Journal of Biomedical Informatics 44 (2011) 576–586

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Supporting effective health and biomedical information retrieval and navigation: A novel facet view interface evaluation

Xiangming Mu *, Hohyon Ryu, Kun Lu

School of Information Studies, University of Wisconsin-Milwaukee, 3210 N Maryland Ave., Milwaukee, WI 53211, USA

ARTICLE INFO

Article history:
Received 11 May 2010
Available online 26 January 2011

Keywords:
Exploratory searching
Biomedical information retrieval systems
Cluster and facet view
Search interface evaluation

ABSTRACT

There is a need to provide a more effective user interface to facilitate non-domain experts' health information seeking in authoritative online databases such as MEDLINE. We developed a new topic cluster based information navigation system called SimMed. Instead of offering a list of documents, SimMed presents users with a list of ranked clusters. Topically similar documents are grouped together to provide users with a better overview of the search results and to support exploration of similar literature within a cluster. We conducted an empirical user study to compare SimMed to a traditional document list based search interface. A total of 42 study participants were recruited to use both interfaces for health information exploration search tasks. The results showed that SimMed is more effective in terms of users' perceived topic knowledge changes and their engagement in user-system interactions. We also developed a new metric to assess users' efforts to find relevant citations. On average, users need significantly fewer clicks to find relevant information in SimMed than in the baseline system. Comments from study participants indicated that SimMed is more helpful in finding similar citations, providing related medical terms, and presenting better organized search results, particularly when the initial search is unsatisfactory. Findings from the study shed light on future health and biomedical information retrieval system and interface designs.

1. Introduction

More people search for health related information through the Internet [1]. The large number of health websites and broad range of topics covered make the web an important resource for people seeking health or biomedical information. Some medical information websites such as PubMed,[1] which traditionally target healthcare practitioners and professional researchers, are now more frequently visited by general users looking for authoritative biomedical information [2,3]. According to the National Library of Medicine, in 2002 one third of PubMed's users were general public users [4]. Studies have showed that general users search differently than practitioners and professionals, and lack the domain knowledge often necessary to search a medical database [4]. It is important for PubMed's interface to accommodate both professional and non-domain expert users.

Many studies have been conducted examining how physicians and clinicians use medical databases [5–8]. Their information needs and concerns are different than general users [4,8]. For example, medical terminology is a main problem for the general users when researching health information [9]. However, few studies have examined how biomedical database interfaces affect general users' health information search. Lau and Coiera have done some research regarding general users' health information seeking behaviors, but their area has focused on biases present when searching for information and methods to debias interfaces and interpretations in order for users to retrieve accurate health information [10,11]. From a different perspective, in this study we explored how clustering technology based interface affects general users' health information access in navigational and exploratory search tasks.

According to Marchionini's model, exploratory searches involve learning and investigating, and are conducted for question investigation and topic exploration [12]. Exploratory searching allows users to find relationships between medical literatures and provides a more holistic view of the information [13]. A study of PubMed queries revealed that most users conduct "informational queries" [4], which are similar to exploratory search in that users are looking for information on topic, not a specific citation. The research challenge is thus how to improve the interface design to support effective exploratory searches in health and biomedical databases [9,14].

In this paper we applied clustering technology in health and biomedical information retrieval to facilitate general users'

* Corresponding author. Fax: +1 414 229 6699.
E-mail address: mux@uwm.edu (X. Mu).

[1] http://www.ncbi.nlm.nih.gov/pubmed/.

1532-0464/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2011.01.008

exploratory searches, which to our knowledge has not been fully examined in previous research. The idea is to organize search results as a list of ranked topic clusters instead of a list of citations. This design offers both an overview of the results and a focus-view of a specific topic by allowing users to further mine an area of interest. Based on this model, we designed and implemented a new prototype cluster-based online search system, which we called SimMed. We evaluated the effectiveness of SimMed compared to a baseline system we developed in an empirical user study. We examined how SimMed influenced users' search behaviors and what new search strategies users applied in explorative health information retrieval.

The remainder of this paper is organized as follows: After a brief literature review for background information, we first introduce the SimMed system, including its user interface, search engine, and topic clustering algorithm. In the methodology section of the evaluation study, we describe the data corpus, study participants, and study procedure. Findings from the study are presented in the results and analysis section. We conclude the paper with discussions and implications.

2. Background

Users search the web to obtain information for knowledge acquisition, known item finding, comprehension and interpretation, comparison, aggregation and integration, verification, question answering, and socialization [12]. These different search tasks can be roughly classified into two main categories: lookup and exploratory searching. While lookup searches are related to finding an item, exploratory searches arise when users ''lack the knowledge or contextual awareness to formulate queries or navigate complex information spaces, [and] the search task requires browsing and exploration'' [15]. The usual strategy for exploratory searching is to submit a tentative query, explore the retrieved information, selectively browse interesting items, and then decide what to do next [12]. Two typical technologies that facilitate exploratory searching are facet view interface and similarity browsing.

2.1. Facet view interface

For most online search systems or digital libraries, the search results are presented as a ranked document list in accordance to the computed degree of relevance. Many studies, however, have demonstrated that such a document ranking list presentation was not effective in supporting users' search result navigation and browsing and suggested a facet-view user interface as a solution [16–19].

For the sake of consistency, we first define a number of related concepts used in this paper. ''Cluster'' refers to a set of grouped documents classified by predefined classification criteria (e.g., we can group a collection of architecture design images into several clusters based on their locations). ''Category'' refers to a set of predefined document types which can be used to judge whether a document belongs or does not belong to a category (e.g., a film is classified as ''Documentary'' based on genre). Presenting a collection of documents by clusters or categories will create a ''facet view'' of these documents. Each cluster or category is referred to as a facet. A category facet usually has a meaningful label, while a cluster facet may not. Document facets can be created manually or automatically [17,20–23]. In terms of presentation, facets can either be listed separately along the search results, or be integrated into the search result list. If there are multiple levels of facets, they are usually organized in a hierarchical tree structure. With a set of facets (also referred to as dimensions, factors, clusters, categories, and variables in literature), the interface offers users a ''fisheye'' view of the search results [24] so the representative documents in each facet can stand out rather than hiding somewhere in a long list. In the meantime, users are able to further explore relevant documents by narrowing down to one or more specific facets. Vivisimo[2] and Clusty[3] are two examples of how the facet view design is implemented in real information retrieval systems.

Organizing information into a hierarchical category structure to facilitate information seeking is a long standing research topic (see a recent review by Perugini [22]). Findings from eye tracking studies and transaction log analysis demonstrated that a substantial portion of users' time was spent on interactions with the result categories [25]. On average, a user spent around 25 s looking at the facets and 50 s looking at the search results for an individual search. In another study, Koshman et al. [26] analyzed the transaction log of the Vivisimo search engine and found that almost half (48.26%) of the post-query records involved displaying the result pages from clicking on a category. The facet view interface was found to be better than the document ranking list one, according to a study by Kules and Shneiderman [18]. Their findings showed that the categorization offered an overview of the result list and encouraged users to explore more deeply and broadly. Users were able to better assess search results and felt more organized [18].

In health information retrieval, Pratt [27] suggested organizing results into categories based on the United Medical Language System (UMLS). Similar documents were dynamically classified into groups in accordance with a predefined terminology model and a query model. The terminology model consisted of the Medical Subjects Headings (MeSH) and the semantic types in the UMLS Semantic Network. The query model provided information about the types of the queries [27]. Based on these models, a search system called DynaCat was developed and its effectiveness was evaluated in a user study. Fifteen patients with breast cancer and their family members participated in the study. The results showed that, as compared to the traditional document ranking list tool, users found significantly more answers in a fixed amount of time and felt more satisfied when using DynaCat [19]. Later studies on PubMed using MeSH browser [14] and a subset of MEDLINE database using UMLS Metathesaurus [9] also concluded the helpfulness of facet category in users' health and biomedical information retrieval.

