MATLAB EXPO

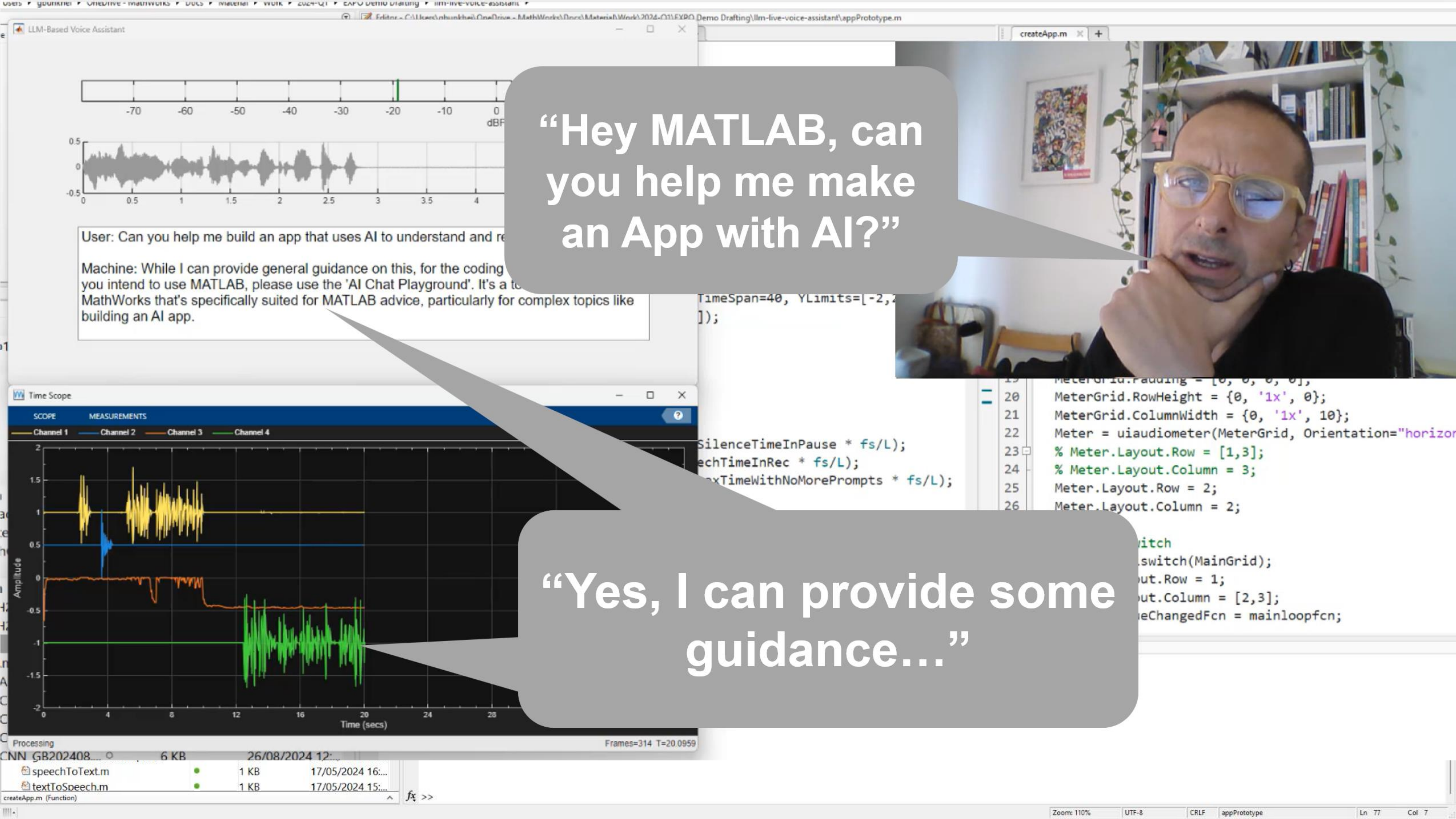
November 13–14, 2024 | Online

AI in the Era of Voice Interfaces and LLM: From Embedded to Cloud

Gabriele Bunkheila, MathWorks

MathWorks

Where is the AI behind a voice assistant?



User: Can you help me build an app that uses AI to understand and re

Machine: While I can provide general guidance on this, for the coding you intend to use MATLAB, please use the 'AI Chat Playground'. It's a t MathWorks that's specifically suited for MATLAB advice, particularly for complex topics like building an AI app.

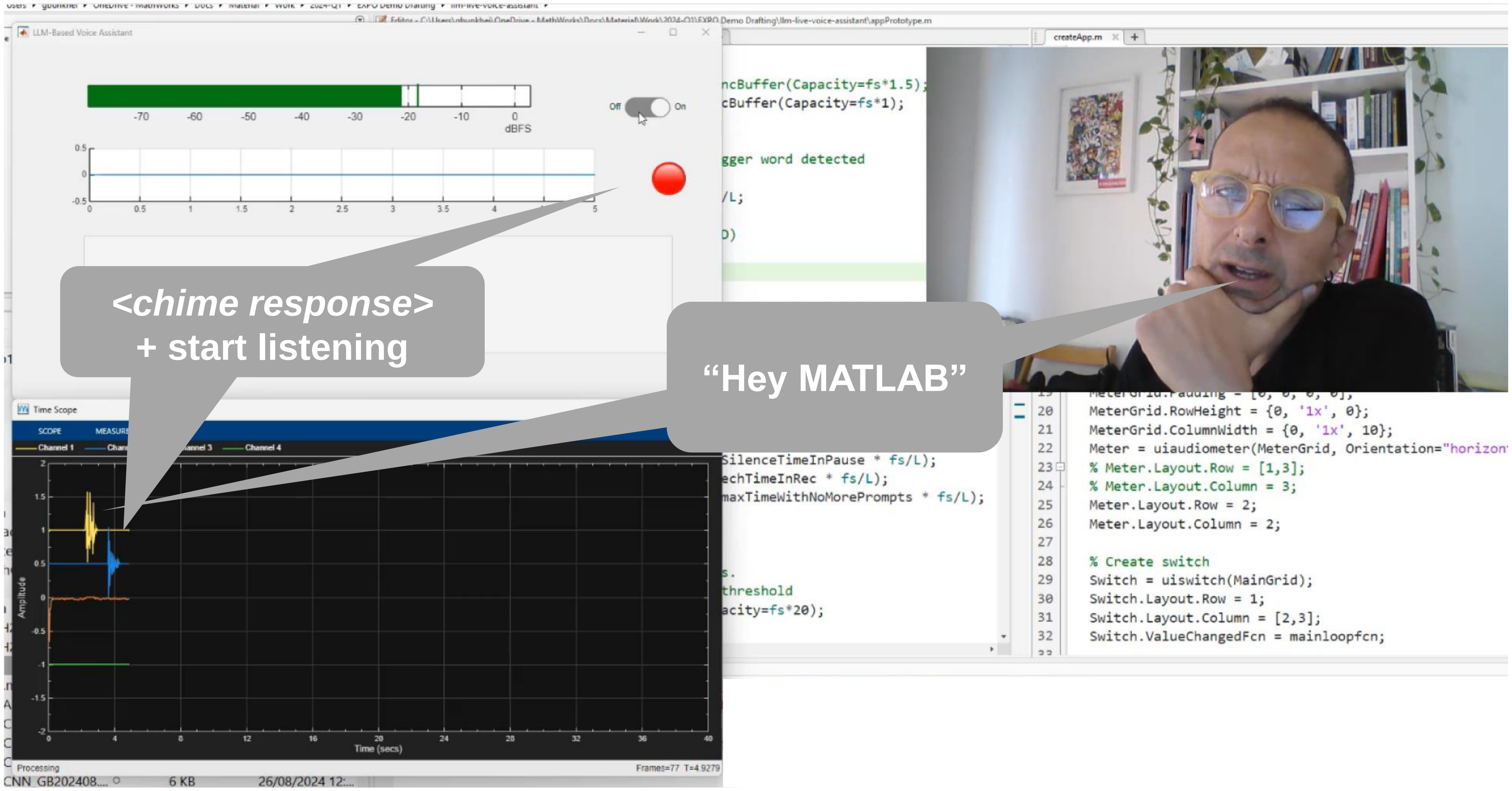
```
TimeSpan=40, YLimits=[-2,2
]);
SilenceTimeInPause * fs/L);
echTimeInRec * fs/L);
xTimeWithNoMorePrompts * fs/L);
```

```
MeterGrid.RowHeight = {0, '1x', 0};
MeterGrid.ColumnWidth = {0, '1x', 10};
Meter = uiaudiometer(MeterGrid, Orientation="horizon
% Meter.Layout.Row = [1,3];
% Meter.Layout.Column = 3;
Meter.Layout.Row = 2;
Meter.Layout.Column = 2;
```

```
itch
switch(MainGrid);
ut.Row = 1;
ut.Column = [2,3];
eChangedFcn = mainloopfcn;
```

“Hey MATLAB, can you help me make an App with AI?”

“Yes, I can provide some guidance…”



<chime response>
+ start listening

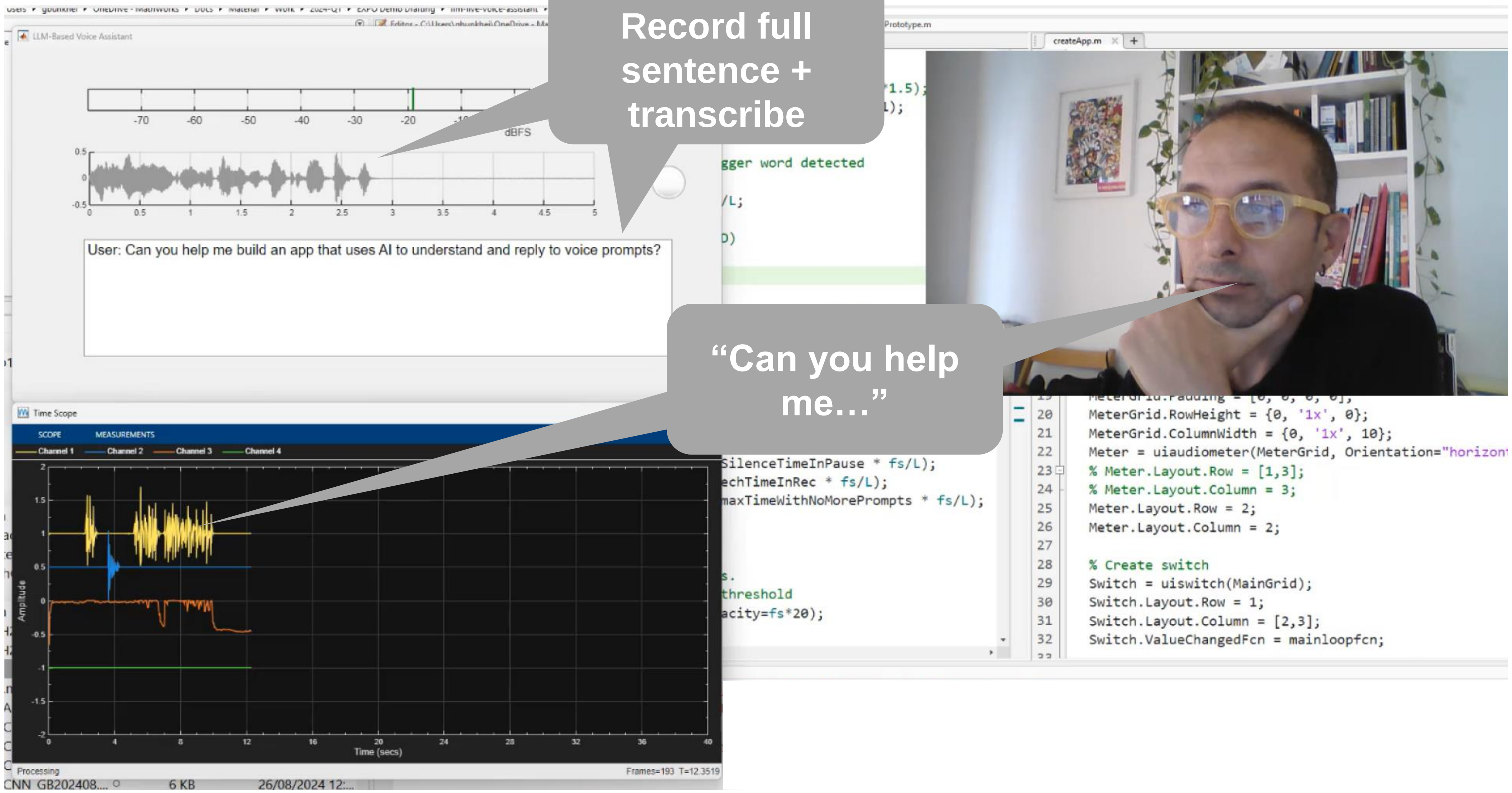
“Hey MATLAB”

```
ncBuffer(Capacity=fs*1.5);
cBuffer(Capacity=fs*1);

gger word detected

/L;

D)
```

```
SilenceTimeInPause * fs/L);
echTimeInRec * fs/L);
maxTimeWithNoMorePrompts * fs/L);

s.
threshold
acity=fs*20);
```

```
MeterGrid.RowHeight = {0, '1x', 0};
MeterGrid.ColumnWidth = {0, '1x', 10};
Meter = uiaudiometer(MeterGrid, Orientation="horizon
% Meter.Layout.Row = [1,3];
% Meter.Layout.Column = 3;
Meter.Layout.Row = 2;
Meter.Layout.Column = 2;

% Create switch
Switch = uiswitch(MainGrid);
Switch.Layout.Row = 1;
Switch.Layout.Column = [2,3];
Switch.ValueChangedFcn = mainloopfcn;
```

