global AI
back together

Para obtener las presentaciones y más recursos del evento.

https://aka.ms/AAe3fz2

Microsoft Speech

MVP Nicolás Nakasone

@nicolasnakasone

BI EXPERT

"*Durante muchos años los humanos hemos aprendido los interfaces de la tecnología y ya es hora de que la tecnología aprenda los interfaces de los humanos.*"

Un poco de conceptos: Human Parity

Human Parity (Paridad Humana):

Es la unidad de medición de la calidad de los modelos de Inteligencia Artificial.

Utilizamos la Paridad Humana como tasa de error, por lo que si un servicio cognitivo se equivoca menos que un humano en uno de los problemas, decimos que ese servicio ya ha superado al ser humano.

98.64%

Neural TTS MOS similarity
to human recording
(August 2018)

**Close-to-human-parity
Speech Synthesis**

*TTS Text To Speech
*MOS: Mean Opinion Score

Cognitive Speech Services general availability [GA]

Speech to Text

Swiftly convert audio to text for natural responsiveness

Robust baseline models in **30** locales

Available via REST and Websockets

Fully tested in a number of Microsoft products (Office, Windows, Cortana, Xbox)

You can further customize speech recognition models for your scenarios

Text to Speech

Enable your applications to talk

A choice of over 80 voices in **49** locales

Available via REST

Customizable for 3 locales

Speech Translation

Easily integrate real-time speech translation to your app

Speech-to-speech translation capabilities with models based on neural machine translation (NMT) technologies

Available in 10+ locales

Customizable

Cognitive Speech Service sample scenarios

Grab your key from Azure Portal and start creating speech functionality your apps can consume effectively.

Transcription

Dictation

Dictate using automatic or manual punctuation

Dictate highly domain specific material using custom speech models

Captioning

Super fast highly accurate caption

Ideal for live events such as presentations, panel discussions, sports events.

Languages 30 + locales

Available Now

Call Centers

High Quality Transcription

Baseline or Custom Models

Batch Transcription API

Submit Requests

Query Status

Download Results

Support for:

Various audio formats

Lengthy audio files

Channel splitting

Conversational Voice Agents

Providing real-time speech capabilities using a plethora of voices

Create a Brand Voice

Create your own custom voice brand in any language

Benefit from latest in Speech synthesis

Languages

Translate speech

Large number of locales

"higher transcription quality, more insightful analytics"

Automotive

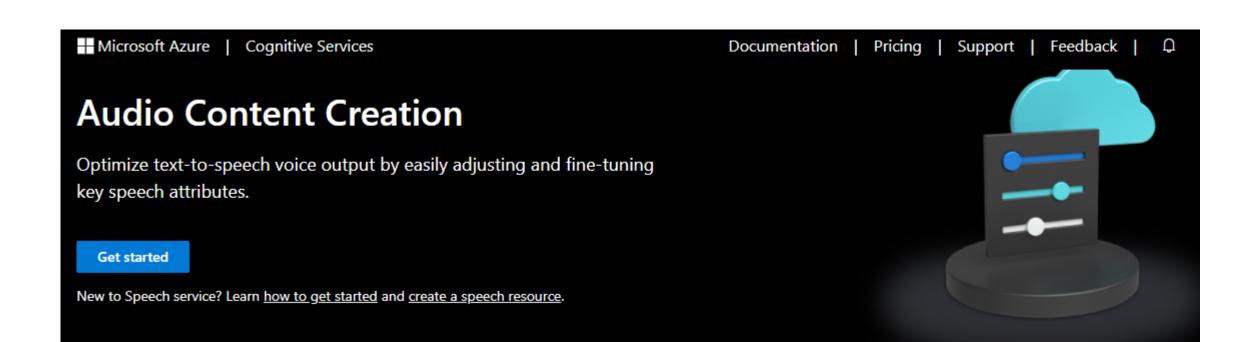Car Assistant, weather, GPS, road conditions etc.

Far Field

Traffic noise

Natural-sounding voices

Speech Recognition

Custom Voice

Microsoft Cognitive Services-Speech-AI

Microsoft named a Leader in IDC MarketScape for WW General-Purpose Conversational AI Platforms

Gracias…!!!   Thanks…!!!