 0.01,
        "negative": 0.01
      },
      "sentences": [
        {
          "text": "Hello world.",
          "sentiment": "neutral",
          "confidenceScores": {
            "positive": 0.07,
            "neutral": 0.91,
            "negative": 0.02
          },
          "offset": 0,
          "length": 12
        },
        {
          "text": "This is some input text that I love.",
          "sentiment": "positive",
          "confidenceScores": {
            "positive": 0.98,
            "neutral": 0.01,
            "negative": 0.01
          },
          "offset": 13,
          "length": 36
        }
      ]
    },
    {
      "id": "2",
      "sentiment": "negative",
      "confidenceScores": {
        "positive": 0.01,
        "neutral": 0.01,
        "negative": 0.98
      },
      "sentences": [
        {
          "text": "I hate this.",
          "sentiment": "negative",
          "confidenceScores": {
            "positive": 0.01,
            "neutral": 0.01,
            "negative": 0.98
          },
          "offset": 0,
          "length": 12
        }
      ]
    }
  ],
  "warnings": []
}
```

```

    },
    {
      "id": "2",
      "sentiment": "positive",
      "confidenceScores": {
        "positive": 0.98,
        "neutral": 0.01,
        "negative": 0.01
      },
      "sentences": [
        {
          "text": "It's incredibly sunny outside!",
          "sentiment": "neutral",
          "confidenceScores": {
            "positive": 0.04,
            "neutral": 0.93,
            "negative": 0.03
          },
          "offset": 0,
          "length": 30
        },
        {
          "text": "I'm so happy.",
          "sentiment": "positive",
          "confidenceScores": {
            "positive": 0.98,
            "neutral": 0.01,
            "negative": 0.01
          },
          "offset": 31,
          "length": 13
        }
      ],
      "warnings": []
    }
  ],
  "errors": [],
  "modelVersion": "2020-04-01"
}

```

Parámetros	
Cabeceras	<i>Content-Type, Ocp-Apim-Subscription-Key</i>
Cuerpo	Colección de documentos a analizar (en JSON)



<i>sentiment</i>	Opinión del documento (<i>positive, neutral, negative, mixed</i>)
<i>confidenceScores</i>	Puntuación asociada a cada posible opinión (de 0 a 1)

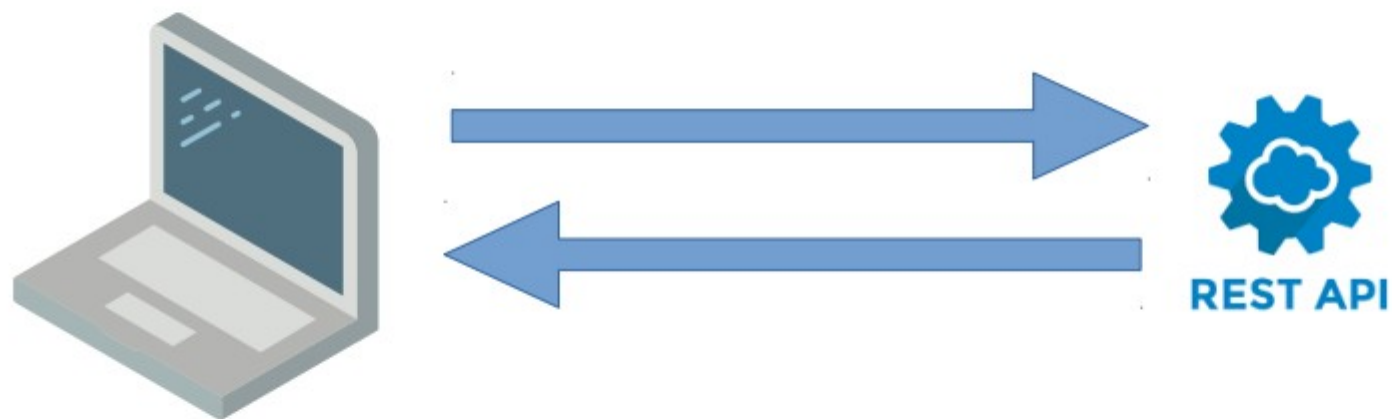
Frases clave (Key Phrases)

- La funcionalidad de obtención de frases clave permite extraer las palabras o grupos de palabras más relevantes del texto. La petición al servicio REST será de tipo POST, y se configurará de la siguiente forma:
 - URL: `https://{endpoint}/text/analytics/v3.0/keyPhrases` (el endpoint lo obtendremos en el portal de Azure).
