

AI on Cloud



Microsoft



AIT

Asian Institute of Technology

SCB ACADEMY

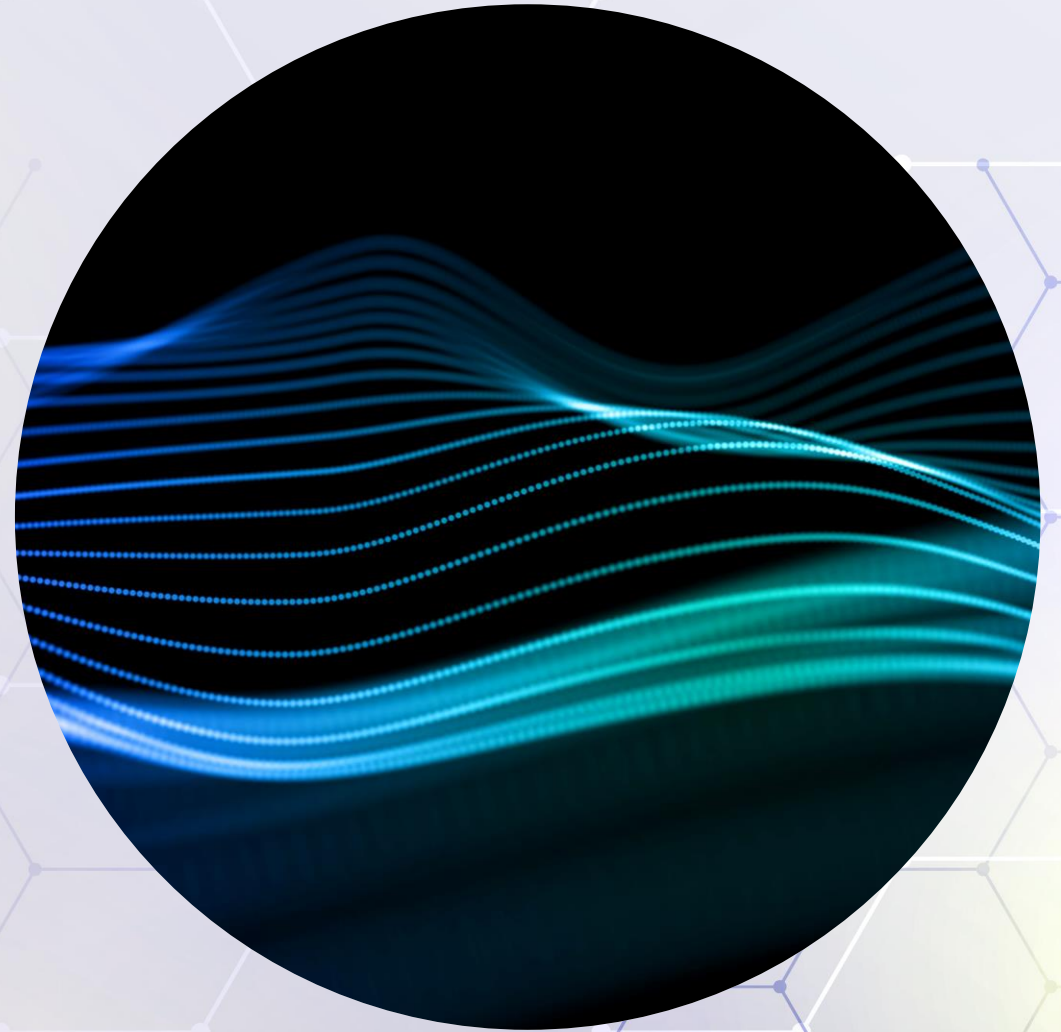


AI on Cloud

Chapter

4

การประมวลผล
ภาษาธรรมชาติ
(Natural Language
Processing (NLP))



AIT
Asian Institute of Technology

SCB ACADEMY

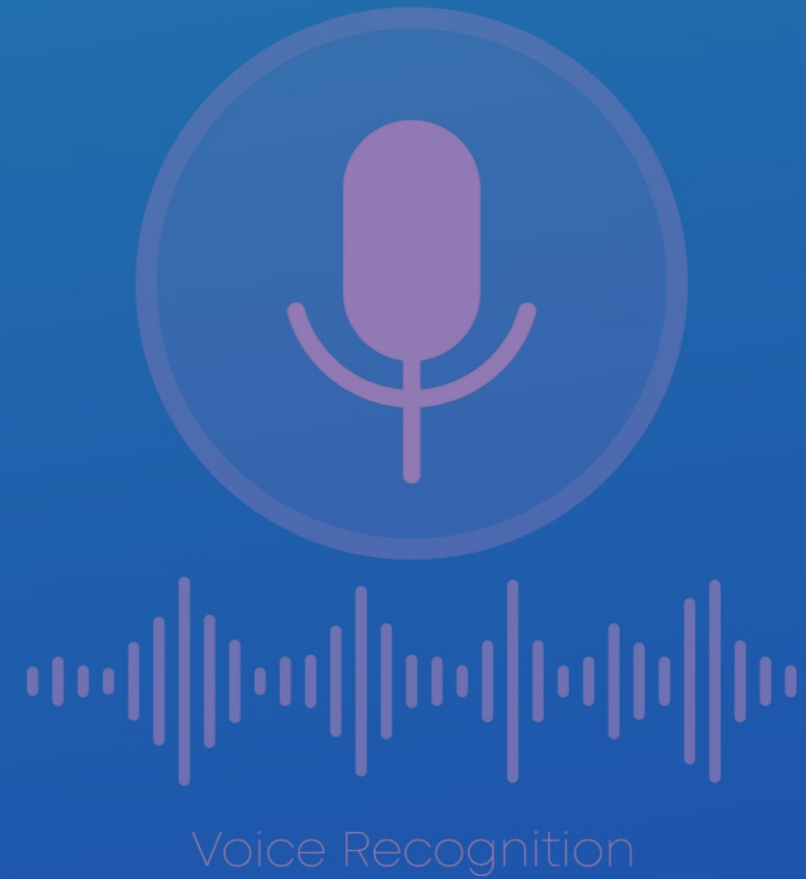


Lab 2: Speech Recognition and Speech Synthesis



Speech Recognition and Speech Synthesis

- **Speech Recognition** – the ability to detect and interpret spoken input.
- **Speech Synthesis** – the ability to generate spoken output.



Speech Recognition

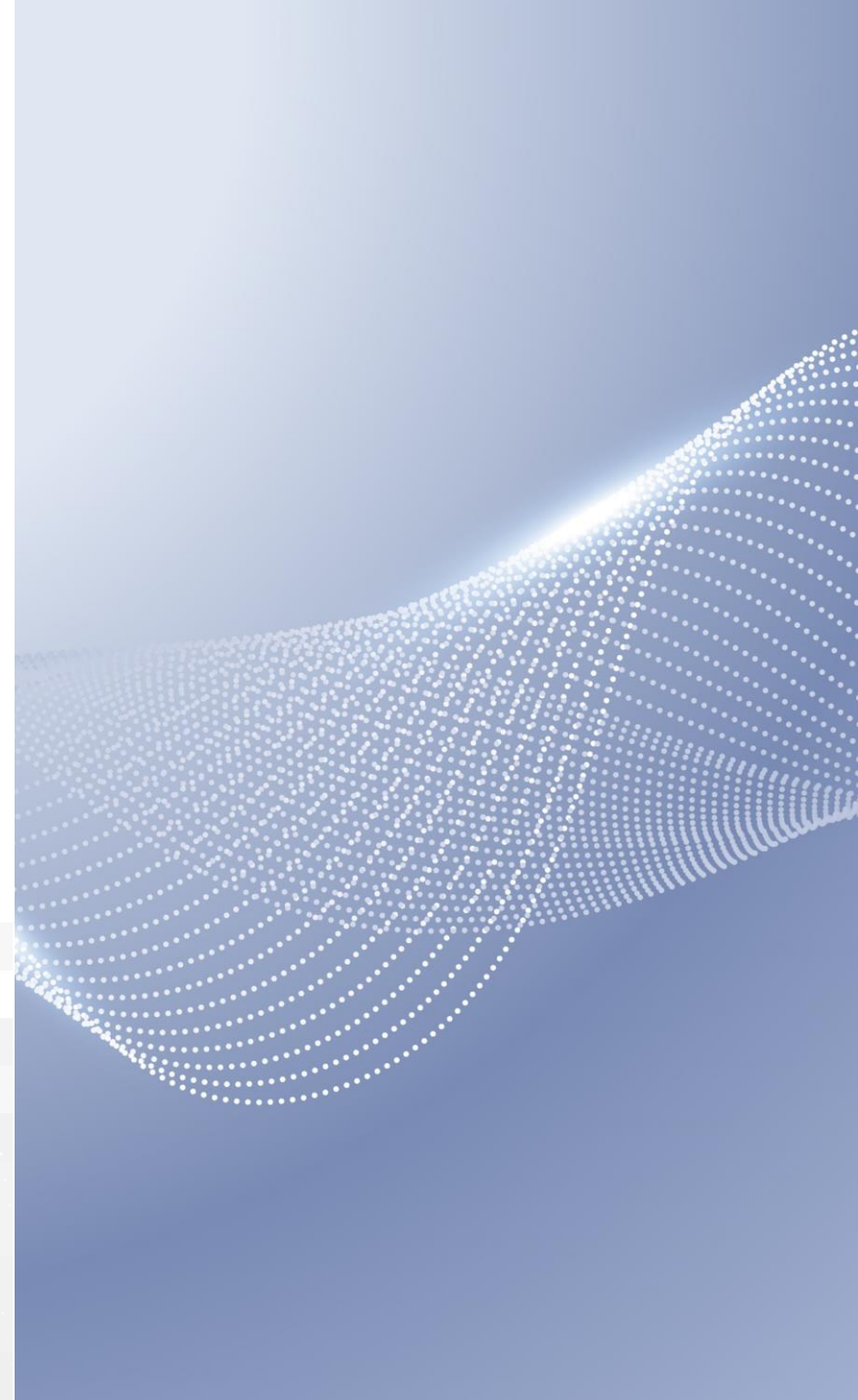
ตัวอย่างการประยุกต์ใช้

- สร้าง CC (Closed Captions) จาก Recorded หรือ Live Videos
- ถอดเสียงจากบทสนทนา หรือบันทึกการประชุม
- พิมพ์ตามคำบอก (Automated note Dictation)
- สั่งงานด้วยเสียง



Speech Synthesis

- Converting text to speech
- โดยส่วนใหญ่จำเป็นต้องระบุ
 1. ข้อความที่ต้องการให้พูด
 2. รูปแบบเสียง



Speech Synthesis

ตัวอย่างการประยุกต์ใช้

- Generating spoken responses to user input e.g., Siri, Alexa, etc.
- Creating voice menus for telephone systems.
- Reading email or text messages aloud in hands-free scenarios.
- Broadcasting announcements in public locations, such as railway stations or airports.



Speech on Azure

- The **Speech-to-Text** API
- The **Text-to-Speech** API



The Speech-to-Text API

- Real-time or Batch Transcription
 - Input: audio source (real-time audio stream from a microphone or an audio file)
 - Output: text
- Based on the Universal Language Model, trained by Microsoft.
- Microsoft-owned and deployed to Microsoft Azure.
- **เหมาะกับการใช้งาน 2 สถานการณ์ Conversational and Dictation.**
- สามารถสร้างและ Train Custom Models เองได้ หาก Pre-built Models ไม่ตอบโจทย์

The Text-to-Speech API

- To Convert text input to audible speech (computer speaker or written to an audio file)
- **Speech synthesis voices**
 - กำหนดรูปแบบเสียงที่ต้องการ
 - สามารถสร้าง custom voices เพื่อใช้งานกับ text-to-speech API ได้



Lab 2: Speech Recognition and Speech Synthesis

- เป้าหมาย : ฟังคำสั่งในการถามเวลา และตอบกลับเป็นเสียง
 - Input: ไฟล์เสียงที่พูดว่า "What time is it?"
 - Output: ไฟล์เสียงคำตอบ
- Tool: *Cognitive Service* Resource



Step 1 Create a *Cognitive Services* Resource

Step 2 Configure and Run a Client Application

ใช้ Cloud Shell UU Azure

```
git clone https://github.com/MicrosoftLearning/AI-900-AIFundamentals ai-900
```

Code สำหรับตัวอย่าง application คือ **Speaking-clock.ps1**

Step 3 Transcribe a spoken input and synthesize an appropriate spoken response.

