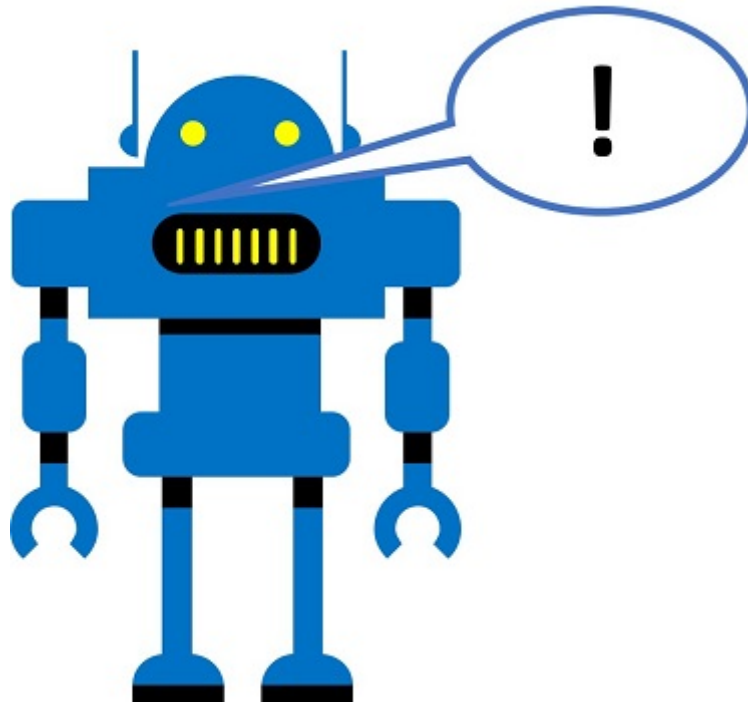


Speech 서비스

점점 더, 우리는 종종 음성 응답에 대한 기대를 가지고 인공지능(AI) 시스템과 대화함으로써 의사소통 할 수 있기를 기대한다.



음성 인식(*Speech Recognition*)(음성어를 해석하는 AI 시스템) 및 음성 합성(*Speech Synthesis*)(음성 반응을 생성하는 AI 시스템)은 음성 지원 AI 솔루션의 핵심 구성 요소이다.

Cognitive Services 리소스 만들기

소리로 된 음성을 해석하고 구두로 응답 할 수 있는 소프트웨어를 구축하기 위해서는 **Speech Cognitive** 서비스를 사용할 수 있는데 이 서비스는 간단한 방법으로 음성언어를 텍스트로 바꿀 수 있고 반대로 음성을 텍스트로 바꿀 수 있다.

이미 만들어진 Cognitive 서비스가없다면 다음과 같은 순서로 Azure 구독에서 **Cognitive Services** 리소스를 생성할 수 있다.

1. 브라우저의 새로운 탭을 열고, Azure portal(<https://portal.azure.com>)을 (<https://portal.azure.com>)을 입력 하고 , Microsoft계정으로 로그인한다..
2. + 리소스 만들기 버튼을 클릭하고, *Cognitive Services* 서비스를 찾은 다음, **Cognitive Services** 리소스를 다음과 같은 내용으로 생성한다.
 - 이름: 유일한 이름을 입력한다(가능하면 영문과 숫자사용).
 - 구독: *Azure* 구독선택.
 - 위치: 가능한 위치(한국 중부 추천):
 - 가격책정계층: 표준 S0
 - 리소스 그룹: 원하는 유리한 이름(가능하면 영문과 숫자사용).
3. 배포가 완료될 때까지 기다린다. 그런 다음 Cognitive Services 리소스로 이동하여 *개요 페이지에서 링크를 클릭하여 서비스 키를 관리한다. 클라이언트 응용 프로그램에서 Cognitive Services 리소스에 연결하려면 엔드포인트와 키가 필요하다.

Cognitive Services 리소스에 있는 키와 엔드포인트 가져오기

Cognitive Services 리소스를 사용하기 위해서는, 클라이언트 응용프로그램에서는 엔드포인트와 인증 키가 필요하다.

1. Azure portal에서, Cognitive Services 리소스를 선택하고 **키 및 엔드포인트** 페이지를 선택한 다음 **키1** 을 복사하여 아래의 **YOUR_COG_KEY**.를 붙여 넣는다.
2. 리소스에 있는 **엔드포인트** 를 복사해서 아래의 **YOUR_COG_ENDPOINT**.에 붙여 넣는다.
3. 리소스에 있는 **위치** 를 복사해서 아래의 **YOUR_COG_REGION**.에 붙여 넣는다.
4. 아래 셀을 선택한 다음 셀 왼쪽에있는 **셀 실행(▶)** 버튼을 클릭하여 아래 코드를 실행한다.

