

# Building a Voice Assistant for Enterprise



Manju Vijayakumar  
Lead Software Engineer, Salesforce

@vmanju  
QConSF, Nov 2018



# Agenda



- Why Voice?
- Demo of Einstein Voice Assistant
- Conversational AI
  - Ecosystem
  - Natural Language Understanding (NLU)
- Challenges
- Future
  - Considerations
  - What's next for NLP and AI

# Voice Recognition - A Story in 3 pictures



VOICE RECOGNITION IS GOING  
TO BE PRETTY MUCH STANDARD

A photograph of a man and a woman standing in a hallway. The man, on the left, is wearing a dark suit and tie, and has his hands clasped in front of him. The woman, on the right, is wearing a dark top. They are both looking towards each other. In the background, there is a window with horizontal blinds, through which some lights are visible. The walls are a light beige color.

ON ANY COMPUTER  
YOU BUY.



A man in a dark suit, white shirt, and red plaid tie is looking up at a woman with a slight smile. The woman has short dark hair and is wearing a dark turtleneck sweater. They appear to be in an indoor setting with warm lighting.

SO YOU CAN BE LIKE,  
"WASH MY CAR," "CLEAN MY ROOM."

# Computing is Evolving

From programmatic to natural interactions



Voice



Touch



Command Line



Point & Click





Deliver an intelligent assistant that leverages Voice and NLU capabilities to understand, and support users in accomplishing their goals

Pilot

salesforce

# EINSTEIN VOICE DEMO





Meet Amy,  
a busy salesperson

Amy needs to  
update Salesforce





9:41 AM

100%



## Einstein Analysis

Met with Chris Hopkins from Acme Corporation we had a great meeting and closed a deal for purchasing merchandise

- Follow up with Chris next week
- Change the deal amount to \$1 million
- Set the closing date to November 15

## CHOOSE YOUR CONTACT



## Chris Hopkins

Name: Chris Hopkins • Account Name:  
Acme • Phone: (212) 555-5555 • Email: ch...



Do not relate to any contact



[Search Manually](#)