For exploratory search, it is important for the search system to effectively support users' navigation of the vast amount of retrieved medical citations [9,13,14,28]. Yamamoto and Takagi [13] developed a health information retrieval system called McSyBi that clustered literature based on bibliographic information. McSyBi allowed users to determine which type of bibliographic feature the documents should cluster on. They evaluated their system quantitatively and qualitatively through statistical evaluations and comparisons with other studies, and found that nonhierarchical clustering provided users with an overview of retrieved documents, while hierarchical clustering allowed users to see more details and relationships among these retrieved documents. Similarly, Zheng et al. [28] created a clustering technology called GOClonto for displaying PubMed search results. Their approach was based on latent semantic analysis and gene ontology. They found GOClonto performed better than a number of other algorithms such as the suffix tree clustering, Lingo, Fuzzy Ants, and the tolerance rough set algorithm.

Despite the success, there are two major limitations for the cluster or category based facet-view user interface: inappropriate categories and labels, and being cognitively expensive. The inappropriate category and category label problem might be caused

2 www.vivisimo.com.
3 www.clusty.com.

by the complex nature of a document. Each document may have multiple labels and may relate to multiple facets. Not all documents can be directly allocated into predefined categories [18]. Confusing categories and clustering labels frustrate users [17–19,29–31]. In such a case, users must determine through trial and error which category label is best related to their search topic. A wrong decision with unrelated categories annoys users [18].

The second drawback is that the added categories or clusters not only occupy precious screen space used for displaying search results, but the facet view is also cognitively expensive: As the search result list is users' main focus, displaying a separated facet or hierarchy could be distracting and overwhelming to users [18]. Users are forced to move their eyes back and forth between the facet structure and the search result list to coordinate two separate cognition processes. This is usually not recommended for effective navigation designs [32]. To overcome these drawbacks, in our proposed system clustered returns are integrated directly into the search result section.

2.2. Similarity browsing

In addition to the facet view interface browsing, another body of related research is similarity browsing. Similarity browsing allows users to further explore more related documents using "similarity links". One benefit is that users can focus on the current result list without moving their eyes to a separated information structure to avoid shifting between two cognitive processes of viewing current result list and viewing their classifications.

Similarity browsing is essentially a relevance-feedback based retrieval interface. Assuming the selected document is relevant, more similar documents will be recommended. The selected document is treated as a very long query and a new result set is created based on this document query. The problem is that a single document query usually does not perform as well as the original query [33]. This is because each document is often about several topics of which only one is the user's search topic. Some early research on similarity browsing systems examined this problem and proposed solutions that incorporate the original user queries for similarity computations [34]. Recently, the research has been focused more on effective document similarity algorithms. For example, Smucker and Allan [35] developed a new query-based similarity model that finds more related documents in the context of the user's search to avoid extraneous topics. The authors indicated that the new model achieved a 23% increase in the arithmetic mean average precision and a 66% increase in the geometric mean average precision.

Many studies have shown that people would like to use the "similarity links" to do further information explorations. For example, transaction log analysis of the Excite search engine by Spink et al. [36] showed that around 5–9.7% queries were related to clicks on "More Like This" links. Lin et al. [29] analyzed a week long search transaction log in PubMed and indicated that 18.5% search sessions included a click on a "related documents" link. Despite its benefits, a key drawback of applying similarity browsing technique in exploratory search tasks is that it does not provide an overview of the different facets of search results.

In this paper we propose a new approach that combines the strengths of facet view and similarity browsing. The query search results are clustered into groups to support the overview. The ranked clusters are displayed directly in the search result list section to facilitate user's further exploration of "similar" health literature. A prototype search system called SimMed was developed based on this new approach. To evaluate its effectiveness and to better understand users search behaviors when using the new search system looking for health and biomedical information, an empirical user study was conducted. Next we will first briefly introduce the new system we developed.

3. SimMed

3.1. Interface

We designed and developed an online retrieval system, called SimMed, as a platform to evaluate the effectiveness of a topic cluster based retrieval system interface.

Fig. 1 is a screen shot of SimMed after sending the query "diabetic gastroparesis treatment." A list of clusters is presented based on their computed relevancies (details are given in Section 3.3). For each cluster, the most relevant document is shown as the "representative document" with title, author, and first three lines of the abstract. The ranking of clusters is ordered by the rankings of those "representative documents." Within each cluster, the titles of three additional top-ranked documents are displayed (the ranking algorithm is given in Section 3.2). The rest of the retrieved result documents in the cluster are collapsed together under the hyperlink "click here to see more relevant documents." The number of total additional documents in a cluster is provided (i.e., "27 related documents"). If the hyperlink is clicked by a user, these hidden documents in the cluster will be expanded (Fig. 1). Another click on the link will collapse these expanded documents and makes users return to the normal SimMed view.

SimMed combines the strengths of both facet and similarity based browsing tools such as Flamenco [37] and the HubMed [38]. The main design purpose of SimMed is to organize the results into a list of clustered groups to give users both over- and focus-view. After scanning the overall scope of the results, user can focus on a specific cluster and further explore related documents. For the user interface, in addition to providing a link to other clustered documents, three top-ranked documents in that cluster presented provide users with a "preview" of other related documents.

3.2. Search engine

The search engine for SimMed is based on the indri engine[4] designed by Carnegie Mellon University and the University of Massachusetts [39], which is part of the Lemur toolkit [40]. The document ranking was computed based on Kullback–Leibler divergence language model [41]. SimMed does not calculate the document clusters on the fly [35], which is always a challenge for many automatic clustering algorithms, particularly when these algorithms are applied on a large data corpus [23]. SimMed presents related documents based on pre-computed clusters based on topic clustering algorithms (see Section 3.3). This approach is also different from using pre-indexed relevant documents such as those in PubMed [29] that provides recommendation of similar documents based on human indexing. We applied the standard stop list and Porter stemming in the data indexing.

3.3. Topic clustering algorithm

There are many topic-based document classification algorithms available [20,21,42]. We used a combined multivariate Bernoulli model-based k-means algorithm for our document topic clustering computation [43,44]. MATLAB from MathWorks[5] was used to implement the algorithm and calculate the topic clusters. Some of MATLAB source codes in TextClust toolbox, which were developed by Zhong and Ghosh [44] for a clustering comparative study, were adapted in our computations.

[4] http://www.lemurproject.org/indri/.

[5] www.mathworks.com.

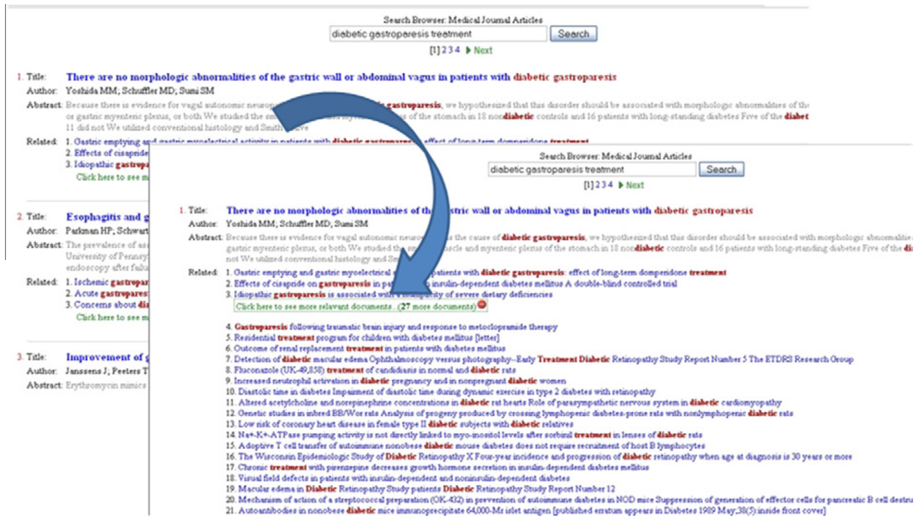

Fig. 1. A screen shot of SimMed user interface after the query ‘‘diabetic gastroparesis treatment’’.

4. Research questions

An empirical study was conducted to compare SimMed to a traditional document ranking list baseline system (Fig. 2). We investigated the following two research questions regarding users' health information exploratory search tasks:

- Compared to the traditional document list baseline system, is the new topic cluster based retrieval system (SimMed) better in terms of users' perceived knowledge change, searching engagement, and satisfaction?
- What are users' search strategies when using SimMed?

5. Methodology

5.1. Data corpus

The data corpus used in SimMed and the study is the OHSUMED test collection [45]. OHSUMED is a MEDLINE subset, consisting of 348,566 citations covering every reference from 270 medical journals over a 5-year period (1987–1991). Each entity in the collection contains several standard fields, including title, abstract, MeSH term list, author, source, and publication type.

5.2. Search tasks

The OHSUMED data set contains 106 search topics. We randomly selected six exploratory search topics for the study (Table 1). The scenario was for general users to search an authoritative health database to find related citations to help them better understand the related topics.

5.3. Participants

We recruited study participants from university students, staff, and their friends. Flyers were posted on campus and solicitation emails were sent to a number of lists. A total of 45 participants were invited to participate in the study in the sequence of their responses. Each participant signed a consent form for the study. At the beginning participants were volunteers. To attract more participants, we later gave small gifts. The last four participants were paid $20 for their participation.

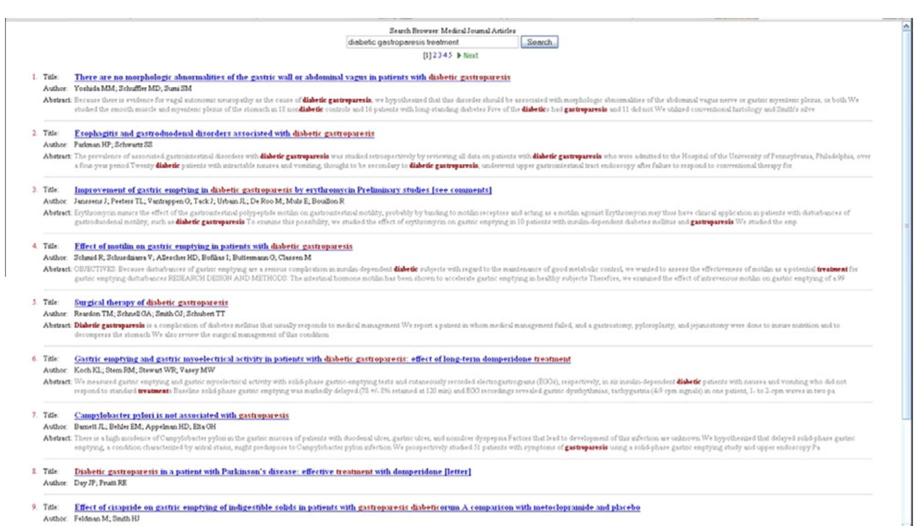

Fig. 2. Document list baseline search system user interface.

Table 1
Descriptions of the six search topics.

Topic ID	Search topic description
1	You are a 50 year old person with COPD You want to know how theophylline should be used for chronic and acute asthma
2	You are a patient with cerebral palsy and depression You want to know about the relationship between cerebral palsy and depression
3	You have diabetic gastroparesis. You want to know what the best treatment is for diabetic gastroparesis
4	You are a patient with a migraine You want to know about the treatment of migraine headaches with beta blockers and calcium channel blockers
5	You are a 60 year old menopausal woman without hormone replacement therapy You want to know if there are adverse effects on lipids when progesterone is given with estrogen replacement therapy
6	You are a 60 year old man with severe malabsorption You want to know about the processes of infiltrative small bowel and information about small bowel lymphoma and heavy alpha chain disease

5.4. Study design and procedure

Two systems were developed and compared in the study. One is the SimMed system (Fig. 1) and the other is a baseline system that provides a traditional document ranking list without clusters (Fig. 2). The baseline system had a similar layout to SimMed. Documents in the result list were presented with the title, author, and the first three lines of the abstract.

We used the within-subject design protocol [46]. That means each participant used one interface (SimMed or the baseline) to search three topics, and then shifted to another one to finish the remaining three. As a result, each person conducted six searches. Latin Square design was used to counterbalance the learning effect on different search topics and the two interfaces [47]. To maximally simulate a real world search environment, no specific trainings on the baseline and SimMed interface were given. A pre-experiment questionnaire was given asking for participants' gender, major, experience using the Internet, and how often they searched for health information online.

Before each search, participants reported their knowledge about the topic using a nine-level Likert scale. The same scale was given immediately after the participant finished the topic search. The difference between them recorded the participant's perceived knowledge change on the search topic. There was no time limitation for participants to complete their search tasks. Participants decided when they should finish searching and answer the after-search questions. The after-search questions were developed based on search topics. For example, for search topic one, the question was: "Please describe how theophylline should be used for chronic and acute asthma."

After a participant finished all six search topics, an overall post-experiment questionnaire was given asking for information about: (1) Which interface they preferred and why; (2) Whether or not the participant clicked and opened the clusters, and why; (3) Under what conditions users would like to click the "related articles" link to expand a cluster; and (4) Suggestions and other comments.

The participants' interactions with SimMed and the baseline system were automatically recorded and stored in a MySQL database as part of the transaction log data. The types of recorded interactions included: query formulations in the search box, document openings, and clicks for expanding or collapsing a cluster.

5.5. Measures

There is no standard for evaluating cluster-based search interfaces such as recall or precision measures [13]. For our study, we used user-centered evaluation approach which is prominent in many interface studies and has been proved effective [9,19,48–51]. Specifically, we evaluated whether users' perceived topic knowledge changed, how much time users have spent interacting with their search results, and which interface they preferred. We used engagement as a variable for effectiveness measure. Although efficiency is widely used for lookup search task evaluation studies [16,19], we believe exploratory searching is better evaluated using engagement – a successful browsing-based search tool should encourage users' engagement by attracting them to spend more time interacting with and exploring the search results [52]. In order to interpret the results as effective engagement instead of user confusion, we also had users report their knowledge change for each topic. We conducted a correlation analysis between the time users spent on topic exploration and the perceived knowledge change reported by users for that topic.

We measured knowledge change using a nine-level Likert scale immediately before and after the search tasks were complete ("Please indicate your current knowledge about this topic (1-No Knowledge; 9-Extremely High Knowledge)"). Time spent on each task was recorded using the computer system. We further applied the Pearson correlation analysis to explore the relationships between users' topic knowledge changes and the time spent. User satisfaction with both interface systems was gathered in post-test questionnaires.

Participant comments were coded using the content analysis approach, which highlighted themes regarding the effectiveness of the SimMed interface. The coding categories were developed after the transcript analysis by a panel of three researchers and two additional usability study specialists (professors in information and library science). The coding was then conducted independently by the three researchers. Any disagreements would be resolved by a group discussion and a consultation of additional specialists.

