An Exploratory Study of Users’ Information Search Performance and Challenges in Using Intelligent Conversational Assistants in Health-related Scenarios

Zhaopeng Xing
Practicum Presentation
PSM in BMHI
4/5/2018
Outlines

• Motivation
  • Introduction
  • Related studies
  • Purpose & scope

• Exploratory study
  • Methods
  • Observations
  • Wrap-up
Motivation
Intelligent Conversational Assistants (ICAs)

A program that assists users to achieve certain results in a dialogic fashion (Dale, 2016; Jurafsky & Martin, 2014).
(Dale, 2016)
Potential but with challenges in healthcare

- Improving information accessibility (Bickmore, Utami, Matsuyama, & Paasche-Orlow, 2016)
- Personalized health information service (Alexa with WebMD)
- Long-tail health issues, such as drugs, alcohol & sex issues (Crutzen, Peters, Portugal, Fisser, & Grolleman, 2011)
- Demanding interactions (Schalkwyk et al., 2010)
- Limited capability of providing satisfying information (Miner et al., 2016)
<table>
<thead>
<tr>
<th>Studies</th>
<th>Topic categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Jiang et al., 2015) (Schalkwyk et al., 2010)</td>
<td>User’s intent and purposes (Preferred topics and scenarios)</td>
</tr>
<tr>
<td>(Guy, 2016) (Jeng, Jiang, &amp; He, 2013, 2016) (Schalkwyk et al., 2010) (Shokouhi, Jones, Ozertem, Raghunathan, &amp; Diaz, 2014) (Lopatovska &amp; Williams, 2018)</td>
<td>Query behaviors &amp; strategies</td>
</tr>
<tr>
<td>(Trippas, Spina, Cavedon, Joho, &amp; Sanderson, 2017, 2018)</td>
<td>Information seeking models and search pattern</td>
</tr>
<tr>
<td>(Trippas, Spina, Cavedon, &amp; Sanderson, 2017) (Avula, Chadwick, Arguello, &amp; Capra, 2018)</td>
<td>Engagement of ICAs</td>
</tr>
</tbody>
</table>
Related studies

Study findings
• Difference of interaction modalities has effects on users’ search behaviors
• The use of voice search is sensitive to the scenario
• Technical issues impacts the search experience, e.g. recognition errors

Methods applied
• Observational studies using mixed method
• Compare with traditional search interface
• Query analysis is not appropriate to use in lab settings
Questions to Figure Out

1. What effects ICAs may have on users’ search performance in healthcare scenarios?
   - Completion time
   - The number of query
   - Perceived usefulness

2. What types of health search tasks and which part of search process may be challenging in using ICAs in contrast with textual search interface?
What types of health search tasks and which part of search experience users may feel challenging using ICAs in contrast with traditional web browser

• Fact-finding task and exploratory search task (Pang, Chang, Pearce, & Verspoor, 2014, 2016)
  • Fact-finding task
  • Exploratory search task
What types of health search tasks and which part of search process users may feel challenging using ICAs in contrast with traditional web browser

• Search subprocess (Marchionini & White, 2007):
  • Express the information needs
  • Examine results
  • Reformulation
  • Make a decision
Study
Methods – Participants & Apparatus

• 10 participants recruited at UNC-CH
  • Native speakers (3 male and 7 female, Age mean = 23, SD = 3.23)
  • (Screen) Have experience of using ICAs on weekly basis (Siri, Alexa, Cortana, Google Assistant)
• Google Chrome app (5 Typers) and Google Allo (5 Talkers)
• Iphone 6s plus, IOS 11.2.6
• Videotaped and Monitor the screen
• In study rooms in HSL at UNC
Methods – Procedure (45min)

• Pre screen with a questionnaire
• Inform consent
• Orientation
• Each participant was assigned three task scripts in random order, one task at one time.
• Post task questionnaire after each task
• Interview
• Search and browse history were deleted from the mobile phone after every participant session
• Search without accounts and the search customization was disabled.
A demo
<table>
<thead>
<tr>
<th>Task Script</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Task 1:</strong> Imagine that you recently have a dry mouth and reduced urinations and realize that you rarely have time for water for the busy schedule. <strong>You want to know the amount of water an adult should drink per day.</strong></td>
</tr>
<tr>
<td><strong>Task 2:</strong> Imagine that after finishing a crazy day in your office, you begin to sneeze and feel extremely fatigue. You feel you get a temperature and a mild headache with chills. You don't want to hang out with your friends as usual but sleep at home. <strong>Now you want to figure out what’s wrong with you and then look up which types of OTC medicine can relieve these symptoms.</strong></td>
</tr>
<tr>
<td><strong>Task 3:</strong> Imagine that you are taking a course, <em>An introduction to Health Policy in U.S.</em>, in this semester. You recently read a news article about American Health Care Act of 2017 (AHCA), which aims to partially repeal and replace Obamacare. You are very interested in this law and expect to learn more details. <strong>Find information about the current state of AHCA and then explore three interesting issues regarding AHCA for a group discussion next week.</strong></td>
</tr>
</tbody>
</table>

