Luware
Nimbus

Create a Trained Speech Recognizer for Luware Nimbus

Step-by-Step Guide

Contents

Introduction

This guide outlines the steps to create and integrate a custom speech recognizer model for Luware Nimbus using Microsoft Azure Speech Services.

Benefits:

- Improved speech recognition accuracy for your specific use case within Luware Nimbus.
- Enhanced captions and transcriptions for virtual assistant interactions.

Here's what you'll learn:

1. Training a Microsoft Speech Recognizer Model
2. Creating Your Training Data
3. Train the Model
4. Test and Deploy Your Model
5. Add the Speech Recognizer to Luware Nimbus

By following these steps, you can create a speech recognizer specifically tailored to your needs within Luware Nimbus, leading to improved accuracy and user experience.

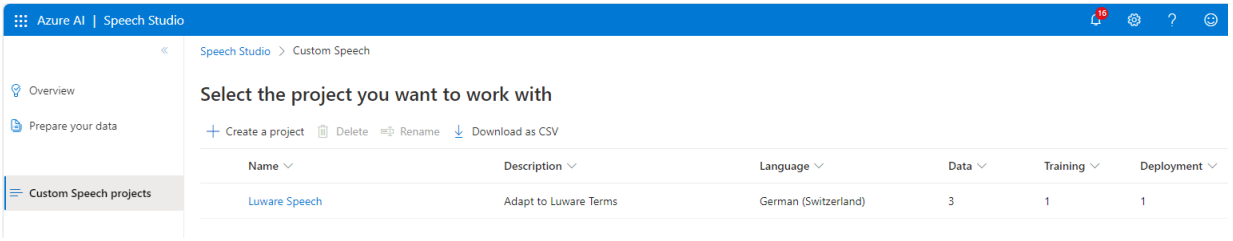
Note: This guide provides a high-level overview. Refer to the official Microsoft documentation for detailed instructions and code samples:

- **Training and testing datasets:** https://learn.microsoft.com/en-us/azure/ai-services/speech-service/how-to-custom-speech-test-and-train

How To Train a Microsoft Speech Recognizer Model

Training a speech model starts in the **Custom Speech section in the Azure Speech Studio**.

To access the Custom Speech area, you need to create a Speech resource in the Azure portal and then sign in to Azure Speech Studio with your Azure credentials. There, you can create a new Custom speech project or use an existing one.



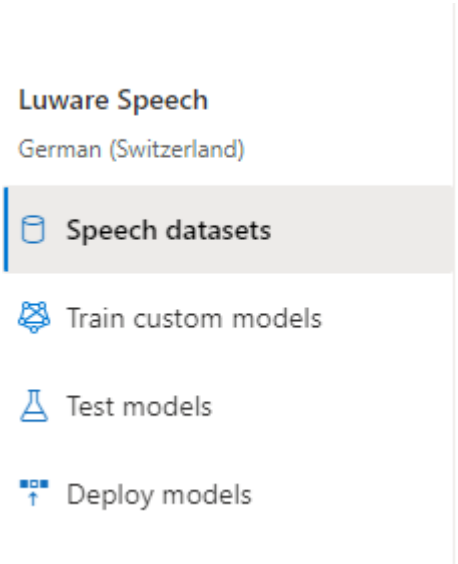
A Custom speech project consists of three components:

- data,
- models,
- and endpoints.

You need to upload your training and testing datasets to the data component, and then use them to train a custom model in the models' component.

You can also evaluate the performance of your model on the testing dataset and compare it with the baseline model.

Once you are satisfied with your model, you can deploy it to an endpoint and use it for speech recognition via the **Speech SDK** or **REST API**.



Creating Your Training Data

The first step in creating your custom speech model is to prepare your training data. Training data consists of audio files and their corresponding transcripts that represent the speech patterns and vocabulary you want your model to recognize. **The quality and quantity of your training data directly affect the accuracy of your custom speech model.**

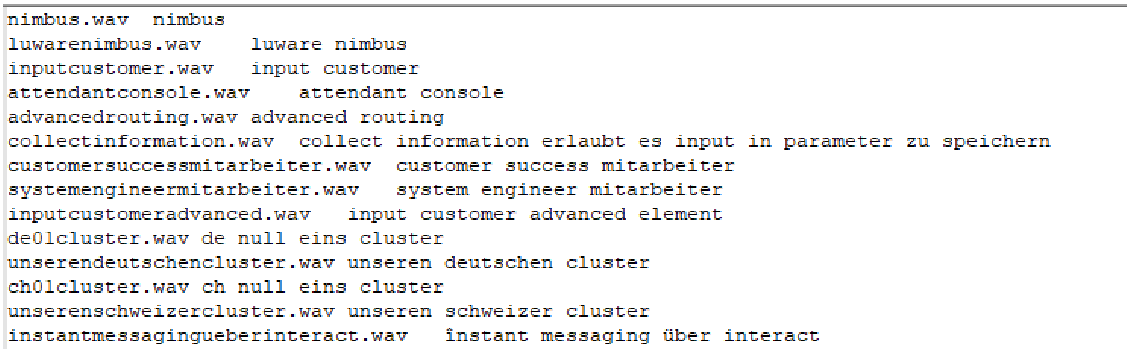
If you want to use the model for your Contact Center, we recommend using the **Audio + human-labelled transcript** data for training or testing. Such data allows you to provide both the original audio file and a manually corrected transcript for each utterance. The transcript should match the audio exactly and include punctuation, capitalization, and special symbols. This way, you can ensure that your model learns from the correct transcription and not from the errors of the automatic speech recognition system.

To Create Your Training Data, Follow These Steps

- Collect audio files that match the characteristics of your target domain, such as **language**, **accent**, **channel**, **background noise**, etc. The audio files should be in **WAV format** with **16 kHz** sampling rate and **16-bit PCM** encoding.

- Transcribe each audio file using the best available method, such as the baseline speech recognition service, a third-party transcription service, or manual transcription. You can also use the transcription feature in the Speech Studio to generate transcripts automatically and edit them as needed.