Previous

Next



# How did Voice Assistant help Amy?



Unstructured data -> Structured data

- Productive
- No system expertise



Accuracy & timeliness of data capture



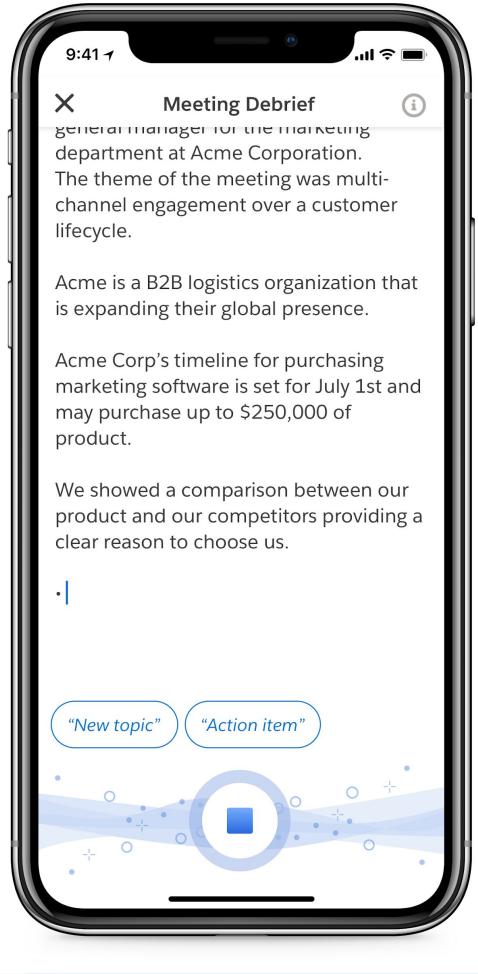
Visible to the team

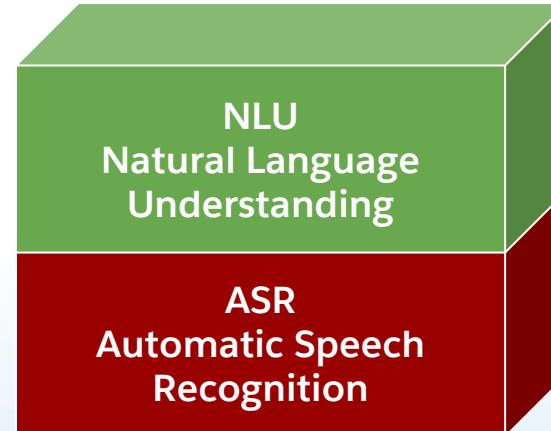
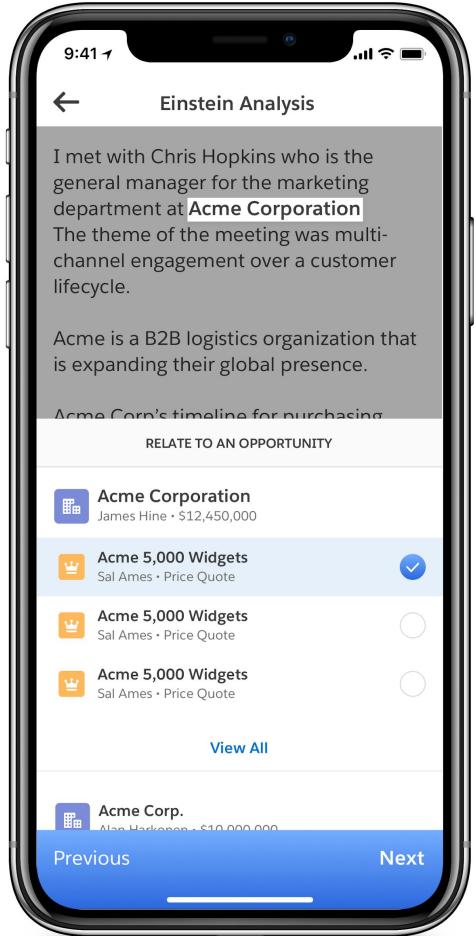


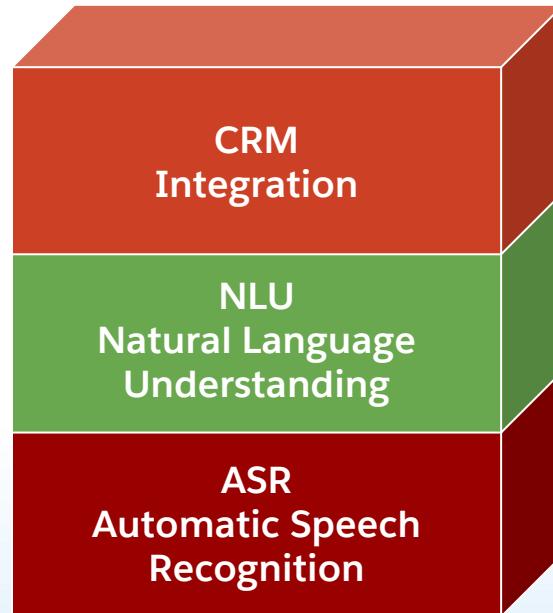
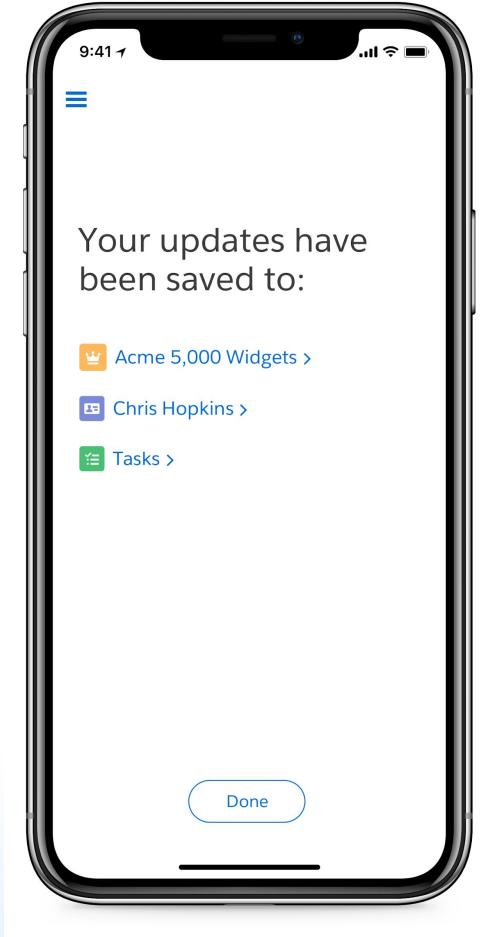


# Building Blocks of Voice Assistant











# Conversational AI Ecosystem



# Conversational AI Ecosystem



Einstein Platform

Automatic Speech  
Recognition  
Models

Natural Language  
Understanding  
Models

Salesforce  
CRM Metadata

# Conversational AI Ecosystem



## Conversational API

Named Entity  
Recognition

Entity  
Resolution

Context  
Management

Text  
Classification

Slot Filling

## Einstein Platform

Automatic Speech  
Recognition  
Models

Natural Language  
Understanding  
Models

Salesforce  
CRM Metadata

# Conversational AI Ecosystem



Einstein Voice  
Assistant

Einstein Voice Bots

Smart Speakers\*

Voice Navigation\*

Conversational API

Named Entity  
Recognition

Entity  
Resolution

Context  
Management

Text  
Classification

Slot Filling

Einstein Platform

Automatic Speech  
Recognition  
Models

Natural Language  
Understanding  
Models

Salesforce  
CRM Metadata

# Conversational AI Service



Conversational API

Named Entity  
Recognition

Entity  
Resolution

Context  
Management

Text  
Classification

Slot Filling

# Conversational AI Service



Conversational API

Named Entity  
Recognition

Entity  
Resolution

Context  
Management

Text  
Classification

Slot Filling

# Named Entity Recognition (NER)



The committee of 30 government and university scientists and engineers, led by McCleese, was asked to recommend to the space agency by the end of this month a rationale and strategy for precursor flights and the sample-return missions.



The	'O'
committee	'O'
of	'O'
...	'O'
McCleese	'PER'
the	'DATE'
end	'DATE'
of	'DATE'
this	'DATE'
month	'DATE'

\*CoNLL format

# Named Entity Recognition (NER)



The committee of 30 government and university scientists and engineers, led by McCleese, was asked to recommend to the space agency by the end of this month a rationale and strategy for precursor flights and the sample-return missions.



**NER7 model recognizes 7 entities:**  
*Person, Organization, Location,  
Date, Time, Money, Percentage*

The	'O'
committee	'O'
of	'O'
...	'O'
McCleese	'PER'
the	'DATE'
end	'DATE'
of	'DATE'
this	'DATE'
month	'DATE'

\*CoNLL format

# What are the entities in the text ?



I met with **Chris Hopkins** who is the general manager for the marketing department at **Acme Corporation**.

PERSON

The theme of the meeting was multi-channel engagement over a customer lifecycle.

ORGANIZATION

Acme Corp's timeline for purchasing marketing software is set for **July 1st** and may purchase up to **\$250,000** of product.

DATE

MONEY

Follow up call with Chris in **two weeks**

DATE

(*two weeks* is normalized to 2018/07/15)