In []:

```
cog_key = 'YOUR_COG_KEY'
cog_endpoint = 'YOUR_COG_ENDPOINT'
cog_region = 'YOUR_COG_REGION'

print('Ready to use cognitive services in {} using key {}'.format(cog_region, cog_key))
```

Cognitive 서비스에 이쓴 Speech 서비스를 이용하기 위해서는 Azure Cognitive Services Speech SDK를 설치해야 한다.K

In []:

```
! pip install azure.cognitiveservices.speech
```

음성인식(Speech recognition)

"전등을 켜라" 또는 "전등을 끄라"와 같은 음성 명령을 수신하는 홈 자동화 시스템을 구축하려고 한다고 가정하자. 응용 프로그램은 오디오 기반 입력(사용자의 음성 명령)을 가져와서 구문을 나누고 분석할 수 있는 텍스트로 변환하여 해석할 수 있어야 한다.

이제 여러분은 몇 가지 말을 글로 옮겨 쓰게 할 준비가 되었다. 입력은 마이크 또는 오디오 파일일 수 있다. 여기서는 오디오 파일을 사용한다.

아래 셀을 실행하여 Speech 서비스의 음성 텍스트 변환(Speech-to-text) 기능을 사용하여 오디오를 글로 변환한다.

In []:

```
import os
import IPython
from azure.cognitiveservices.speech import SpeechConfig, SpeechRecognizer, AudioConfig

# 오디오 파일로부터 음성 명령을 가져온다.
file_name = 'light-on.wav'
audio_file = os.path.join('data', 'speech', file_name)

# 음성 인식을 설정한다
speech_config = SpeechConfig(cog_key, cog_region)
audio_config = AudioConfig(filename=audio_file) # 기본인 마이크 대신에 파일로 바꾼다
speech_recognizer = SpeechRecognizer(speech_config, audio_config)

# 일회성의 동기화되는 요청을 이용하여 음성명령을 글로 표현한다.
speech = speech_recognizer.recognize_once()

# 오디오를 실행하고 텍스트로 받아 쓴 명령을 보여준다.
IPython.display.display(IPython.display.Audio(audio_file, autoplay=True),
                        IPython.display.HTML(speech.text))
```

file_name 변수의 값을 *light-off.wav*로 바꾸고 셀을 다시 실행한다. 두개의 파일의 내용이 텍스트로 바르게 쓰여져야 한다.

음성합성(Speech synthesis)

이제 여러분은 스피치 서비스가 어떻게 말을 텍스트로 옮겨 쓸 수 있는지 보았을 것이다. 하지만 그 반대는 어떨까? 어떻게 텍스트를 음성으로 변환할 수 있을까?

흠 자동화 시스템이 조명을 켜는 명령을 해석했다고 가정해 보자. 적절한 응답은 구두로 명령을 확인하는 것일 수 있다(실제로 명령된 작업을 수행하는 것!).

In []:

```
import os
import IPython
from azure.cognitiveservices.speech import SpeechConfig, SpeechSynthesizer, AudioConfig

# 음성으로 보내질 텍스트를 가져온다.
response_text = 'Turning the light on.'

# 음성 학습 서비스를 설정한다
speech_config = SpeechConfig(cog_key, cog_region)
output_file = os.path.join('data', 'speech', 'response.wav')
audio_output = AudioConfig(filename=output_file) # Use a file instead of default (speakers)
speech_synthesizer = SpeechSynthesizer(speech_config, audio_output)

# 텍스트를 음성으로 바꾼다
result = speech_synthesizer.speak_text(response_text)

# 추출된 오디오 파일을 실행한다.
IPython.display.display(IPython.display.Audio(output_file, autoplay=True),
                        IPython.display.Image(data=os.path.join("data", "speech", response_text
                                                                .lower() + '.jpg'))))
```

response_text 변수값을 *Turning the light off.*(마지막에 마침표 포함)로 바꾸고 셀을 실행해서 결과를 들어보자.

심화학습

노트북에서 예제로 사용된 것은 매우 간단한 Speech Cognitive 서비스이다.

Speech 서비스내 음성 텍스트 변환 서비스에 대해서 좀 더 알아보고 싶다면 여기를 참고하라 [음성 텍스트 변환\(speech-to-text\)](https://docs.microsoft.com/azure/cognitive-services/speech-service/index-speech-to-text) (<https://docs.microsoft.com/azure/cognitive-services/speech-service/index-speech-to-text>).

그리고 Speech 서비스 내 텍스트 음성 변환 서비스에 대해서 알아보고 싶다면 다음을 참고하라 [텍스트 음성 변환 서비스\(text-to-speech\)](https://docs.microsoft.com/azure/cognitive-services/speech-service/index-text-to-speech) (<https://docs.microsoft.com/azure/cognitive-services/speech-service/index-text-to-speech>).

In []: