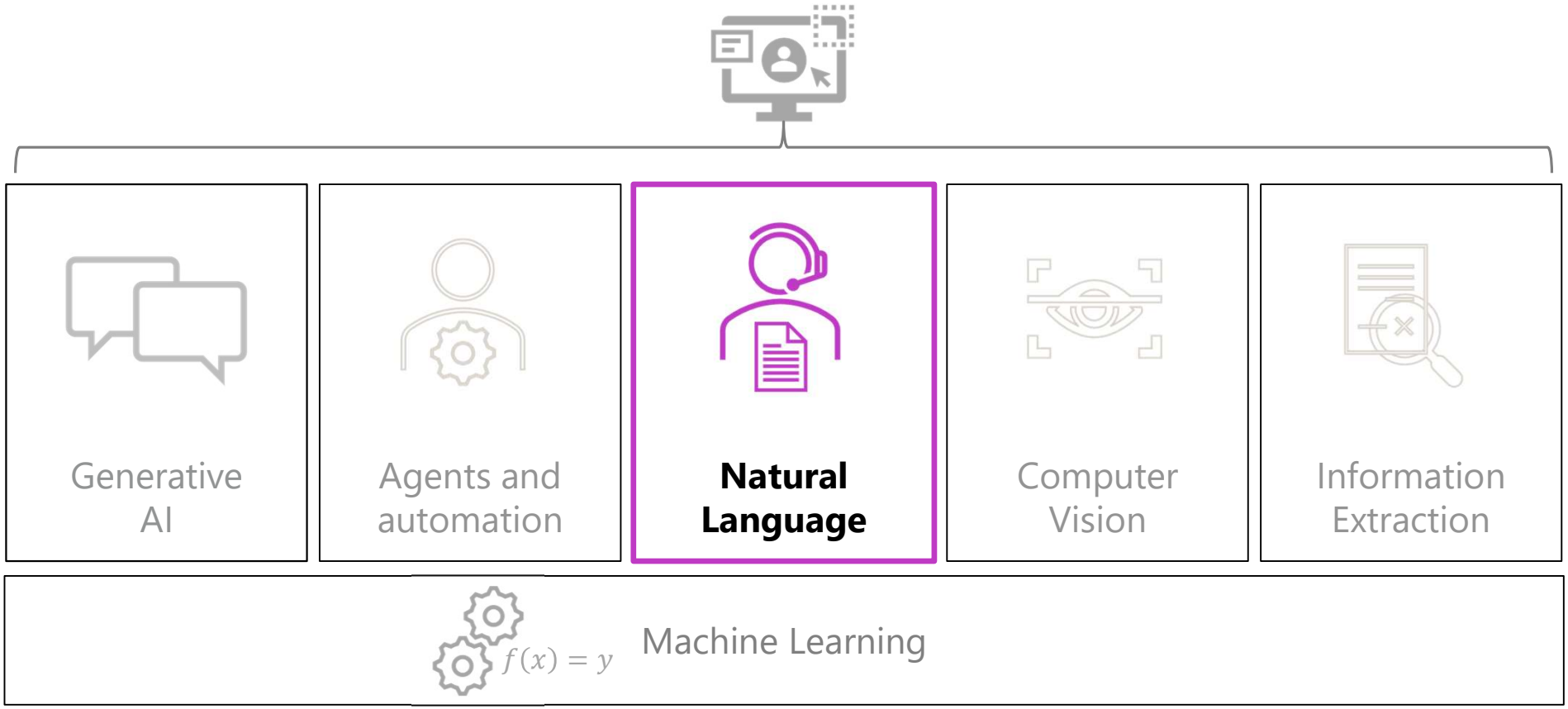
Microsoft

Introduction to AI in Azure:
Natural Language Processing

Natural Language Processing in context



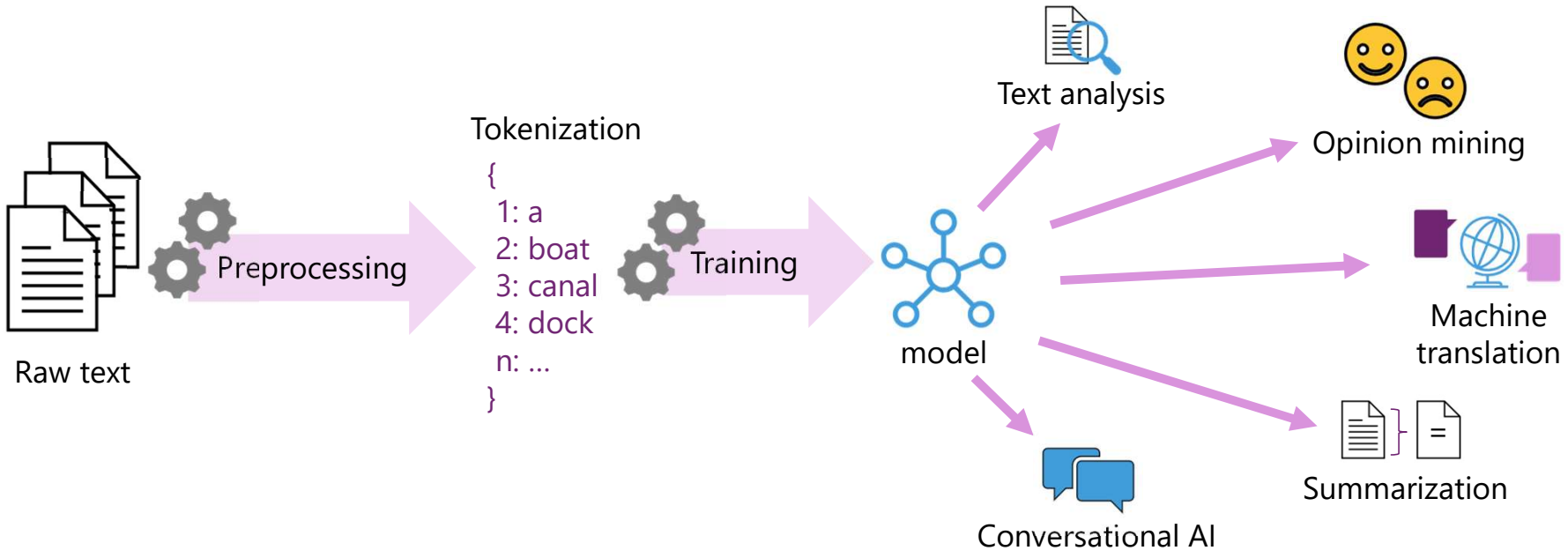
Agenda

- Introduction to natural language processing concepts
- Get started with natural language processing in Microsoft Foundry
- Get started with speech in Microsoft Foundry

Introduction to natural language processing concepts

https://aka.ms/mslearn-nlp

What is natural language processing?



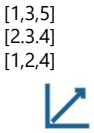
Techniques for language modeling and text analysis

Statistical techniques

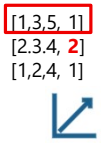
- *Classification*: Use algorithms like Naïve Bayes and Logistic Regression to classify text based on the presence of specific tokens (words). For example, implementing an email spam filter based on terms that often appear in junk mail but not in "regular" email.

- *Term Frequency / Inverse Document Frequency*: A statistical technique that compares the frequency of a specific term in individual documents with its frequency across a full *corpus* of documents. Used to determine the core *subject* of an individual document

Semantic modeling techniques

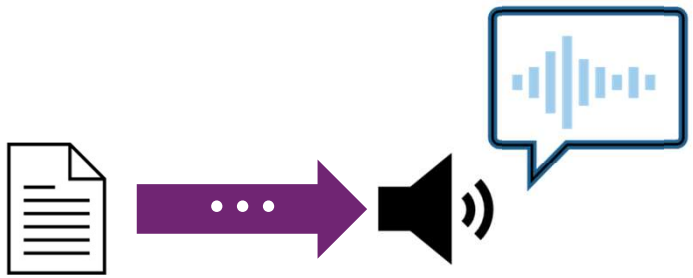


- *Transformer models*: Represent language tokens as vector-based *embeddings* that encapsulate semantic relationships. Related terms have similar vector directions. Used for tasks like *translation* and *next-word-completion*.