Our last form of measurement to triangulate the effectiveness of SimMed was through the inversed minimum clicks (IMC), which calculated the average efforts of reaching relevant documents in terms of number of user clicks. Details about IMC will be given in Section 6.5.

5.6. Validity

The search task descriptions and survey questions were developed by the researchers. Then they were reviewed by a panel of four external people (two usability study specialists, one medical librarian, and one nurse) for content validity. A pilot study with five potential users was also conducted. Three suggestions for minor revision (two typo errors and one better rewording) were adopted.

6. Results and analysis

A total of 45 participants were invited to participate in the study. After the data collection, 42 sets were valid. We discarded two data sets due to technical problems in data recording and one for not completing the experiment. Among the 42 participants, nine of them were male (21.43%) and 33 were female (78.53%). Most participants searched the Internet daily (95.24%), one weekly (2.38%) and one occasionally (2.38%). Regarding their history of searching for health information on the Internet, none reported "daily," three (7.14%) reported "weekly," 38 (90.48%) reported "occasionally," and one (2.38%) reported "never."

Table 2
Participants' knowledge changes (before and after search).

Topic ID	Before	After	Change (percentage)	
Baseline				
T1	1.38	3.25	1.87	(135.51%)
T2	1.54	1.88	0.34	(22.08%)
T3	1.21	2.83	1.62	(133.88%)
T4	1.88	3.71	1.83	(97.34%)
T5	1.63	3.38	1.75	(107.36%)
T6	1.17	2.17	1	(85.47%)
Average	1.47	2.87	1.40	(95.46%)
SimMed				
T1	1.5	3.89	2.39	(159.33%)
T2	1.72	2.11	0.39	(22.67%)
T3	1.22	2.94	1.72	(140.98%)
T4	1.72	3.72	2	(116.28%)
T5	1.94	4.28	2.34	(120.62%)
T6	1.06	2.33	1.27	(119.81%)
Average	1.53	3.21	1.69	(110.37%)

Table 3
Time and percentages spent on interface components (in seconds and in percentage).

Topic ID	Result list		Individual document		Cluster		Total
Baseline							
T1	76.38	(35.22%)	140.48	(64.78%)			216.86
T2	178.95	(88.53%)	23.19	(11.47%)			202.14
T3	123.05	(54.57%)	102.45	(45.43%)			225.5
T4	120.1	(42.54%)	162.24	(57.46%)			282.34
T5	63.05	(37.74%)	104	(62.26%)			167.05
T6	149.05	(70.64%)	61.95	(29.36%)			211
Average	118.43	(54.46%)	99.05	(45.54%)			217.48
SimMed							
T1	90	(44.90%)	96.38	(48.09%)	14.05	(7.01%)	200.43
T2	185.52	(64.32%)	54.48	(18.89%)	48.43	(16.79%)	288.43
T3	70.95	(48.91%)	52.6	(36.26%)	21.5	(14.82%)	145.05
T4	97.24	(46.98%)	87.81	(42.42%)	21.95	(10.60%)	207
T5	92.14	(37.81%)	117.29	(48.13%)	34.24	(14.05%)	243.67
T6	227.91	(62.90%)	86.86	(23.97%)	47.59	(13.13%)	362.36
Average	127.29	(52.78%)	82.57	(34.24%)	31.29	(12.98%)	241.16

6.1. Knowledge changes

Table 2 lists the average knowledge (before and after the search) and knowledge changes for the six search topics reported by 42 participants (the baseline and SimMed). Even though an ANOVA analysis with knowledge change as dependent variable did not show statistically significant difference between the two interfaces ($F_{1,240} = 0.082$; $\alpha = 0.775 > 0.05$), we noted that, however, on average participants reported more knowledge change when using SimMed: From 1.53 (before search) to 3.21 (after search) with 110.37% increase. This is 20.19% higher than the reported knowledge change for the baseline: 1.47 (before search) to 2.87 (after search) with 95.46% increase. For individual topics, we noted that the knowledge change with SimMed was consistently greater than that with the baseline (Fig. 3).

Our transaction log data indicated that 28 participants (66.67%) clicked the cluster expanding links at least once. A further inspection revealed that the average knowledge change for those who clicked the cluster expanding links was 272.73% more than those who did not (2.00 vs. 0.54). This may indicate that additional exploration in topic clusters might help to better understanding and gaining more knowledge about the search topics.

6.2. Time spent

For SimMed, there were three components users could interact with. In addition to browsing the ranked result list or opening an individual document to view, users could also expand a cluster and browse the additional documents included. Table 3 lists the time (in seconds and in percentage) participants spent interacting with these three different components in SimMed and the baseline. On average, participants spent 10.89% more time interacting with SimMed than with the baseline (241.16 s vs. 217.48 s). The difference was statistically significant at a 0.05 confidence level ($F_{1,240} = 4.393$; $\alpha = 0.037 < 0.05$). In other word, users spent significant more time interacting and using SimMed for these search tasks.

As an individual component, we noted that the result list was the one on which users spent the most time (54.45% for the baseline and 52.78% for SimMed). As compared to the baseline, participants spent relatively less time on opening and viewing documents (16.48 s less on average) in SimMed. Instead, 12.98% of users' search time was used to interact with clusters. The average interaction time spent by those who clicked the cluster expanding links was 406.01 s, which was significantly higher (53.86%) than the time spent by those who did not (263.88 s). In other word, the clusters provided in SimMed encourage users to spend more time interacting with the search results, which benefits users in exploratory searches [12].

We also conducted the nonparametric Spearman correlation analysis between the time spent on a search task and the perceived knowledge change for that task. The results indicated a positive ($r = 0.178$) correlation (two tailed) and it is statistically significant ($\alpha = 0.004 < 0.01$, $N = 252$). That means the more time users spent interacting with the system, the higher score users reported their topic knowledge changes. This indicates SimMed encourages effective user exploration instead of user confusion.

In Fig. 3 we noted that participants had the least knowledge improvement for search topic 2 ("finding the relationship of

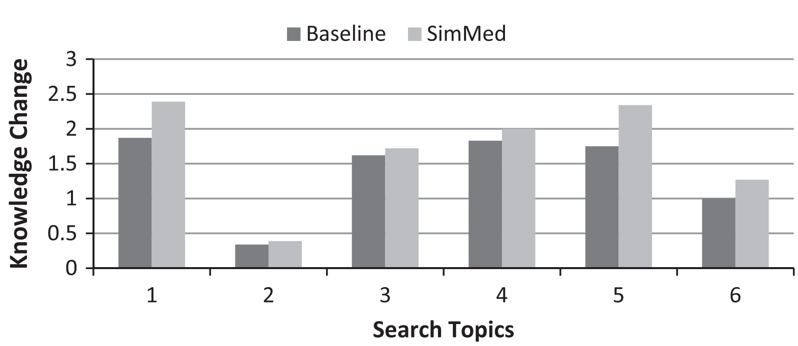

Fig. 3. Knowledge change for each of the six search topics.

cerebral palsy and depression") for both the baseline and SimMed. But participants spent 42.69% more time on topic 2 when using SimMed than when using the baseline. After a further inspection, we found that even though time spent on browsing the result list was relatively similar, participants used 48.43 s, the highest among the six search topics, to explore the clusters in SimMed. In addition, users spent 31.29 s more to open documents when using SimMed. One possible explanation is that when query search did not provide satisfactory results, SimMed attracted users to further interact with and explore the search result. This is consistent with Käki's finding that a cluster interface helps users when document ranking fails [17]. Participant comments presented in Section 6.4 further confirm this explanation.

6.3. User-system interactions and patterns

It is interesting to examine users' interactions with the search system to better understand their search behaviors and strategies. There were three main user-system interactions observed in SimMed: (1) Searching and Browsing (SB), including formulating a query in the search box, sending a query to SimMed to get a list of ranked documents, and browsing the result list; (2) Opening a Document (OD), including opening an interesting document to view the detail; and (3) Expanding/Collapsing a cluster (EC), including expanding a cluster to browse and collapsing an expanded cluster. Table 4 lists the average number of interactions in one topic search. We found that participants had an average of 27.86 interactions per topic search when using SimMed, which was 14.60% more than those using the baseline. The main difference came from the unique expanding/collapsing cluster interaction in SimMed, which included expanding a cluster, scanning the list of additional titles in the cluster, and collapsing a cluster to return to the normal result list view. This result confirmed that SimMed encouraged users to further explore the search results by viewing additional related documents in clusters.

It is also interesting to explore the shifts from one interaction to another. Table 5 lists the six shifting patterns we observed in SimMed as well as their frequencies.

We noted that the main shift occurred between the traditional SB and OD interaction, which accounts for 53.24% of the total 494 shifts. The next most active shift pattern occurred between SB and EC, which accounts for 28.34% of the total shifts. The remaining 18.42% shifts were between OD and EC. Even though the traditional shift between forming a query and Opening a Document to view was still the primary pattern observed, the shifts involving expanding a cluster attributes to 46.76% of the total. It is relatively easier to interpret the SB–OD shift: users start with a query based on their information needs, then navigate the result list (SB), open an interesting document for details (OD), then reformulate or create a new query based on feedbacks gained, and so on. But to better understand those EC involved interaction shifts (EC–OD and EC–SB), we need further analyze participants' responses to the post-experiment survey questions (questionnaires).

Table 4
Average number of interactions for a search topic in terms of SB, OD, and EC.

	Searching and Browsing (SB)	Opening Document (OD)	Expanding/ Collapsing (EC)	Total
SimMed	17.00	6.95	3.90	27.86
Baseline	17.10	7.21	0	24.31

Table 5
Interaction shifts and their frequencies when using SimMed.

Interaction shifts	Frequency
Searching and Browsing (SB) → Opening a Document (OD)	156 (31.58%)
Opening a Document (OD) → Searching and Browsing (SB)	107 (21.66)
Searching and Browsing (SB) → Expanding a Cluster (EC)	80 (16.19%)
Expanding a Cluster (EC) → Searching and Browsing (SB)	60 (12.15%)
Opening a Document (OD) → Expanding a Cluster (EC)	46 (9.31%)
Expanding a Cluster (EC) → Opening a Document (OD)	45 (9.11%)
Total	494 (100%)

6.4. Participant responses

Additional data came from the questionnaires filled out by 42 valid participants. Among them, 35 (83.33%) preferred SimMed and the rest, 16.67%, favored the baseline. SimMed was reported to be more effective, efficient and helpful. Specifically, participants commented that SimMed was more flexible by providing additional navigation choices and was an effective system to manage search results.

We coded the positive participant comments into seven categories (Table 6). If a comment could be included in multiple categories, we classified it into the most salient one. We used "others" category for comments that could not be put into the other six categories, including two "N/A" comments. The coding was conducted independently by the three researchers and only one disagreement was found, which was resolved by a group discussion and a consultation of another two user study specialists.

In general, SimMed helped users better organize the search results, find more relevant documents, choose the right medical terms for queries, and give more search choices. Accordingly, SimMed was believed to be more effective and efficient for the exploratory search tasks.

Among the seven participants who favored the baseline interface, five gave their comments. Two comments said they did not realize the differences between the two interfaces: "Actually I didn't really pay attention nor notice any difference – I was only paying attention to trying to submit search words." One comment said that the baseline interface was simple and "less crowded." Another comment mentioned that the availability of the abstract for each title was helpful. The last pro-baseline comment simply mentioned that "I found more relevant results using this [the baseline] interface."

In the questionnaire we asked the participants to specify why they expanded a cluster to navigate more "related articles". Not surprising, the most common reason we found was to find more relevant citations. For example:

> *"Hoping they'd offer relevant documents. As long as the articles actually are related, it can be very helpful and possibly allow you to find something you would have otherwise missed"*

> *"Because once you find one that is relevant, it would be useful to know others like that"*

Another reason frequently mentioned by participants was that SimMed allowed them to explore more documents without reformulating a query:

> *"I used the 'related articles' option to scan through other articles that might be of interest to the search being conducted so I didn't have to refine my search right away or go through the other pages of results"*

Also asked in the questionnaire was under what conditions the participants would like to click the "related articles" link to expand a cluster. A total of 31 participants answered this question. The other eleven participants either said "I didn't [use it]" or skipped this question. Except for "others" and "under all conditions," four main conditions were identified from the coded comments

Table 6
Coded user positive comments on SimMed with examples and comment numbers.

Description	Examples	Number
Better organized search results	"I preferred that last interface [SimMed]. When I did find the information I was looking for any other related articles could easily be accessed without have to go through pages of documents. It makes the results more manageable" "[I like] the related articles information and the drop-down menu that I could close and open. It allowed me to control the amount of information displayed on the screen, making it less overwhelming"	6 (17.14%)
More choices and options	"I liked having more options to peruse for my results sort of put immediately in front of me" "I liked the related articles option, because it provided more options"	6 (17.14%)
Helpful in finding more relevant documents	"I found an article that I thought would be helpful so I hoped it would lead me to other possibilities" "It provided the possibility of looking at the related article around a given article. It could have been used to further the search process"	6 (17.14%)
Helpful in finding related medical terminologies	"The related links at least clustered articles together, supplying new keywords I could use in searches" "The related articles gave a good idea of whether or not I was using the proper search term. Sometimes by glancing at related articles I would get a new idea as to what sort of words I should be using to search"	5 (14.29%)
More efficient	"I was able to expand on articles that were close to my search topic and find my answer more quickly" "I thought it brought more information up on the front screen and allowed me to drill down to relevant information quickly. It also allowed me to skip those results which did not relate to my search"	4 (11.43%)
More effective	"Sometimes I was only able to find relevant information by looking at the related articles" "I actually found most of my answers through clicking on the related clusters"	4 (11.43%)
Others	"N/A" "I think it is helpful during searches to have relevant information in the interface especially when you have no background knowledge on the topic"	4 (11.43%)
Total		35 (100%)

(Table 7): when a relevant document was found; when the search results were unsatisfactory; when the search topic was unfamiliar; and when relevant documents failed to provide enough information. We separated the category "when relevant documents failed to provide enough information" from a more general category "when the search results were unsatisfactory" because three users specifically mentioned this condition.

There were 14 participants who did not expand a cluster when using SimMed. Two main reasons were concluded from their comments on why they did not use it. The first was that they thought it was unnecessary: "Because there was enough information already to read over – I didn't need more, but less." Another was that they did not notice it: "I didn't see it. I was having such a hard time finding relevant articles, related articles wouldn't have helped."

Table 7
Conditions under which users would like to expand a cluster.