# Conversational AI Service



Conversational API

Named Entity  
Recognition

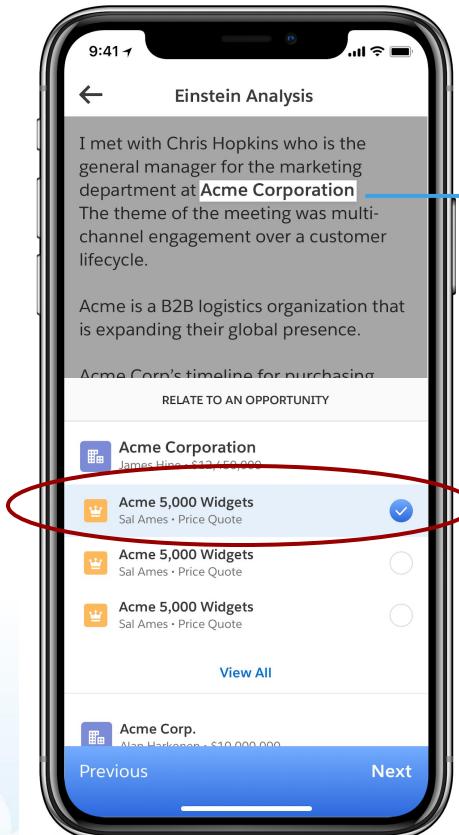
Entity  
Resolution

Context  
Management

Text  
Classification

Slot Filling

# Entity Resolution - Is this entity in my CRM ?



Salesforce  
CRM DB



Records matched  
for 'Acme'

Send records to user  
to disambiguate

# Conversational AI Service



Conversational API

Named Entity  
Recognition

Entity  
Resolution

**Context  
Management**

Text  
Classification

Slot Filling

# Context Management - What data do we have so far ?



A screenshot of a mobile application interface showing a list of items. The first item is "Acme Corporation" with a building icon, followed by three items labeled "Acme 5,000 Widgets" each with a crown icon. The first "Acme 5,000 Widgets" item has a checked checkbox icon to its right, while the others have empty circles.



```
{  
  "context": {  
    "Organization": {  
      "id": "001XXXX",  
      "name": "Acme Corp"  
    },  
    ...  
  }  
}
```



I met with **Chris Hopkins** who is the general manager for the marketing department at Acme Corporation.



A screenshot of a mobile application interface showing a list of two employees. The first employee is "Chris Hopkins" with a building icon, and the second is "Sam Quintana" with a building icon. Both have checked checkbox icons to their right.



Do we have organization in the context?

# Conversational AI Service



Conversational API

Named Entity  
Recognition

Entity  
Resolution

Context  
Management

**Text  
Classification**

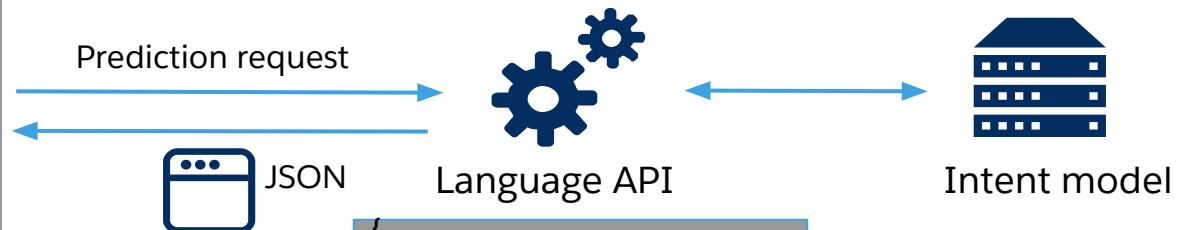
Slot Filling

# Text Classification - What are the intents ?



Acme Corp's timeline for purchasing  
Marketing software is set for July 1st and may purchase up to \$250K of product

Follow up call with Chris in two weeks



```
{  
  "probabilities": [  
    {  
      "label": "CREATE",  
      "probability": 0.9904295  
    },  
    {  
      "label": "UPDATE",  
      "probability": 0.009345241  
    },  
    ...  
  ]  
}
```

# Conversational AI Service



Conversational API

Named Entity  
Recognition

Entity  
Resolution

State  
Management

Text  
Classification

**Slot Filling**

# Slot Filling - What are the slots for each action item ?



A screenshot of a Salesforce interface. On the left, there's a sidebar with a crown icon and the word "UPDATES". Below it, there are two sections: "Close Date" with the value "July 1st, 2018 from Oct 12, 2019" and "Amount" with the value "\$250,000 from \$200,000". On the right, there's a "TASKS" section with a green checklist icon. It contains "Due Date: 10/7/18" and the task description "Follow up call with Chris in 2 weeks". Each entry has a small edit pencil icon to its right.

Fill in the **date** and **money** slots for Update action

Fill in the **date** slot and **person** slot for Create Task action.

Here, date is **normalized**:  
*In 2 weeks => 10/7/18*



# Challenges



# Data challenges



How do you make it  
work for every  
customer schema ?