(Hendahewa & Shah, 2015); (Pang, Chang, Pearce, & Verspoor, 2014); (Wildemuth & Freund, 2012)
## Data Collection

### Measurements

<table>
<thead>
<tr>
<th></th>
<th>Quantitative</th>
<th>Qualitative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information search performance</td>
<td>Completion time (video review)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The number of queries (Query and chat logs)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Users’ perceived usefulness of obtained information (5 point scale likert)</td>
<td>Semi-structured interview</td>
</tr>
<tr>
<td>Perceived difficulties</td>
<td>Overall difficulties of each task (5 point scale likert)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Difficulties of going thru every single sub-process (5 point scale likert)</td>
<td></td>
</tr>
</tbody>
</table>
Observations – Completion Time

- (T1) Four talkers used information given by answer box
- (T3) Talkers had more visits to webpages (3.6 v.s 1.8 webpages)
- (T3) Back-and-forth check between task script and webpage than the typers (3.4 times vs 1.2 times)
Observations – Query number

(T2) Talkers reissued query partially because of interruption or forgetting to tap speak button 6/23
(T3) All talkers did slide the suggested query tags but only 1 talker tapped it.
(T3) All typers used recommended queries at least once
(T3) 4 typers reformulated their query by slightly modifying queries
How useful was the information you obtained using GA/GC

Users' Perceived Usefulness

- Task3_Typers
- Task3_Talkers
- Task2_Typers
- Task2_Talkers
- Task1_Typers
- Task1_Talkers

Legend:
- Very useful
- Useful
- Neutral
- Useless
- Very useless
How difficult was it to complete this task using GA/GC

Users' Perceived Difficulties

- Task3_Typers
- Task3_Talkers
- Task2_Typers
- Task2_Talkers
- Task1_Typers
- Task1_Talkers

The number of users rated it over 3

<table>
<thead>
<tr>
<th>Sub-process in task 3</th>
<th>Talkers</th>
<th>Typers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiate query</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Examine results</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Reformulate</td>
<td>3 (difficult)</td>
<td>0</td>
</tr>
<tr>
<td>Make a decision</td>
<td>3 (2 difficult, 1 very difficult)</td>
<td>1 (difficult)</td>
</tr>
</tbody>
</table>
Talkers said:

• Prefer to use voice search in health fact-finding tasks (All takers)
• A preference for a list of results in exploratory search task (All talkers)
• Have to think about how to formulate query effectively (All talkers)
• Suggested query is useless (all talkers)
• Tend to avoid rewording queries in exploratory tasks (Talker1, 2, 4, 5)
• Seldom use voice search for health information (Talker 1, 3, 4, 5)
• Answer box is useful in quick judgement of information usefulness (Talker 1, 2, 5)
Some other interesting observations

• Didn’t notice copy-pasting behaviors on typers’ sessions.
• Speech recognition errors occurred 2 times overall but interruption ruined users search experience
• Humanoid moment excited users (“so sweet”)
  • I am sneezing and tired with a headache
    - don't try to talk if it hurts, but you can keep trying
• Funny moment made users laugh to cry
  • My symptoms are mild headache chills, fatigue and sneezing
    - Ok I will remember that
Conclusion

• Potential but with rooms for improvements

• Google assistant has great advantage in health fact-finding tasks and needs improvements in addressing exploratory search tasks on
  • facilitate users’ query reformulation process
  • provide indications of useful results properly
Future study

- **Research topics:** query logs analysis, existing information behavior model still works
- **Methods:** Wizard of Oz, physiological measurement
- **Subjective:** the elder had more difficulties in finding health related information ([Czaja, Sharit, & Nair, 2008; Sharit, Hernandez, Nair, Kuhn, & Czaja, 2011; Trewin et al., 2012](https://www.youtube.com/watch?v=19RXpg65CbA))
- **Different input modality:** Alexa medical AI with WebMD
Thanks
Citation


Citation


Citation


