

# A Conversational Agent-based Clinical Trial Search Engine

**Dina Utami, Barbara Barry, Timothy Bickmore**  
Northeastern University  
College of Computer and Information Science  
360 Huntington Ave, WVH 202, Boston, MA  
{dinau,bbarry,bickmore}@ccs.neu.edu

**Michael Paasche-Orlow**  
Boston University School of Medicine  
801 Massachusetts Ave  
Boston, MA  
mpo@bu.edu

## ABSTRACT

The design and evaluation of a web-based search engine for cancer clinical trials is described. The search task is framed as a conversation with an animated agent in order to make it accessible to individuals with low health and computer literacy. Preliminary evaluation comparing the agent to a conventional keyword-based search interface indicates that the agent is at least as effective as the conventional interface, and users are significantly more satisfied with it.

## Author Keywords

Embodied Conversational Agent; Clinical Trial

## ACM Classification Keywords

H.3.3 Information Storage and Retrieval: Information Search and Retrieval.; H.5.2 Information Interfaces and Presentation: User Interfaces.

## INTRODUCTION

Medical information seeking is one of the most common search tasks on the web [17]. However, these critical resources remain out of reach for a significant portion of the population. Fully 36% of US adults have inadequate health literacy, the ability to acquire and act on information related to healthcare [12]. This deficiency not only implies low “domain knowledge” [15], but typically low reading literacy, low numeracy [14], and is related to low computer literacy (in our work we have found a significant correlation between health literacy and computer literacy [4]).

In this work we have focused on developing a web-based search interface to allow individuals with low health and computer literacy to find cancer-related research clinical trials. One study found that 85% of cancer patients were unaware that there were clinical trials they could participate in [10], and other studies have demonstrated numerous barriers individuals face in finding trials they are eligible for [13]. Several web-based search engines have been developed to increase participation in clinical trials by allowing users to more easily find trials for which they may be interested in volunteering. However, these search

engines may be difficult for individuals with low health and computer literacy to navigate. In a pilot usability study, we found that a start-of-the-art web-based search engine (NCI, [1]) was unusable by participants with inadequate health literacy, who were unable to complete any of the 3 standardized tasks given (participants with adequate health literacy completed 1.25 tasks on average) [3].

## RELATED WORK

Several studies have investigated the use of standard keyword-based search interfaces for users with low domain knowledge [16], who speak a foreign language [11], who are children [9], or older adults [2], all of which share characteristics with our task and population. These studies have demonstrated that even the simplest keyword-based search interfaces are out of the reach of many users, and that special design considerations—such as simplifying results [2] and providing language and interaction support [9]—are important for disadvantaged users.

## DESIGN

The overall task our system supports is finding one or more cancer-related clinical trials that a user is eligible for and interested in participating in from the National Cancer Institute’s (NCI) database of 10,730 active trials.

## Conversational Agent Interface

Based on our experience in developing several health counseling dialogue systems for patients with low health literacy [4], we designed the overall interaction as a dialogue with an embodied conversational agent [6]. The agent speaks using synthetic speech, generated from an augmented-transition-network-based dialogue model, template-based text generation, and a dynamically-updated user model, accompanied by conversational nonverbal behavior (hand gestures, facial displays, gaze, etc.) animated in synchrony with the speech [5] (Figure 1). The agent also manipulates artifacts it is discussing with the user, in this case documents that represent aspects of the clinical trials being discussed. User inputs to the conversation are primarily restricted to multiple-choice selection of utterances from a list that is dynamically updated during each turn of the conversation. The interaction is thus system-initiated at the dialogue adjacency-pair level (e.g., agent question / user response), but user initiative is provided by allowing the user to select topics of conversation and ask questions at pre-defined points in the dialogue by selecting them from pre-defined lists. We have successfully used this interface modality

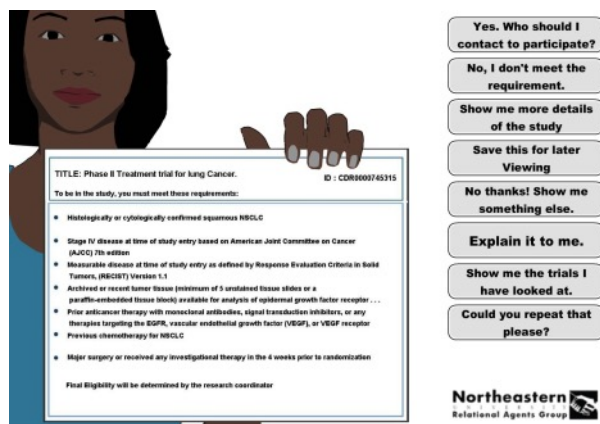


Figure 1. Conversational Agent Interface

with over a thousand patients in clinical trials, including hundreds who have low health literacy and many who have never touched a computer before. The resulting system could be characterized as a “fully faceted” search interface, in which users are never asked to recall and type text, but are always scaffolded with the range of possible inputs they can make.

