

ETRI 음성인식 및 실습

[1] 음성 출력/합성 현황

- TTS 기법 정리 : <https://theaisummer.com/text-to-speech/>
- STT 기법 정리 : <https://theaisummer.com/speech-recognition/>
- 응용분야와 목적에 따라, 대기업 클라우드 사용과 직접제작 인식기 사용 중 선택
: 사용목적, 유연성, 인식 성능, 비용

[2] 한글 음성 합성 능력 비교

- 2021년 1월 기준 음성합성 Cloud Service 비교 :
<https://www.atlantis-press.com/journals/ijndc/125951137/view>

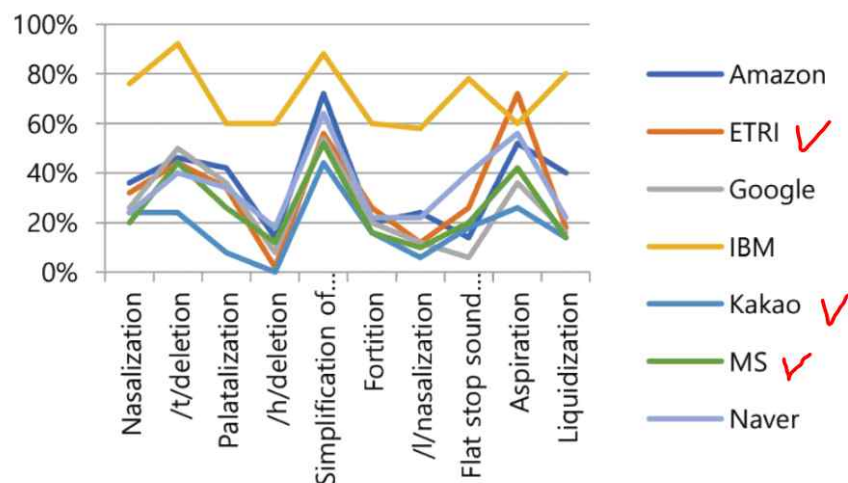


Figure 6 Comparison of WER% of phonological rules by company.

[3] ETRI Open API Key 발급 및 관리

- https://aiopen.etri.re.kr/key_main.php

안녕하세요. 김병기 님,

공공 인공지능 오픈 API·DATA 서비스에 관심을 가져주셔서 감사드립니다.

회원님께서 신청하신 API Key 정보는 아래와 같습니다.

신청자 이메일: pkkim@silla.ac.kr

API Key: 3ab9d53f-cce1-4d82-b207-[REDACTED]

발급일: 2021-09-06 21:04:02

이제부터 발급받으신 API Key를 사용하여 인공지능 Open API 서비스를 이용하실 수 있습니다.

발급된 API Key는 제3자가 이용할 수 없습니다.

서비스 사용에 대한 자세한 내용은 'ETRI 인공지능 Open API 서비스'의 이용약관
(http://aiopen.etri.re.kr/service_list.php)를 꼭 참고해주세요.

감사합니다.

관리자 드림

[4] etri 음성인식 API 사용

- https://aiopen.etri.re.kr/guide_recognition.php

Python

python 3.0을 기준으로 작성되었습니다.

HTTP 통신을 위해 urllib3 라이브러리를 사용하여 제공하고 있습니다. Python 3.0 이하의 버전에서 예제를 실행하기 위해서는 별도로 urllib3의 설치が必要です. 설치에 대한 설명은 <https://pypi.python.org/pypi/urllib3> 를 참고하시기 바랍니다. urllib3 라이브러리에 대한 자세한 설명은 <https://urllib3.readthedocs.io/en/latest/> 에서 확인하실 수 있습니다.

```
#!/usr/bin/env python
#-*- coding:utf-8 -*-
import urllib3
import json
import base64
openApiURL = "http://aiopen.etri.re.kr:8000/WiseASR/Recognition"
accessKey = "YOUR_ACCESS_KEY"
audioFilePath = "AUDIO_FILE_PATH"
languageCode = "LANGUAGE_CODE"

file = open(audioFilePath, "rb")
audioContents = base64.b64encode(file.read()).decode("utf8")
file.close()

requestJson = {
    "access_key": accessKey,
    "argument": {
        "language_code": languageCode,
        "audio": audioContents
    }
}

http = urllib3.PoolManager()
response = http.request(
    "POST",
    openApiURL,
    headers={"Content-Type": "application/json; charset=UTF-8"},
    body=json.dumps(requestJson)
)

print("[responseCode] " + str(response.status))
print("[responBody]")
print(str(response.data, "utf-8"))
```

다음은 파라미터에 대한 설명입니다.

Field 명	타입	필수 여부	설명
access_key	String	O	API 사용을 위해 ETRI에서 발급한 사용자 API Key 입니다.
argument	Object	O	API 사용 요청 시 분석을 위해 전달할 내용입니다.
language_code	String	O	음성인식의 입력 음성 언어 코드입니다. 요청할 수 있는 언어 코드는 아래와 같습니다. korean : 한국어 음성인식 코드 english : 영어 음성인식 코드 japanese : 일본어 음성인식 코드 chinese : 중국어 음성인식 코드 spanish : 스페인어 음성인식 코드 french : 프랑스어 음성인식 코드 german : 독일어 음성인식 코드 russian : 러시아어 음성인식 코드 vietnam : 베트남어 음성인식 코드 arabic : 아랍어 음성인식 코드
audio	String	O	음성인식을 할 녹음된 음성파일의 base64-encoded string 입니다.

[5] 대기업 클라우드 음성인식 API 사용의 한계

- 다국어(한국어+영어 등)에 대한 한계
- 신조어나 변화에 대한 개인적 학습이 불가능
- 인식결과와 신뢰도에 대한 자료 부재

[6] 음성합성 실습

```
md etri
```

```
cd etri
```

```
conda create -n etri pyAudio matplotlib ffmpeg
```

```
conda activate etri
```

```
pip install speechrecognition gtts playsound==1.2.2 pydub librosa
```

실행

```
python etri.py ko
```