

Going Beyond the Basics



Jared Rhodes

INDEPENDENT CONSULTANT

@qimata www.jaredrhodes.com



Overview



Batch Transcription

Customize Acoustic Models

Customize Language Models

Customize Pronunciation

Accuracy Tests

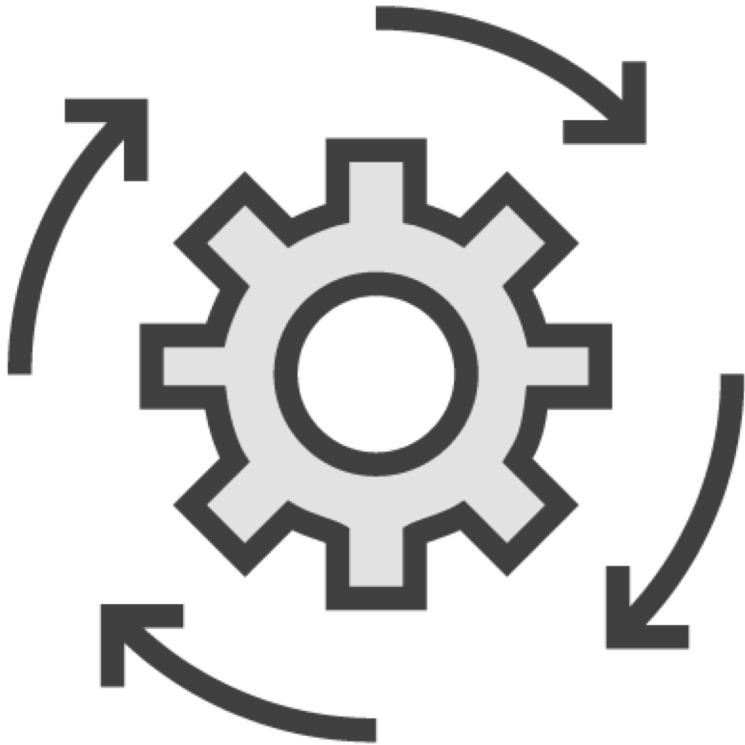
Explore the Devices SDK



Batch Transcription



Batch Processing Use Cases



Call Center Call Processing

Audio Book Transcription

Voice Chat Text Logging

Many more

REST API Batch Features



Batch Processing Requests



Query Status



Downloading Transcriptions



Supported Audio

WAV

MP3

OGG



Configuration Fields



recordingsUrl

models

locale

name

description

properties

Properties Fields

**Profanity
FilterMode**

PunctuationMode

**AddWordLevel
Timestamps**



Configuration Fields

```
{  
  "recordingsUrl": "https://...",  
  "models": ["optionalId", "optionalId"],  
  "locale": "en-US",  
  "name": "Pluralsight Transcribe",  
  "description": "Demo",  
  "properties": { ... }  
}
```

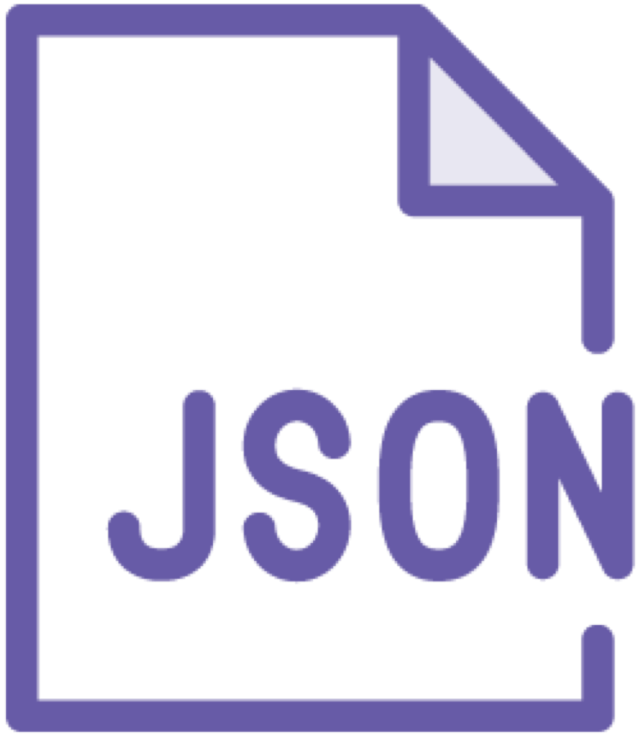


Properties

```
"properties":  
{  
  "ProfanityFilterMode": "Masked",  
  "PunctuationMode": "DictatedAndAutomatic",  
  "AddWordLevelTimestamps" : "True"  
}
```