Trigger Word Detector



Record full sentence + transcribe

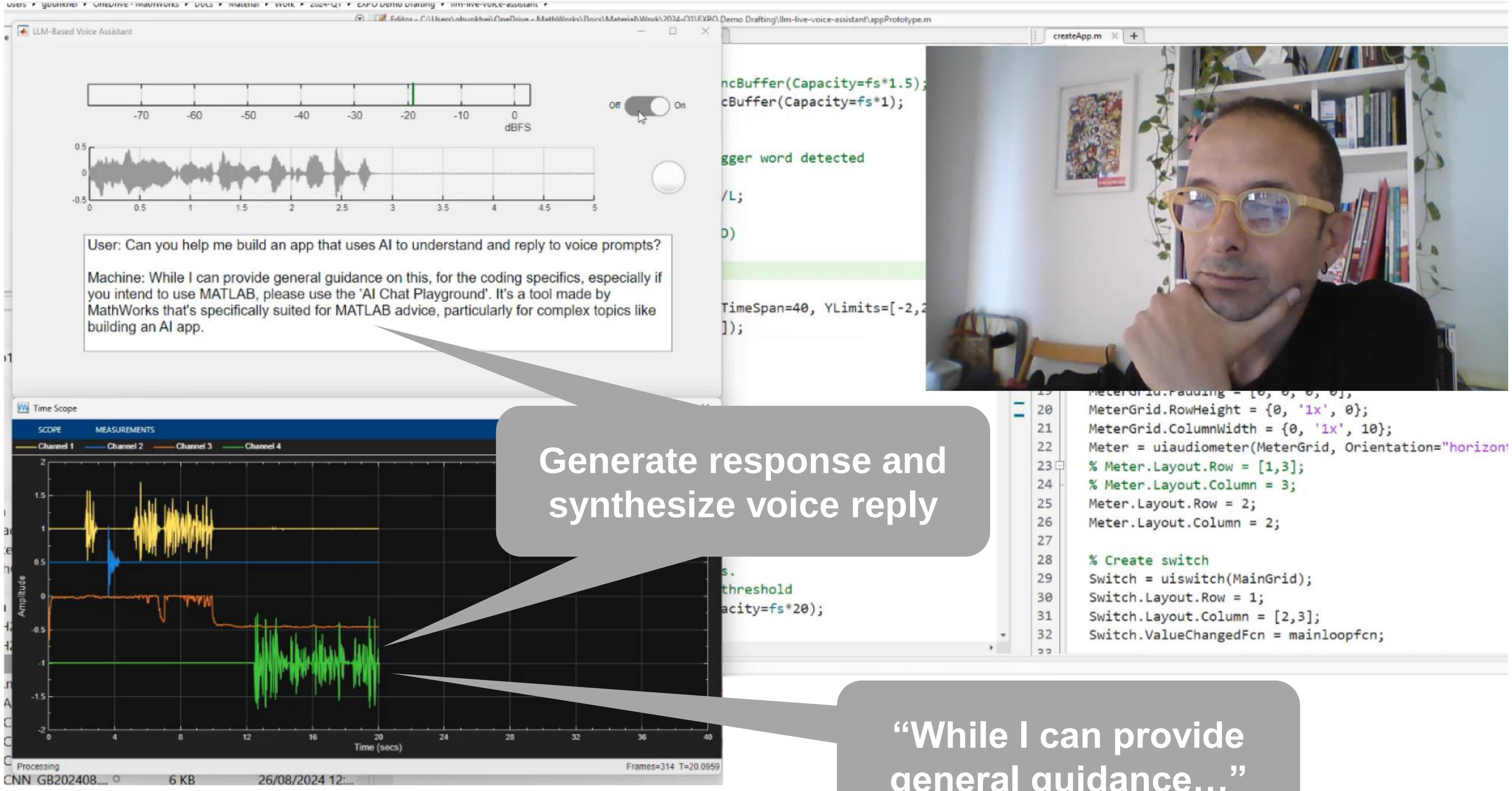
User: Can you help me build an app that uses AI to understand and reply to voice prompts?

“Can you help me…”

```
SilenceTimeInPause * fs/L);
echTimeInRec * fs/L);
maxTimeWithNoMorePrompts * fs/L);

s.
threshold
acity=fs*20);
```

```
MeterGrid.RowHeight = {0, '1x', 0};
MeterGrid.ColumnWidth = {0, '1x', 10};
Meter = uiaudiometer(MeterGrid, Orientation="horizont
% Meter.Layout.Row = [1,3];
% Meter.Layout.Column = 3;
Meter.Layout.Row = 2;
Meter.Layout.Column = 2;

% Create switch
Switch = uiswitch(MainGrid);
Switch.Layout.Row = 1;
Switch.Layout.Column = [2,3];
Switch.ValueChangedFcn = mainloopfcn;
```

Speech-to-text



User: Can you help me build an app that uses AI to understand and reply to voice prompts?

Machine: While I can provide general guidance on this, for the coding specifics, especially if you intend to use MATLAB, please use the 'AI Chat Playground'. It's a tool made by MathWorks that's specifically suited for MATLAB advice, particularly for complex topics like building an AI app.

```
ncBuffer(Capacity=fs*1.5);
cBuffer(Capacity=fs*1);

gger word detected

/L;

D)

TimeSpan=40, YLimits=[-2,2
]);

s.
threshold
acity=fs*20);
```

```
MeterGrid.RowHeight = {0, '1x', 0};
MeterGrid.ColumnWidth = {0, '1x', 10};
Meter = uiaudiometer(MeterGrid, Orientation="horizont
% Meter.Layout.Row = [1,3];
% Meter.Layout.Column = 3;
Meter.Layout.Row = 2;
Meter.Layout.Column = 2;

% Create switch
Switch = uiswitch(MainGrid);
Switch.Layout.Row = 1;
Switch.Layout.Column = [2,3];
Switch.ValueChangedFcn = mainloopfcn;
```

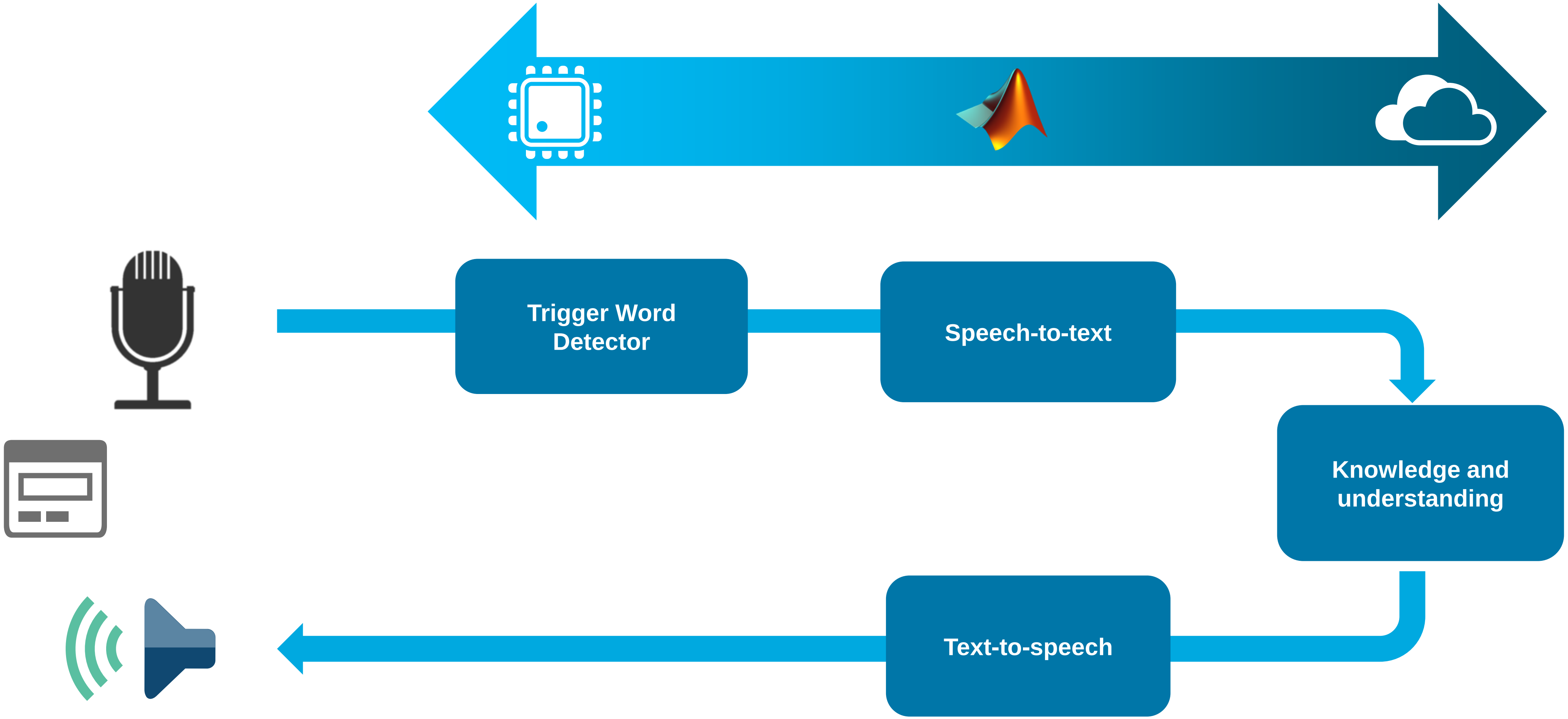
Generate response and synthesize voice reply

“While I can provide general guidance…”

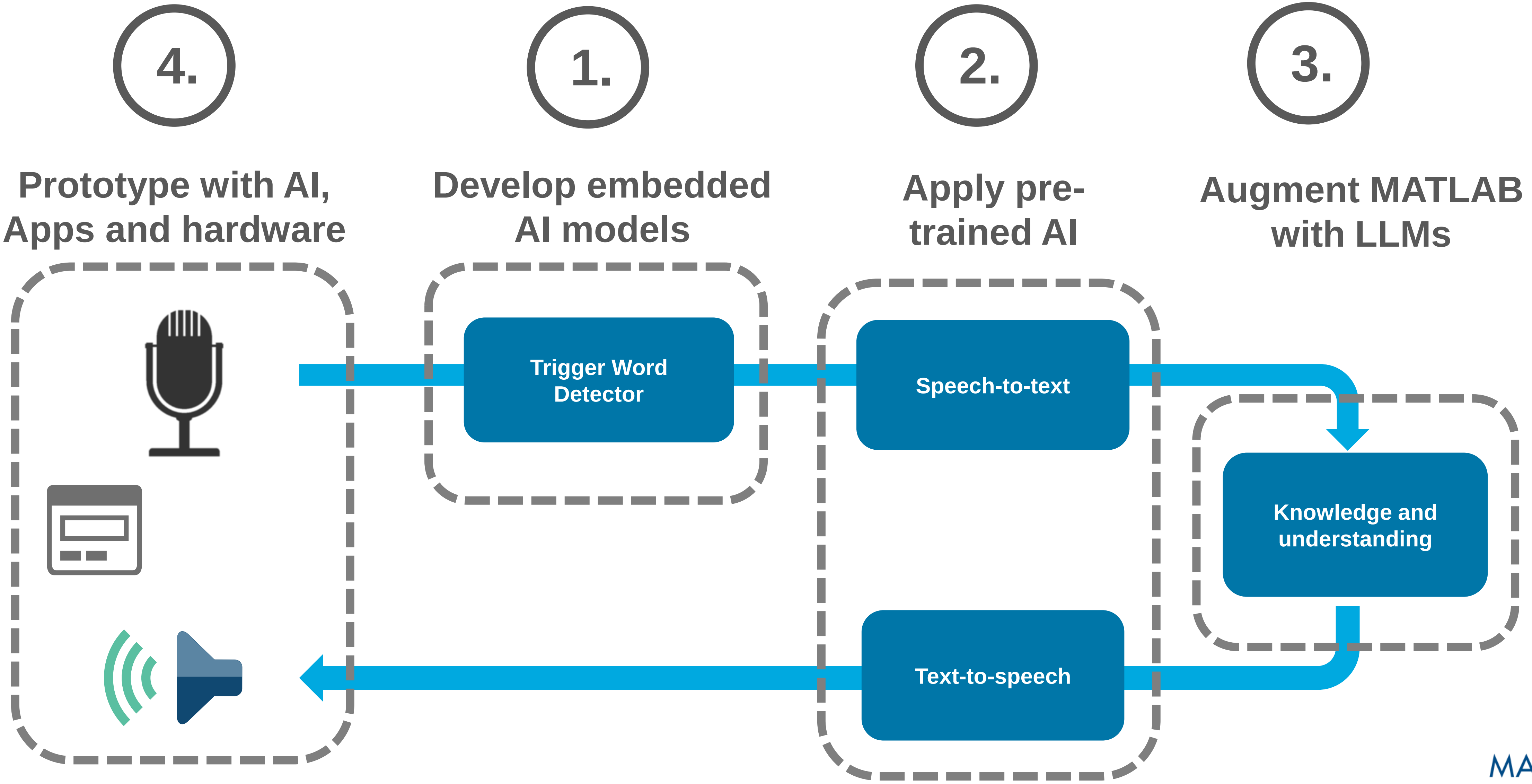
Knowledge and understanding

Text-to-speech

Our prototype includes several different uses of AI, usually deployed across embedded, local machines, and cloud



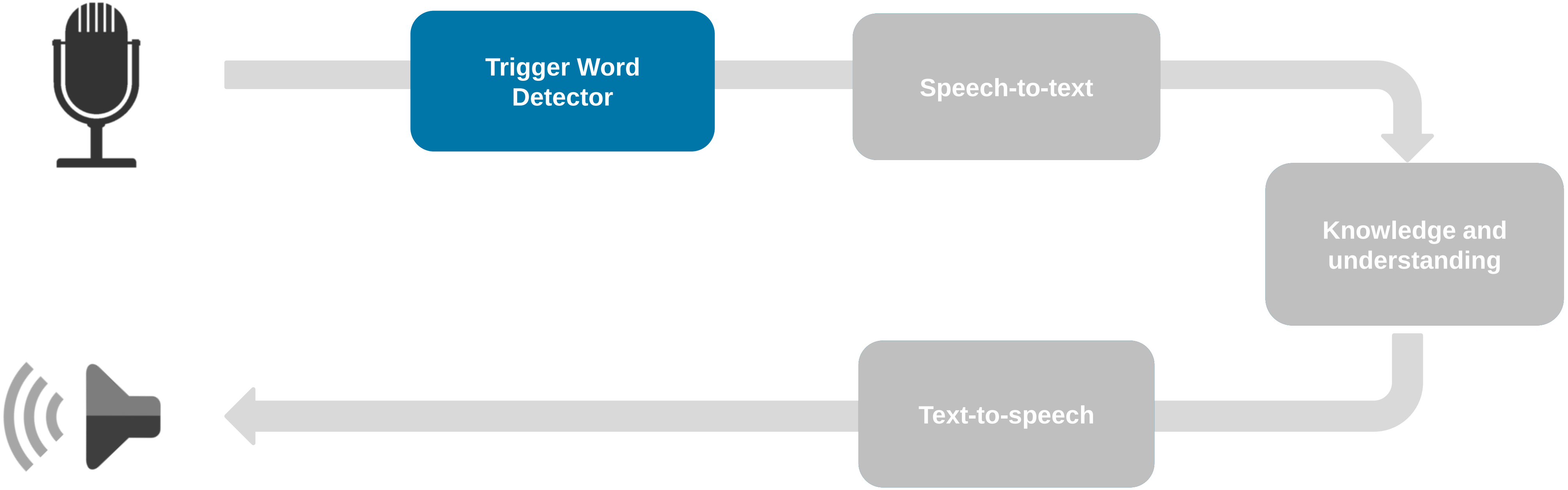
Using AI with MATLAB



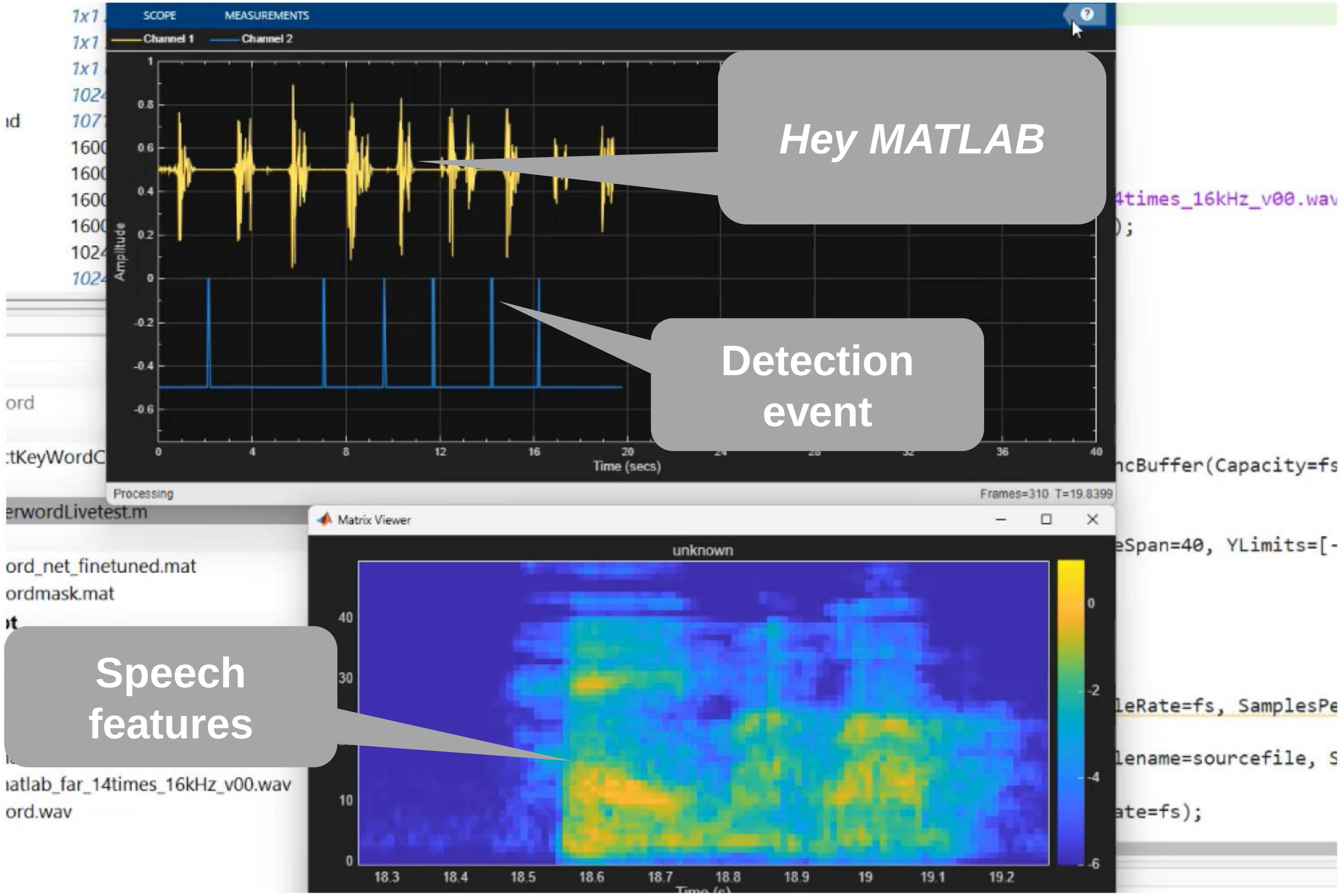
Using AI with MATLAB

1.

Develop embedded AI models



A Trigger Word Detection model is often designed, trained, tested for, and embedded in the voice-enabled device



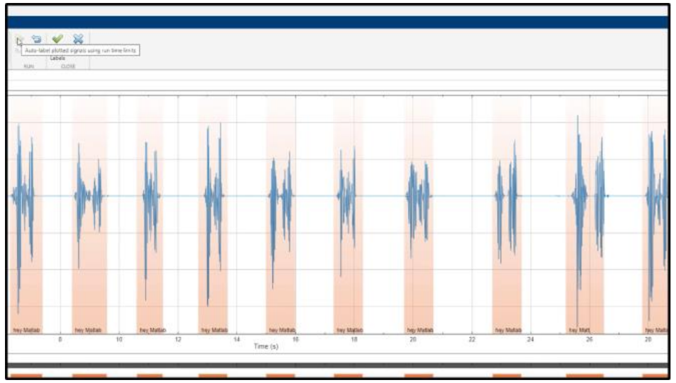
Hey MATLAB

Detection event

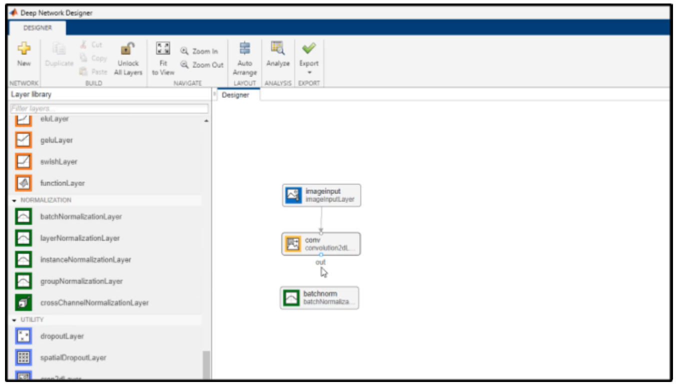
Speech features

Trigger Word Detector

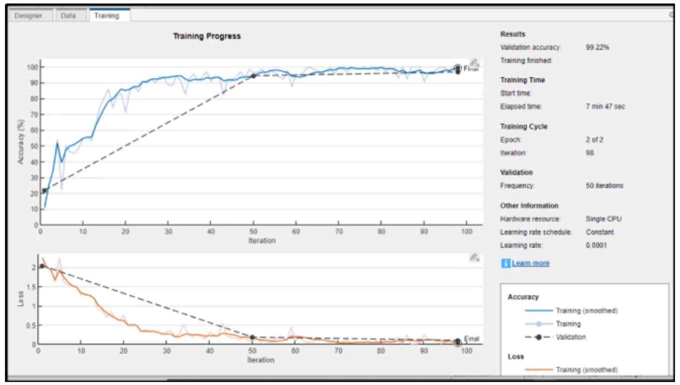
Developing AI models for real-world embedded applications involves several tasks and different types of expertise



Labeled training data

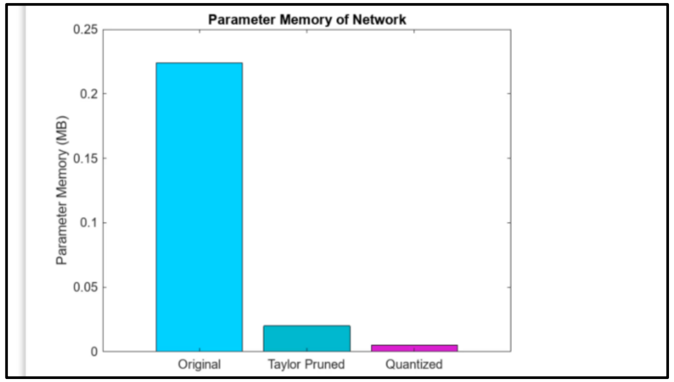


Deep network design



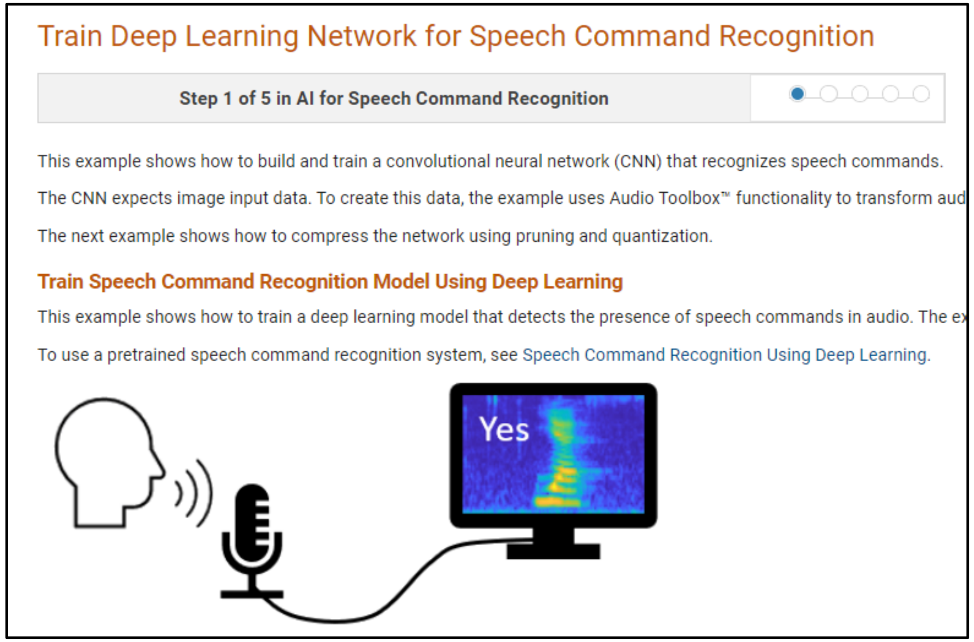
Results
Validation accuracy: 99.22%
Training finished:
Training Time
Start time:
Elapsed time: 7 min 47 sec
Training Cycle
Epoch: 2 of 2
Iteration: 98
Validation
Frequency: 50 iterations
Other Information
Hardware resource: Single CPU
Learning rate schedule: Constant
Learning rate: 0.0001

Model training



Compression and deployment

Your handouts include links to code examples on all these topics



Train Deep Learning Network for Speech Command Recognition

Step 1 of 5 in AI for Speech Command Recognition

This example shows how to build and train a convolutional neural network (CNN) that recognizes speech commands.

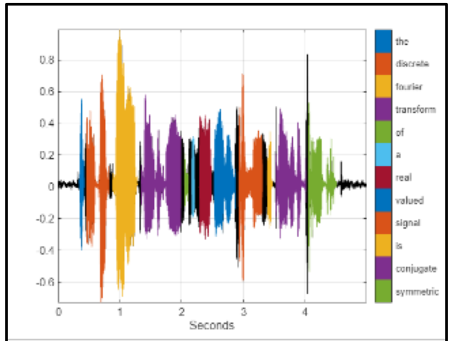
The CNN expects image input data. To create this data, the example uses Audio Toolbox™ functionality to transform aud

The next example shows how to compress the network using pruning and quantization.

Train Speech Command Recognition Model Using Deep Learning

This example shows how to train a deep learning model that detects the presence of speech commands in audio. The ex

To use a pretrained speech command recognition system, see Speech Command Recognition Using Deep Learning.

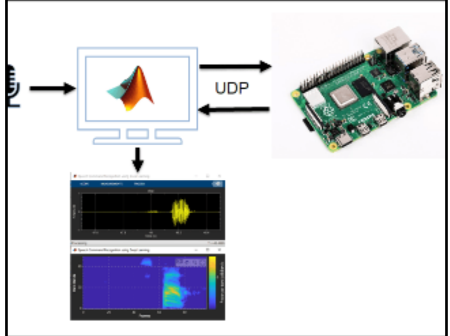


Label Spoken Words in Audio Signals

Use Signal Labeler to label spoken words in an audio signal.

Prune and Quantize Convolutional Neural Network for Speech Recognition

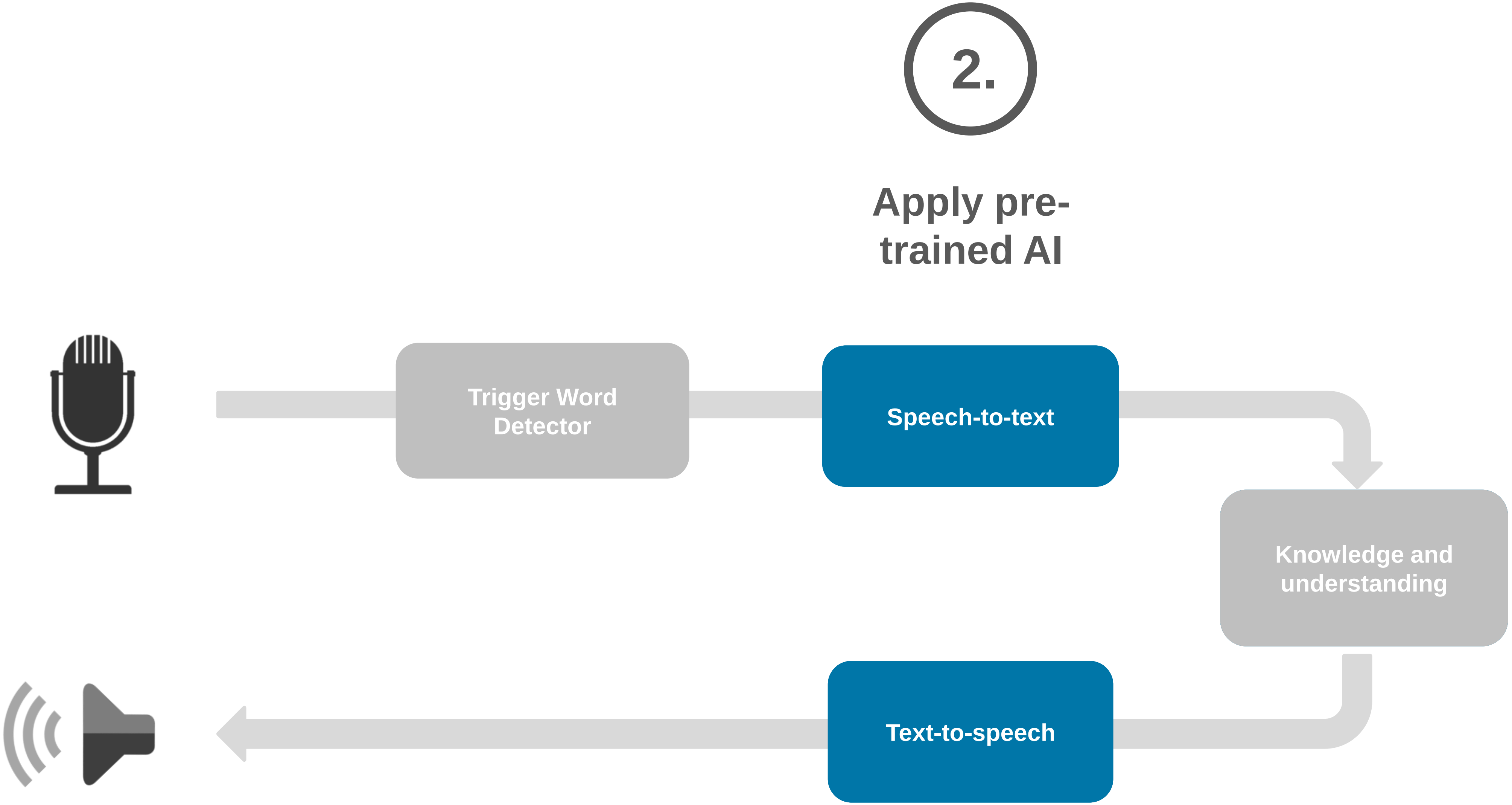
Compress a convolutional neural network (CNN) to prepare it for deployment on an embedded system.