 - Cabeceras:
 - Content-Type: `application/json` o `text/json`.
 - Ocp-Apim-Subscription-Key: clave para acceder a la API, que obtendremos del portal de Azure.
 - Cuerpo:
 - Los documentos a analizar en JSON, con el mismo formato que en la funcionalidad anterior.

- En el cuerpo de la respuesta recibiremos la lista de frases clave para cada uno de los documentos enviados en la petición.

```
{
  "documents": [
    {
      "id": "1",
      "keyPhrases": [
        "world",
        "input text"
      ],
      "warnings": []
    },
    {
      "id": "2",
      "keyPhrases": [
        "monde"
      ],
      "warnings": []
    }
  ]
}
```

Parámetros	
Cabeceras	<i>Content-Type, Ocp-Apim-Subscription-Key</i>
Cuerpo	Colección de documentos a analizar (en JSON)



<i>keyPhrases</i>	Lista de frases clave (palabras o grupos de palabras) detectadas.
-------------------	---

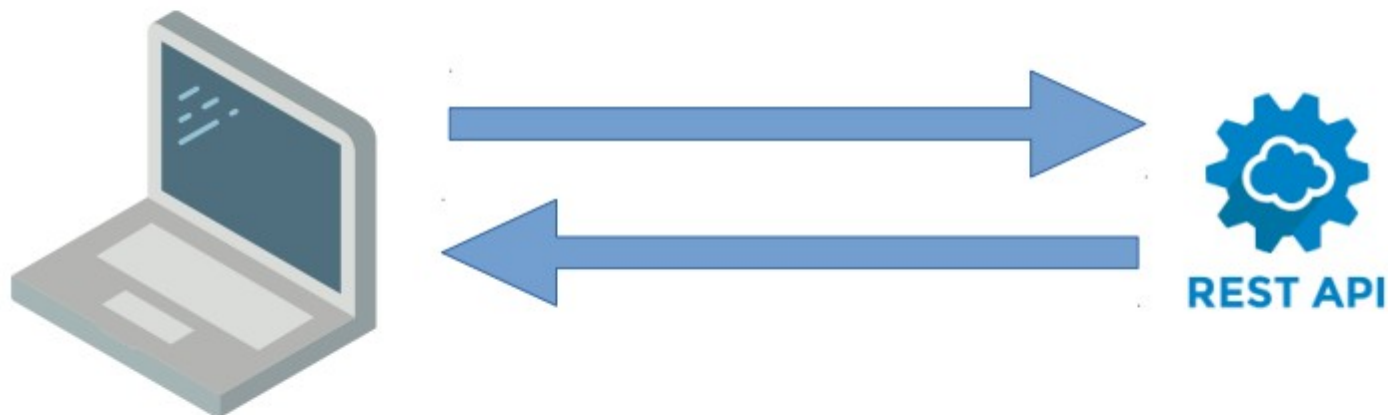
Reconocimiento de entidades (Named Entity Recognition)

- Esta funcionalidad permite detectar en un texto entidades pertenecientes a una serie de categorías predefinidas. La lista de categorías se puede consultar en la documentación de Text Analytics. La petición al servicio REST será de tipo POST, y se configurará de la siguiente forma:
 - URL: `https://{endpoint}/text/analytics/v3.0/entities/recognition/general` (el endpoint lo obtendremos en el portal de Azure).

- Cabeceras:
 - Content-Type: application/json o text/json.
 - Ocp-Apim-Subscription-Key: clave para acceder a la API, que obtendremos del portal de Azure.
- Cuerpo:
 - Los documentos a analizar en JSON, con el mismo formato que en las funcionalidades anteriores.
- La respuesta se recibirá en formato JSON, y contendrá para cada documento, la lista de entidades reconocidas. Para cada una de ellas se incluirá la categoría asociada (y en algunos casos también una subcategoría), la confianza del reconocimiento y la posición del texto en el documento.