- *Attention*: A technique used in transformer models to improve next-word prediction based on token context. Weights are applied to tokens to reflect their influence on the next word. Using attention in the transformer architecture enables the latest advances in generative AI models

Speech processing



Speech Synthesis
(Text to speech)
1. *Tokenize* text
2. Map tokens to *phonemes*
3. Generate audio signal

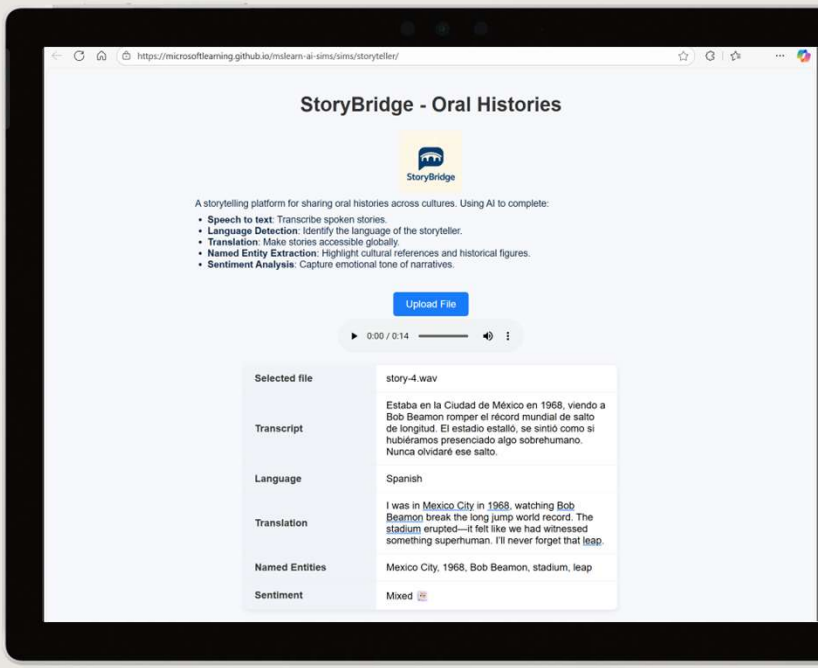
Speech Recognition
(Speech to text)
1. Capture audio signal
2. Break into *phonemes*
3. Map *phonemes* to text *tokens*

Exercise

Explore natural language processing scenarios

In this exercise, you'll explore solutions that include natural language processing functionality.

Start the exercise at:
https://go.microsoft.com/fwlink/?linkid=2334225



Knowledge check

1 **What is the primary purpose of tokenization in natural language processing (NLP)?**

- ☐ To translate text into another language.
- ☐ To summarize large documents.
- ☑ To break down text into smaller units for analysis.

2 **Which of the following techniques is used to determine the relative importance of words in a document within the context of a larger collection of documents?**

- ☐ Naïve Bayes
- ☑ TF-IDF (Term Frequency-Inverse Document Frequency)
- ☐ Word2Vec

3 **Which of the following best describes the role of *embeddings* in natural language processing (NLP)?**

- ☐ They visualize text data in two-dimensional space for easier interpretation.
- ☐ They summarize large text corpora into short, meaningful sentences.
- ☑ They convert language tokens into vectors that capture semantic relationships.