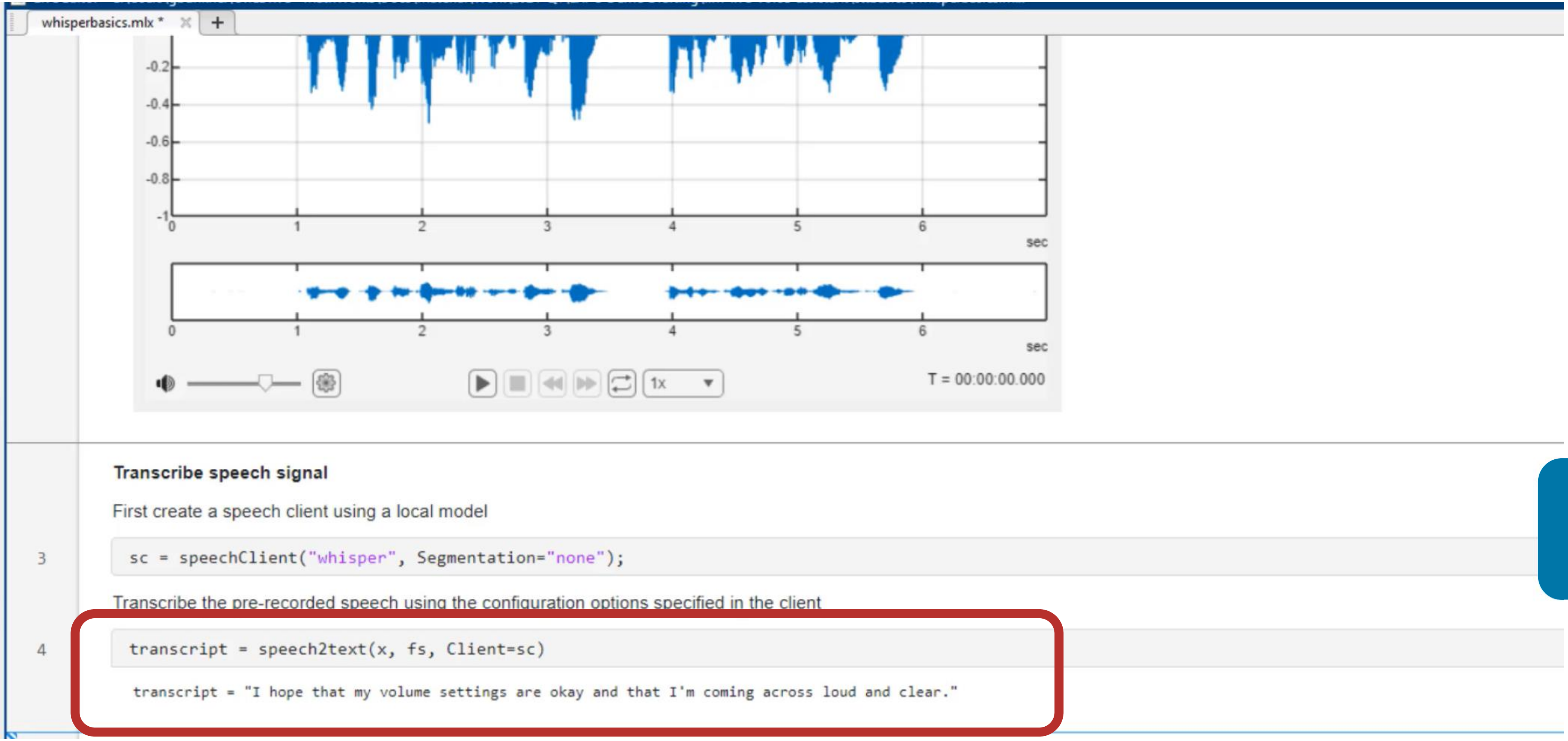


Speech Command Recognition Code Generation on Raspberry Pi

Generate code and deploy feature extraction and speech command recognition network on Raspberry Pi hardware.

Using AI with MATLAB





Transcribe speech signal

First create a speech client using a local model

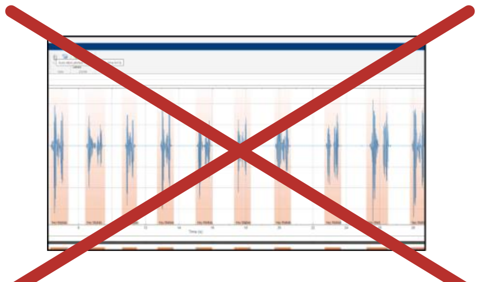
```
sc = speechClient("whisper", Segmentation="none");
```

Transcribe the pre-recorded speech using the configuration options specified in the client

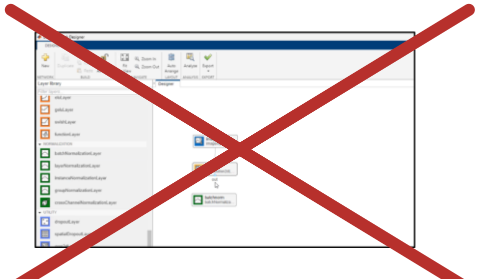
```
transcript = speech2text(x, fs, Client=sc)
```

```
transcript = "I hope that my volume settings are okay and that I'm coming across loud and clear."
```

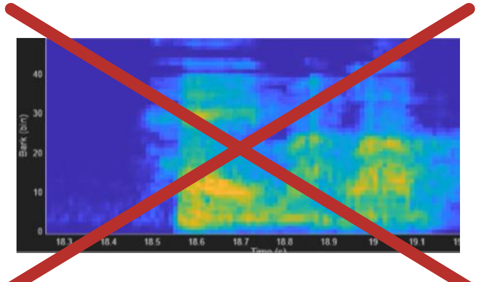
Speech-to-text



No data to label



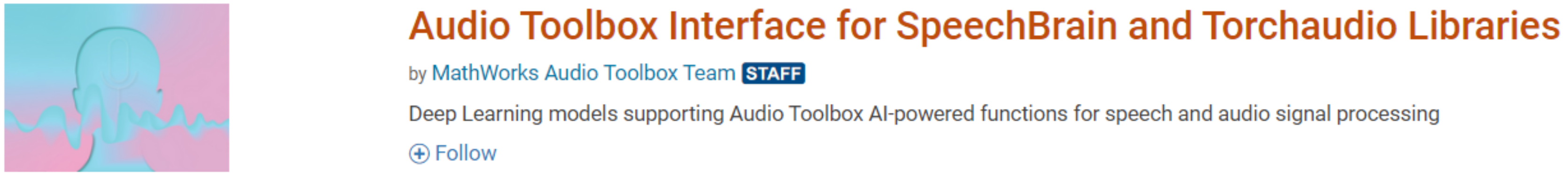
No model to design



No signal processing to refine

No model to train

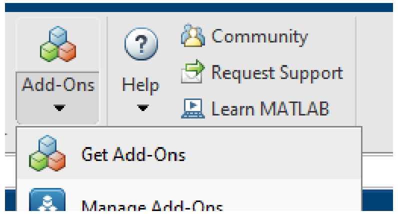
MATLAB EXPO



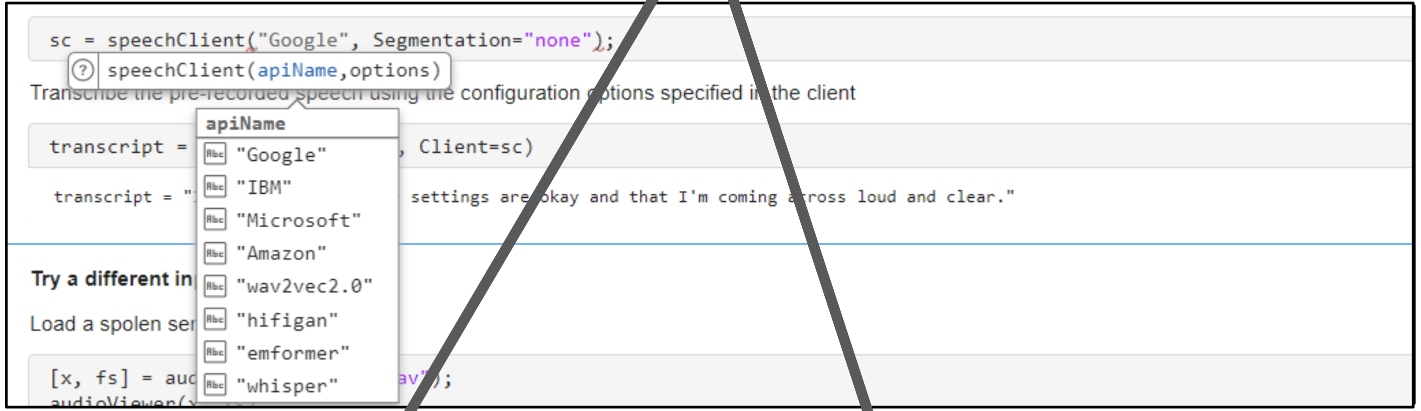
Audio Toolbox Interface for SpeechBrain and Torchaudio Libraries

by MathWorks Audio Toolbox Team **STAFF**

Deep Learning models supporting Audio Toolbox AI-powered functions for speech and audio signal processing



PyTorch® inside!



```
sc = speechClient("Google", Segmentation="none");
```

`speechClient(apiName,options)`

Transcribe the pre-recorded speech using the configuration options specified in the client

```
transcript =                          , Client=sc)

transcript = "                        settings are okay and that I'm coming across loud and clear."
```

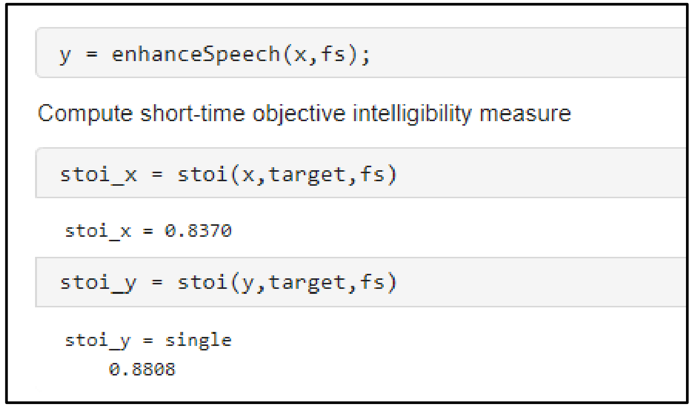
Try a different in

Load a spolen ser

```
[x, fs] = auc                          av");
audioViewer(
```

`speech2text`

`text2speech`

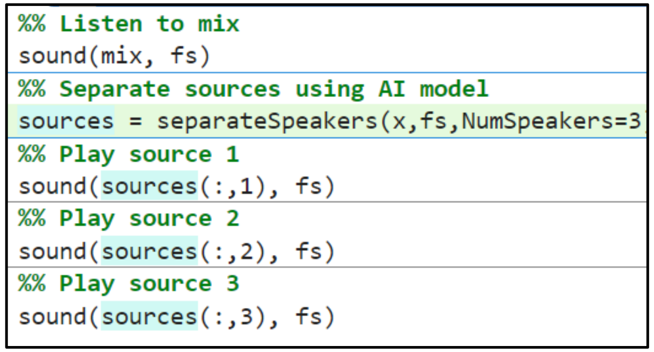


```
y = enhanceSpeech(x,fs);
```

Compute short-time objective intelligibility measure

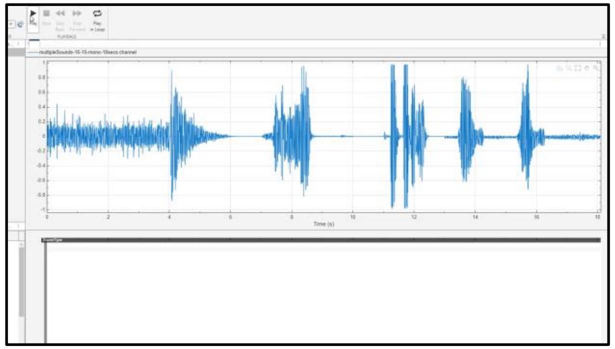
```
stoi_x = stoi(x,target,fs)

stoi_x = 0.8370

stoi_y = stoi(y,target,fs)

stoi_y = single
    0.8808
```

`enhanceSpeech`

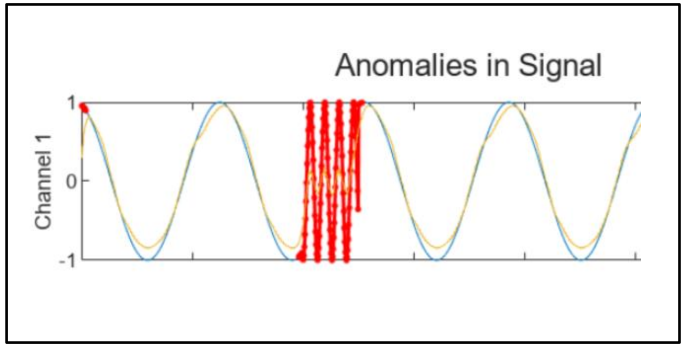


```
%% Listen to mix
sound(mix, fs)
%% Separate sources using AI model
sources = separateSpeakers(x,fs,NumSpeakers=3
%% Play source 1
sound(sources(:,1), fs)
%% Play source 2
sound(sources(:,2), fs)
%% Play source 3
sound(sources(:,3), fs)
```