```
"documents": [
{
  "id": "1",
  "entities": [
    {
      "text": "Seattle",
      "category": "Location",
      "offset": 26,
      "length": 7,
      "confidenceScore": 0.80
    },
    {
      "text": "last week",
      "category": "DateTime",
      "subcategory": "DateRange",
      "offset": 34,
      "length": 9,
      "confidenceScore": 0.8
    }
  ],
  "warnings": []
},
{
  "id": "2",
  "entities": [
    {
      "text": "Microsoft",
      "category": "Organization",
      "offset": 10,
      "length": 9,
      "confidenceScore": 0.99
    }
  ],
  "warnings": []
}
]
```

Parámetros	
Cabeceras	<i>Content-Type, Ocp-Apim-Subscription-Key</i>
Cuerpo	Colección de documentos a analizar (en JSON)



<i>entities</i>	Lista de entidades reconocidas, indicando para cada una la categoría y la posición en el texto.
-----------------	---

Traducción (Translator)

- Cognitive Services incluye un completo conjunto de servicios relacionados con la traducción englobados bajo el nombre de Translator. Los servicios de Translator se dividen en tres categorías:
 - Traducción de texto: es la funcionalidad básica de traducción. Además de la traducción entre dos idiomas también se incluye la transliteración (conversión entre alfabetos), la búsqueda de una palabra en el diccionario o la detección automática del idioma del texto.
 - Traducción de documentos: este servicio permite traducir un documento manteniendo el formato del documento original. Es compatible con PDF, HTML y los formatos de Microsoft Office.
 - Traducción personalizada: de forma análoga al servicio Custom Vision en el análisis de imágenes, Custom Translator permite definir modelos personalizados de traducción para dominios específicos.

- Cabe destacar que la funcionalidad de Translator solamente se ofrece a través de la API REST (no se ha implementado un SDK). Para este tema, nos centraremos únicamente en la funcionalidad básica de traducción de texto.
- La petición al servicio REST será de tipo POST, y se configurará de la siguiente forma:
 - URL: <https://api.cognitive.microsofttranslator.com/translate>
 - Parámetros:
 - api-version: utilizaremos la versión 3.0.
 - from: lenguaje de origen del texto. Es opcional, si no se indica se detectará de forma automática.
 - to: lenguaje de salida de la traducción. Se puede traducir a la vez a varios lenguajes repitiendo el parámetro to.

– Cabeceras:

- Content-Type: application/json en UTF8.
- Ocp-Apim-Subscription-Key: clave para acceder a la API, que obtendremos del portal de Azure.

– Cuerpo:

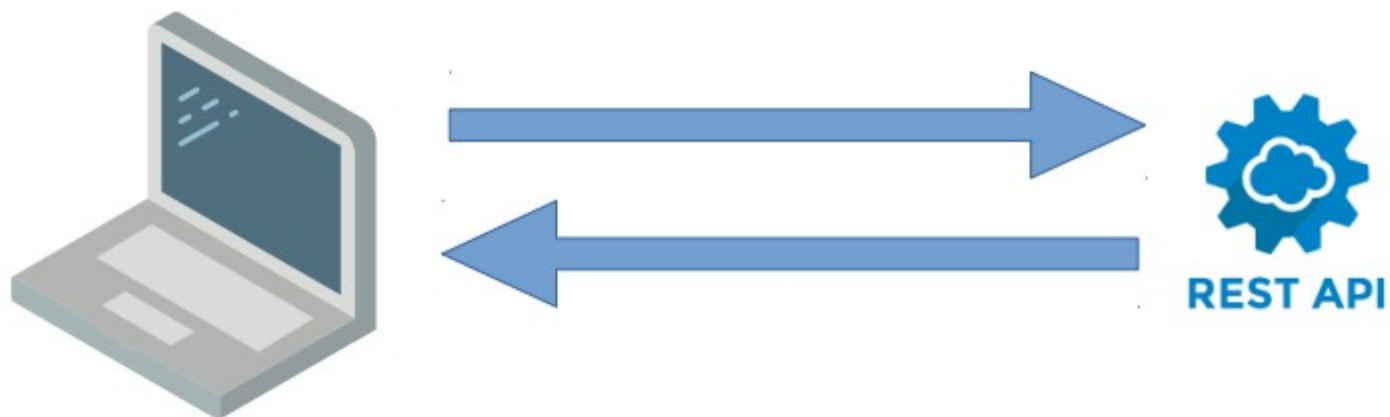
- Lista de textos a traducir en JSON, por ejemplo:

```
[  
  {"Text":"Hello, world!"}  
]
```

- La respuesta se recibirá en formato JSON, y contendrá para cada texto enviado en la petición, una lista con las traducciones en los idiomas solicitados.