Status Fields



Id / Name / Description

CreatedDateTime / LastActionDateTime

Status / StatusMessage

Locale

RecordingsUrl / ResultsUrl

Customized Acoustic Model



Scenarios



Customized for
Environment



Customized for
Speaker



Customized for
Equipment

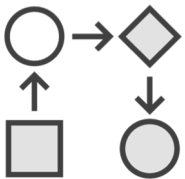
Steps for Creating a Custom Acoustic Model



Prepare the Data



Import the Acoustic Dataset



Create the Model



Acoustic Data



WAV (RIFF) Audio Format

8 KHz or 16 KHz Sampling Rate

Single-Channel (Mono) Audio

10-12 seconds Duration

Single Utterance

Unique File Names

Single folder no subdirectories

Long silence with background noise

speech01.wav speech recognition is awesome

speech02.wav the quick brown fox jumped all over the place

speech03.wav the lazy dog was not amused

Transcripts

UTF-8



Normalizing Text



ASCII Only



Lowercase



Remove Punctuation



Expand Numbers



Normalizing Text

**Remove
Abbreviations**

**Write Out
Numeric Strings**

**Transcribe as
Announced**

**Leave Pronounced
as Words**

**Change
Pronounced as
Letter**



Accuracy Tests



Reference
Transcription



Recognition



Word Error Rate



Demo



Generate the dataset

Upload the dataset

Create the custom model

Use the custom model



Customized Language Model





The Patient Has Had Urticaria for the Past Week

The Patient Had a Well-healed Herniorrhaphy Scar

I Failed to Account for the Appropriate Eigen Vectors

The Patient Went to the Ophthalmologist for Phthisis

Solar Spectroscopy Based on the Hypothesis of Elemental Dissociation



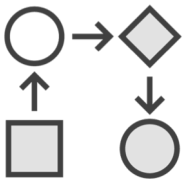
Steps for Creating a Custom Language Model



