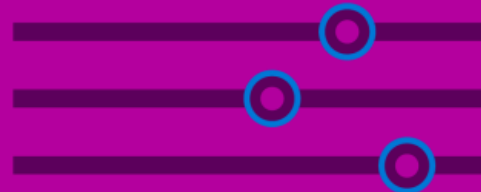


#FusionAIDay



SPEECH SERVICES

Jamie Dalton, @daltskin



**MICROSOFT SPEECH
HAS HIGHER
ACCURACY THAN
HUMANS**



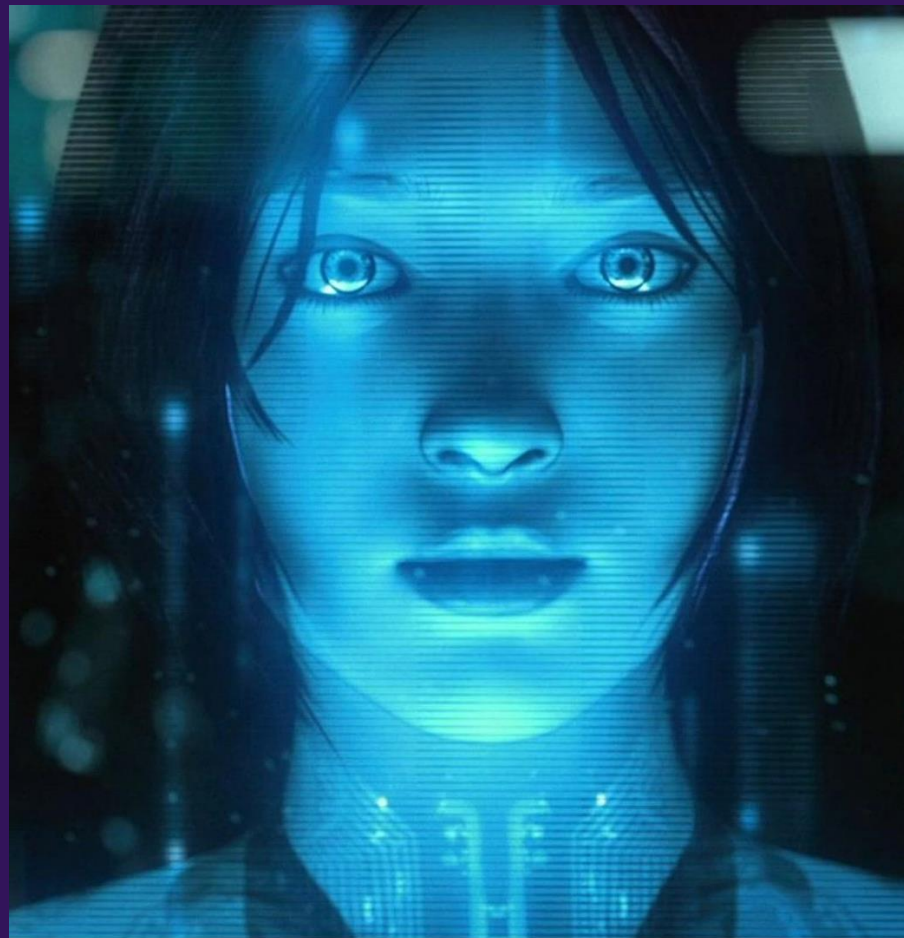
#FusionAIDay



THE SWITCHBOARD TEST



- ▶ Industry standard speech recognition test
- ▶ Word error rate of 5.1%
- ▶ Less than human transcriptionists
- ▶ Lowest ever recorded
- ▶ aka.ms/SpeechSwitchboardTest





Vision



Language



Speech



Search



Knowledge

Computer vision
Face
Emotion
Content Moderator
Video
Video Indexer

Custom

Vision Service

Text analytics
Spell check
Web language model
Linguistic analysis
Translator Text

Custom

Language
Understanding

Translator Speech
Speaker recognition
Bing Speech

Custom

Speech Service

Web search
Image search
Video search
News search
Autosuggest
Entity search

Custom

Search

Academic knowledge
Entity linking service
Knowledge exploration
Recommendations
QnA maker

Custom

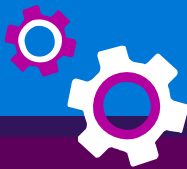
Decision Service

+6 Cognitive Services Labs APIs

COGNITIVE SERVICES



BING SPEECH



- ▶ Speech Recognition
- ▶ Text to Speech
- ▶ Real-time, partial results



TRANSLATOR SPEECH



- ▶ Transcribe and translate
- ▶ 60 languages
- ▶ Built on Microsoft Translator Hub



SPEAKER RECOGNITION



- ▶ Speaker identification
- ▶ Users enrol to the service
- ▶ Audio verification



DEMO

SPEECH DEMO



#FusionAIDay



CUSTOM SPEECH



- ▶ Built on Bing Speech
- ▶ Custom Language Models
- ▶ Custom Acoustic Environments



#FusionAIDay



DEMO

MCDONALDS DRIVE THROUGH



#FusionAIDay



DEMO

STARSHIP COMMANDER

<https://customers.microsoft.com/en-us/story/human-interact-cognitive-services>

#FusionAIDay



IN SUMMARY



- ▶ We have the best speech recognition
- ▶ Several Cognitive services available
- ▶ Custom speech for specific applications



CONTACT ME


JAMIE DALTON

blogs.msdn.microsoft.com/jamiedalton

@daltskin

Jamie.Dalton@microsoft.com





| | | |
|-------|-------------------------------------|----------------|
| 10:15 | The Microsoft AI Platform | Martin Kearn |
| 10:45 | Computer Vision Services | Frances Tibble |
| 11:30 | Break | |
| 11:45 | Knowledge Services | Martin Kearn |
| 12:30 | Speech Services | Jamie Dalton |
| 13:00 | Lunch | |
| 14:00 | Language Services | Frances Tibble |
| 14:45 | Bots and Conversational Apps | Jamie Dalton |
| 15:30 | Break | |
| 15:45 | Machine Learning | Mason Cusack |
| 16:45 | Wrap-up | Martin Kearn |