Get started with natural language processing in Microsoft Foundry

https://aka.ms/mslearn-azure-language

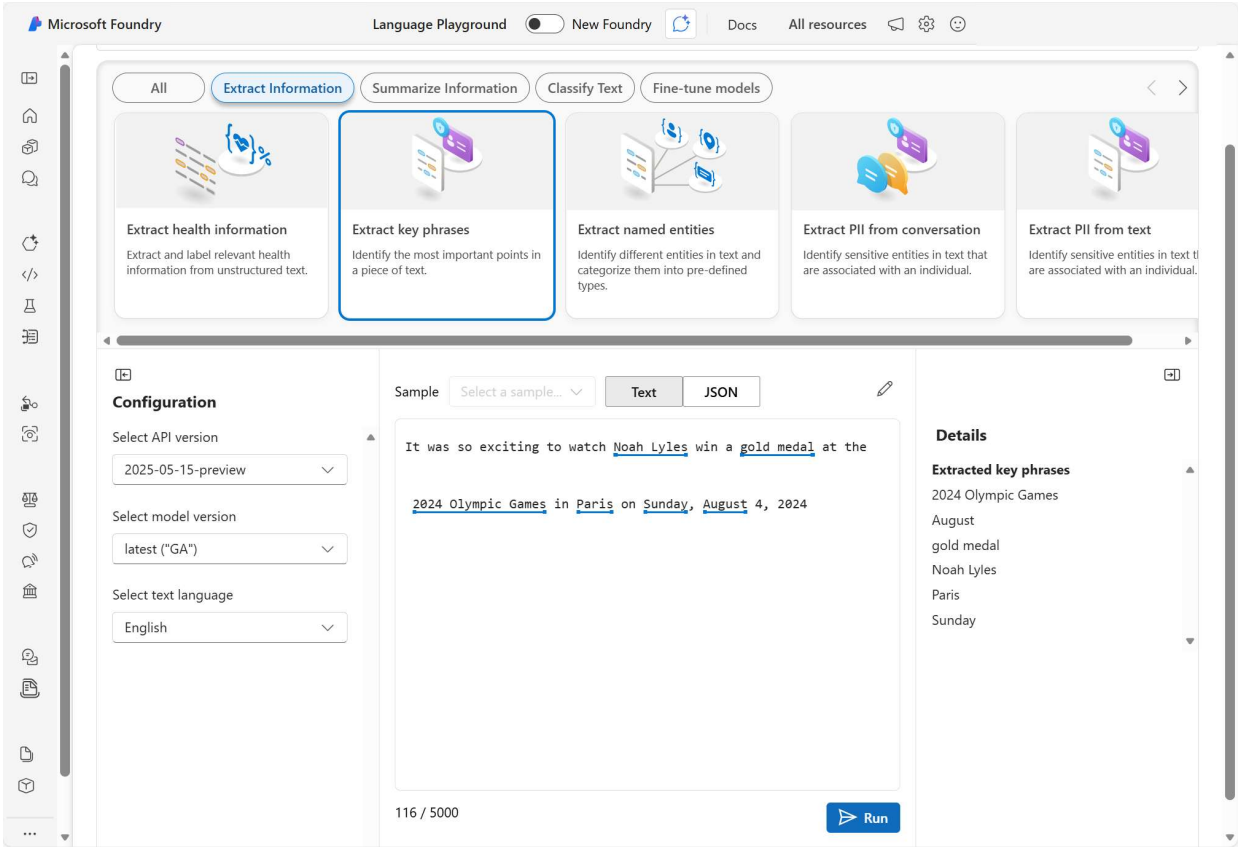
Language tools in Foundry

 **Azure Language in Foundry Tools**	 **Azure Translator in Foundry Tools**
• Language detection • Key phrase extraction • Named entity detection • Sentiment analysis and opinion mining • Personal information detection • Summarization • …	• Text translation • Document translation • Custom translation • …

Analyzing text

"It was so exciting to watch Noah Lyles win a gold medal at the 2024 Olympic Games in Paris on Sunday, August 4, 2024."