- Create a TXT – File, that combines the WAV files with the according transcription. **The name of the WAV file must be separated by a tab (\t)**. The transcription is only done with lower case letters. There is a guidance for different languages available.



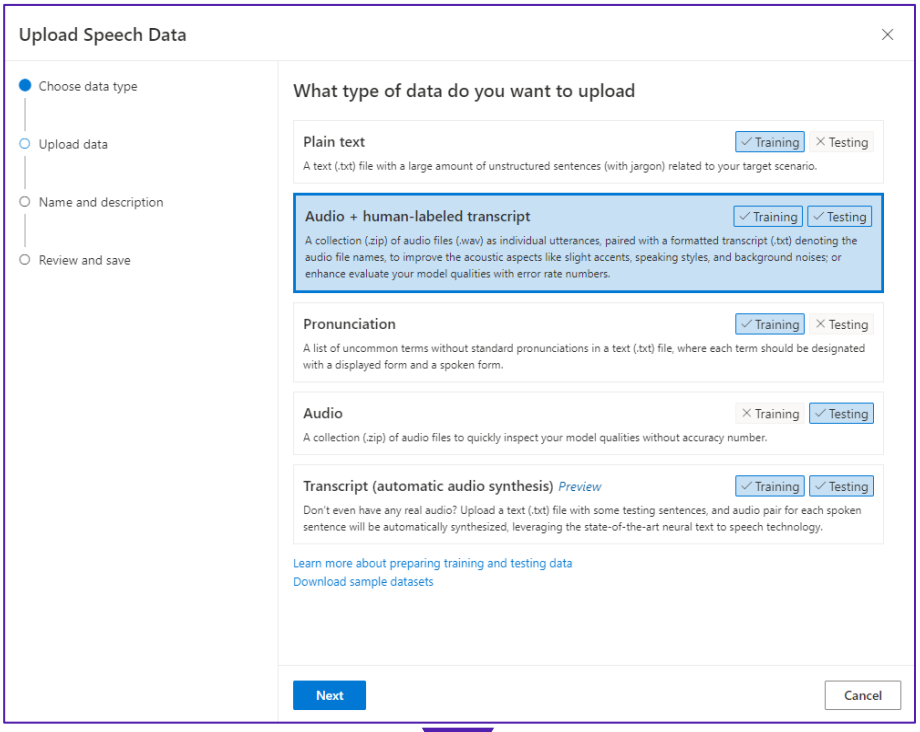
```
nimbus.wav  nimbus
luwarenimbus.wav    luware nimbus
inputcustomer.wav   input customer
attendantconsole.wav    attendant console
advancedrouting.wav advanced routing
collectinformation.wav  collect information erlaubt es input in parameter zu speichern
customersuccessmitarbeiter.wav  customer success mitarbeiter
systemengineermitarbeiter.wav   system engineer mitarbeiter
inputcustomeradvanced.wav   input customer advanced element
de01cluster.wav de null eins cluster
unserendeutschencluster.wav unseren deutschen cluster
ch01cluster.wav ch null eins cluster
unserenschweizercluster.wav unseren schweizer cluster
instantmessagingueberinteract.wav   înstant messaging über interact
```

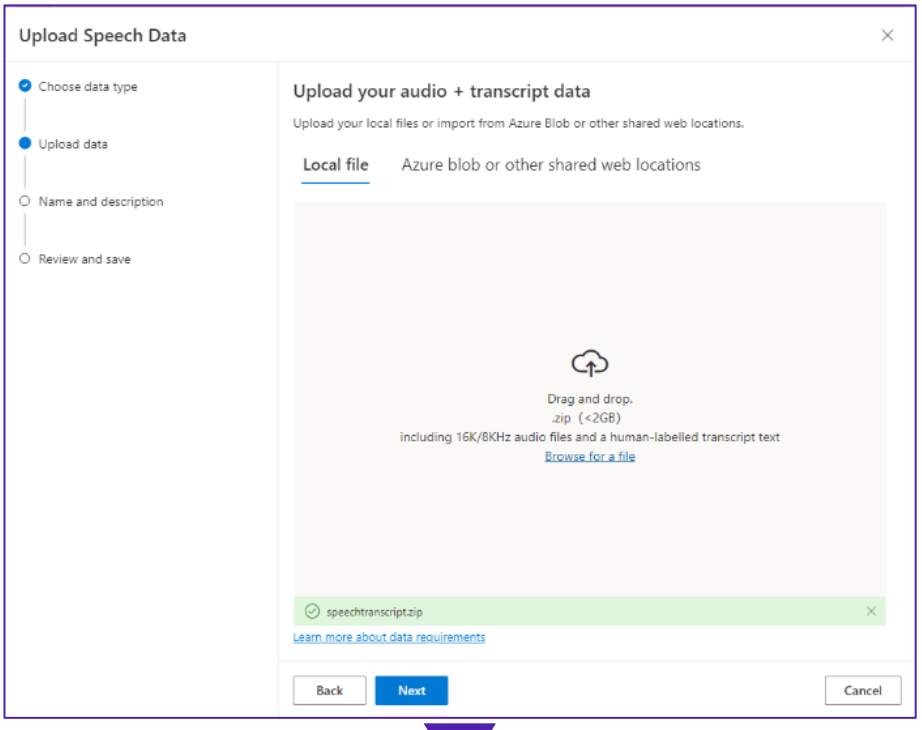
- Review and correct the transcripts to make sure they are accurate and consistent. You should also format the transcripts according to the guidelines provided in the documentation, such as **using brackets for non-speech events, using commas for pauses, and using quotation marks for direct speech**.