```
[
  {
    "translations": [
      {
        "text": "Hallo Welt!",
        "to": "de"
      },
      {
        "text": "Salve, mondo!",
        "to": "it"
      }
    ]
  }
]
```

Parámetros	<i>from, to, api-version</i>
Cabeceras	<i>Content-Type, Ocp-Apim-Subscription-Key</i>
Cuerpo	Lista de textos a traducir en JSON (máximo 100 elementos, 10.000 caracteres en total).

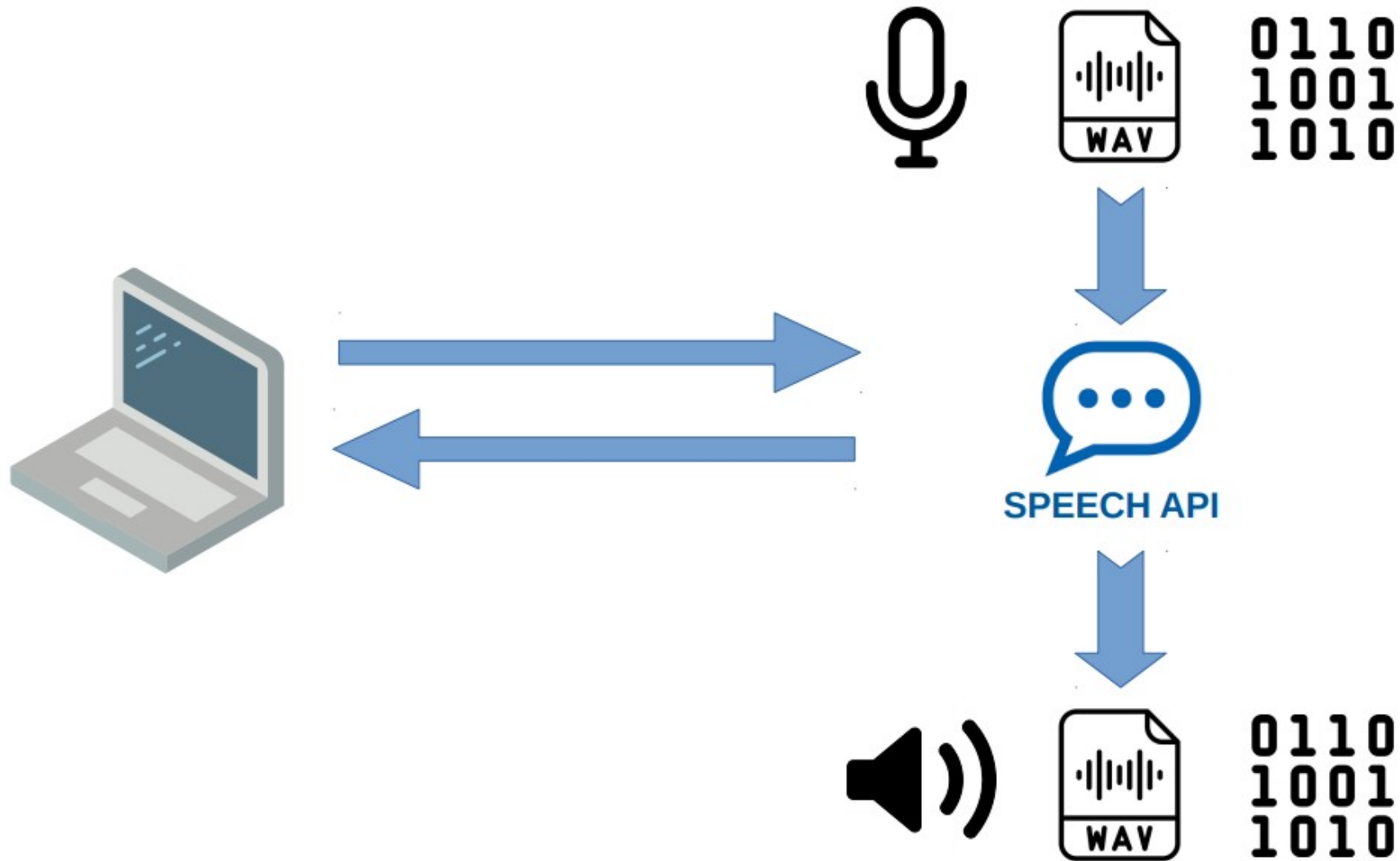


<i>text</i>	Texto traducido
<i>to</i>	Código del idioma de destino

Servicios de voz (Speech)

- Una de las cuatro categorías principales de servicios en Cognitive Services es la de servicios de voz. La oferta incluye los siguientes servicios:
 - Voz a texto: se utiliza para obtener el texto presente en una fuente de audio. Se incluye la posibilidad de realizar transcripciones en tiempo real, e incluso de definir modelos de habla personalizada para adaptar el servicio a un dominio específico.