Description	Examples	Number
When users found a relevant document	"I used them when the first given article was extremely relevant to my search. In such a case the related articles were also helpful" "When I felt I found an article that matched what I was trying to locate. I expanded the related articles with hopes to identify familiar terms associated with some of the more technical terms used"	10 (32.26%)
When the search results were unsatisfactory	"The regular search results didn't have a lot of relevant articles, so I thought I would search for related articles on the one result that seemed to fit my search criteria" "When I thought that they were relevant to what I was looking for and didn't feel like the article given through the search was necessarily relevant"	8 (25.81%)
When the search topic was unfamiliar	"I needed the related sites to guide me through incidents that I have absolutely no prior knowledge of" "They were particularly useful when I was very unfamiliar with the subject"	3 (9.68%)
When the relevant documents failed to provide enough information	"When the original article is close to answering my question but does not have enough detail" "I used them when the initial article was not enough to fully answer my questions"	3 (9.68%)
Under all conditions	"I used it in all situations" "In all the searches"	5 (16.13%)
Others	"Just out of curiosity" "I don't understand the question"	2 (6.45%)
Total		31 (100%)

6.5. Relevant document accessibility

In addition to quantitative and qualitative analysis, we also compared SimMed to the baseline in terms of the cost to reach relevant documents. This helped to triangulate our findings.

One design goal of SimMed is to improve the accessibility to relevant documents. An effective way to measure the accessibility is the number of clicks required to reach a relevant document [31]. For example, if a relevant document appears in the first page of a result list and can be reached without any clicking, the number of clicks for this document is zero. If a relevant document appears in the fourth page of the result list, a user needs to click "next page" button at least three times to access it. The least number of clicks to reach that document is then three. In this paper we defined this variable as "required minimum clicks," or RMC.

For each search topic in the OHSUMED dataset, we had a list of all relevant documents [45]. After a user query was submitted, we computed the RMC numbers of these relevant documents as well as their average, which we referred to as average RMC, or ARMC. The value of ARMC is the average minimum clicks a user needs to access relevant documents for a search topic. To some degree the value of ARMC reflects the average cost to find relevant documents.

We used standard queries for ARMC computing. For the six search topics, the standard queries were created by choosing the key concepts in each topic description. For example, the first search topic was "You are a 50 year old person with COPD. You want to know how theophylline should be used for chronic and acute asthma". The standard query was created to include three key terms – COPD, theophylline, and "chronic and acute asthma". We discarded all supporting terms such as "should, between, about, etc." and age. In this way we created six standard queries and computed their ARMCs for both SimMed and the baseline.

Table 8 lists the average required minimum clicks (ARMC) number for each of the six search topics as well as their mean for SimMed and the baseline. We noted that on average, a user needed 1.37 clicks to access a relevant document in SimMed, but 13.61 clicks in the baseline. Document 88,238,705, for example, was a relevant document to the search topic one. In the baseline system, it ranked 357 and thus needed 35 clicks to be reached. In SimMed, however, the document was ranked 20 in the second cluster and could be reached with one expanding click from the first page. One explanation is SimMed has the capability to display more documents in one page. But we believe that the primary reason attributed to such a sharp difference was that the cluster structure offers a two-layer hierarchical display to replace the traditional linear ranking list display. Mathematically, reaching a leaf node along a hierarchical tree structure is on average faster than along a linear ranking list.

One drawback for measuring the cost of accessing a relevant document using the average required minimum clicks (ARMC) was that it was heavily skewed favoring documents ranked far behind (with larger RMC values). A document ranked 357, for example, will be weighted much more than a document ranked three or four during the calculating of the ARMC. To overcome this problem and give more weight to highly ranked documents, we revised our formula and used inversed minimum clicks (IMC):

$$\mathrm{IMC} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{N_{\mathrm{RMC}}(i)+1}$$

where N_{RMC} (i) is the required minimum click (RMC) number for the ith relevant document and N denotes the total number of relevant documents for a search topic. The higher the IMC value is, the easier to access relevant documents. Table 9 lists the value of IMC for each of the six search topics as well as their mean. We noted that on average SimMed was significantly superior to the baseline in terms of the IMC (0.522 vs. 0.278) ($t = 5.33$, df = 5, $\alpha = 0.003 < 0.01$). In other words, it was much easier for users to access a relevant document in SimMed than in the baseline.

Table 8
Average required minimum clicks (ARMC) number for six search topics and their mean.

	T1	T2	T3	T4	T5	T6	Mean
Baseline	10.03	17.00	4.75	15.86	17.25	16.75	13.61
SimMed	1.35	1.40	0.80	1.33	1.75	1.56	1.37

Table 9
Inversed minimum clicks (IMC) for each of the six search topics and their mean.

	T1	T2	T3	T4	T5	T6	Mean
Baseline	0.278	0.246	0.636	0.254	0.159	0.096	0.278
SimMed	0.533	0.517	0.667	0.520	0.433	0.464	0.522

7. Discussion

In this study we compared SimMed with the traditional document ranking list interface for health and biomedical information exploratory search tasks. Both quantitative and qualitative data, as well as the relevant document accessibility measured by Inverted Minimum Clicks (IMC), were analyzed. In quantitative data analysis, we found that users' perceived topic knowledge changes when using SimMed were on average 20.19% higher than when using the baseline. Also, users were observed to engage more in user-system interactions in SimMed. Users spent significantly more time (10.89%) interacting with SimMed than with the baseline. An action analysis further supported this conclusion. An average of 3.55 more actions per topic search were observed when using SimMed. Cluster related operations attributed to most of these additional engagements. Furthermore, the user-system interaction component and shift pattern analysis also indicated that 46.76% of interaction shifts involved the cluster expanding operation.

Qualitative analysis from post-experiment questionnaires indicated that SimMed was more effective for the health information exploration tasks. We noticed that many participants frequently mentioned the helpfulness of SimMed in providing a better interface to find related terms. One challenge for general users (or non-domain expert users) who search for health information online is the vocabulary mismatch that occurs between consumers' natural language and the terminology used in professional medical literature [53–56]. With relatively low domain knowledge regarding health and biomedical information, general users need assistance from the search system to better understand the related concepts and to optimize their search queries [57]. Studies have demonstrated that it is significantly difficult for general users to form effective queries using the traditional document ranking list interface [58]. In addition, users may not be able to discern how documents are related to each other, or know that unfamiliar terms contained in multiple documents discuss the same information [59]. By clustering similar documents into topic clusters with relevant keywords highlighted, SimMed helps in facilitating non-expert users' information Searching and Browsing.

We codified user responses to find when they would expand a cluster to view more related documents. It is not surprising that users would click and open a cluster to further explore a topic when a relevant document was found. But we also noticed three other conditions under which users would like to open a cluster: when the relevant document failed to provide enough information, when the search results were unsatisfactory, and when the search subject was unfamiliar. These conditions were in line with Pratt's [27] research findings that categorization tools were more helpful when document ranking fails. From this perspective, we believe that SimMed also provides users with a better way to organize and more choices to manipulate the search results. When the results from the initial query were unsatisfactory, users took advantage of the additional cluster feature to do further exploration.

When navigating the search results, users may use two browsing strategies: depth- and breadth-first browsing [60]. With depth-first strategy, users examine each item in the list starting from the top and decide immediately whether to open an interesting document to explore details. With breadth-first strategy, users look ahead at a number of list entities and then revisit the most promising ones for more details. SimMed supported both browsing strategies well. First, after a query search, in addition to the cluster representative document, the titles of three more top-ranked documents in the cluster were presented to improve the overview and facilitate the breadth-first browsing. Secondly, if users clicked the "more relevant documents" link, additional documents in the cluster were presented. It facilitated user's depth-first browsing by allowing users to drill down on a specific topic.