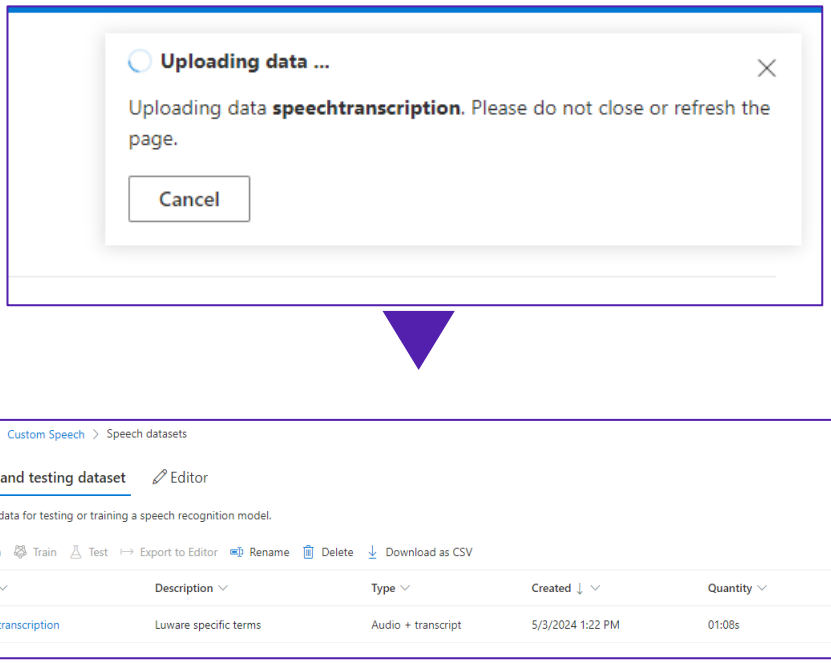
- Create a ZIP file, which contains the TXT file as well as the WAV files.

- Upload the ZIP file to the data component in the Speech Studio. You should also **provide some metadata for each file**, such as the locale, the speaker ID, and the transcription type.

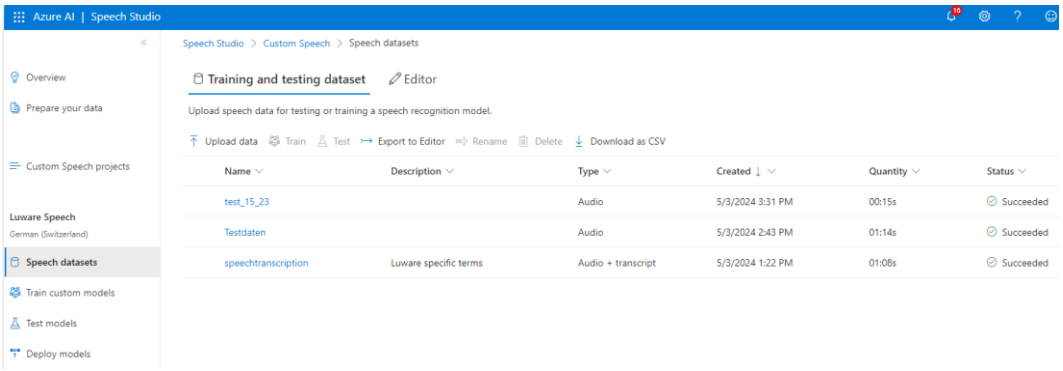
You can **use the import feature** in the Speech Studio to upload your data and fill in the metadata automatically.







- Label your data as training or testing data. You should use at least 80% of your data for training and the rest for testing. You can also create multiple datasets for different purposes, such as **tuning**, **validation**, or **evaluation**. Furthermore, you can use the dataset feature in the Speech Studio to create and manage your datasets.

