Creating Task and Enterprise-Specific Intelligent Assistants for Healthcare

Charles R. Jankowski Jr., Ph. D.

Conversational Interactions Conference 2017
January 31, 2017

charles.jankowski@cloudminds.com
Question Answering is Everywhere!

• 2017 dubbed the “Year of the Bot”

OK Google, who’s the 44th president of the United States?

Show me hotels in Seattle for this weekend that cost less than $300.

Alexa, what’s the score of the Giant’s game?

On the way to my brother’s house, I need to pick up some cheap wine that goes well with lasagna.
QA for your Doctor

• Before/after a colonoscopy exam
• Actual questions from deployed app

- Can I drink coffee?
- What can I eat today?
- Can I brush my teeth the day of the exam?
- Can I eat pancakes for breakfast four days before my procedure?
- How do I find out what time my appointment is?
- Can I take my chemotherapy pills?
- What if I throw up the bowel prep?
- Can I still get the colonoscopy if I have a cold?
So what’s different?
So what’s different?

<table>
<thead>
<tr>
<th>General QA System</th>
<th>Health Domain QA System</th>
</tr>
</thead>
<tbody>
<tr>
<td>General population</td>
<td>Patients of particular Provider</td>
</tr>
<tr>
<td>Domain-invariant content</td>
<td>Content can vary significantly across Providers and domains</td>
</tr>
<tr>
<td>Time-invariant content (sometimes)</td>
<td>Content significantly different depending on date/time relative to procedure</td>
</tr>
<tr>
<td>Content pre-approval not needed</td>
<td>Content approval needed by Provider</td>
</tr>
<tr>
<td>Less sensitivity to content development/update time</td>
<td>Extreme sensitivity to time to develop content for new Provider/domain</td>
</tr>
<tr>
<td>Automatic content extraction</td>
<td>Accurate content approved by Provider</td>
</tr>
<tr>
<td>Lower impact of errors</td>
<td>Higher impact of errors</td>
</tr>
</tbody>
</table>
What do I need to know?

• When is my exam?
• What kind of exam am I having?
• Who is my practice and doctor?
• What’s in this food?
• What ingredients can I eat, and when?
• What medications can I take, and when?
• What can I do, and when?
• What do I do if I have particular side effects, and when?
Rapid Content Development

• Responses can differ based on Provider/domain
  – For colonoscopy, stop eating 2 v. 1 day before exam based on different Doctor
  – GI cares about detailed food intake, other domains not so much

• Need to structure content to make it very easy to tweak existing content

• For new Providers in same domain, can be mere days to generate new tweaked content

• Important for efficiency of deployment
Incorporating Meta Information

• Answer depends on:
  – Question
  – Domain (e.g., medical specialty)
  – Provider
  – Procedure (e.g., colonoscopy, sigmoidoscopy)
  – Current time relative to exam

• Time breakpoints can differ slightly between Providers
  – App matches existing Provider content

• Must have content structure that allows very rapid cloning and tweaking of answer rules for non-question factors
Content Workflow

• Custom services aspect
• Content mirrors existing Provider documentation
• Need to be able to match existing content in days
• Start with templates
  – Domain org (e.g., ACG for colonoscopy) standards
  – Hierarchy of templates
• Need process for Provider approval/editing!
Ontologies

• For rapid content generation/tweaking, need established ontologies of terms

• Based on
  – General
    • WordNet, VerbNet
    • Custom, provider-specific
  – Specific
    • Medication: NDC list (updated quarterly)
    • Foods: Fast-food menus
    • Amazon Mechanical Turk
Making it work: ASR

• Different types of requests
  – Foods, drink, activities
  – Single medication ("Can I take ___?")
  – Medication list
  – General FAQ
  – Vocal navigation (e.g., “Next”, “Back”, “Done”)
  – Story/video search

• Various SR options
  – Network: Nuance NDEV, Dragon Medical, Google, Bing
  – On-device: PocketSphinx, Sensory
Making it work: ASR

• SR engines differ widely in performance
  – 10-40% WER
• Solution: leverage multiple SR engines
  – Different engine based on domain
  – Combine outputs from various engines
  – Different engine based on latency requirements
• Infrastructure needed for metric evaluation
  – Extensive application text logging
  – Logging of all spoken waveforms
  – Offline running of speech engines
  – User voting of answer usefulness
  – Transcription of logged waveforms
  – Extensive test sets and automatic evaluation process
Field Experience

• SR performance good, even long foods/meds lists
  – Varies somewhat depending on domain/engine
• Correct answer performance > 80-90%
  – Continuous offline running of large test sets
• User voting of answer usefulness mostly positive
• Able to use predictive analytics to predict patient no-shows
  – Significant cost savings for Provider
Biography

• Charles R. Jankowski Jr., Ph.D.
  – Nuance, 1998-2011, Speech Scientist, Manager/Senior Manager, Director
  – Performance Technology Partners (PTP) 2012-2013
  – 22otters, Director of Speech and Natural Language, 2013-2016
  – CloudMinds Technologies, Director of NLP Applications, 2016--present