22otters: A Targeted, Customizable, Intelligent Personal Assistant for Patient Care

Charles R. Jankowski Jr., Ph. D.
Ann Thyme-Gobbel, Ph. D.

Mobile Voice Conference
April 12, 2016

charles_jankowski@22otters.com
ann@22otters.com
22otters

- Platform for interactive communication with Patients and Caregivers for acute and chronic medical conditions and procedures
- Supplement to current paper-based instructions
- Provider-approved multi-channel content
- Multi-channel: SMS, Voice Outbound, Mobile app (iOS and Android), Voice inbound soon
- Care calendar, step-by-step training/coaching
- Question Answering system component of Mobile app
22otters Mobile App

- SUNDAY, AUGUST 16
  5. Stop Taking Supplements
  6. Get MiraLAX® and Magnesium Citrate

- MONDAY, AUGUST 17
  7. Stop Eating High Fiber Foods

- THURSDAY, AUGUST 20
  8. Stop Eating Solid Foods and Dairy Products
  9. Mix Your Laxative

Stop taking fiber supplements

Are you okay to stop taking these supplements?

What's your question?

CLEAR LIQUIDS:
- Water
- Black coffee
- Tea no milk
- Clear juice
  - Apple
  - White grape
  - White cranberry
# 22otters Question Answering Requirements

<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</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>
Speech Recognition

• Different types of requests
  – Foods, drink, activities, single medication (“Can I take ___?”)
  – Medication list (“What medications are you current taking?”)
  – General FAQ
  – Vocal navigation (e.g., “Next”, “Back”, “Done”)
  – Story/video search (“Show me the MRI video”)

• Various SR options

• SR engines differ widely in performance: 10-40% WER

• Solution: leverage multiple SR engines
  – Patent pending algorithm
  – Different engine based on domain
  – Combine outputs from various engines
Rapid Custom Content Development

• Answer depends on:
  – Question
  – Domain, Provider, Procedure
  – Current time relative to exam (vary between providers)

• 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

• 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
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
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
Does age influence input modality?

- Data with DOB for users at a GI clinic
  - 54 users
  - Ages 28 - 83
  - 164 questions total
  - Users asked questions using speech-only, typing-only, or both
Observations

• Users favored one mode over another
  – Some used both; most were speech → type following reco error
• Typing strongly preferred by youngest user segment
• Speech preferred by older users but even the oldest users typed
Preferred input mode

Speech preferred by
c. 1/2 of older users
c. 1/8 of younger users

(Mechanical Turk)
## Post-procedure Survey Results

<table>
<thead>
<tr>
<th>5-point Likert scale or Yes/No</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%  Easy to use</td>
</tr>
<tr>
<td>88%  (Very) helpful in preparing for the colonoscopy</td>
</tr>
<tr>
<td>90%  Info was (very) helpful &amp; easy to understand</td>
</tr>
<tr>
<td>81%  Voice-over was (very) helpful</td>
</tr>
<tr>
<td>68%  App reminded me of something I would have forgotten</td>
</tr>
<tr>
<td>87%  I used written patient instruction in addition</td>
</tr>
</tbody>
</table>
## Pre/Post-Launch Usage Data

### Cancellations

<table>
<thead>
<tr>
<th></th>
<th>Preventable reason</th>
<th>Unknown reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-launch</td>
<td>14%</td>
<td>33%</td>
</tr>
<tr>
<td>Post-launch</td>
<td>7%</td>
<td>17%</td>
</tr>
</tbody>
</table>

### Procedure Through-Put

<table>
<thead>
<tr>
<th></th>
<th>Colonoscopy</th>
<th>Double</th>
<th>EGD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-launch</td>
<td>91%</td>
<td>90%</td>
<td>84%</td>
</tr>
<tr>
<td>Post-launch</td>
<td>93%</td>
<td>93%</td>
<td>91%</td>
</tr>
</tbody>
</table>

### Lost Billable Hours

<table>
<thead>
<tr>
<th></th>
<th>15-minute increments left empty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-launch</td>
<td>32%</td>
</tr>
<tr>
<td>Post-launch</td>
<td>8%</td>
</tr>
</tbody>
</table>
Biographies

• 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-current

• Ann Thyme-Gobbel, Ph.D.
  – Ph.D. in Cognitive Science and Linguistics 1993
  – Nuance, 1999-2012, Senior Principal VUI Designer
  – Lab126 (Amazon), Senior UX Designer, 2012-2013
  – 22otters, Head of UX/UI Experience & Design, 2013-current