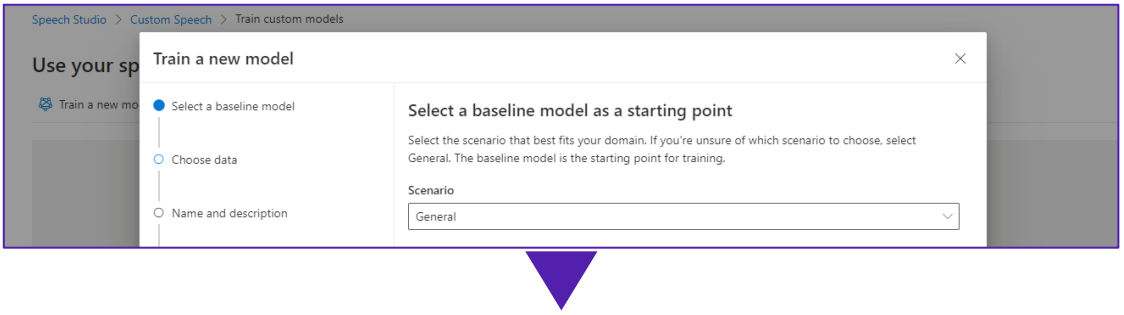


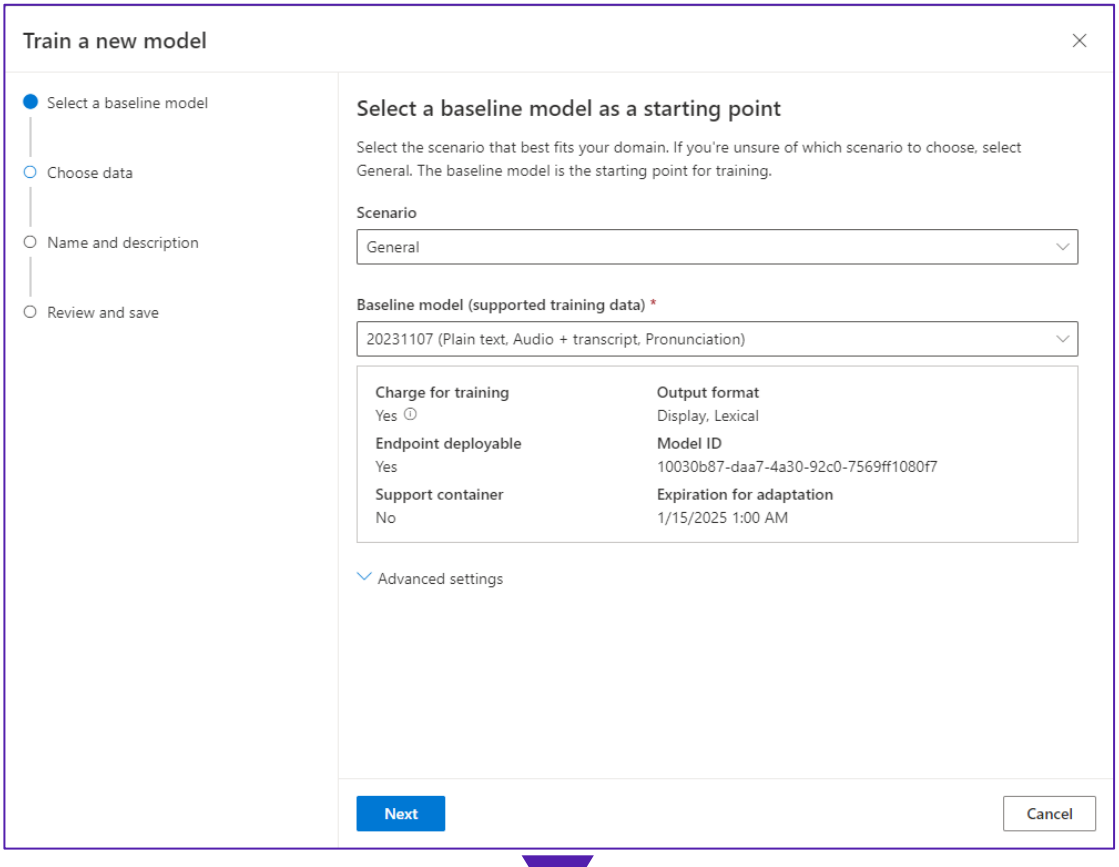
Train the Model

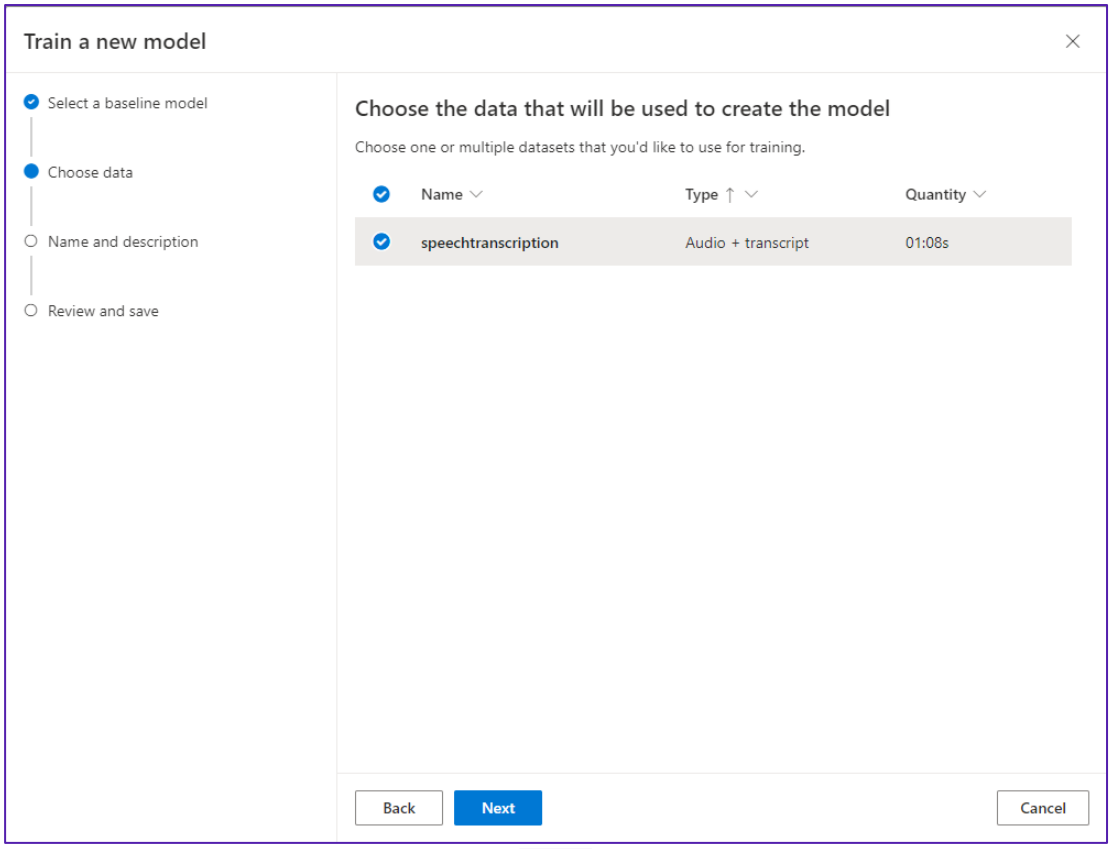
Switch to the **train custom models** tab to train the model.

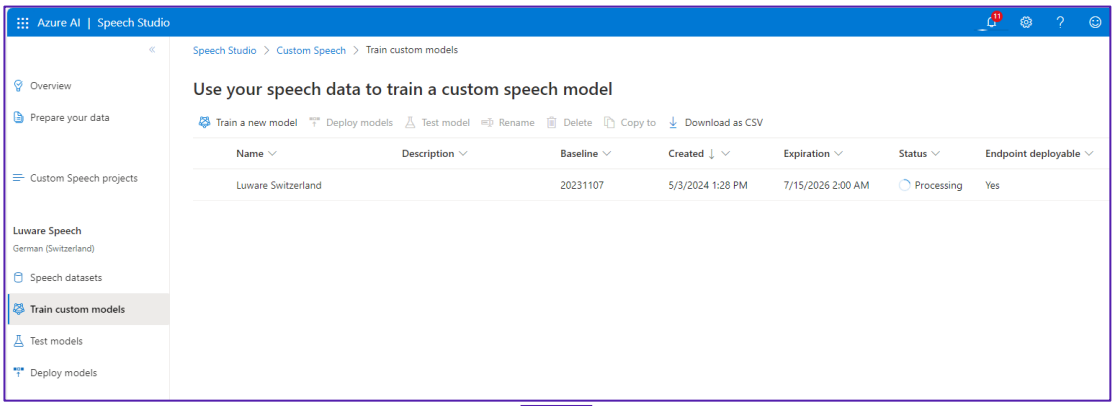
To train your custom model, follow these steps:

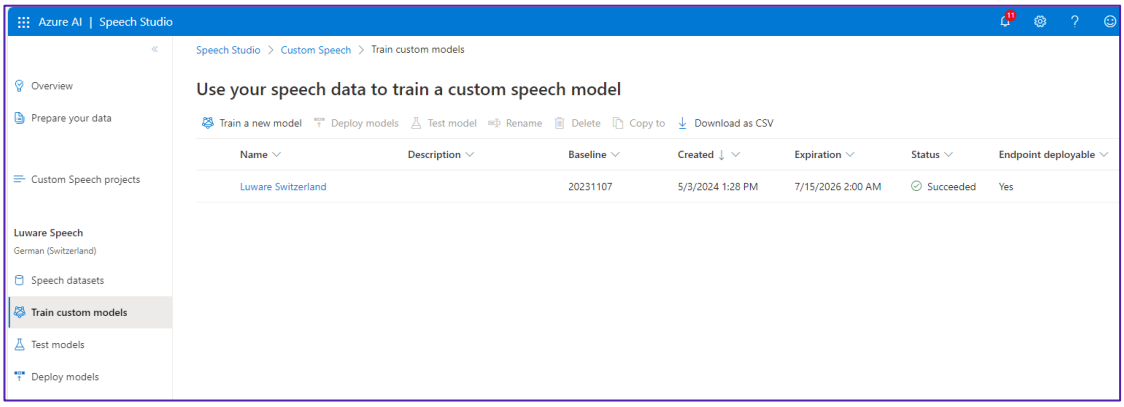
- Select the dataset you want to use for training from the dropdown menu.
- Choose a base model that matches the language and domain of your data. You can also select a custom model that you have previously trained as a base model.
- Specify a name and a description for your new custom model.
- Optionally, adjust the training settings, such as the number of epochs, the learning rate, and the dropout rate.
- Click on the Train button and wait for the training process to complete. You can monitor the progress and the performance metrics on the dashboard.
- Once the training is done, you can download or deploy your custom model for testing or production purposes. You can also compare your custom model with the base model or other custom models on the Evaluate tab.





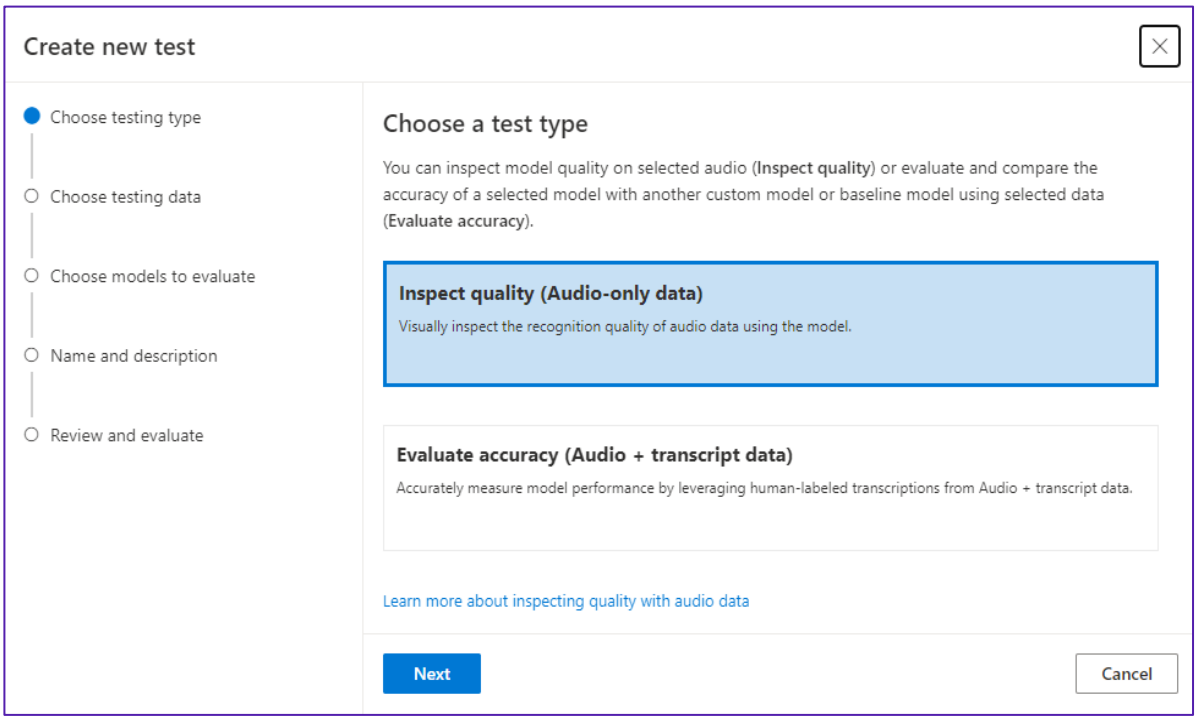

Train a new model
Select a baseline model
Choose data
Name and description
Review and save
Choose the data that will be used to create the model
Choose one or multiple datasets that you'd like to use for training.
Name
Type
Quantity
speechtranscription
Audio + transcript
01:08s
Back
Next
Cancel

