



SPEECH CHARACTERISTICS

**Enhancing Speech Dialogue with
Emotional Intelligence**

Brian Garr, CRO
Brian.Garr@govivace.com

From Mid 90's to today – How far did we come?

- I can speak my phone number and the credit card number to an IVR (and expect it to get that right)
- Speech transcription tending towards “Human Level” accuracy. Most of the errors are around “uh”/“um” “a”/“the” kind of confusions.
- Dictation and Live closed captioning are real, if somewhat error prone.

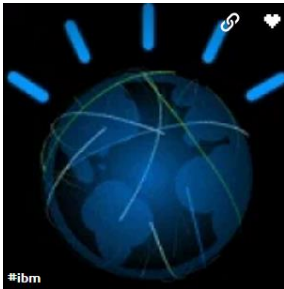
Now, speech is part of a dialog



Would you like
to order a pizza?

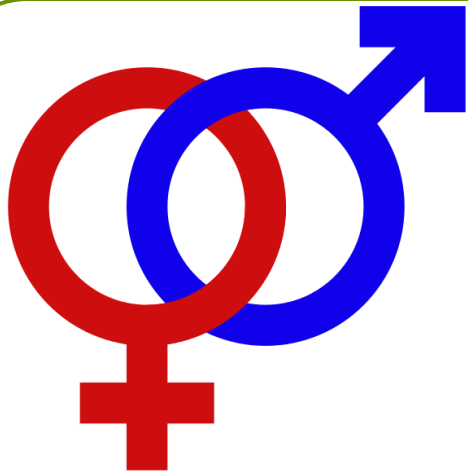


Book me a flight
from MIA to ROM



Diagnose
the MRI

What can help us in a dialog?

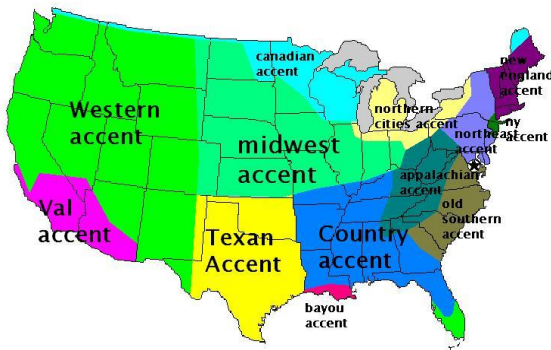


- Address with the proper salutation
- Offer gender appropriate incentives
- Use in disambiguation algorithms



- Block inappropriate transactions
- Offer age appropriate incentives
- Use in disambiguation algorithms

What can help us in a dialog?

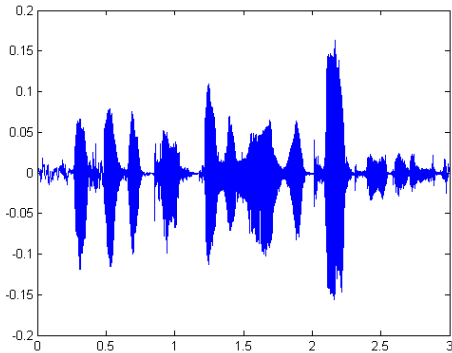


- Use colloquial jargon
 - Pop vs Soda
 - Ridy-bob vs seesaw
 - Scaper vs critter



- Transfer call to supervisor
- Offer incentives
- Use in disambiguation algorithms