 - Texto a voz: permite sintetizar una voz artificial que hable un texto. Se pueden personalizar distintos aspectos de la voz (diferentes hablantes, estilo de habla, pausas,...).
 - Traducción de voz: permite traducir directamente de una fuente de audio en un idioma a un audio en un idioma diferente.
 - Reconocimiento del hablante: es un servicio análogo al de identificación que vimos en el tema de Face, pero en este caso se identifican las personas por medio del habla.
 - Asistentes de voz: incluye servicios orientados a la creación de interfaces de usuario de aplicaciones basadas en la voz.

- En el caso de los servicios relacionados con Speech se recomienda de forma general el uso del SDK en lugar del acceso directo a la API REST.
- De hecho, existen funcionalidades que solamente son accesibles utilizando el SDK.
- Por esta razón, el SDK de Speech tiene un grado de actualización mayor que el del resto de servicios, y está disponible para más lenguajes y plataformas.



Recursos

- Página principal de Text Analytics

<https://azure.microsoft.com/es-es/services/cognitive-services/text-analytics>

- Text Analytics en Microsoft Docs

<https://docs.microsoft.com/es-es/azure/cognitive-services/text-analytics>

- Referencia de la API Text Analytics

<https://westus2.dev.cognitive.microsoft.com/docs/services/TextAnalytics-v3-0>

- Ejemplos en GitHub

<https://github.com/Azure-Samples/cognitive-services-quickstart-code>

- Página principal de Translator

<https://azure.microsoft.com/es-es/services/cognitive-services/translator>

- Translator en Microsoft Docs

<https://docs.microsoft.com/es-es/azure/cognitive-services/translator>

- Página principal de Speech

<https://azure.microsoft.com/es-es/services/cognitive-services/speech-services>

- Speech en Microsoft Docs

<https://docs.microsoft.com/es-es/azure/cognitive-services/speech-service>

- Ejemplos de Speech en GitHub

<https://github.com/Azure-Samples/cognitive-services-speech-sdk/tree/master/quickstart>

Creación del recurso Text Analytics en el portal de Azure

- Para poder utilizar la API de análisis de texto deberemos crear un recurso de tipo Text Analytics en el portal de Azure.
- En la ventana de creación del recurso se solicitará la siguiente información: •
 - Suscripción: deberemos escoger nuestra suscripción de Azure para estudiantes.
 - Grupo de recursos: es un contenedor que permite agrupar recursos de Azure relacionados entre sí. Se puede crear un único grupo de recursos para todos los recursos que iremos creando en el curso, o uno para cada uno de los temas.
 - Región: establece en cuál de las regiones de Azure se creará el recurso. Podemos dejar la región por defecto (normalmente Este de EE. UU.) aunque en alguna ocasión es posible que necesitemos seleccionar otra región (algunos servicios no están disponibles en todas las regiones para la suscripción de Azure para estudiantes).

- Nombre: el nombre que le queremos dar al recurso. Este nombre también se utilizará para formar la URL del punto de conexión a la API.