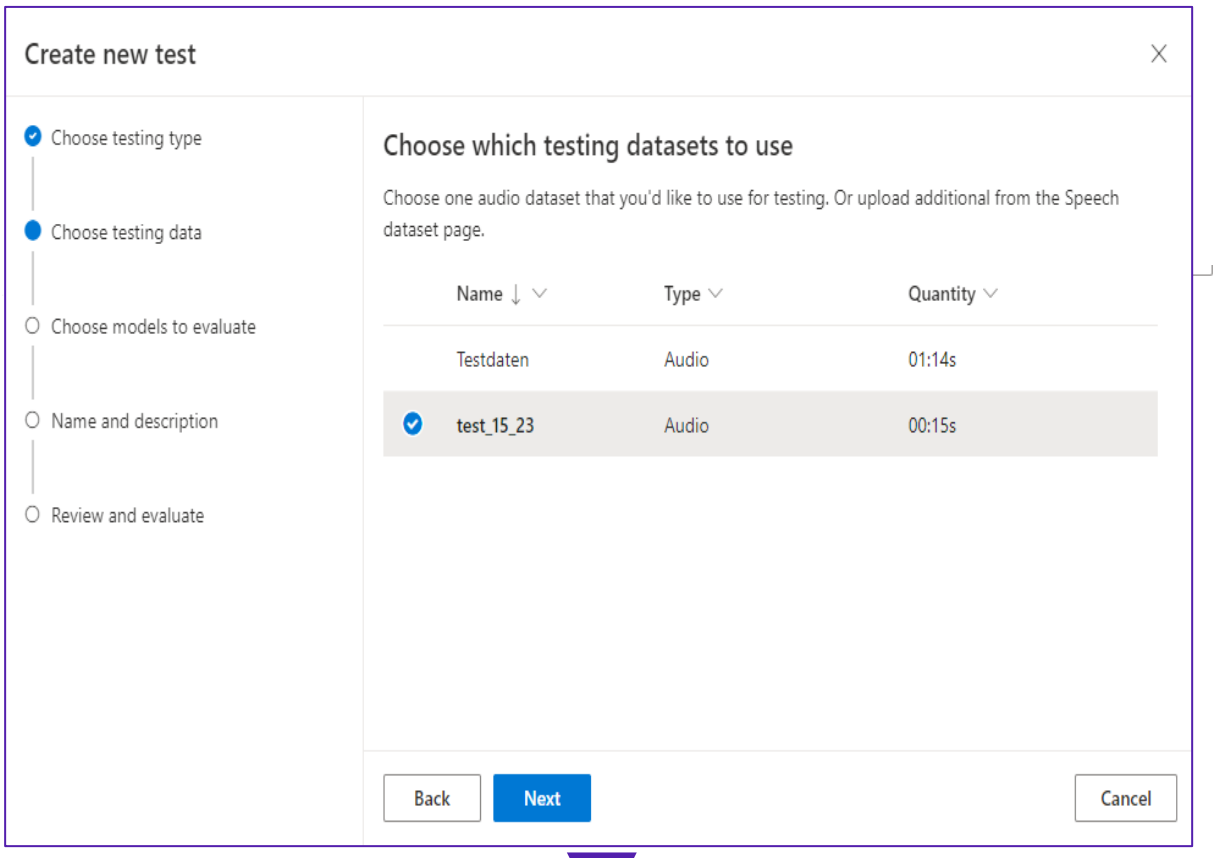

Azure AI | Speech Studio
Speech Studio > Custom Speech > Train custom models
Use your speech data to train a custom speech model
Overview
Prepare your data
Custom Speech projects
Luware Speech
German (Switzerland)
Speech datasets
Train custom models
Test models
Deploy models
Train a new model
Deploy models
Test model
Rename
Delete
Copy to
Download as CSV
Name
Description
Baseline
Created
Expiration
Status
Endpoint deployable
Luware Switzerland
20231107
5/3/2024 1:28 PM
7/15/2026 2:00 AM
Processing
Yes

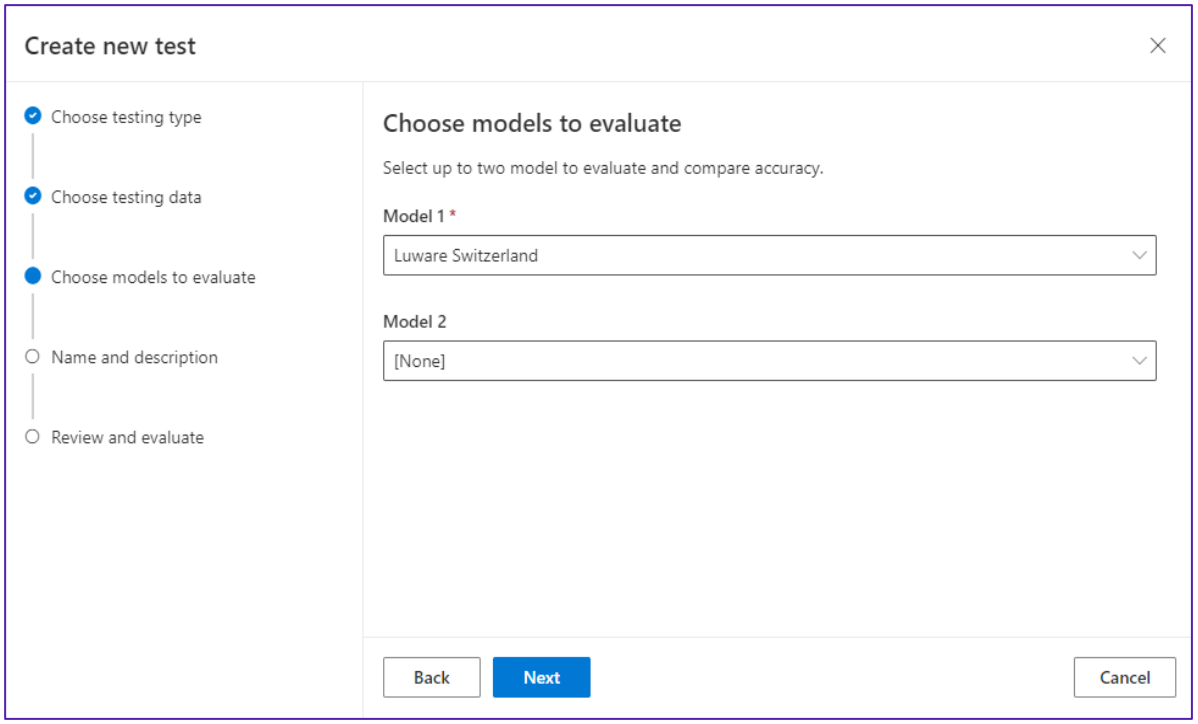

Azure AI | Speech Studio
Speech Studio > Custom Speech > Train custom models
Use your speech data to train a custom speech model
Overview
Prepare your data
Custom Speech projects
Luware Speech
German (Switzerland)
Speech datasets
Train custom models
Test models
Deploy models
Train a new model
Deploy models
Test model
Rename
Delete
Copy to
Download as CSV
Name
Description
Baseline
Created
Expiration
Status
Endpoint deployable
Luware Switzerland
20231107
5/3/2024 1:28 PM
7/15/2026 2:00 AM
Succeeded
Yes