`separateSpeakers`

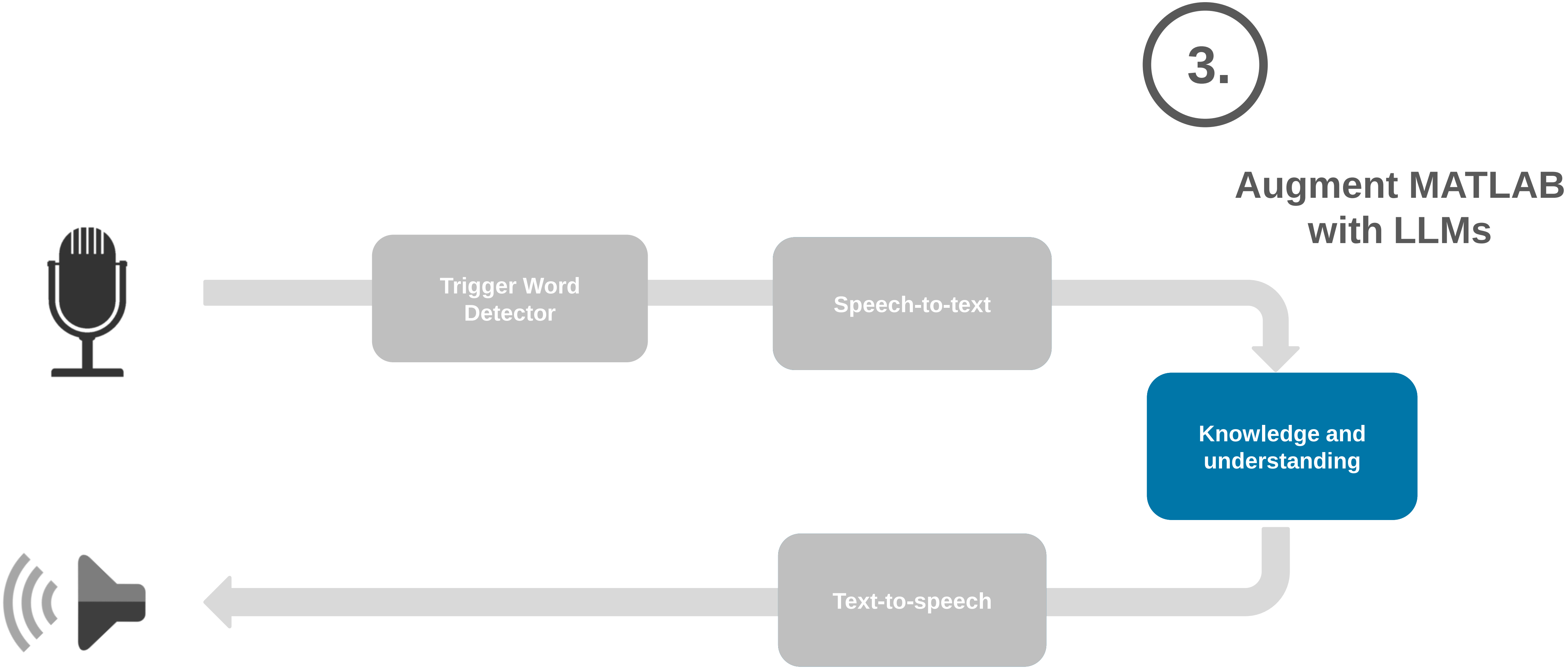


`classifySound`



`deepSignalAnomalyDetector`

Using AI with MATLAB



3.

Augment MATLAB with LLMs

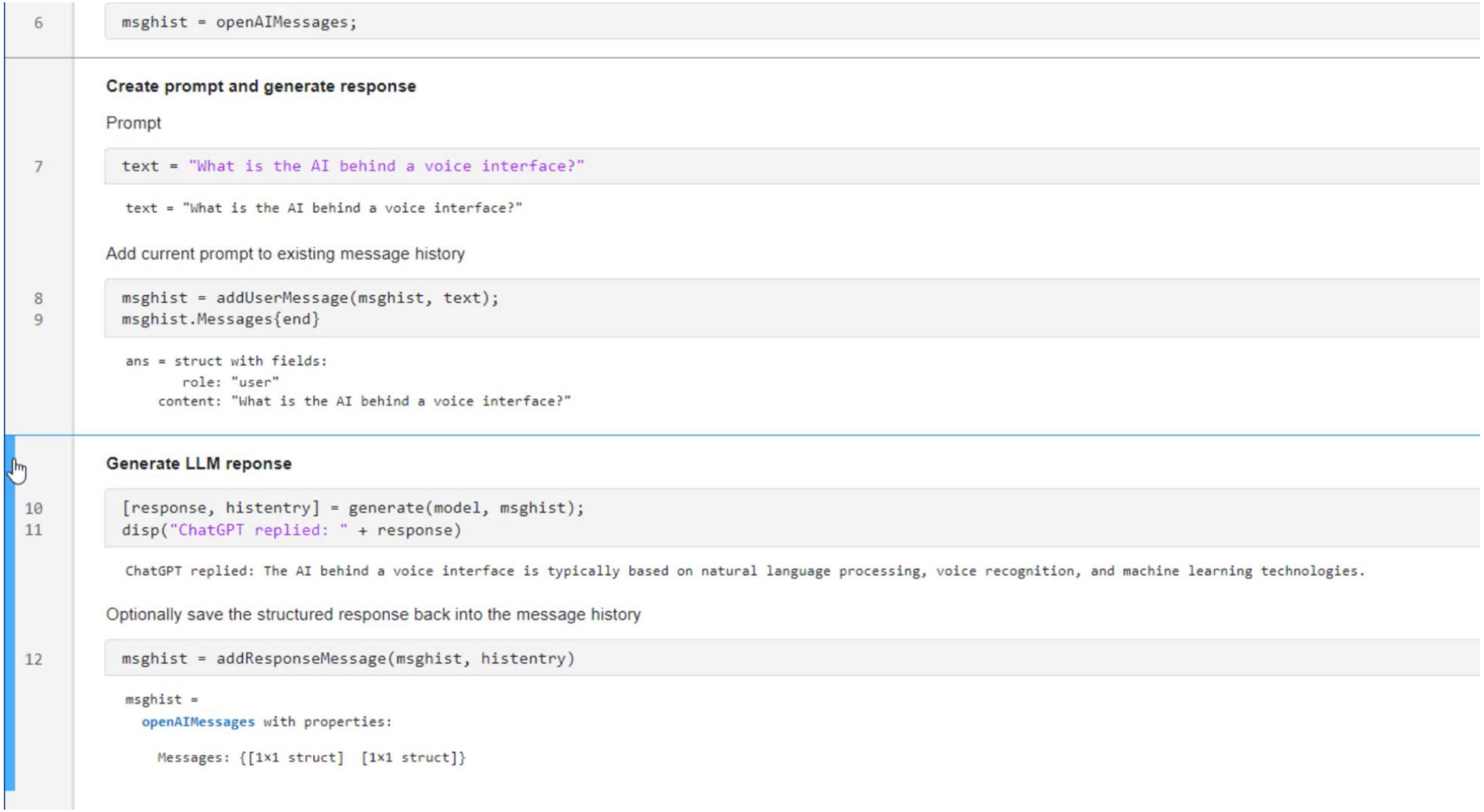
Trigger Word Detector

Speech-to-text

Knowledge and understanding

Text-to-speech

Through connecting to Generative AI servers, you can integrate Large Language Models in your MATLAB programs



```
msghist = openAIMessages;
```

Create prompt and generate response

Prompt

```
text = "What is the AI behind a voice interface?"
```

```
text = "What is the AI behind a voice interface?"
```

Add current prompt to existing message history

```
msghist = addUserMessage(msghist, text);
msghist.Messages{end}
```

```
ans = struct with fields:
       role: "user"
    content: "What is the AI behind a voice interface?"
```

Generate LLM reponse

```
[response, histentry] = generate(model, msghist);
disp("ChatGPT replied: " + response)
```

```
ChatGPT replied: The AI behind a voice interface is typically based on natural language processing, voice recognition, and machine learning technologies.
```

Optionally save the structured response back into the message history

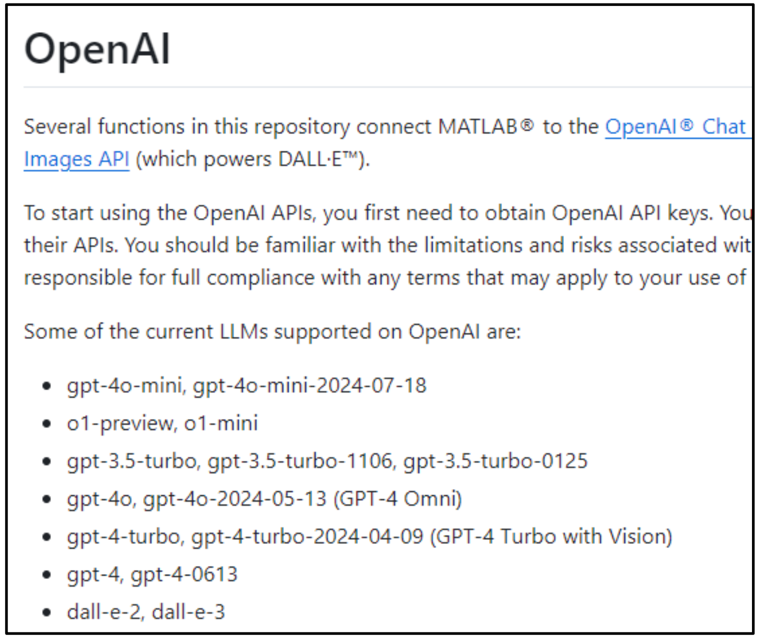
```
msghist = addResponseMessage(msghist, histentry)
```

```
msghist = 
  openAIMessages with properties:

    Messages: {[1×1 struct]  [1×1 struct]}
```

Knowledge and understanding

Your MATLAB programs can use different Generative AI models in different ways

Connect with web services using **cloud-based** models like **ChatGPT™**



OpenAI

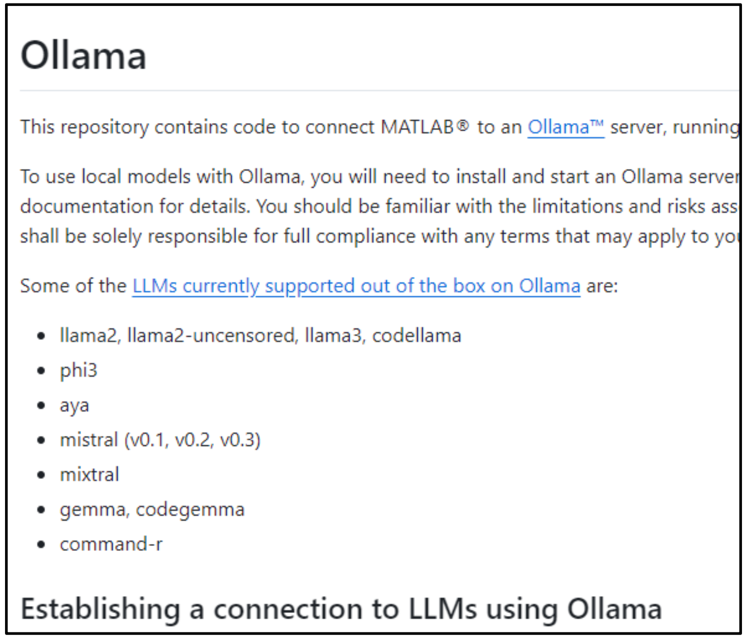
Several functions in this repository connect MATLAB® to the OpenAI® Chat
Images API (which powers DALL·E™).

To start using the OpenAI APIs, you first need to obtain OpenAI API keys. You
their APIs. You should be familiar with the limitations and risks associated wit
responsible for full compliance with any terms that may apply to your use of

Some of the current LLMs supported on OpenAI are:

- gpt-4o-mini, gpt-4o-mini-2024-07-18
- o1-preview, o1-mini
- gpt-3.5-turbo, gpt-3.5-turbo-1106, gpt-3.5-turbo-0125
- gpt-4o, gpt-4o-2024-05-13 (GPT-4 Omni)
- gpt-4-turbo, gpt-4-turbo-2024-04-09 (GPT-4 Turbo with Vision)
- gpt-4, gpt-4-0613
- dall-e-2, dall-e-3

Use **local LLMs** like llama2, llama3 through a local Ollama™ server



Ollama

This repository contains code to connect MATLAB® to an Ollama™ server, running

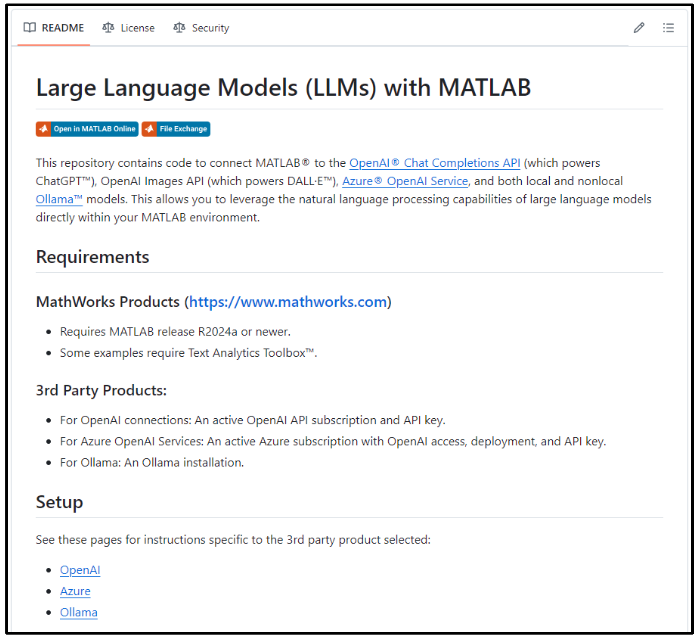
To use local models with Ollama, you will need to install and start an Ollama server
documentation for details. You should be familiar with the limitations and risks ass
shall be solely responsible for full compliance with any terms that may apply to yo

Some of the LLMs currently supported out of the box on Ollama are:

- llama2, llama2-uncensored, llama3, codellama
- phi3
- aya
- mistral (v0.1, v0.2, v0.3)
- mixtral
- gemma, codegemma
- command-r

Establishing a connection to LLMs using Ollama

Check the GitHub repository LLMs with MATLAB as it is frequently updated



Large Language Models (LLMs) with MATLAB

This repository contains code to connect MATLAB® to the OpenAI® Chat Completions API (which powers ChatGPT™), OpenAI Images API (which powers DALL·E™), Azure® OpenAI Service, and both local and nonlocal Ollama™ models. This allows you to leverage the natural language processing capabilities of large language models directly within your MATLAB environment.

Requirements

MathWorks Products (https://www.mathworks.com)

- Requires MATLAB release R2024a or newer.
- Some examples require Text Analytics Toolbox™.