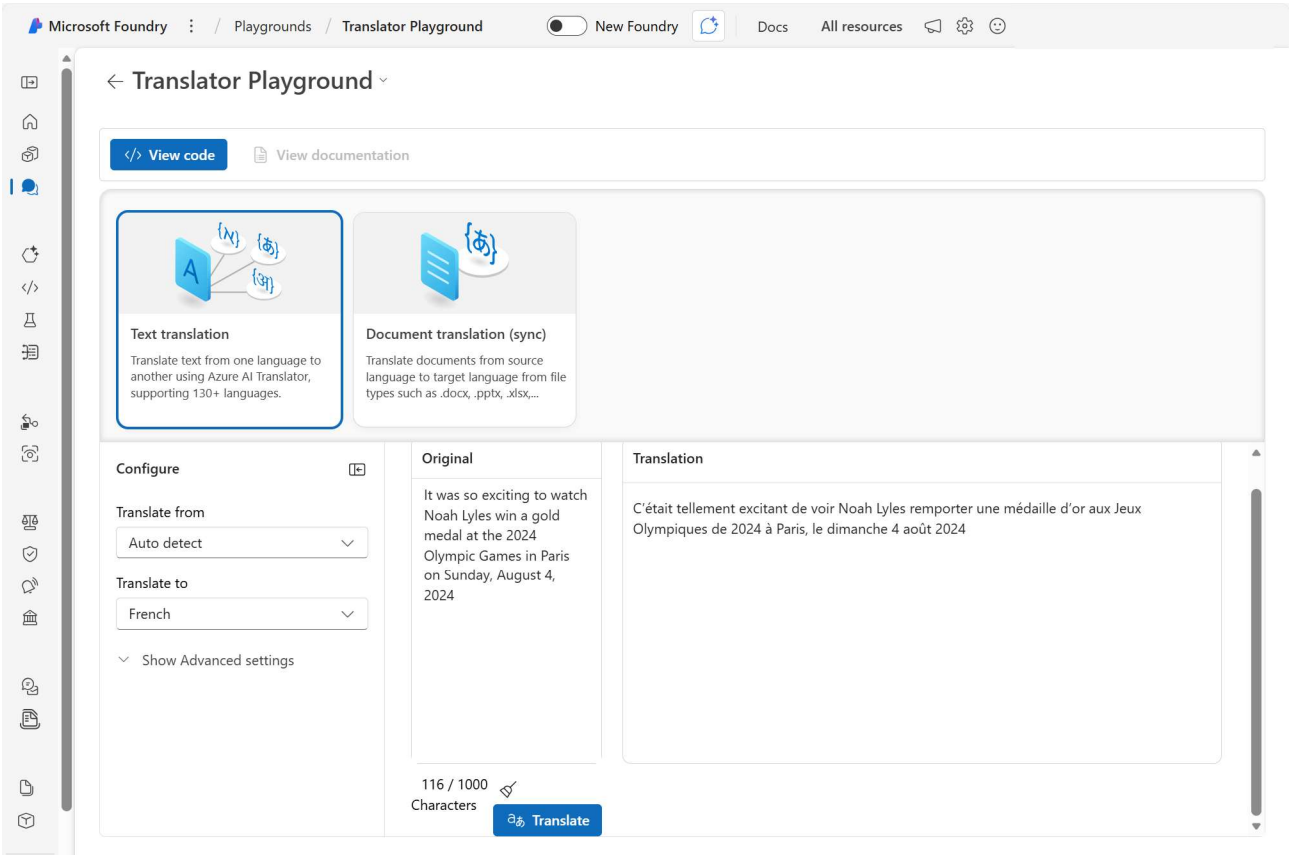
- **Language:** English
- **Sentiment:** Positive
- **Key Phrases:**
  - 2024 Olympic Games
  - gold medal
  - Noah Lyles
  - Paris
  - …
- **Entities:**
  - *Person*: Noah Lyles
  - *Event*: Olympic Games
  - *City*: Paris
  - *Date*: Sunday, August 4, 2024
  - …



Translating text

"It was so exciting to watch Noah Lyles win a gold medal at the 2024 Olympic Games in Paris on Sunday, August 4, 2024."

"C'était tellement excitant de voir Noah Lyles remporter une médaille d'or aux Jeux olympiques de 2024 à Paris le dimanche 4 août 2024."

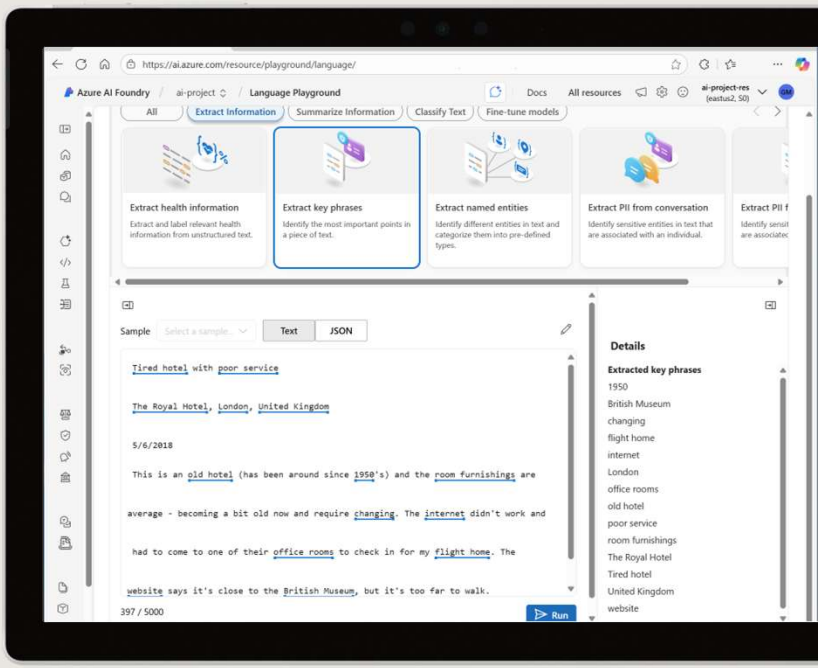


Exercise – If time permits

Analyze text in Microsoft Foundry

In this exercise, you'll use Microsoft Foundry to explore Azure Language capabilities.