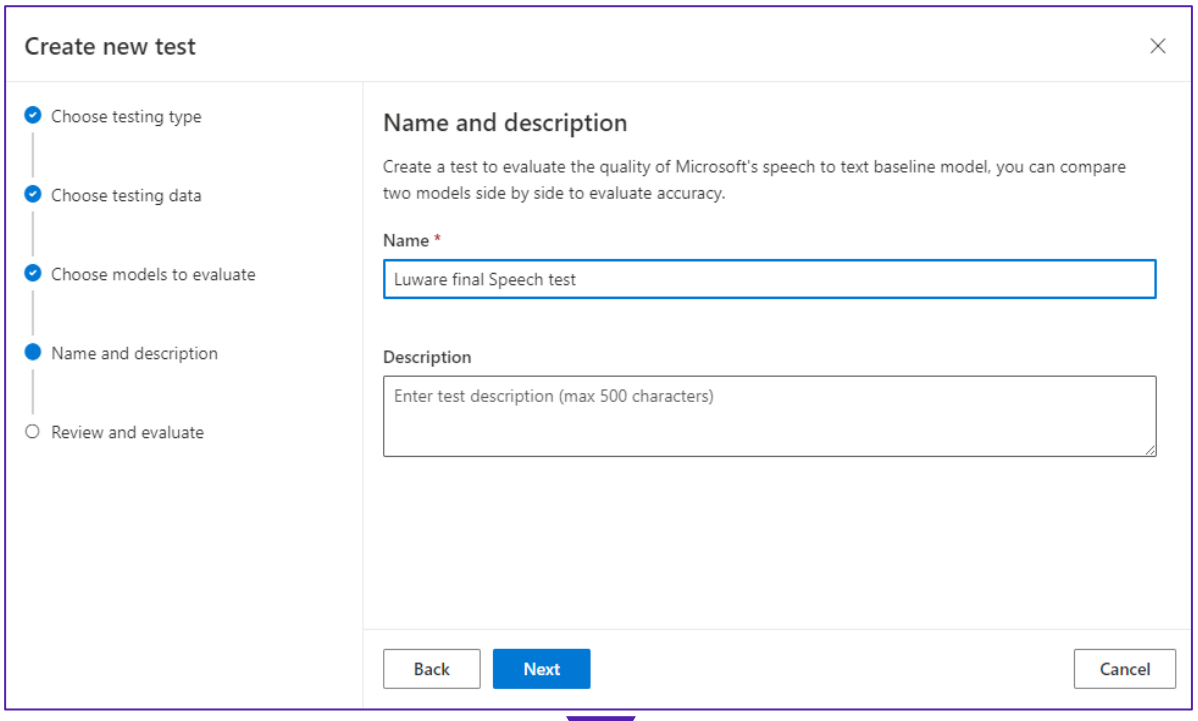
Test and Deploy Your Model

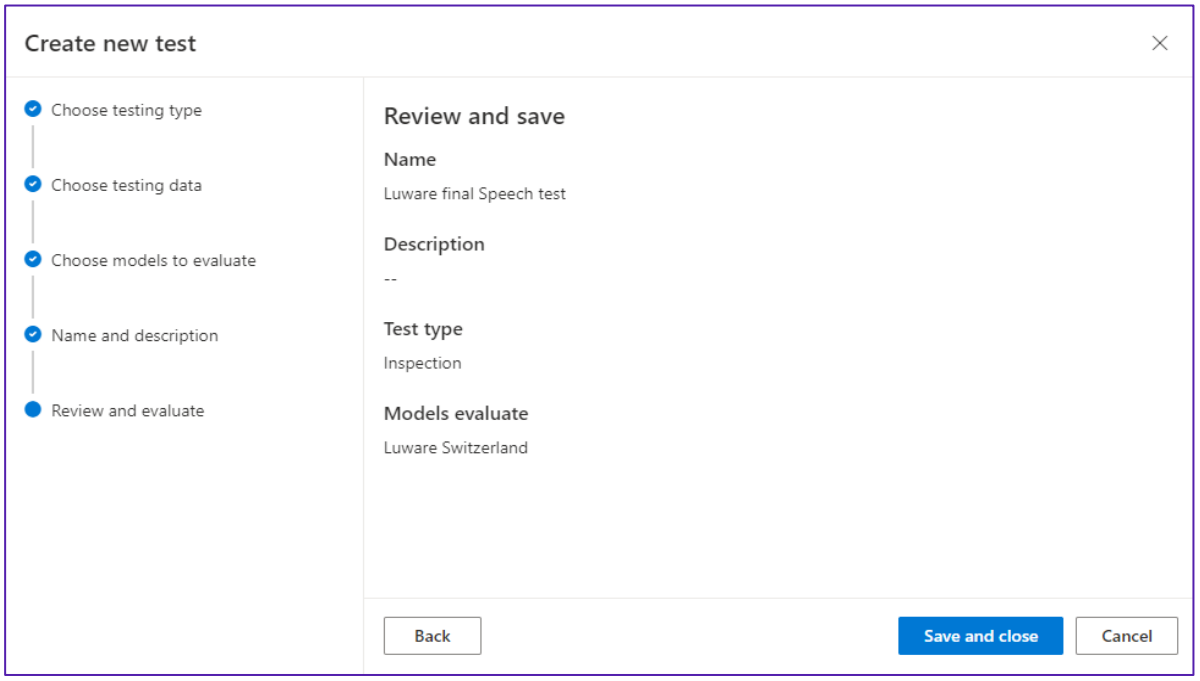
Once the model has been trained, you can now test the model with test data.