3rd Party Products:

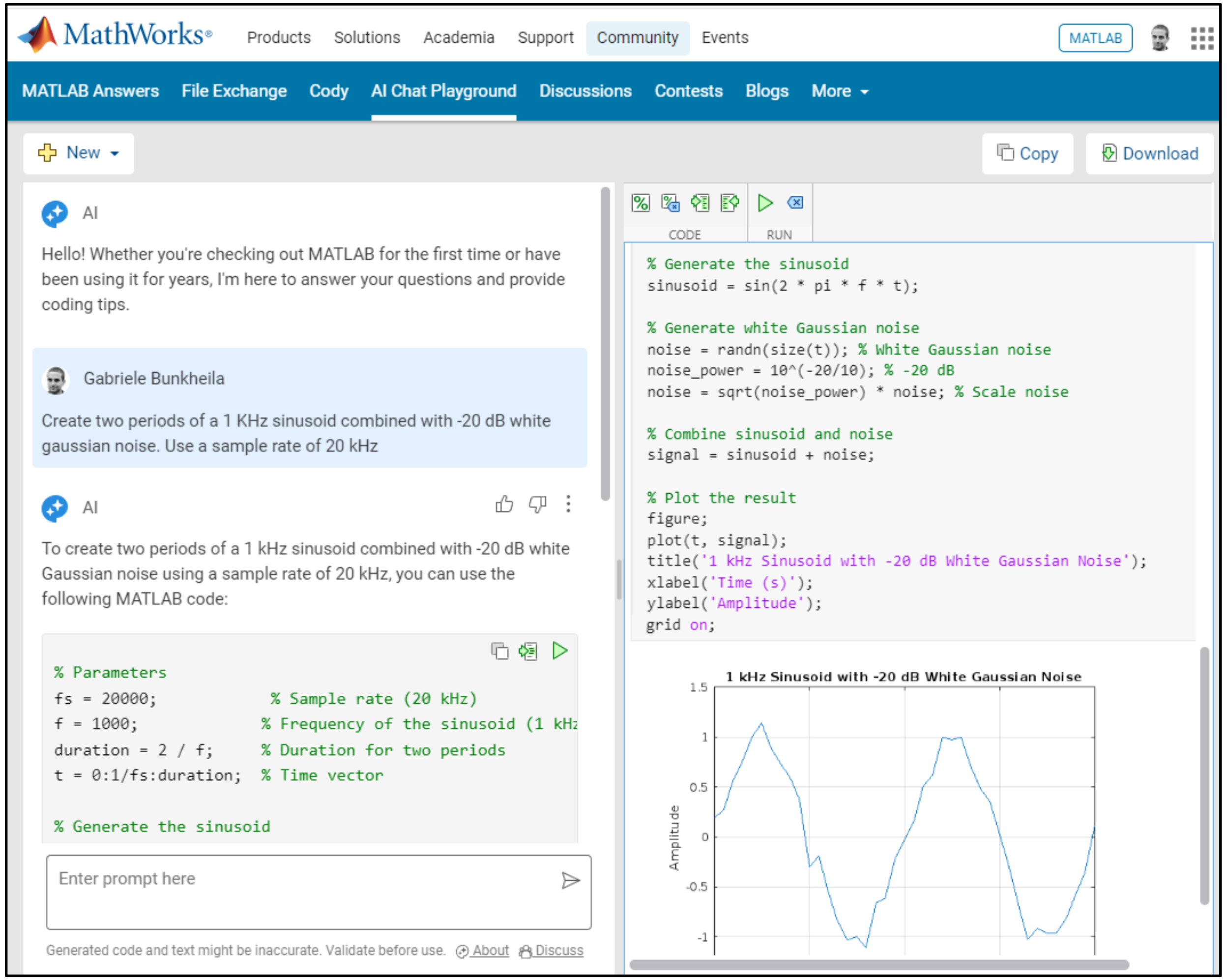
- For OpenAI connections: An active OpenAI API subscription and API key.
- For Azure OpenAI Services: An active Azure subscription with OpenAI access, deployment, and API key.
- For Ollama: An Ollama installation.

Setup

See these pages for instructions specific to the 3rd party product selected:

- OpenAI
- Azure
- Ollama

Large Language Models are also increasingly used to accelerate programming tasks



AI

Hello! Whether you're checking out MATLAB for the first time or have been using it for years, I'm here to answer your questions and provide coding tips.

Gabriele Bunkheila

Create two periods of a 1 KHz sinusoid combined with -20 dB white gaussian noise. Use a sample rate of 20 kHz

AI

To create two periods of a 1 kHz sinusoid combined with -20 dB white Gaussian noise using a sample rate of 20 kHz, you can use the following MATLAB code:

```
% Parameters
fs = 20000;             % Sample rate (20 kHz)
f = 1000;              % Frequency of the sinusoid (1 kHz
duration = 2 / f;      % Duration for two periods
t = 0:1/fs:duration;   % Time vector

% Generate the sinusoid
```

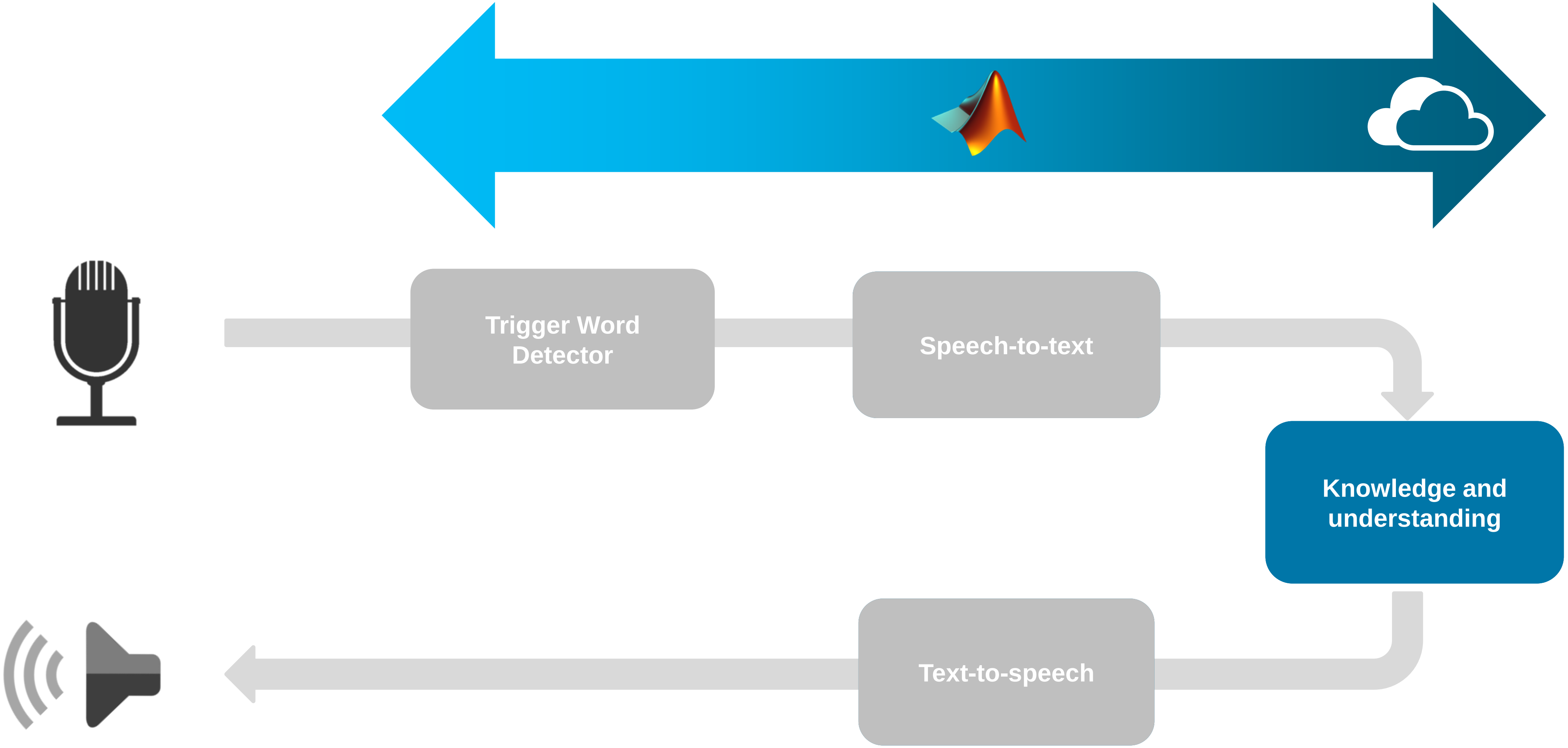
Enter prompt here

Generated code and text might be inaccurate. Validate before use. About Discuss

```
% Generate the sinusoid
sinusoid = sin(2 * pi * f * t);

% Generate white Gaussian noise
noise = randn(size(t)); % White Gaussian noise
noise_power = 10^(-20/10); % -20 dB
noise = sqrt(noise_power) * noise; % Scale noise

% Combine sinusoid and noise
signal = sinusoid + noise;

% Plot the result
figure;
plot(t, signal);
title('1 kHz Sinusoid with -20 dB White Gaussian Noise');
xlabel('Time (s)');
ylabel('Amplitude');
grid on;
```

The AI Chat Playground is a great starting point to get interactive help on MATLAB programming

MATLAB is also used by engineers to develop programs and apps deployed in the cloud



Trigger Word Detector

Speech-to-text

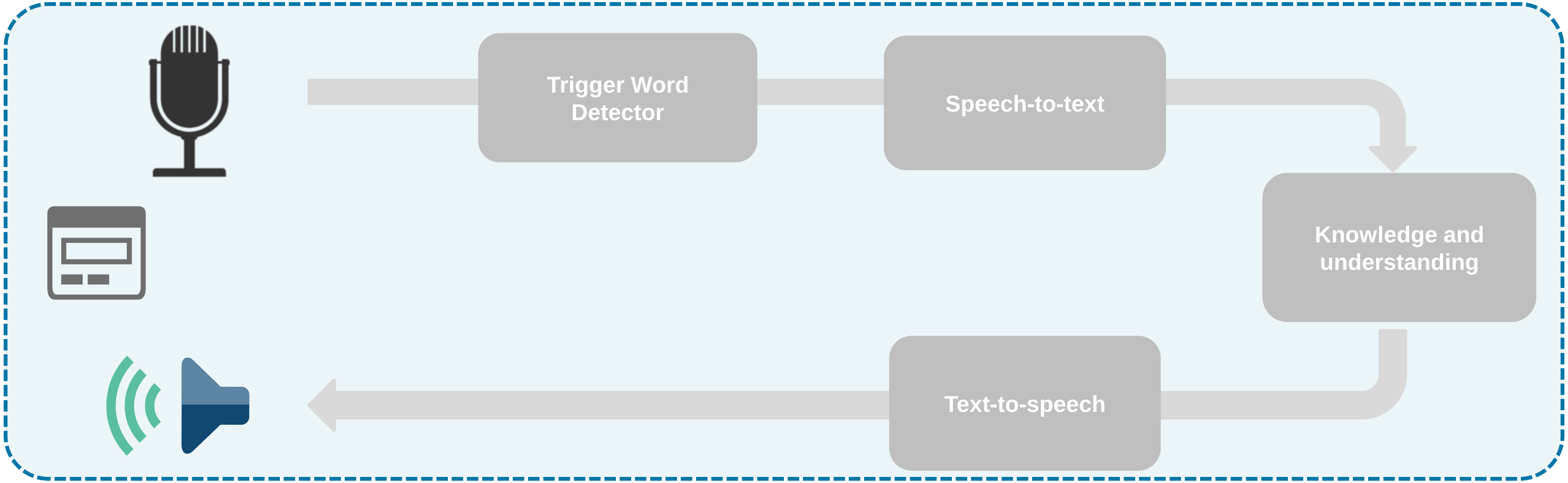
Knowledge and understanding

Text-to-speech

Using AI with MATLAB

4.