Prepare the Data



Import the Acoustic Dataset



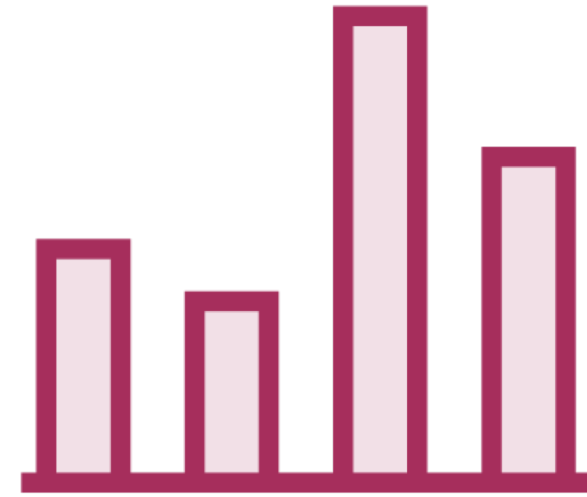
Create the Model



Data Preparation



Do not have to be grammatically correct



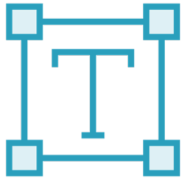
Increase occurrences to increase weight



Data Tips



Avoid Repeating Characters



No Special Characters



No URIs



Demo



Create custom language data

Upload the custom language data

Create the custom language model

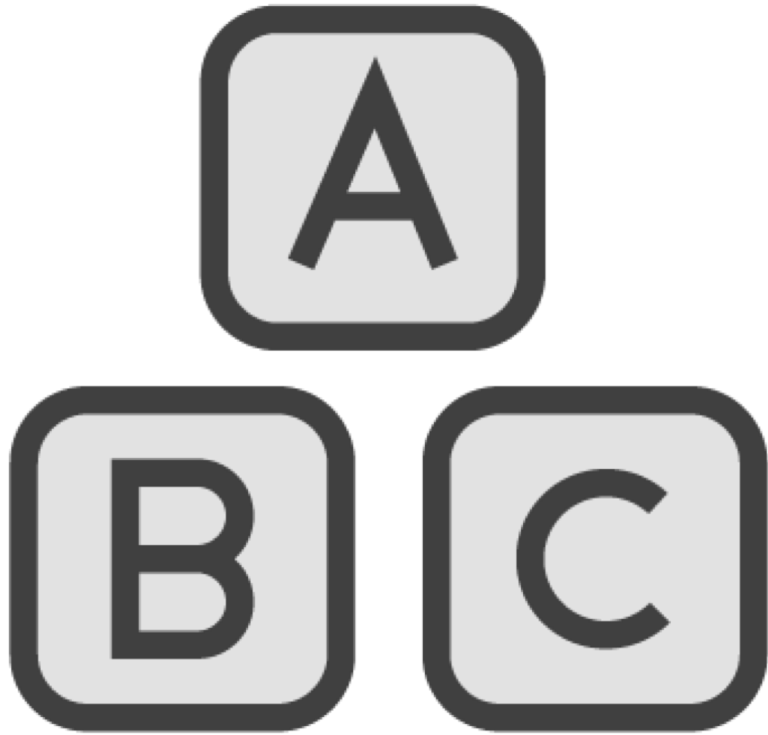
Use the custom language model



Customized Pronunciation



Customized Pronunciation



L8R

IEEE

DMARDs

DAL



Display form <Tab> Spoken form <Newline>

Pronunciation File



Pronunciation File Requirements

Display Form

Custom Word

A Term

An Acronym

Compound Words that Combine
Existing Words

Alternative Pronunciations for Common
Words

Spoken Form

Lowercase

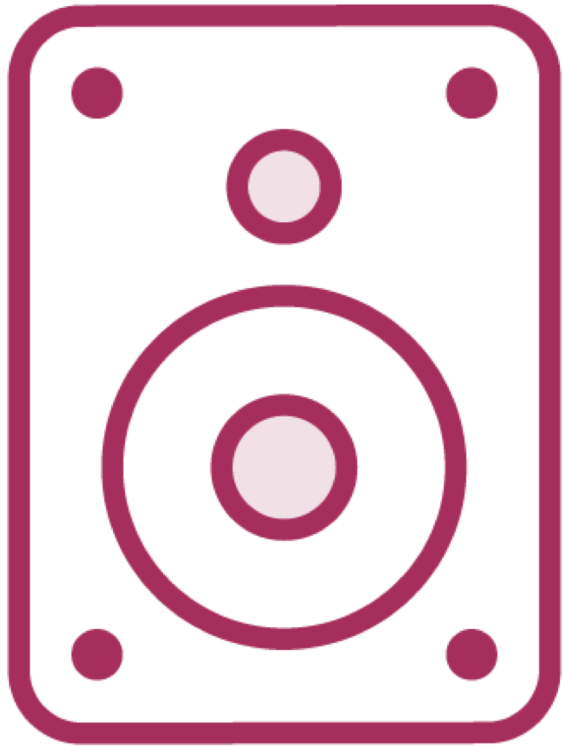
No Tab Character

Forbidden Characters Allowed



Speech Devices SDK





Drive-thru Ordering Systems

In-store or In-home Assistants

Smart Speakers



Speech Devices SDK



**Purpose-built
Microphone
Array**



Rapidly Test



**Easily
Integrate Your
Device**



**Exceptional
User Experience**



Review



Batch Transcription

Customize Acoustic Models

Customize Language Models

Customize Pronunciation

Accuracy Tests

Explore the Devices SDK