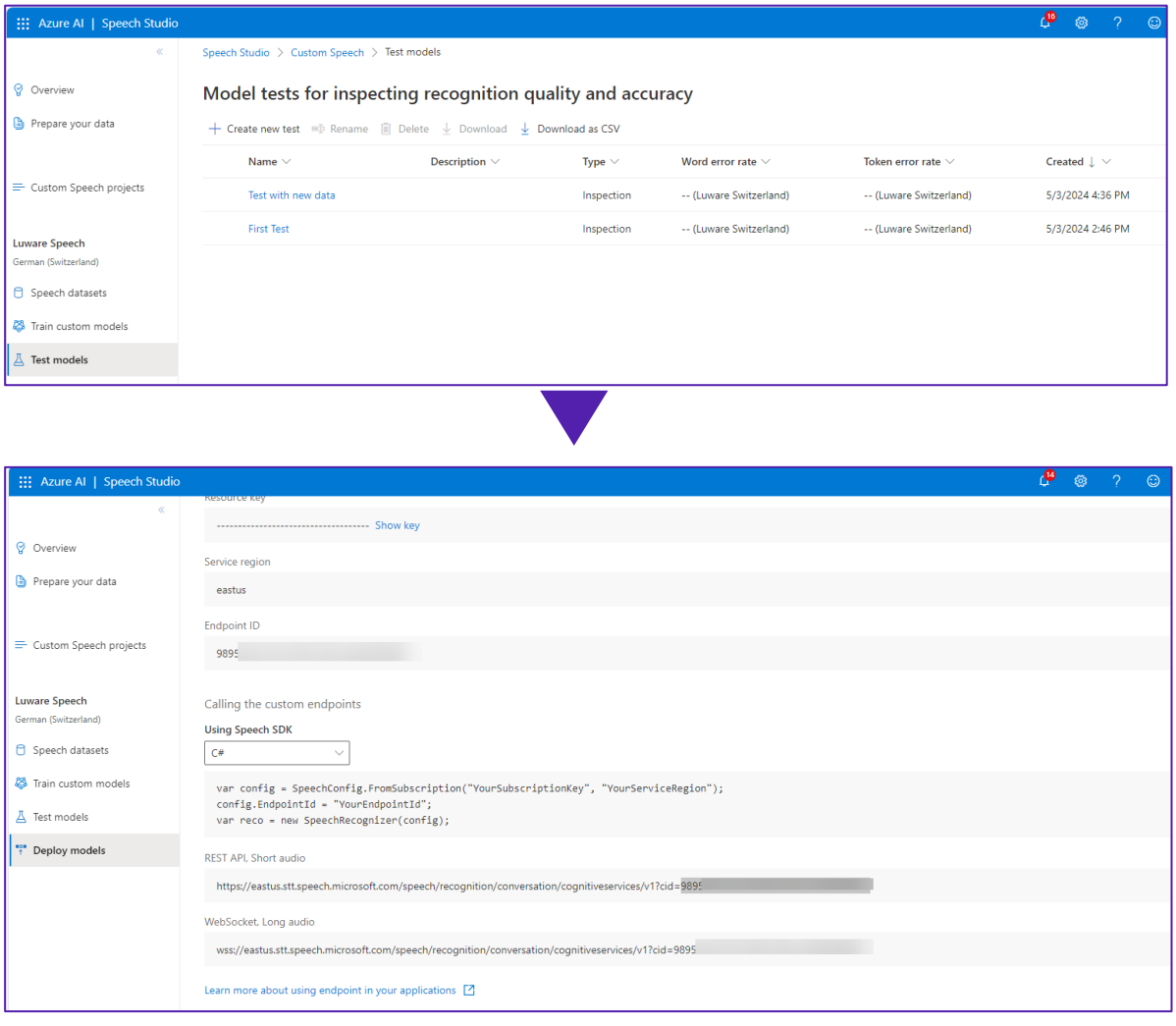

Create new test
Choose testing type
Choose testing data
Choose models to evaluate
Name and description
Review and evaluate
Choose models to evaluate
Select up to two model to evaluate and compare accuracy.
Model 1 *
Luware Switzerland
Model 2
[None]
Back
Next
Cancel


Create new test
Choose testing type
Choose testing data
Choose models to evaluate
Name and description
Review and evaluate
Name and description
Create a test to evaluate the quality of Microsoft's speech to text baseline model, you can compare two models side by side to evaluate accuracy.
Name *
Luware final Speech test
Description
Enter test description (max 500 characters)
Back
Next
Cancel



Now you can switch to the Tab **deploy models** and deploy the model. As a result, you will get the new model.

To enable the new model in Luware Nimubs, you need to **copy the Key** as well as the **region**, the model is created in.



Add the Speech Recognizer to Luware Nimbus

You have successfully created and deployed your custom speech model.

To use it in Luware Nimbus, follow these steps:

- Go to the Azure portal and select your Custom Speech resource.
- On the left pane, click on Keys and Endpoint. You will see two keys and an endpoint URL. Copy one of the keys and the region name from the URL. For example, if the URL is [URL], then the region name is westeurope.
- Open the Luware Nimbus Admin Portal and navigate to Settings > Virtual Assistants.
- Click on the + button to create a new Speech Recognizer. Give it a name and select the organizational unit that will use it.
- Paste the key and the region name that you copied from Azure into the corresponding fields.
- Click on Save to create the Speech Recognizer.
- Now go to Service Settings > Virtual Assistant and check the boxes for Captions and/or Transcription, depending on your preference.
- Select the Speech Recognizer that you just created from the drop-down menu.
- Click on Save to apply the changes.

Congratulations! You've successfully integrated your custom speech model into Luware Nimbus.

To test its performance:

- Initiate a call with a virtual assistant.
- Observe the generated captions and/or transcriptions.

For ongoing monitoring:

- Visit the Azure portal to track the model's quality and accuracy metrics.

By following these steps, you can assess the effectiveness of your custom speech model in real-world scenarios and make necessary adjustments.

solutions@luware.com
+41 58 404 28 00
www.luware.com