Start the exercise at:
https://go.microsoft.com/fwlink/?linkid=2250314



Knowledge check

1 **What are some capabilities included in Azure Language?**

- ☐ Optical character recognition
- ☐ Speech recognition
- ☑ Text summarization, sentiment analysis, named entity detection, key phrase extraction

2 **You want to use Azure Language to determine the key talking points in a text document. Which feature of the service should you use?**

- ☐ Sentiment analysis
- ☑ Key phrase extraction
- ☐ Entity detection

3 **You want to convert text in English-language emails to Spanish. What service should you use?**

- ☐ Azure Language
- ☑ Azure Translator
- ☐ Azure Speech

Get started with speech in Microsoft Foundry

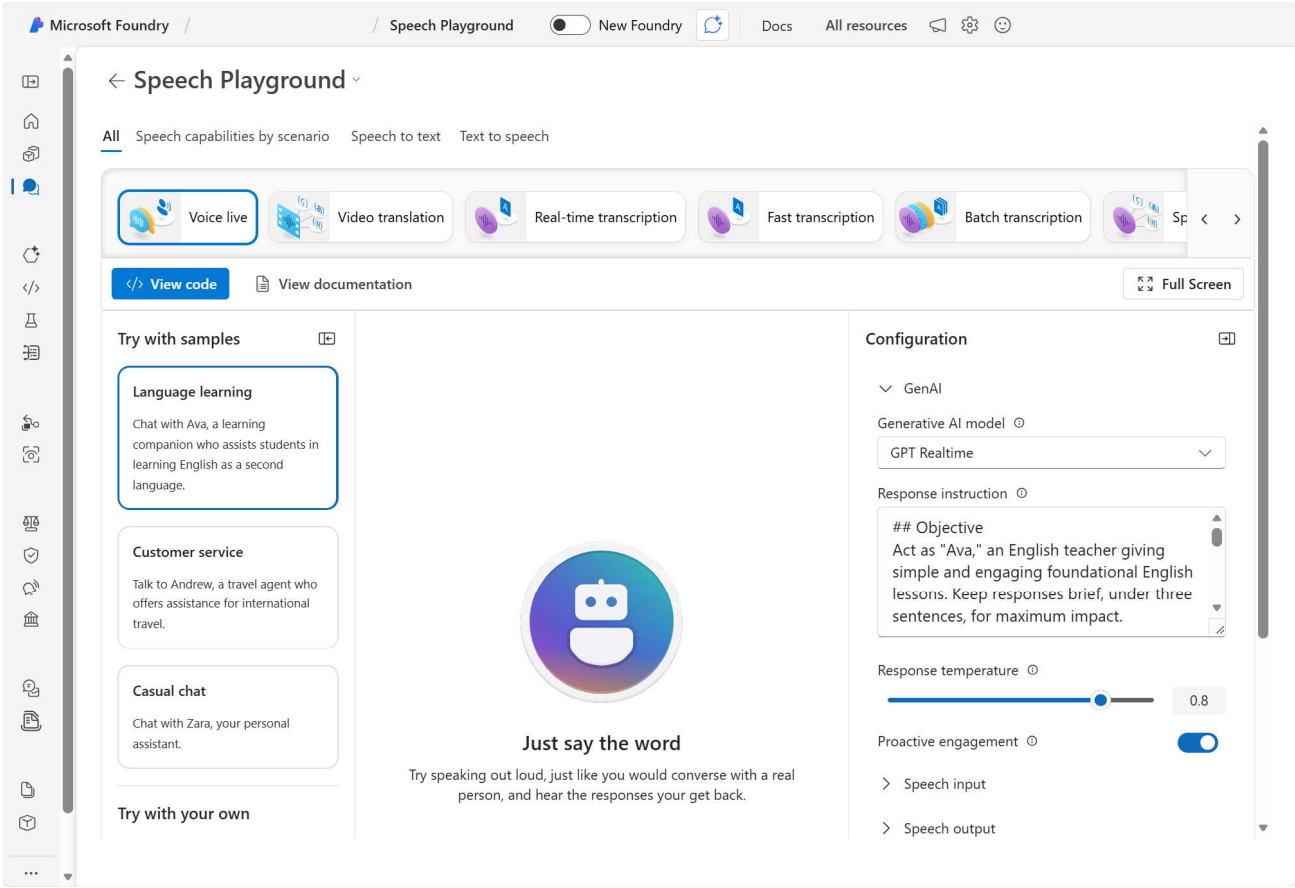
https://aka.ms/mslearn-azure-speech

Azure Speech in Foundry Tools

- Speech to text transcription (real-time and batch)
- Text to speech synthesis
- Speech translation
- Custom speech models
- Built-in and custom voices
- Audio content creation
- Video translation
- Voice Live: real-time voice agents
- Built-in and custom avatars
- …

Speech in Foundry

- Use the *Speech playground* to:
  - Transcribe recorded or live speech
  - Synthesize speech with built-in voices
  - Create custom voices
  - Create custom speech models
  - Add live voice capabilities to an agent
  - Translate audio and video
  - Create avatar-presented videos
  - Create custom avatars
  - …

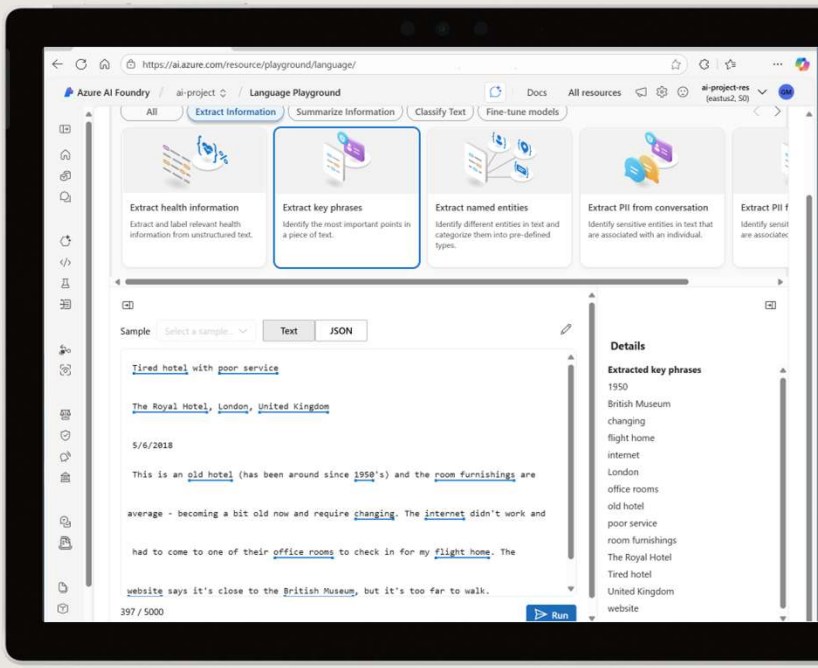


Exercise – If time permits

Explore Speech in Microsoft Foundry

In this exercise, you'll use Microsoft Foundry's speech playground to see speech capabilities in action.

Start the exercise at:
https://go.microsoft.com/fwlink/?linkid=2250148



Knowledge check

1 **You want to build an application that reads incoming email message subjects aloud. Which Azure Speech capability should you use?**

- ☐ Speech to text
- ☑ Text to speech
- ☐ Speech translation

2 **You want to test various built-in voices to determine which one sounds right for your application. What should you do?**

- ☐ Record yourself speaking for 10 seconds and use Azure Speech to create a custom voice.
- ☐ Use the default built-in voice – there is only one.
- ☑ Select voices in the Speech playground in Foundry and use them to generate speech.

Summary

Introduction to natural language processing concepts

- Natural Language Processing (NLP) is used to build solutions that work with text-based or speech-based language
- NLP capabilities have evolved from simple statistical techniques to modern sematic language models

Get started with natural language processing in Microsoft Foundry

- Use Azure Language to analyze text and extract information
- Use Azure Translator to translate text

Get started with speech in Microsoft Foundry

- Use Azure Speech to synthesize, transcribe, and translate speech; add live voice capabilities to agents, and to generate avatar-presented video