THE NEXT GENERATION OF SPEECH RECOGNITION:

IVR IN THE AGE OF SIRI AND WATSON
INTRODUCTION

HOWARD LEE, CEO OF SPOKEN COMMUNICATIONS
SIRI’S CAPABILITIES

• Basic commands or search in a limited environment
SIRI’S CAPABILITIES

• Select a gender and accent
  • English, French, German
• Select a language
  • Chinese – Hong Kong Cantonese, Chinese Mandarin, and Taiwanese Mandarin
  • English – United States, United Kingdom, Canada, and Australia
  • French – Canada, France, and Switzerland
  • German – Germany and Switzerland
  • Italian – Italy and Switzerland
  • Japanese
  • Korean
  • Spanish – Mexico, Spain, and United States
• Ask any question
  • Searches Google, Bing, Wikipedia, Twitter – so can get a second opinion on your answer
• Displayed in Siri’s interface rather than through third party apps
• Give commands
  • Set a timer
  • Turn airplane mode on/off
  • Turn Bluetooth on/off
  • Make the screen brighter
  • Go to twitter settings
  • Show me my privacy settings
  • Read me my text messages
  • Call a contact
  • Enable wifi
  • Open mail settings
  • Turn on do not disturb
• Interact with friends through Twitter
  • Ask "what's going on?" to discover what's trending on Twitter
  • Follow a hashtag on Twitter
  • Post to Twitter or Facebook
• Play back voicemail
• Hear missed calls
• Control iTunes radio
• Play and pause music
• Create reminders
• Choose a movie
• Find out the score of a sports game
• Launch apps
• Check the weather
• Get directions
• Schedule a meeting
• Send an email
• Find a contact
• Set an alarm
WATSON’S CAPABILITIES

This 2-word phrase means the power to take private property for public use; it’s OK as long as there is just compensation.

Eminent domain
- 98%
the electric company
capitalist economy
- 9%
- 5%
WHAT IS WATSON?

• Artificial intelligence

• Question answering computing system
  • natural language processing
  • information retrieval
  • knowledge representation
  • automated reasoning
  • machine learning

• Access to 200 million pages (4 terabytes) of structured and unstructured data
HOW DOES WATSON RUN?

- 90 IBM Power 750 servers, each of which uses a 3.5 GHz POWER7 eight core processor, with 4 threads per core.
- In total, 2,880 POWER7 processor cores and has 16 terabytes of RAM
- Can process 500 gigabytes per second
- Estimated hardware cost of $3 million
WATSON'S KEY INNOVATIONS

• Ability to quickly execute thousands of proven language analysis algorithms simultaneously to find the correct answer

• Hypothesis generation and evaluation - by applying advanced analytics to weight and evaluate a panel of responses based on only relevant evidence

• Evidence-based learning - to improve based on outcomes to get smarter with each iteration and interaction
HOW DO SIRI AND WATSON AFFECT
THE MODERN CONTACT CENTER?
WHICH METHOD DO YOU PREFER?

- Forrester survey
- 79% phone
- 10% automated phone system
- 8% online phone call

In a Forrester survey, an overwhelming majority of consumers prefer to contact customer service by phone compared to other options.
ONLINE PHONE BEATS EMAIL

- Forrester survey
- 79% phone
- 10% automated phone system
- 8% online phone call
72% OF INTERACTIONS STILL TAKE PLACE OVER THE PHONE

More Companies Offer Phone Support

For multi-channel contact centers, phone support is still by far the most popular offering.

CONTACT CENTER INTERACTIONS BY TYPE

<table>
<thead>
<tr>
<th>Contact Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phone (agent)</td>
<td>72%</td>
</tr>
<tr>
<td>Phone (self-service)</td>
<td>12%</td>
</tr>
<tr>
<td>Email</td>
<td>10%</td>
</tr>
<tr>
<td>Letter</td>
<td>2%</td>
</tr>
<tr>
<td>Text chat</td>
<td>2%</td>
</tr>
<tr>
<td>Fax</td>
<td>1%</td>
</tr>
<tr>
<td>Other</td>
<td>1%</td>
</tr>
</tbody>
</table>
HOW ARE THOSE PHONE INTERACTIONS DIFFERENT FROM 10 YEARS AGO?
CHANNEL CHANGE: ALL ABOUT VOICE
CHANNEL CHANGE: ALL ABOUT VOICE

VOICE
WEB
EMAIL
SOCIAL
SMS
CHANNEL CHANGE: ALL ABOUT VOICE

VOICE
WEB
EMAIL
SOCIAL
SMS
CHANNEL CHANGE: ALL ABOUT VOICE
THE GOAL FOR NATURAL LANGUAGE PROCESSING: COMPLEX COMMANDS

• “I want to upgrade my flight to Japan tomorrow using my miles. If it’s under $500, just do it.”
VOICE AS INTERFACE

- The real value is in Watson as a user interface
- Watson/Siri will win as voice interface, not on the database side
- Few players left on the speech rec field: IBM, Google Voice, Nuance, Microsoft
SECRET INGREDIENT FOR NATURAL LANGUAGE SUCCESS

JUST ADD HUMANS
PINPOINTED HUMAN ASSISTANCE IMPROVES SPEECH RECOGNITION

Speech

“Well, uh, you overcharged me on my last bill”

Language model

Dynamic data based on ongoing caller requests

Text conversion

“? You overcharged me on my last bill”

Categorization

Categories calls based on intuitive statistical models

Knowledge Base
- Billing
- Sales
- Customer Service
- Outage

Route

To billing

Silent Guides make corrections to utterances

Knowledge Base
My account was charged $100 this month instead of $50. I have a question about my bill. My bill is wrong. You overcharged me. You’re saying I owe an extra $50.
INCREASED COMPLEXITY

80% OF CALLS ARE NO LONGER THE SAME 3 QUESTIONS
HOW TO MODERNIZE YOUR CONTACT CENTER

• Don’t assume the 80% model still applies to people who have searched online first

• Action: track what your top 100 questions are and categorize them
HOW TO MODERNIZE YOUR CONTACT CENTER

- Action: track the queries and call types your contact center actually handles rather than what the old data predicts.
- Action: train for the long tail of customer service.

![Diagram showing the long tail of demand](image)
HOW TO MODERNIZE YOUR CONTACT CENTER

• Action: improve natural language capabilities by adding humans to the IVR interaction
QUESTIONS?

www.spoken.com