- Plan de tarifa: determina el nivel de servicio y la facturación asociada. En nuestro caso, seleccionaremos el plan gratuito (F0).

Una vez creado el recurso, nos dirigiremos a la sección Claves y punto de conexión del recurso para obtener la URL a la que deberemos hacer las peticiones (extremo) y la clave de autenticación de la API (se ofrecen dos claves pero solo es necesario utilizar una de ellas). Estos dos valores (la URL y la clave) son los que necesitaremos para utilizar el servicio desde nuestro programa.

Creación del recurso Translator en el portal de Azure

- Para poder utilizar el servicio de traducción deberemos crear un recurso de tipo Translator en el portal de Azure.
- En la ventana de creación del recurso se solicitará la siguiente información: •
 - Suscripción: deberemos escoger nuestra suscripción de Azure para estudiantes.
 - Grupo de recursos: es un contenedor que permite agrupar recursos de Azure relacionados entre sí. Se puede crear un único grupo de recursos para todos los recursos que iremos creando en el curso, o uno para cada uno de los temas.
 - Nombre: el nombre que le queremos dar al recurso. En este caso, la URL del extremo será común para todos los usuarios y no se formará a partir de este nombre.
 - Región: establece en cuál de las regiones de Azure se creará el recurso. En este servicio, se recomienda utilizar la región Global.
 - Plan de tarifa: determina el nivel de servicio y la facturación asociada. En nuestro caso, seleccionaremos el plan gratuito (F0).

Creación del recurso Speech en el portal de Azure

- Para poder utilizar el servicio de voz crear un recurso de tipo Speech en el portal de Azure.
- En la ventana de creación del recurso se solicitará la siguiente información:
 - Suscripción: deberemos escoger nuestra suscripción de Azure para estudiantes.
 - Grupo de recursos: es un contenedor que permite agrupar recursos de Azure relacionados entre sí. Se puede crear un único grupo de recursos para todos los recursos que iremos creando en el curso, o uno para cada uno de los temas.
 - Nombre: el nombre que le queremos dar al recurso. En este caso el nombre del recurso tampoco influirá en la formación de la URL del extremo.
 - Región: establece en cuál de las regiones de Azure se creará el recurso de aprendizaje. Podemos dejar la región por defecto (en este caso Centro-Sur de EE. UU.) aunque en alguna ocasión es posible que necesitemos seleccionar otra región (algunos servicios no están disponibles en todas las regiones para la suscripción de Azure para estudiantes).
 - Plan de tarifa: determina el nivel de servicio y la facturación asociada. En nuestro caso, seleccionaremos el plan gratuito (F0).

SDK de Text Analytics para C#

- Para poder utilizar el SDK de Text Analytics en nuestro proyecto será necesario añadir el paquete NuGet Azure.AI.TextAnalytics. Añadiremos la versión 5.0.0 (que es la que utiliza la versión 3.0 de la API de Text Analytics). Una vez añadido el paquete, haremos referencia al siguiente espacio de nombres en nuestro código:

```
using Azure.AI.TextAnalytics;
```

- La clase principal del SDK es la siguiente:
 - TextAnalyticsClient: la utilizaremos como punto de conexión a la API de Text Analytics. Requiere la URL y la clave obtenidas del portal de Azure. Contiene métodos para las diferentes funcionalidades de la API.

SDK de Speech para C#

- Para poder utilizar el SDK de los servicios de voz añadiremos el paquete NuGet Microsoft.CognitiveServices.Speech. Una vez añadido, haremos referencia a los siguientes espacios de nombres en nuestro código:

```
using Microsoft.CognitiveServices.Speech;  
using Microsoft.CognitiveServices.Speech.Audio;
```

- Las clases principales del SDK son las siguientes:
 - SpeechConfig: se utiliza para suministrar los datos de la suscripción (en este servicio será necesaria la clave y la región, pero no el extremo).
 - AudioConfig: nos permite configurar el origen y el destino de los servicios de voz, que podrá ser un dispositivo (micrófono o altavoces), un fichero WAV o un stream en memoria.
 - SpeechRecognizer: cliente de la API de voz a texto.
 - SpeechSynthesizer: cliente de la API de texto a voz.