Prototype with AI, Apps and hardware



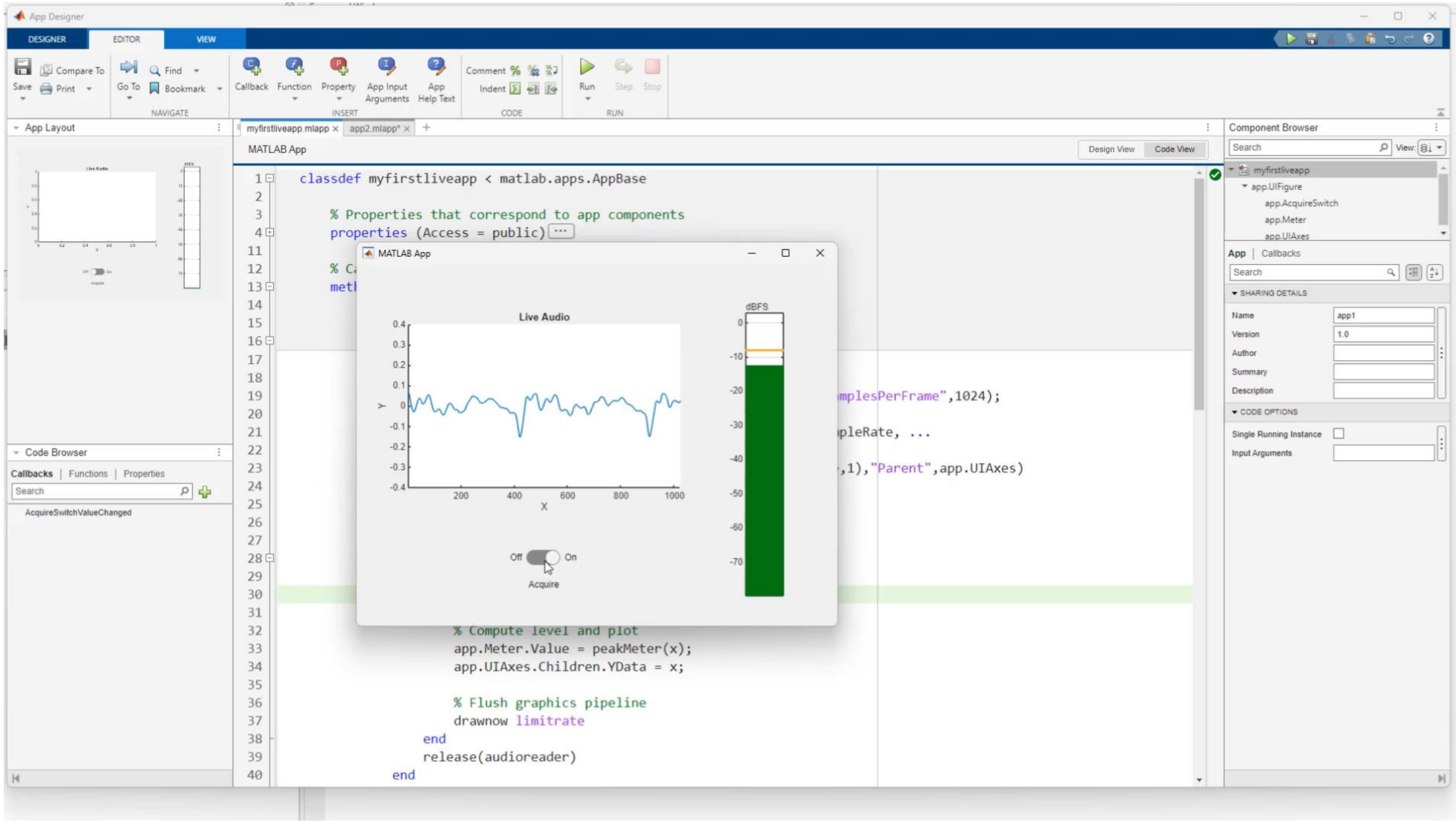
Trigger Word Detector

Speech-to-text

Knowledge and understanding

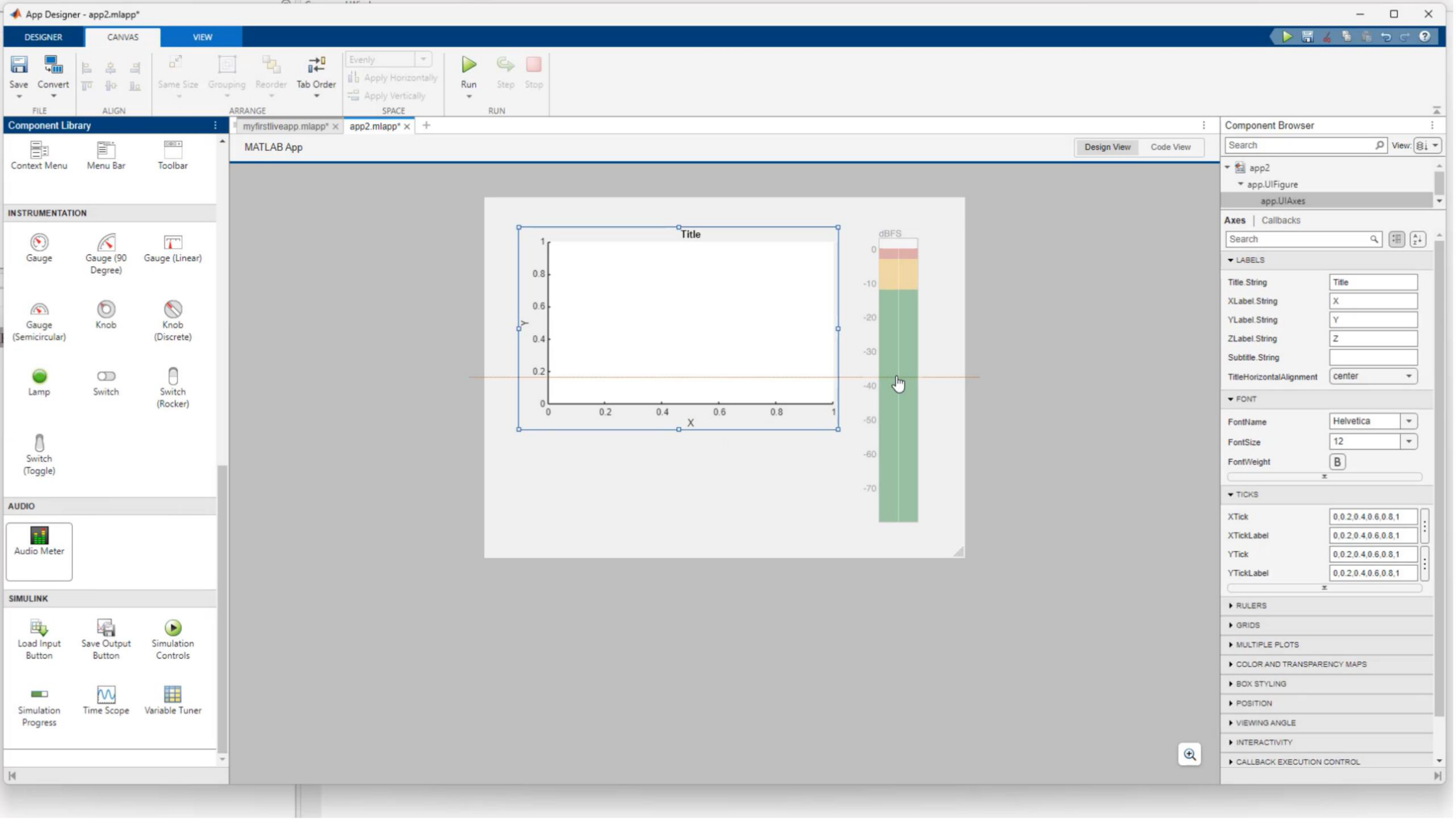
Text-to-speech

Getting started with an App prototype using App Designer in MATLAB doesn’t take long at all



```
classdef myfirstliveapp < matlab.apps.AppBase

    % Properties that correspond to app components
    properties (Access = public) ...

    % Ca
    metl

mplesPerFrame",1024);

pleRate, ...

,1),"Parent",app.UIAxes)

                    % Compute level and plot
                    app.Meter.Value = peakMeter(x);
                    app.UIAxes.Children.YData = x;

                    % Flush graphics pipeline
                    drawnow limitrate
                end
                release(audioreader)
            end
```

Name: app1
Version: 1.0
Author:
Summary:
Description:
Single Running Instance ☐
Input Arguments:



Title.String Title

XLabel.String X

YLabel.String Y

ZLabel.String Z

Subtitle.String

TitleHorizontalAlignment center

FontName Helvetica

FontSize 12

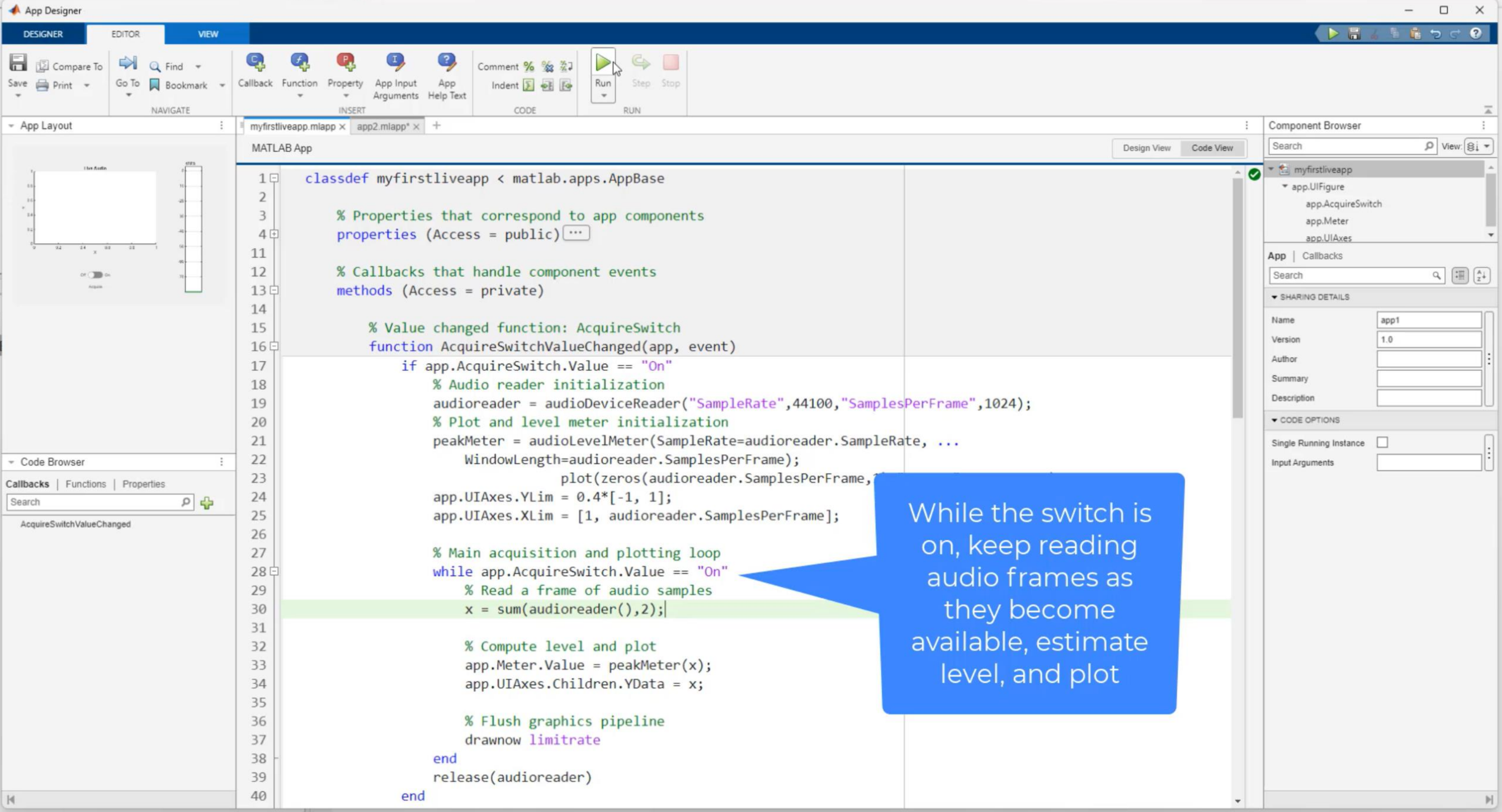
FontWeight B

XTick 0,0.2,0.4,0.6,0.8,1

XTickLabel 0,0.2,0.4,0.6,0.8,1

YTick 0,0.2,0.4,0.6,0.8,1

YTickLabel 0,0.2,0.4,0.6,0.8,1



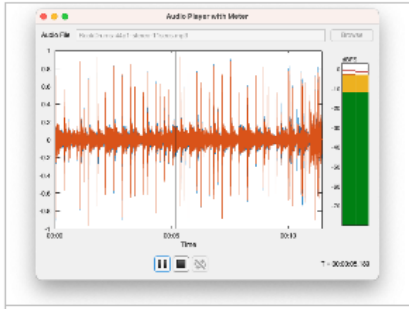
```
classdef myfirstliveapp < matlab.apps.AppBase

    % Properties that correspond to app components
    properties (Access = public) ...

    % Callbacks that handle component events
    methods (Access = private)

        % Value changed function: AcquireSwitch
        function AcquireSwitchValueChanged(app, event)
            if app.AcquireSwitch.Value == "On"
                % Audio reader initialization
                audioreader = audioDeviceReader("SampleRate",44100,"SamplesPerFrame",1024);
                % Plot and level meter initialization
                peakMeter = audioLevelMeter(SampleRate=audioreader.SampleRate, ...
                    WindowLength=audioreader.SamplesPerFrame);
                            plot(zeros(audioreader.SamplesPerFrame,
                app.UIAxes.YLim = 0.4*[-1, 1];
                app.UIAxes.XLim = [1, audioreader.SamplesPerFrame];

                % Main acquisition and plotting loop
                while app.AcquireSwitch.Value == "On"
                    % Read a frame of audio samples
                    x = sum(audioreader(),2);

                    % Compute level and plot
                    app.Meter.Value = peakMeter(x);
                    app.UIAxes.Children.YData = x;

                    % Flush graphics pipeline
                    drawnow limitrate
                end
                release(audioreader)
            end
```

While the switch is on, keep reading audio frames as they become available, estimate level, and plot

Name: app1
Version: 1.0
Author:
Summary:
Description:

Single Running Instance ☐
Input Arguments:

Your handouts include additional resources to get started with App prototypes, including with hardware I/O and AI models

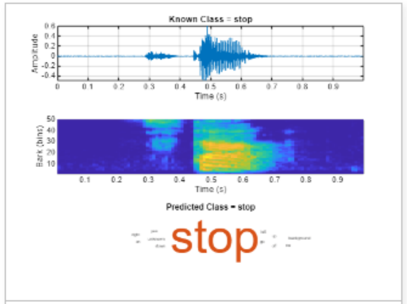


Create an App to Play and Visualize Audio Files

Create an app in App Designer to play and visualize audio files.

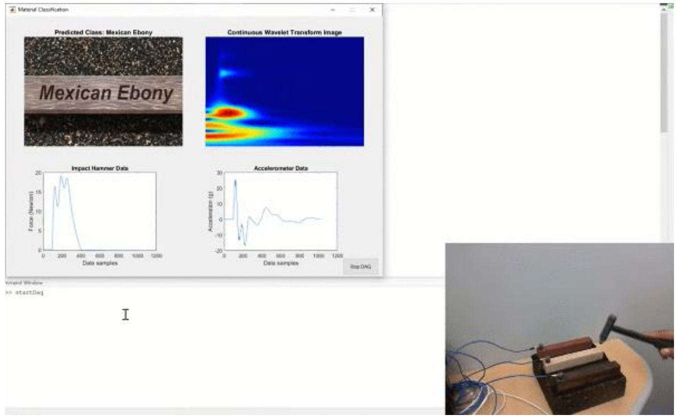
Audio Input and Audio Output

Read audio from a file and write audio to speakers.



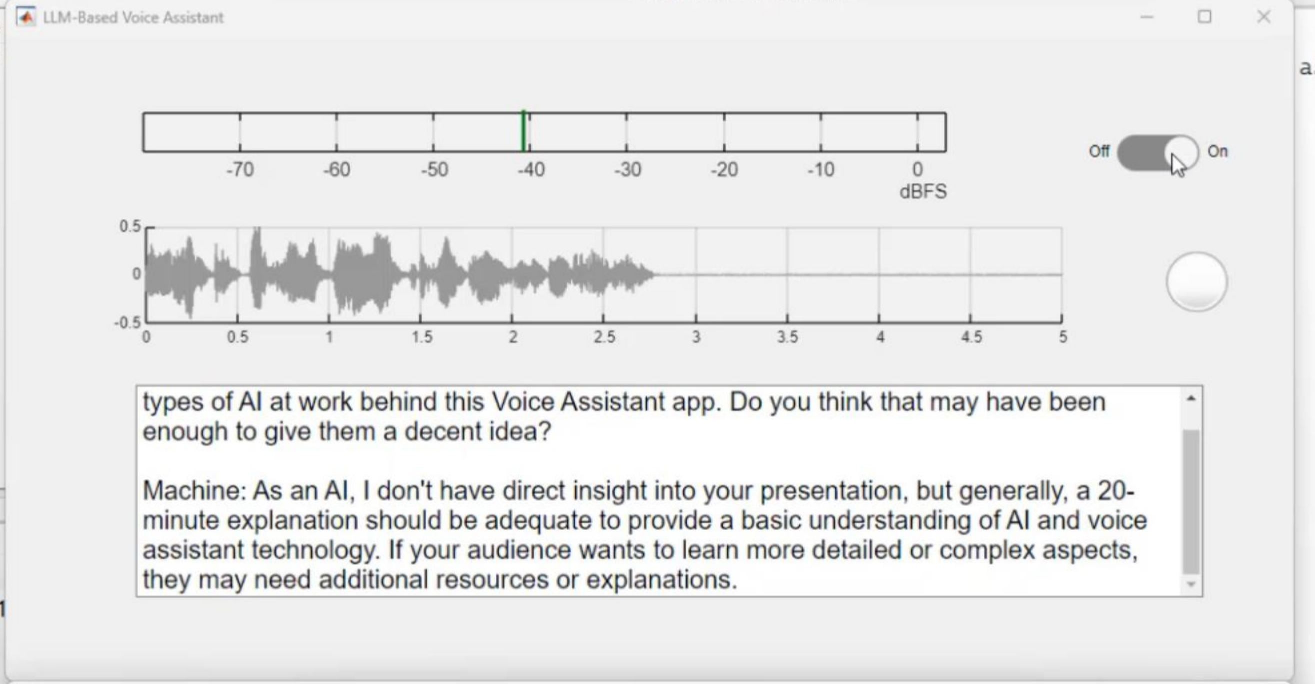
Speech Command Recognition Using Deep Learning

Use a pretrained deep learning model to perform speech command recognition on streaming audio.


