Standards for Virtual Agents

Deborah Dahl, Ph.D.
Conversational Technologies
Conversational Interaction Conference
San Jose, January 30-31, 2017
Virtual agents help users with:

- Getting information – weather, traffic, sports scores...
- Managing daily life – making lists, ordering, shopping, reminders, exercise
- Playing music
- Interacting with the Internet of Things
- Enterprise agents help with customer support or shopping

Interactions are based on spoken natural language
Examples

- Apple Siri
- Amazon Alexa
- Google Assistant
- Microsoft Cortana
- SoundHound
- Hound
- Samsung Viv
- Semantic Machines
<table>
<thead>
<tr>
<th>Application Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apple Sirikit</td>
</tr>
<tr>
<td>Amazon Alexa Skills kit</td>
</tr>
<tr>
<td>Actions on Google</td>
</tr>
<tr>
<td>Microsoft Cortana Skills Kit</td>
</tr>
<tr>
<td>SoundHound Houndify</td>
</tr>
<tr>
<td>Api.ai (Google)</td>
</tr>
<tr>
<td>Wit.ai (Facebook)</td>
</tr>
<tr>
<td>Microsoft LUIS</td>
</tr>
<tr>
<td>Nuance Mix</td>
</tr>
<tr>
<td>Recast.ai</td>
</tr>
<tr>
<td>RASA (open source)</td>
</tr>
<tr>
<td>MITIE (open source)</td>
</tr>
<tr>
<td>Clarity Labs Lucida (open source)</td>
</tr>
</tbody>
</table>
What’s inside? components of a generic spoken language system

- Speech recognition
- Natural language understanding
- Dialog management
- Applications
- Output generation
- User
Can we also use these for virtual agents?
Potential common formats and protocols

- Speech recognition
- Natural language understanding
- Statistical Language Models
- Dialog management
- Output generation
- Dialog scripting
- Applications
- Generic application API
- Potential common formats and protocols

- EMMA?
- SSML?
- SCXML?
- PLS?
How to find, access, and communicate with other virtual agents?
Potential common formats for tools

- Application design
- Development tools
- Annotation formats
- Visualization tools
- Testing and Logging
- Analysis and reporting
A Community Group at the W3C
Exploring new ideas for standard formats and protocols
- https://www.w3.org/community/voiceinteraction/
- Develop proposals for new standards
- Speech recognition
- Structured formats for commonly used concepts – time, location, a person
- Dialog
- Interoperability between virtual agents
- Chairs are: Brian Susko, Marco Kerwitz, and Deborah Dahl
- 16 more participants – join!
- Sunrise session Tuesday at SpeechTEK
About Community Groups at the W3C

- Easy to explore ideas with lightweight process
- Open to everyone
- No fee
- Public
- Balanced IPR policy
Summary

• There are currently many virtual agents and virtual agent development tools
• All with their own formats and API’s
• Development effort will be reduced if the communities agree on formats for scripting components and component output
• Voice Interaction Community Group talking about ideas