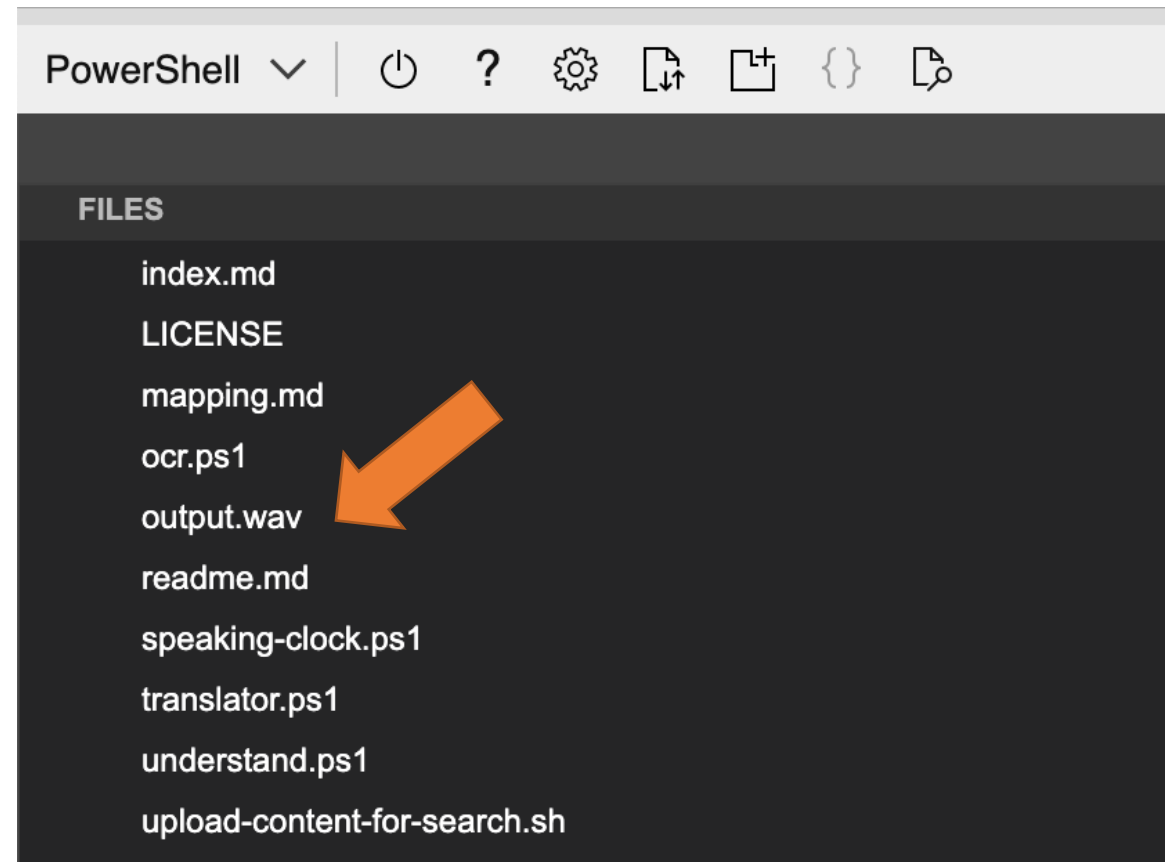
```
./speaking-clock.ps1
```

```
Analyzing audio...
```

```
You said 'What time is it?'
```

```
Synthesizing speech...
```

```
Response saved in output.wav
```





End of Lab2

- Speech-to-text (S2T)
- Text-to-speech (T2S)

Thank you