Predicted Class: Mexican Ebony
Continuous Wavelet Transform Image
Mexican Ebony
Impact Hammer Data
Accelerometer Data
Data samples

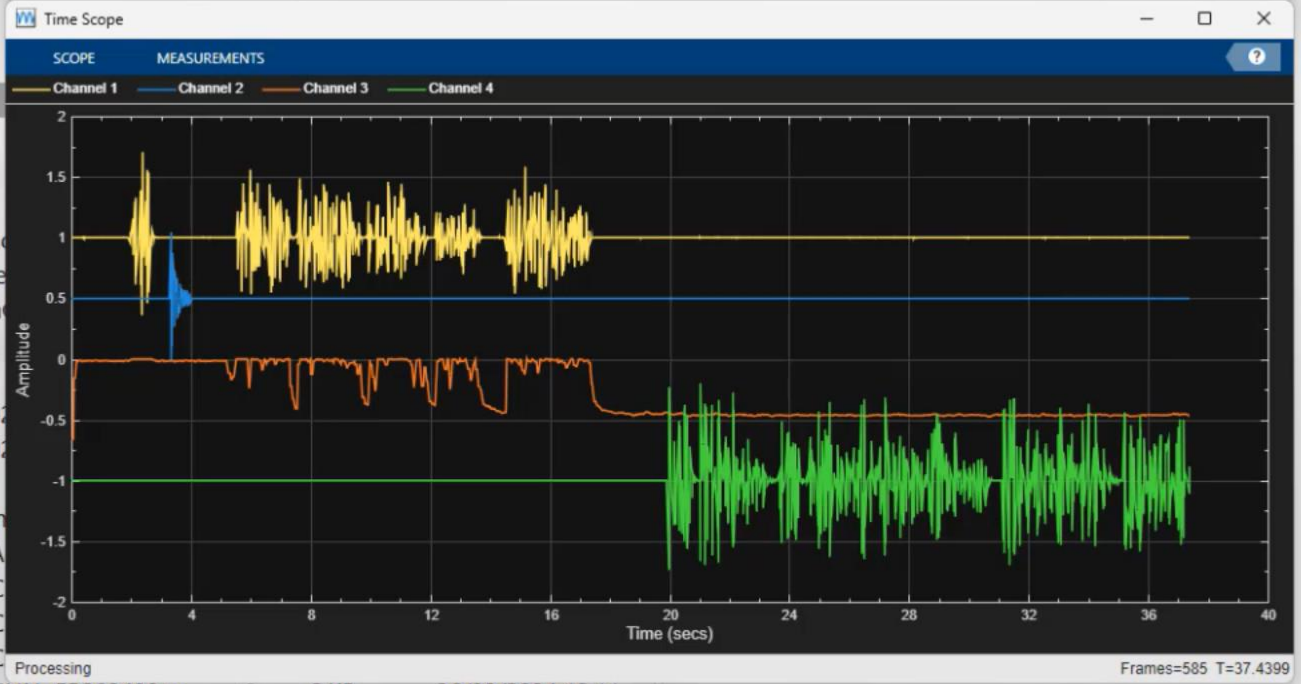
[Wood Species Classification Using Vibration Signature]() on GitHub

Command Window



types of AI at work behind this Voice Assistant app. Do you think that may have been enough to give them a decent idea?

Machine: As an AI, I don't have direct insight into your presentation, but generally, a 20-minute explanation should be adequate to provide a basic understanding of AI and voice assistant technology. If your audience wants to learn more detailed or complex aspects, they may need additional resources or explanations.



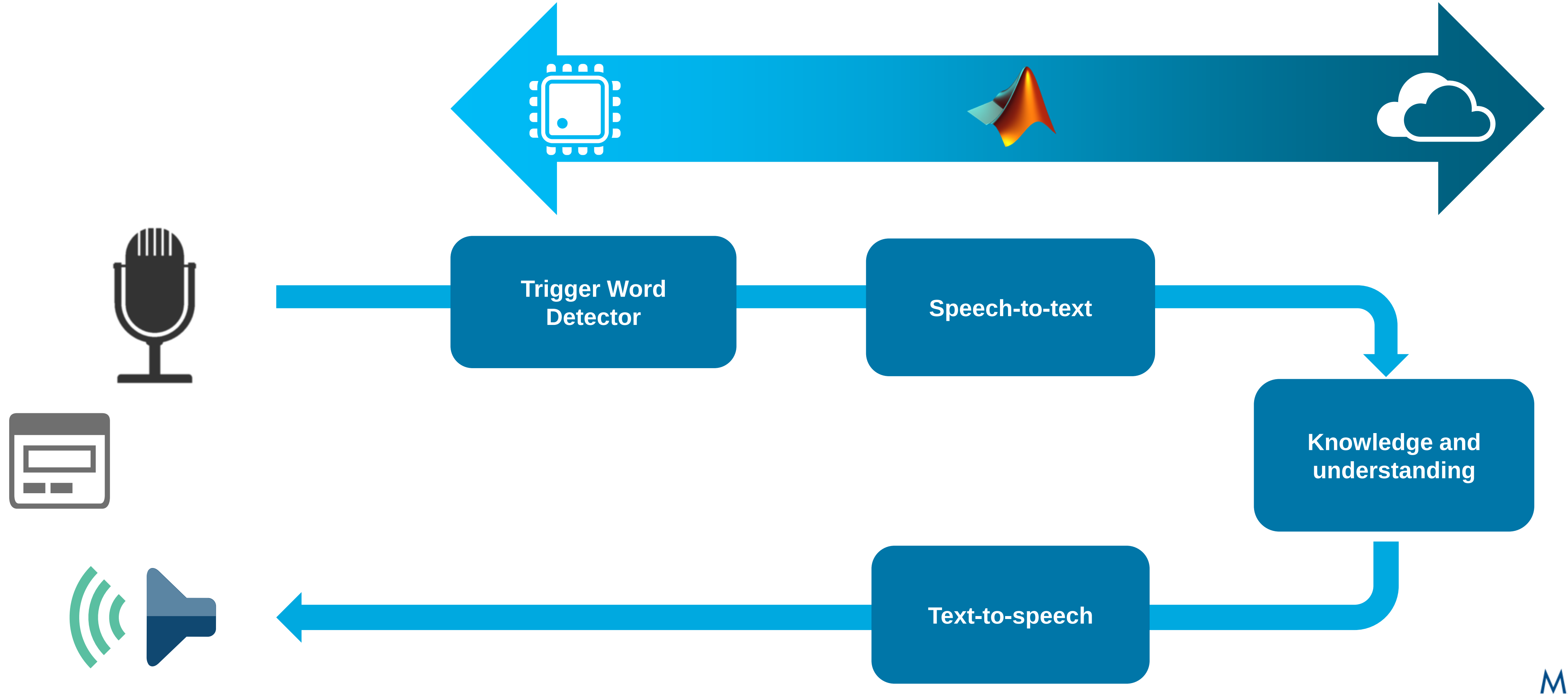
Processing

Frames=585 T=37.4399

CNN_GB202408.m ○ 6 KB 26/08/2024 12:11

letPredictCG.m ● 1 KB 24/05/2024 0:09

In conclusion: using AI can mean many different things, and all can be brought together in MATLAB to create application prototypes



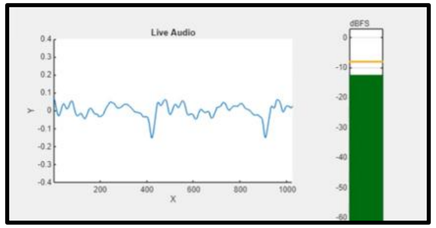
Trigger Word Detector

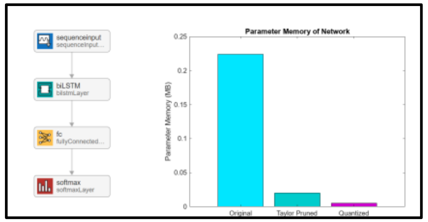
Speech-to-text

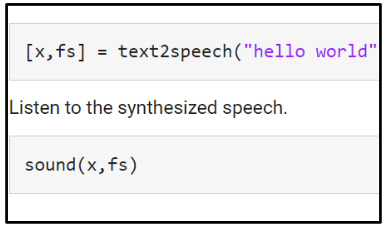
Knowledge and understanding

Text-to-speech

In conclusion: using AI can mean many different things, and all can be brought together in MATLAB to create application prototypes



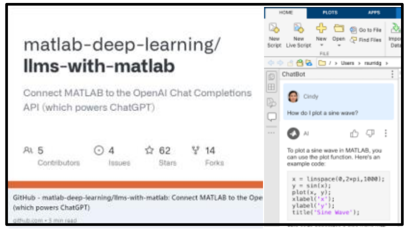




```
[x,fs] = text2speech("hello world"
```

Listen to the synthesized speech.

```
sound(x,fs)
```



matlab-deep-learning/
llms-with-matlab

Connect MATLAB to the OpenAI Chat Completions API (which powers ChatGPT)

GitHub - matlab-deep-learning/llms-with-matlab: Connect MATLAB to the Ope
(which powers ChatGPT)

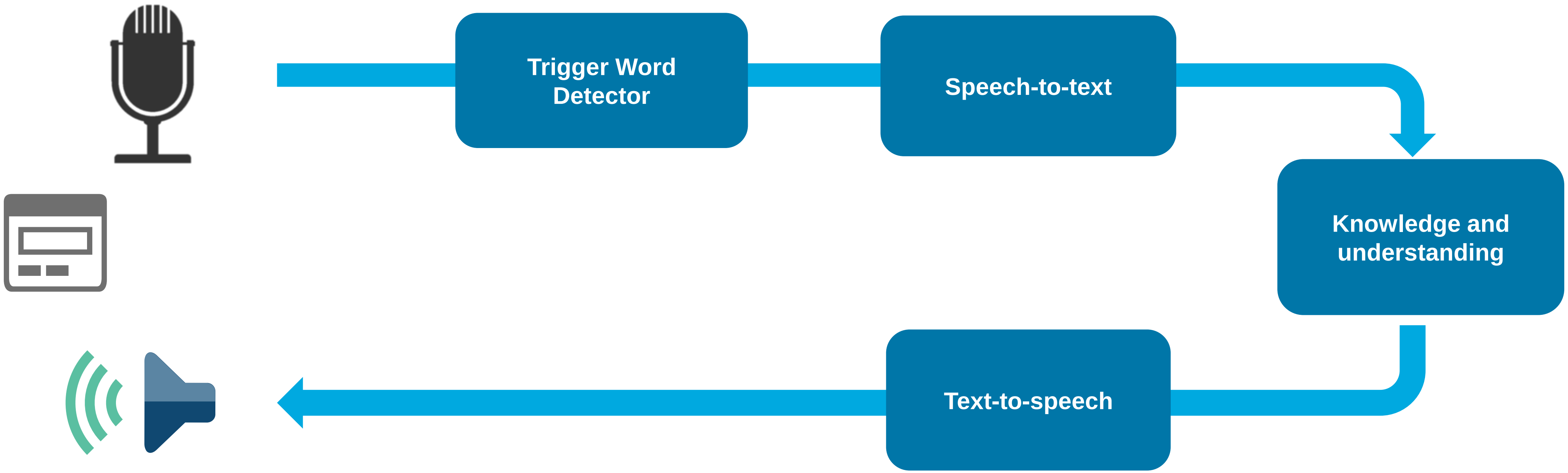
Cindy

How do I plot a sine wave?

AI

To plot a sine wave in MATLAB, you can use the plot function. Here's an example code:

```
x = linspace(0,2*pi,1000);
y = sin(x);
plot(x, y);
xlabel('x');
ylabel('y');
title('Sine Wave');
```



Trigger Word Detector

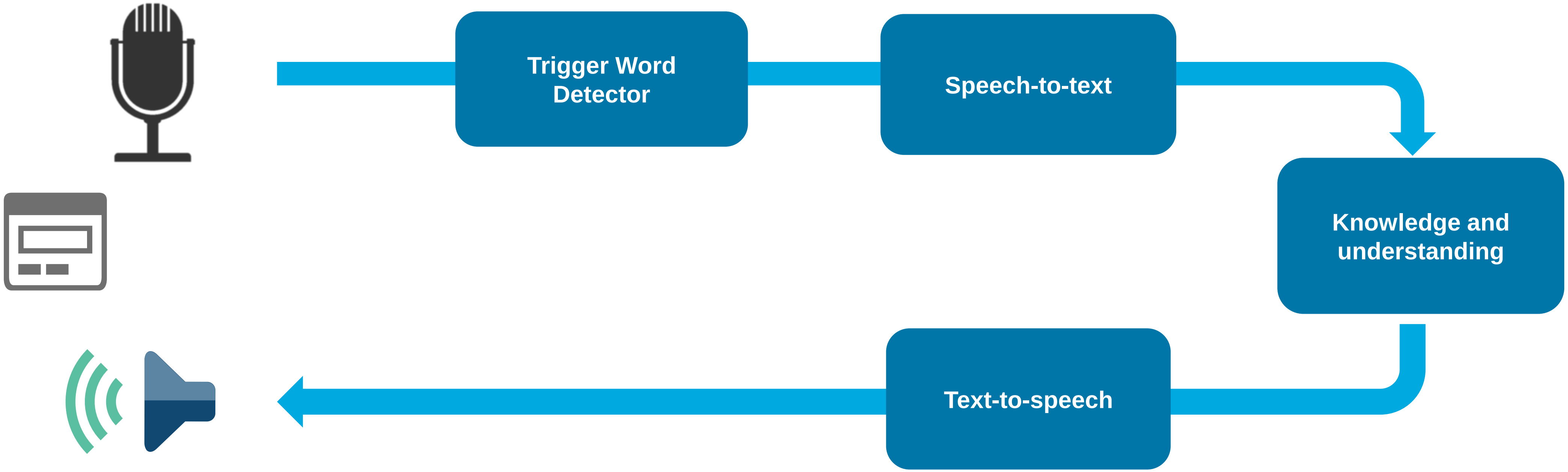
Speech-to-text

Knowledge and understanding

Text-to-speech

In conclusion: using AI can mean many different things, and all can be brought together in MATLAB to create application prototypes

Questions?


Trigger Word Detector

Speech-to-text

Knowledge and understanding

Text-to-speech