Twenty-eight participants (66.67%) clicked the links and further expanded the clusters during the study. As compared to those who did not click the links, these 28 participants reported significant more topic knowledge changes and they spent much more time interacting with the search results. This implies that the retrieved documents organized in clusters were useful and helpful. For the 14 participants who did not expand clusters, nine of them (64.29%) still preferred SimMed to the baseline. Their comments showed that the additional three titles in each cluster are helpful. For example, one participant mentioned that she had "more relevant answers on the first pages." In addition, users also appreciated the control and options offered by SimMed. One user commented that "More choices just in case one would turn up to be relevant" and "at least made it seem like I had some leverage by suggesting related topics."

In summary, SimMed helps users foraging for health information by providing a cluster-based overview of the search results and additional exploration opportunities through useful and relevant links. According to the Information Foraging Theory [61], we believe SimMed was more attractive to users because of its information dense clusters. When users' initial search fails and they are not satisfied with the search results, SimMed gives an information "scent" encouraging users to expand to more relevant documents in the current information "patch." Currently PubMed interface supports "Also try" and "Related links" functions to provide another type of "scent" to facilitate users' information foraging. If a query of "breast cancer" is issued, the "Also try" will provide "metastatic breast cancer," "breast cancer risk," "breast cancer stem," "triple negative breast cancer," and "breast cancer review" for users to further explore the related topics. After a document link is clicked, some more related links based on similarity to the selected document are provided under "Related links." According to Lin et al., these links were accessed by roughly a fifth of all search sessions [29]. In addition to expanding the search scope, PubMed also supports "Titles with your search terms" function to allow users to effectively limit their search results. In SimMed we provide a different approach to encourage information foraging. We clustered the user's search results into groups based on similarity. These clusters provide users an overview of their search results. Users are encouraged to further explore related documents within a cluster by clicking the "Click here to see more relevant documents" link.

In this paper we introduced the concept of Inverted Minimum Clicks (IMC) to theoretically analyze the average cost to access relevant documents. The results indicated that on average significantly less clicks were needed to reach a relevant document using SimMed than using the baseline. This makes SimMed particularly effective and efficient for high recall search tasks (more exhaustive searches are needed). The improved relevant document accessibility was also supported by evidence from our analysis of search topic three and four, in which users spent less time on SimMed but obtained higher topic knowledge changes. Comments from users further confirmed such benefits. For example, "I thought it brought more information up on the front screen and allowed me to drill down to relevant information quickly. It also allowed me to skip those results which did not relate to my search."

As we mentioned previously, no training or tutorials were given for the new SimMed interface. In the study we found that users had no problem using the new cluster-based interface for their searches. We did not receive any complaints about SimMed from the participants' comments. Even though we did not ask questions in the study about the simplicity or difficulty of using SimMed, several participants indicated that it was more enjoyable to use SimMed than to use the baseline. For example, one participant commented that "It made it easier to search because the database was being helpful in providing me related documents. It was a more pleasurable searching experience."

We also noticed a number of limitations of the study. One is that we only examined the effectiveness of SimMed design implemented based upon one clustering algorithm with one particular clustering number, and we only compared it to one type of baseline. It would be interesting to further investigate the best clustering algorithm in a specific retrieval context and compare its performance with other similar cluster or facet view interfaces.

Second, we chose to use OHSUMED dataset because of the large number of citations, the exploratory search tasks, and a set of ground truth for these tasks. We acknowledge that OHSUMED is not current but given the time and labor constraints, we were unable to create a dataset to specifically match our user group. Third, the tasks provided by OHSUMED asked study participants to imagine a given scenario, and their engagement with the system may have been affected by the level of role-playing. In hindsight, simpler, more relevant search tasks could have yielded better results. In the future we can further explore users' search behaviors in fact finding lookup search tasks using more straightforward yes/no questions such as those used in studies by Westbook et al. [7] and Lau and Coiera [11].

The third limitation is that this study focused on health information exploratory tasks by general users with post-secondary education. Those users represent only a portion of the population that uses PubMed. In the future it would be informative to also look at other user groups' exploratory searches using SimMed.

8. Conclusion

> *"Research tools critical for exploratory search success involve the creation of new interfaces that move the process beyond predictable fact retrieval"* [12].

The purpose of this study is to develop a new cluster ranking based health and biomedical information retrieval system and examine its advantages and disadvantages for exploratory search tasks. The new system, called SimMed, groups topic similar documents together in order to provide users with both over- and focus-views. Users not only were able to rapidly scan the result list organized by clusters for an overall understanding of the scope and nature of the results, but could also focus on an interesting topic and further explore related documents.

The results from our evaluation study with 42 participants indicated that the new interface is more effective and helpful than the traditional ranking list interface. Users spent significantly more time interacting with SimMed and reported more topic knowledge increases. User comments indicated that SimMed was helpful in finding more relevant documents, organizing search results, and providing related medical terms. This conclusion is particularly true when the initial search is unsatisfactory.

The novelty of our study is to develop a new cluster-based health and biomedical information retrieval system to facilitate non-domain expert users exploring medical information using authoritative databases such as MEDLINE. By supporting different views of the search results and richer user-system interactions, SimMed provides an effective alternative to the traditional document ranking list based search result presentation. In addition, the findings from our usability study reviewed the search strategies employed by non-domain expert users that can be helpful in future system and interface designs. Finally, our evaluation metrics, especially the proposed inversed minimum clicks (IMC) index, can be useful for performance evaluations on other similar health and biomedical information retrieval systems.

Acknowledgments

The authors thank all participants for their participations. We also thank Natalie Hazen Burclaff for her many valuable suggestions and help in proofreading the paper.