Smarter systems for today



SPEAKER VERIFICATION
SPEAKER IDENTIFICATION



Age
Accent

CHARACTERISTICS



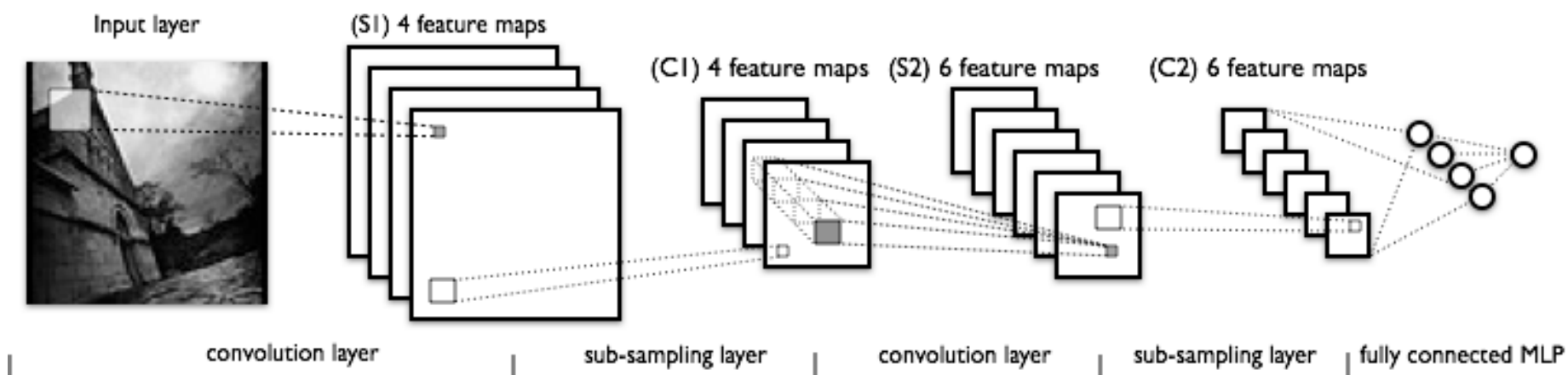
BETTER OUTCOMES?

Emotion Detection

- Paul Ekman – basic emotions
 - Happy
 - Sad
 - Angry
 - Disgust
 - Fear
 - Surprise
- These emotions are universal across cultures. Facial expressions change similarly as these emotions occur.

Convolutional Neural Networks

- Sparse Connectivity – local receptive field
- Shared Weights – same filter applied across the signal
- Pooling for dimension reduction



Let's analyze some examples

- We will play an audio example
- Audience will need to guess the following
 - ▣ Gender
 - ▣ Language -English, Hindi, Thai, German, French, Russian, Arabic, ...
 - ▣ Age
 - ▣ Accent (only for english) -Eastern, NY, CA, TX, Canada, Germany, China, India, France, Russian
 - ▣ Level of Arousal (or Neutrality)
- Compete with computer estimates

The guessing game



The guessing game



	M/F	Lang	Acc	Acc	Acc	Acc	Age	EState
				2	3	4		
Real	M	En	TX				25	NA
Est	M	En	TX .78	IND .6	East 0.59	CDN 0.49	30	Neutral 0.74



The guessing game



	M/F	Lang	Acc	Acc 2	Acc 3	Acc 4	Age	EState
Real	M	En	TX				25	NA
Est	M	En	TX 78	IND 6	East 0.50	CDN 0.40	30	Neutral 0.74
	M/F	Lang	Acc	Acc 2	Acc 3	Acc 4	Age	EState
Real	M	En	can				47	NA
Est	M	En	can .68	china .58	CA/ 0.56	TX/ 0.54	46	Dominant 0.739



The guessing game



	M/F	Lang	Acc	Acc 2	Acc 3	Acc 4	Age	EState
Real	M	En	TX				25	NA
Est	M	En	TX 0.78	IND 0.6	East 0.50	CDN 0.40	30	Neutral 0.74