The overall search experience is thus framed as an extended conversation, in which the user is first interviewed about their requirements and preferences, then shown candidate

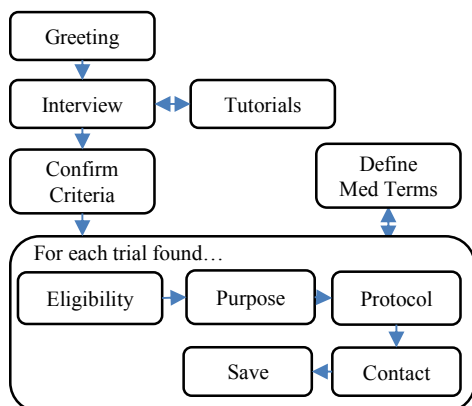


Figure 2. Typical Dialogue Flow

trials, with the agent providing as much scaffolding—through tutorials, explanations and suggestions—along the way as possible. Given that clinical trial descriptions can be very complex and tedious for users to read, we err on the side of eliciting as much information as possible from them prior to search, and displaying information about a trial in stages, revealing only the details they need at each point of their evaluation. The overall flow of a typical conversation is shown in Figure 2.

## DEFINING SEARCH CRITERIA

To define the search criteria the agent elicits from the user, we leveraged qualitative findings from our usability study [1]. Participants in this study were asked to choose between two clinical trial descriptions and then asked to explain their rationale. Coding transcripts of explanations using

grounded theory [7] revealed information seeking practices and deliberation themes. We cross-referenced the resulting list of search criteria preferences elicited from users with the clinical trial schema in NCI database. We found that some user criteria were readily indexable (age, sex, cancer type, geographic location, trial type and phase, medication use), some could be inferred through text classification of fields in the database (pain tolerance, invasiveness tolerance, time commitment), and some comprised a heterogeneous collection of individual user beliefs or personal facts that could not be generalized into search criteria. In our system the indexable search criteria are used to search the database and inferred criteria are used to sort results.

## SEARCH INTERFACE FEATURES

Guided by design principles and observations of information seeking of pilot users, the agent has the following capabilities to scaffold search:

- **Dictionary:** The NCI website provides users with a dictionary of medical terms, but as a separate module from the search engine. In our user studies of this website we observed that finding definitions often distracted users from their main search task. We integrated a dictionary with our search functionality: while the agent explains search results to users, she automatically extracts difficult terms from the text and offers to explain them.
- **Read-aloud:** Users are able to ask the agent to read aloud, and repeat if necessary, any clinical trial text or definition. While speaking aloud the agent holds up a visualization of the text, enabling users to read along.
- **Simplified title:** Clinical trial title can be very long, complex, and hard to remember. We simplified display titles using the phase and type of the trial, and cancer type the trial is for (e.g., replacing “A Study of MGCD265 Given With Erlotinib ...” with “Phase II Treatment Trial for Lung Cancer”).
- **Education modules:** The agent offers to explain several concepts underlying clinical trials, such as voluntariness and risk, at appropriate points in the search.
- **Eligibility confirmation:** The agent periodically displays users’ specified eligibility criteria and allows them to revise their choices.
- **Bookmarking:** Users can save trials for review in later interactions.
- **Summary of views:** The agent periodically displays a list of trials viewed during the search session, with bookmarked trials highlighted.
- **Query refinement:** If a search returns no results or the user exhausts the list of indexed trials, the agent suggests ways to modify their search criteria.

- Levels of detail: Each clinical trial description is presented in three levels with associated displays. The user can opt to view only the title and eligibility of a trial (See Figure 1), then opt to view the trial research purpose, before opting to view details of the trial procedure.

**PRELIMINARY EVALUATION**

To evaluate our system, we conducted a between subjects randomized experiment, comparing our search engine (AGENT) to the conventional keyword-based search engine (CONTROL) developed by the National Cancer Institute (NCI, [1]). Participants were recruited from a pool of cancer patients from across the literacy spectrum.

We asked participants to perform two search tasks. In the first, participants were told to search for a clinical trial for themselves. In the second they were asked to search for a trial for someone else with specified eligibility criteria (age, cancer type, trial type, geographic location), as a standardized test.

Self-report scale measures were obtained after participants completed each task (Table 1). Participants were also asked to recall the number of trials they examined and the number that met their criteria. We also captured the clinical trial ID that the participants found (if any) and the time needed to complete each task. Health literacy was assessed using the REALM instrument [8].

**Results**

The study is ongoing. To date 40 participants have completed the protocol; 21 in the AGENT condition, 19 in CONTROL. Participants had a mean age of 58.6 years (sd 10.1), were 68% male, and 39% low health literacy. 17% reported never using a computer before and another 34% reported only using one a few times. 34% reported never using

a web-based search engine and another 29% reported only using one a few times. Due to the small number of participants and lack of normality in many measures, all statistical tests are non-parametric. Table 1 summarizes the results.

**Task 1 Results**

Most participants started this task without a clear idea of what they were looking for. 38% of users in the AGENT condition found a trial they were interested in, compared to only 21% of users in the CONTROL condition, although this difference was not significant. Among participants who found trials, those in the AGENT group felt that the trials they found matched their criteria to a greater degree than those in the CONTROL group, 4.1 vs. 2.9,  $p=.07$ , approaching significance. There were no significant differences between groups in the number of trials participants reported examining, nor the percent of these trials they felt met their criteria. Participants in the AGENT group reported being more likely to actually enroll in the study they found compared to those in the CONTROL group,  $p=0.08$ , approaching significance. Participants in the AGENT group spent significantly longer on the task compared to those in the CONTROL group, 18.2 vs. 9.8 minutes,  $p<.05$ . However, there were no significant differences in self-reported attitudes towards the amount of time the search took. All participants in the AGENT group were significantly more satisfied with the experience compared to those in the CONTROL group, based on ratings of overall satisfaction, frustration, and positive affect, all  $p<.05$ .

**Task 2 Results**

For the standardized task, 33.3% of participants in the AGENT condition found a clinical trial that satisfied the given criteria, compared to 26.3% of participants in the CONTROL condition (the differences are not significant). As in Task 1, participants spent significantly longer on the task in the AGENT condition compared to CONTROL, 9.1

Measure	Task 1			Task 2				
	A	C	p	A	C	p		
Found a trial (only correct trials counted for Task 2)(% of participants)	38.1	21.1	.24	33.3	26.3	.89		
Elapsed time (minutes)	18.2	9.8	<b>.02</b>	9.1	5.6	<b>.01</b>		
Number of trials examined (self report)	3.4	3.9	.46	2.4	6.8	.66		
Percent of examined trials that meet criteria (self report)(% of trials)	57.2	36.5	.23	72.9	66.7	.76		
<b>Scale Self Report Items</b>	<b>Anchor 1</b>	<b>Anchor 7</b>						
To what degree did you know what you wanted in a trial?	Not at all	Exactly	2.8					
To what degree did the trial match what you were looking for?	Not at all	Exactly	4.1	2.9	.07	5.9	5.2	.36
How likely are you to sign up for the trial that you found?	Not Likely	Very Likely	4.2	3.5	.08			
How much time do you feel it took to use the system?	Too Little	Too Much	4.3	4.7	.29	4.1	4.9	.10
How satisfied were you with the clinical trial search system?	Not at all	Very	5.0	2.8	<b>.001</b>	4.9	2.8	<b>.002</b>
How frustrated do you feel right now?	Not at all	Very	1.8	3.7	<b>.002</b>	2.5	3.8	<b>.05</b>
How pleased do you feel right now?	Not at all	Very	5.1	2.8	<b>.001</b>	4.9	2.8	<b>.002</b>

**Table 1. Study Measures and Results (A=AGENT; C=CONTROL; tests are  $X^2$  or Mann-Whitney U)**

vs. 5.5 minutes, although there were no significant differences in self-reported attitudes towards the time the task took. Also, as in Task 1, participants in the AGENT group were significantly more satisfied with the experience compared to those in the CONTROL group, based on ratings of overall satisfaction, frustration, and positive affect.

