Best Practices in Developing Conversational Systems
Conversational Interaction Conference, 2020, San Jose
Jyo Gadewadikar, jgadewadikar@deloitte.com
Vatamaja, vatamaja@deloitte.com
Key Takeaways –

I. To create an Engaging User Experience for achieving a great degree of Containment in Voice Virtual Assistants with interaction that mimics a Human Conversational Style, remains a progressive area of research and currently a big challenge.

II. In this document we present some of the critical Actionable Leading Design Practices for Complex Conversational Systems.
How to make Virtual Assistants Conversational

Selected Design techniques and user experience practices for complex conversational systems

Design Call Flow
- AI driven call listening and variation capture
- Stochastic call flow design

Customize based on Context
- Utterance Collection = f (Machine Learning + Crowdsourcing)
- Customized Response = f (Sentiment + Intent + Preferences)

Select Vendor/Technology
- Feature and Functionality Evaluation
- Non-Functional Evaluation

Develop, and Train/Test/Tune Iteratively
- Continuously Improve

User Experience and Leading Design Practices extensively focus on actionable methods/process, techniques and procedure to make the conversation approach ‘natural’ style.
Call Listening and Variation Capture

Methods to design product are critical initial steps

1. Analyze Business Requirements
   Conduct need analysis, identify requirements and project goals, map process flow

2. Design Preliminary Conversation Structure

3. Develop Sequential Mapping
   Develop tabulation to connect nodes to intent, enabling improvements for Virtual Assistant

4. Analysis & Call Transcription
   Listen to calls, identify & discard irrelevancies, transcribe

5. Structured Node Mapping
   Map nodes to intent

6. Accommodate New Intents
   Continuously identify and supplement new intents to node structure

7. Record Frequency of Utterances & Intents
   Aggregate individual utterances and intents and record frequency of each

8. Align Analysis with Product Design
   Utilize node structure to inform Voice Virtual Agent product design and development
Stochastic Call Flow

True Human Conversations are varied hence designing a conversation shall not be straightforward or linear
Utterance Collection – Machine Learning

A representative set of training data is crucial in helping NLP engines understand

**Data**
- Transcribe Audio files to text, perform speaker separation
- Mask PII, tabulate utterances (sentences uttered by callers are “utterances”)

**Preprocess**
- Filter caller utterances based on word-of-interest. Break down the utterances to sentences by periods for completeness
- Apply additional parts-of-speech tagging to the sentences for additional contextual information

**Embedding**
- Convert text data to its vector representation
- This step will make the data easily consumable by machine learning models

**Clustering**
- Use clustering techniques to group similar sentences together
- Based on domain knowledge, select the best number of clusters to keep

**Topic Modeling**
- Apply topic modeling techniques to find out the topic of each cluster
  - Options
    - Latent Semantic Analysis (LSA)
    - Latent Dirichlet Allocation (LDA)
    - Term frequency–inverse document frequency (TF-IDF)

**Output**
- With sentences grouped by topics, we have an output that serves as a foundation for “intent and utterances dictionary”
- Such “dictionary” is added to the conversational AI to understand various topics callers utter
Utterance Collection – Crowdsourcing

Populate utterances to enhance base natural language understanding engine

- A large group, thinking individually in the shoes of callers, generate possible utterances.
- Remove duplicates
- Label intents
- Correctly classified
- Mis-classified
- Not classified

Ignore

Examine training data

Can be mapped to existing intents

Add utterances to existing intents

Cannot be mapped to existing intents

Group utterances to new intents & Modify Product

Automated

Update the NLU Engine and Re-test
Customize Responses – f(Sentiment + Intent + Preferences)

Ability to be empathetic is critical

“I would like to reschedule my flight.”

“I just broke up with my girlfriend.. I’m rebooking my flight to leave early.”

“My flight was canceled but I have an important meeting to get to.”

Caller’s Intent: Reschedule an airline ticket

Intent I₁ → Response R₁
Intent I₂ → Response R₂
Intent I₃ → Response R₃

Intent I₁

Emotion E₁ → Response R₁₁
Emotion E₂ → Response R₁₂
Emotion E₃ → Response R₁₃

Neutral: “Sure. One moment.”
Sad: “I’m sorry about that. Let me help you book another flight.”
Anxiety: “No worries. We’ll get you there. Let’s look at your options.”
## Select Vendor/Technology

### Features’ Evaluation

<table>
<thead>
<tr>
<th>Category</th>
<th>Feature</th>
<th>Vendor 1</th>
<th>Vendor 2</th>
<th>Vendor 3</th>
<th>Vendor 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel</td>
<td>Telephonic Integration</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Omni-channel Experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voice</td>
<td>Speech-to-Text</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Text-to-Speech</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SSML</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Speaker Recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AI Engine</td>
<td>Intent Recognition and Mapping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Entity Extraction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dialogue Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Context Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Agent Training</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Agent Testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fallback Intent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sentiment Analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multilingual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pre-built Agents</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Small Talk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System Integration</td>
<td>Serverless Backend</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Knowledge Extension</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>Security</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Monitoring &amp; Reporting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hand-off to Human Agent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Differentiated**

The feature is mature, distinctive from the standard functionality, and ready for production.

**Standard**

The feature has basic capability of fulfilling what is required and can be deployed as part of the Conversational AI system.

**Developing**

Capability not available or related feature is in preview/alpha/beta (Not covered by SLA or ready for production deployment).
Continuous Improvement

- Develop
- Intent, Utterance, Flow Updates
- Human Supervision
- Natural Language Processing
- Test
- Deploy
- Collect Missed Utterances/Call Failure Analyses
Questions

Thank you!

Jyo Gadewadikar, PhD
jgadewadikar@deloitte.com
Chief Product Officer – Conversational AI
Analytics and Cognitive – Applied AI
Deloitte Consulting
30 Rockefeller Plaza, New York, NY, 10112
Mobile: +1 646 734 6510
www.deloitte.com

Vatatmaja, PhD
vatatmaja@deloitte.com
Chief Technology Officer – Conversational AI
Analytics and Cognitive – Applied AI
Deloitte Consulting
1100 Walnut Street, Kansas City, MO 64106
USA
Mobile: +1 913 653 9185
www.deloitte.com