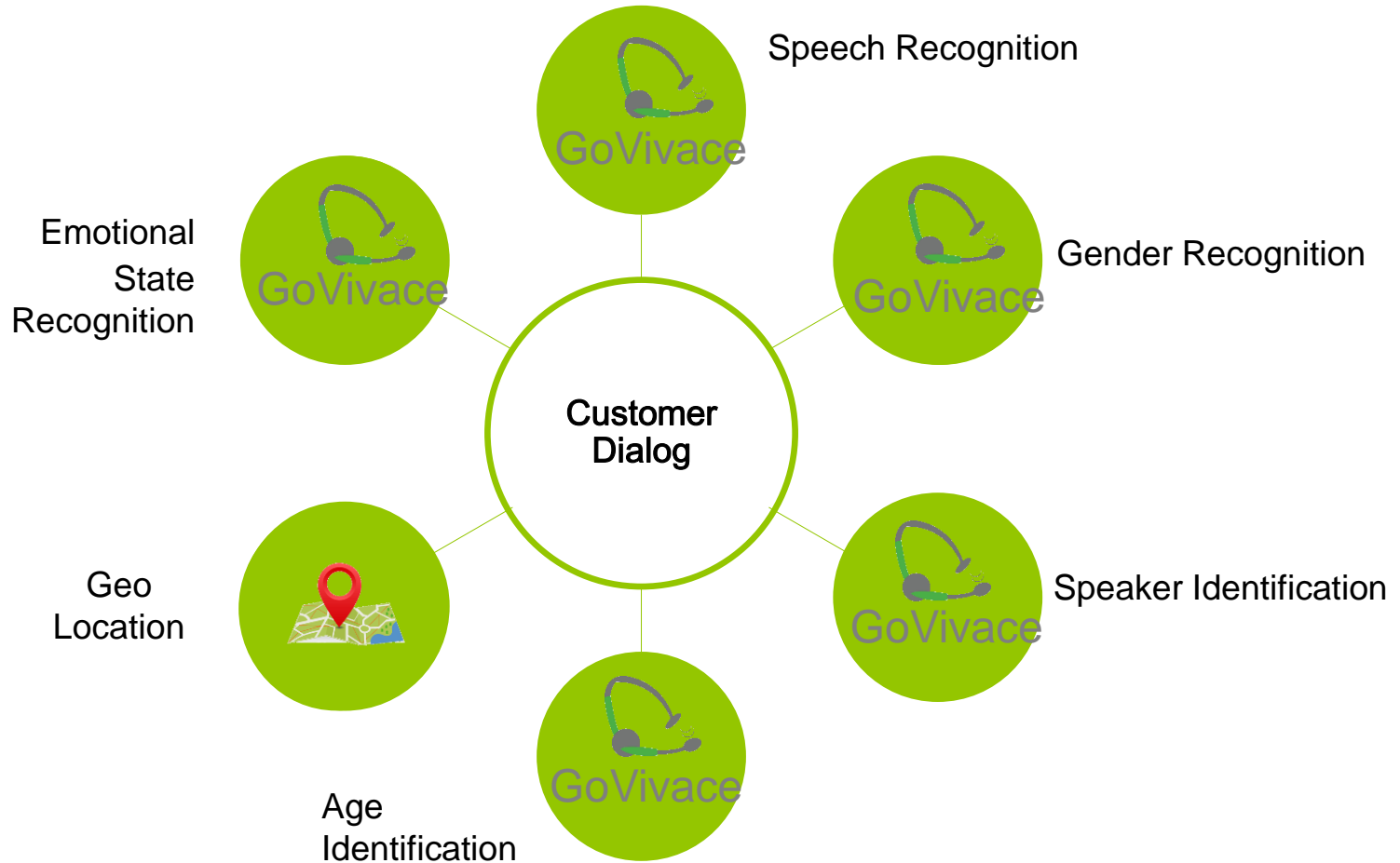


	M/F	Lang	Acc	Acc 2	Acc 3	Acc 4	Age	EState
Real	M	En	can				47	NA
Est	M	En	can .68	chin a	CA/ 0.56	TX/ 0.54	46	Dominant



	M/F	Lang	Acc	Acc 2	Acc 3	Acc 4	Age	EState
Real	F	En	Chi n				30	NA
Est	F	En	Chi n .71	RU/ .70	TX/0 .58	NY/ 0.51	29	Neutral 0.80

The Complete Conversation



Questions?