## DISCUSSION

The preliminary results indicate that the agent interface is at least as effective as the conventional interface in helping users find clinical trials. However, users are significantly more satisfied with the agent interface compared to the standard, despite the fact that it takes significantly longer to perform a standardized search task. There are several reasons why the agent interface takes significant longer: the time required to hear spoken prompts; the number of questions asked to obtain search criteria during interview; and social dialog, tutorials, and other off-task talk. However, users in our target demographic are clearly happy to spend the extra time in order to obtain better results; in the browsing task they chose to spend almost twice as long finding a trial compared to those using the conventional interface.

Most of the users in our study had previously been involved in clinical trials and are thus not representative of the general population of individuals with cancer since,; they may have had higher than average background knowledge.. However, participants were randomly assigned to the two trial arms, therefore, both groups should be equivalent in this regard.

Our finding suggests that agent-based interface could potentially be a good alternative to standard mouse and keyboard interface, especially for users who have never used computer before. However, this approach will only work for constrained search, for which the search criteria are pre-defined.

## ACKNOWLEDGMENTS

Thanks to Ramesh Manuvinakurike, Juan Fernandez, Lazlo Ring, and Maryam Aziz with their help developing the system, and to Lori Henault and our collaborators at Virginia Commonwealth University for their assistance in running the study. This work was funded by NIH National Cancer Institute grant R01CA158219.

## REFERENCES

1. Atkinson, N., Saperstein, S., Massett, H., Leonard, C., Grama, L. and Manrow, R. Using the Internet to search for cancer clinical trials: A comparative audit of clinical trial search tools. *Contemporary Clinical Trials*, 29 (2008), 555-564.
2. Aula, A. and Käki, M. Less is more in Web search interfaces for older adults. *First Monday*, 10, 5 (2005).
3. Bickmore, T., Paasche-Orlow, M., Aziz, M. and Barry, B. Health Literacy and Usability of Clinical Trial Search Engines. In *Proc. Health Literacy Annual Research Conference (HARC)* (2012).
4. Bickmore, T., Pfeifer, L., Byron, D., Forsythe, S., Henault, L., Jack, B., Silliman, R. and Paasche-Orlow, M. Usability of Conversational Agents by Patients with Inadequate Health Literacy: Evidence from Two Clinical Trials. *Journal of Health Communication* 15, Suppl 2 (2010), 197-210.
5. Bickmore, T., Schulman, D. and Shaw, G. DTask & LiteBody: Open Source, Standards-based Tools for Building Web-deployed Embodied Conversational Agents. In *Proc. Intelligent Virtual Agents* (2009).
6. Cassell, J., Sullivan, J., Prevost, S. and Churchill, E. *Embodied Conversational Agents*. The MIT Press, City, 2000.
7. Corbin, J. and Strauss, A. *Basics of qualitative research: techniques and procedures for developing grounded theory*. Sage, Thousand Oaks, CA, 2008.
8. Davis, T., Long, S., Jackson, R., et al. Rapid estimate of adult literacy in medicine: a shortened screening instrument. *Fam Med*, 25 (1993), 391-395.
9. Gossen, T., Nitsche, M. and Nurnberger, A. Knowledge Journey: A Web SEArch Interface for Young Users. In *Proc. HCIR* (2012).
10. Interactive, H. Misconceptions and lack of awareness greatly reduce recruitment for cancer clinical trials. *Health Care News*, 1 (2001), 1-3.
11. Komlodi, A., Jozsa, E., Koles, M. and Hercegfı, K. Search Quality Differences in Native and Foreign Language Searching. In *Proc. HCIR* (2011).
12. Nielsen-Bohlman, L., Panzer, A. and Kindig, D. *Health Literacy: A Prescription to End Confusion*. The National Academies of Sciences, 2004,
13. Ross, S., Grant, A., Counsell, C., Gillespie, W., Russell, I. and Prescott, R. Barriers to participation in randomised controlled trials: a systematic review. *J Clinical Epidemiology*, 52 (1999), 1143-1156.
14. Rothman, R., Montori, V. and Cherrington, A. Perspective: the role of numeracy in health care. *J Health Commun*, 13 (2008), 583-595.
15. White, R., Dumais, S. and Teevan, J. Characterizing the influence of domain expertise on Web search behavior. In *Proc. WSDM* (2009).
16. Zhang, X., Liu, J., Yuan, X., Cole, M., Belkin, N. and Liu, C. Do users with different domain knowledge select different sets of documents? In *Proc. HCIR* (2012).
17. Zickuhr, K. *Generations 2010. Pew Internet & American Life project.*, 2010.