Heterogenous  
database

AccountID	Name	Phone

AccountID	Name	Phone	Bank Account

- Customers can define custom schemas
- Schemas are not consistent



# Data challenges



The screenshot shows a mobile-style interface for 'Einstein Analysis'. At the top, it says 'SFDC' with signal strength, '1:11 PM', and '100% battery'. The main heading is 'Einstein Analysis'. Below it, a message says 'We think you mentioned an account...'. The text content is: 'I met with Chris Hopkins who is the general manager for the marketing department at Acme Corporation Inc. The theme of the meeting was multi-channel engagement over a customer lifecycle.' A blue callout bubble originates from the word 'Acme' in this text. Below this, there's a section titled 'RELATE TO AN OPPORTUNITY' with a list of results:

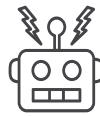
- Acme Corporation (James Hine - \$12,450,000)
- Acme 5,000 Widgets (Sal Ames - Price Quote)
- Acme 5,000 Widgets (Sal Ames - Price Quote)
- Acme 5,000 Widgets (Sal Ames - Price Quote)

At the bottom, there are buttons for 'View All', 'Acme Corporation', 'Previous', and 'Next'.

Which Acme Corp.  
did you mean ?

- Lots of duplicates
- Identify the most relevant 'Acme'
- Affects user experience

Inconsistent  
data



# Automatic Speech Recognition is not perfect



DOMAIN SPECIFIC  
JARGON

AUDIO  
ENVIRONMENT

ACCENTS &  
LINGUISTIC PROFILES





# Named Entity Recognition is not perfect



Named Entity Recognition is easy for  
humans but hard for machines



# Named Entity Recognition is not perfect



Today, **JP Morgan** and I spoke about...



Is **JP Morgan** a company or a person ?

..the **san juan** center is led by a team of scientists..



Cannot identify **san juan** as a location due to case sensitivity

..**Man joy** and I met today at Starbucks to discuss..



“Manju” misspelled as “Man Joy”. Misspelled pronouns are hard to catch

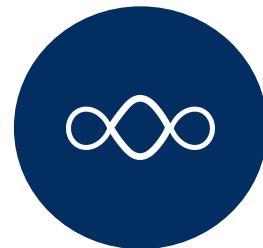


# Future Considerations



## Optimized Models

- Configurable
- Normalized



## Feedback

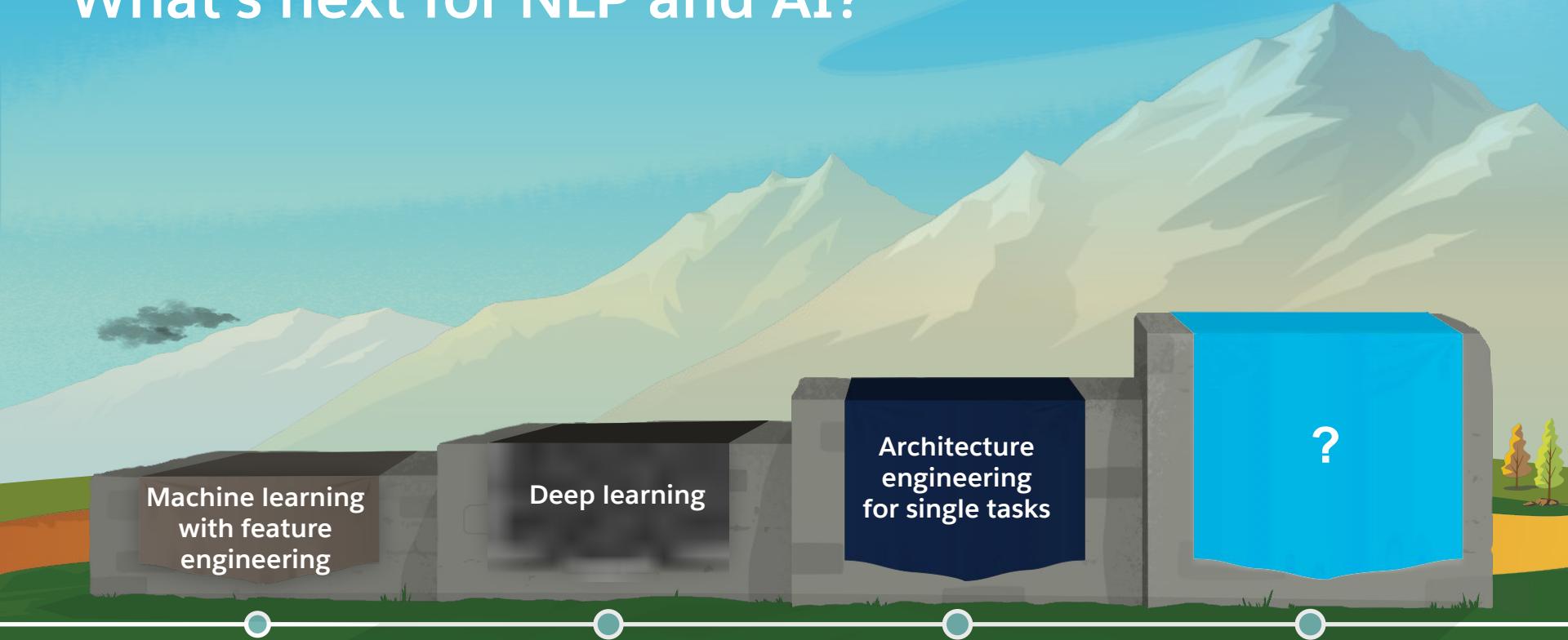
- Capture Feedback
- Retrain Models



## Voice

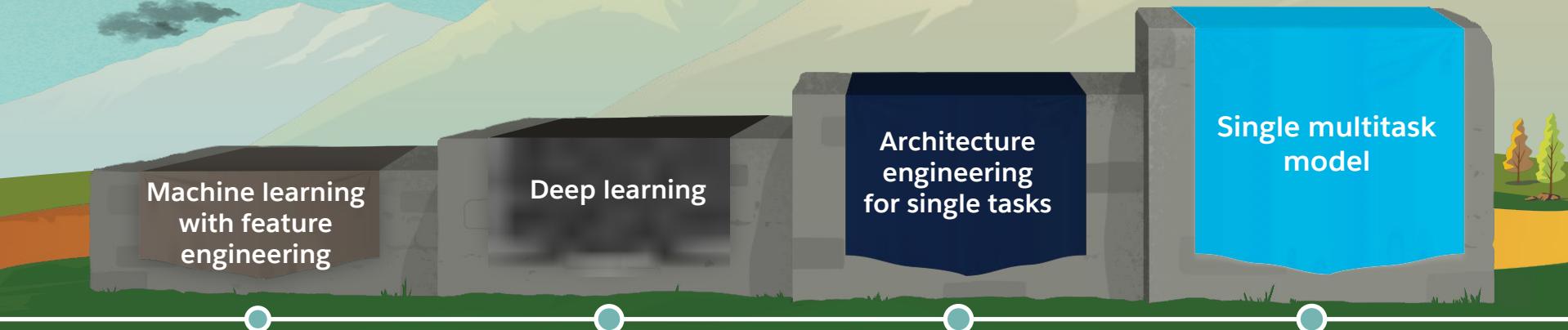
- Guided user experience
- Multi channel

# What's next for NLP and AI?



# What's next for NLP and AI?

To learn more: [decaNLP.com](https://decaNLP.com)



# Voice Recognition - A Story in 3 pictures



VOICE RECOGNITION IS GOING  
TO BE PRETTY MUCH STANDARD



ON ANY COMPUTER  
YOU BUY.

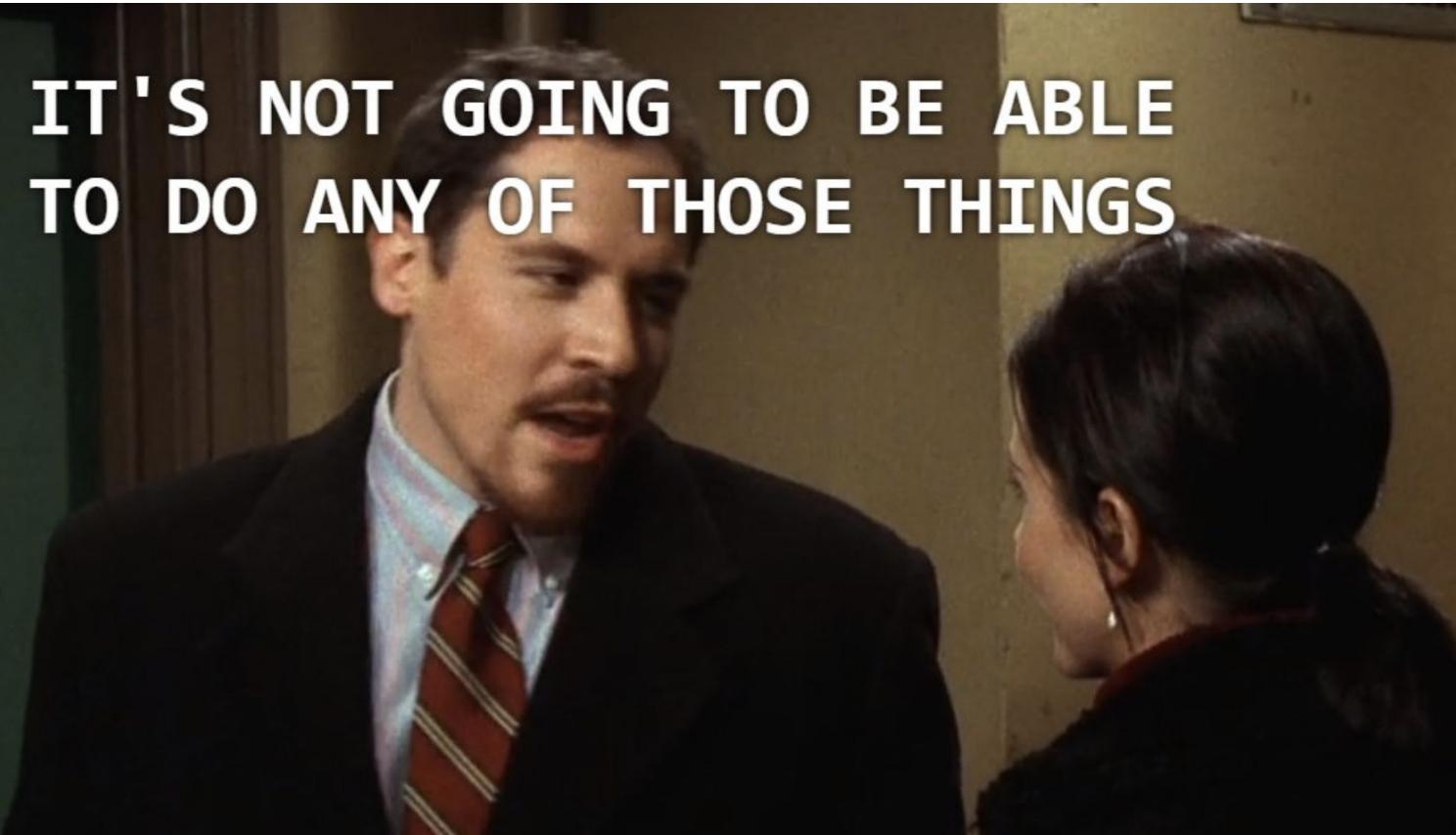


SO YOU CAN BE LIKE,  
"WASH MY CAR," "CLEAN MY ROOM."

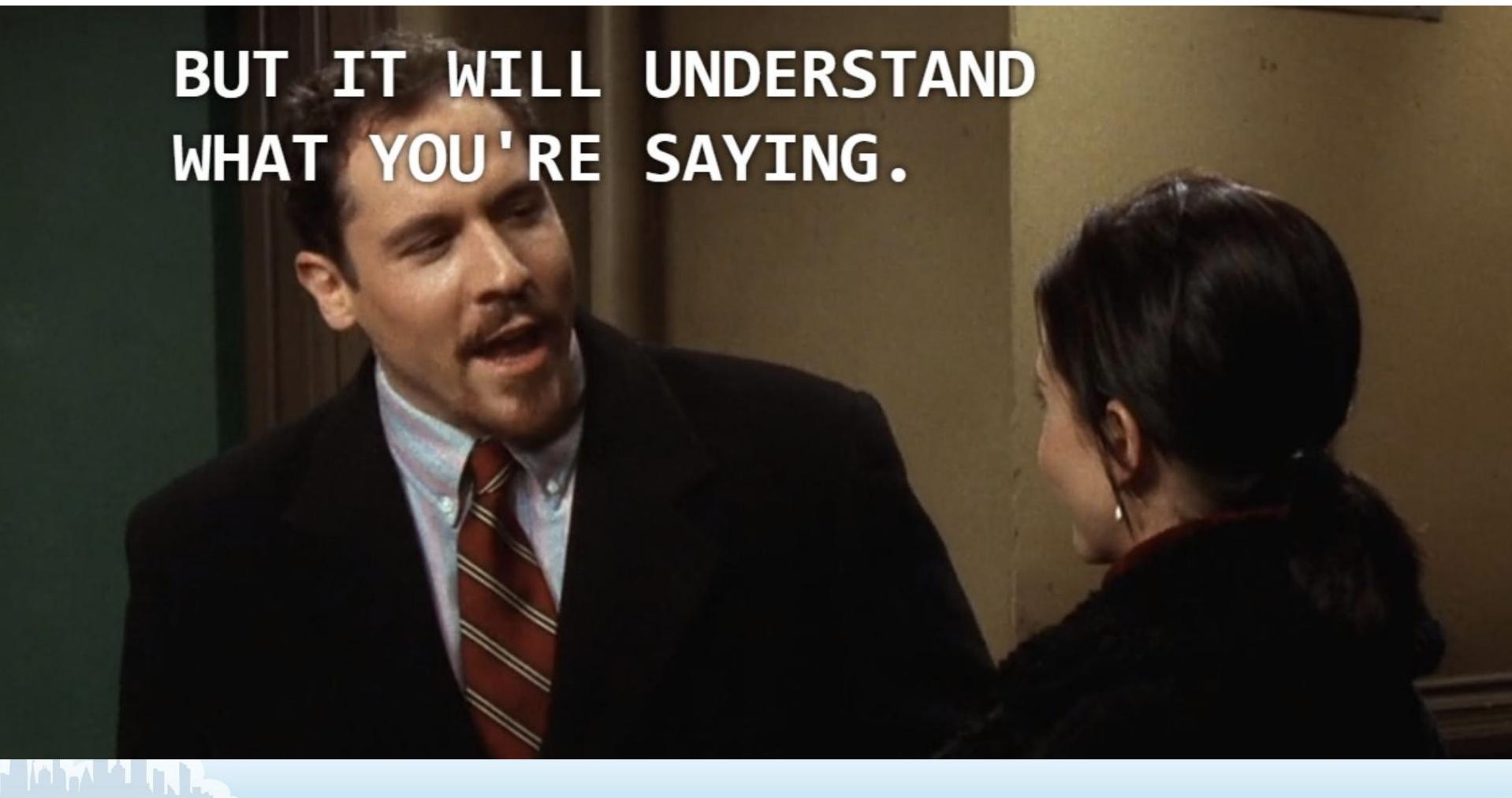
# Voice Recognition - The Complete story



IT'S NOT GOING TO BE ABLE  
TO DO ANY OF THOSE THINGS

A medium shot from a movie scene. A man with a mustache, wearing a dark suit, light blue shirt, and a red striped tie, is speaking. He is looking towards the right side of the frame. A woman with dark hair, seen from the side and wearing a dark jacket, is facing him. They appear to be in an office or hallway setting with wood paneling in the background.

BUT IT WILL UNDERSTAND  
WHAT YOU'RE SAYING.



# Key Takeaways



- 📖 Language understanding is AI Complete.
- 👤👤 Focus on solving customer pain points in your domain.
- 🔊 Voice will become the new User Interface.

# thank you



# Resources



[Einstein.ai](#) - published papers, research etc.

[Einstein.ai/careers](#) - We are hiring!



@vmanju