References

[1] Fox S. Online health search. Pew Research Center's Internet & American Life Project; 2006. <http://www.pewinternet.org/Reports/2006/Online-Health-Search-2006.aspx> [cited 5.6.10].
[2] Lu Z, Wilbur WJ. Improving accuracy for identifying related PubMed queries by an integrated approach. J Biomed Inform 2009;42(5):831–8.
[3] Revere D, Turner AM, Madhavan A, Rambo N, Bugni PF, Kimball A, et al. Understanding the information needs of public health practitioners: a literature review to inform design of an interactive digital knowledge management system. J Biomed Inform 2007;40(4):410–21.
[4] Herskovic JR, Tantaka LY, Hersh W, Bernstam EV. A day in the life of PubMed: analysis of a typical day's query log. J Am Med Inform Assoc 2007;14(2):212–20.
[5] Vibert N, Ros C, Le Bigot L, Ramond M, Gatefin J, Rouet J-F. Effects of domain knowledge on reference search with the PubMed database: an experimental study. J Am Soc Inform Sci Technol 2009;60(7):1423–47.
[6] Westbrook JI, Gosling AS, Coiera EW. Do clinicians us online evidence to support patient care? A study of 55,000 clinicians. J Am Med Inform Assoc 2004;11(2):113–20.
[7] Westbrook JI, Coiera EW, Gosling AS. Do online information retrieval systems help experienced clinicians answer clinical questions? J Am Med Inform Assoc 2005;12(3):315–21.
[8] Revere D, Turner AM, Madhavan A, Rambo N, Bugni PF, Kimball AM, et al. Understanding the information needs of public health practitioners: a literature review to inform design of an interactive digital knowledge management system. J Biomed Inform 2007;40:410–21.
[9] Mu X, Ryu H, Lu K. Search strategies on new health information retrieval system. Online Inform Rev 2010;34(3):440–56.
[10] Lau AYS, Coiera EW. Do people experience cognitive biases while searching for information? J Am Med Inform Assoc 2007;14:599–608.
[11] Lau AYS, Coiera EW. Can cognitive biases during consumer health information searches be reduced to improve decision making? J Am Med Inform Assoc 2009;16:54–65.
[12] Marchionini G. Exploratory search: from finding to understanding. Commun ACM 2006;49(4):41–6.
[13] Yamamoto Y, Takagi T. Biomedical knowledge navigation by literature clustering. J Biomed Inform 2007;40(2):114–30.
[14] Tang M. Browsing and searching in a faceted information space: a naturalistic study of PubMed users' interaction with a display tool. J Am Soc Inform Sci Technol 2007;58(13):1998–2006.
[15] White RW, Kules B, Drucker SM, Schraefel MC. Supporting exploratory search. Commun ACM 2006;49(4):37.
[16] Chen H, Dumais S. In: Bringing order to the web: automatically categorizing search results. CHI '00: proceedings of the SIGCHI conference on human factors in computing systems; The Hague, The Netherlands. New York (NY, USA): ACM; 2000. p. 145–52.
[17] Käki M. Findex: search result categories help users when document ranking fails. In: CHI '05: proceedings of the SIGCHI conference on human factors in computing systems; Portland, Oregon, USA. New York (NY, USA): ACM; 2005. p. 131–40.
[18] Kules B, Shneiderman B. Users can change their web search tactics: design guidelines for categorized overviews. Inform Process Manage 2008;44(2):463–84.
[19] Pratt W, Fagan L. The usefulness of dynamically categorizing search results. J Am Med Assoc 2000;7(6):605–17.
[20] Qi X, Davison BD. Web page classification: features and algorithms. ACM Comput Surv 2009;41(2):1–31.
[21] Carpineto C, Osinski S, Romano G, Weiss D. A survey of web clustering engines. ACM Comput Surv 2009;41(3):1–38.
[22] Perugini S. Supporting multiple paths to objects in information hierarchies: faceted classification, faceted search, and symbolic links. Inform Process Manage 2010;46(1):22–43.
[23] Cutting DR, Karger DR, Pedersen JO, Tukey JW. Scatter/gather: a cluster-based approach to browsing large document collections. In: SIGIR '92: proceedings of the 15th annual international ACM SIGIR conference on research and development in information retrieval; Copenhagen, Denmark. New York (NY, USA): ACM; 1992. p. 318–29.
[24] Furnas GW. Generalized fisheye views. In: CHI '86: proceedings of the SIGCHI conference on human factors in computing systems; Boston, Massachusetts, United States. New York (NY, USA): ACM; 1986. p. 16–23.
[25] Kules B, Capra R, Banta M, Sierra T. What do exploratory searchers look at in a faceted search interface? In: JCDL '09: proceedings of the 9th ACM/IEEE-CS joint conference on digital libraries; Austin, TX, USA. New York (NY, USA): ACM; 2009. p. 313–22.
[26] Koshman S, Spink A, Jansen BJ. Web searching on the vivisimo search engine. J Am Soc Inform Sci Technol 2006;57(14):1875–87.
[27] Pratt W. Dynamic organization of search results using the UMLS. J Am Med Inform Assoc 1997(Suppl. S):480–4.
[28] Zheng H, Borchert C, Kim H. GOClonto: an ontological clustering approach for conceptualizing PubMed abstracts. J Biomed Inform 2010;43(1):31–40.
[29] Lin J, DiCuccio M, Grigoryan V, Wilbur WJ. Navigating information spaces: a case study of related article search in PubMed. Inform Process Manage 2008;44(5):1771–83.
[30] Hearst MA. Clustering versus faceted categories for information exploration. Commun ACM 2006;49(4):59–61.
[31] Zamir O, Etzioni O. Grouper: a dynamic clustering interface to web search results. In: WWW '99: proceedings of the eighth international conference on World Wide Web; Toronto, Canada. New York, NY, USA: Elsevier North-Holland, Inc.; 1999. p. 1361–74.
[32] Furnas GW. Effective view navigation. In: CHI '97: proceedings of the SIGCHI conference on human factors in computing systems; Atlanta, Georgia, United States. New York (NY, USA): ACM; 1997. p. 367–74.
[33] Wilbur W, Coffee L. The effectiveness of document neighboring in search enhancement. Inform Process Manage 1994;30(2):253–66.
[34] Thompson RH, Croft WB. Support for browsing in an intelligent text retrieval system. Int J Man Mach Stud 1989;30(6):639–68.
[35] Smucker MD, Allan J. Find-similar: similarity browsing as a search tool. In: SIGIR '06: proceedings of the 29th annual international ACM SIGIR conference on research and development in information retrieval; Seattle, Washington, USA. New York (NY, USA): ACM; 2006. p. 461–8.
[36] Spink A, Wolfram D, Jansen MBJ, Saracevic T. Searching the web: the public and their queries. J Am Soc Inform Sci Technol 2001;52(3):226–34.
[37] Hearst M. Design recommendations for hierarchical faceted search interfaces. ACM SIGIR workshop on faceted search; 2006.
[38] Eaton AD. HubMed: a web-based biomedical literature search interface. Nucleic Acids Res 2006;34:745–7.
[39] Metzler D, Croft WB. Combining the language model and inference network approaches to retrieval. Inform Process Manage 2004;40(5):735–50.
[40] Ogilvie P, Callan JP. In: Experiments using the lemur toolkit. Text REtrieval conference (TREC); 2001.
[41] Cover TM, Thomas JA. Elements of information theory. New York: Wiley; 1991.
[42] Han J, Kamber M, Tung AKH. Spatial clustering methods in data mining: a survey. In: Miller HJ, Han J, editors. Geographic data mining and knowledge discovery. New York: Taylor & Francis; 2001.
[43] Larkey LS, Croft WB. Combining classifiers in text categorization. In: SIGIR '96: proceedings of the 19th annual international ACM SIGIR conference on research and development in information retrieval; Zurich, Switzerland. New York (NY, USA): ACM; 1996. p. 289–97.
[44] Zhong S, Ghosh J. Generative model-based document clustering: a comparative study. Knowl Inform Syst 2005;8(3):374–84.
[45] Hersh W, Buckley C, Leone TJ, Hickam D. OHSUMED: an interactive retrieval evaluation and new large test collection for research. In: SIGIR '94: proceedings of the 17th annual international ACM SIGIR conference on research and development in information retrieval, Dublin, Ireland. New York (NY, USA): Springer-Verlag New York, Inc.; 1994. p. 192–201.
[46] Martin D. Doing psychology experiments. 3rd ed. Wadsworth; 2007. p. 150–73.
[47] Raghavarao D, Padgett LV. Block designs: analysis, combinatorics, and applications. New Jersey: World Scientific; 2005.
[48] Hearst M, Pedersen P. Reexamining the cluster hypothesis: scatter/gather on retrieval results. In: Proceedings of 19th annual international ACM/SIGIR conference, Zurich; 1996.
[49] Karahoca A, Bayraktar E, Tatoglu E, Karahoca D. Information system design for a hospital emergency department: a usability analysis of software prototypes. J Biomed Inform 2010;43(2):224–32.
[50] Hyun S, Johnson SB, Stetson PD, Bakken S. Development and evaluation of nursing user interface screens using multiple methods. J Biomed Inform 2009;42(6):1004–12.
[51] Yu H, Lee M, Kaufman D, Ely J, Osheroff JA, Hripcsak G, et al. Development, implementation, and a cognitive evaluation of a definitional question answering system for physicians. J Biomed Inform 2007;40(3):236–51.
[52] O'Brien HL, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. J Am Soc Inform Sci Technol 2008;59(6):938–55.
[53] Zeng Q, Kogan S, Ash N, Greenes RA, Boxwala AA. Characteristics of consumer terminology for health information retrieval. Methods Inform Med 2002;41(4):289–98.
[54] Soergel D, Tse T, Slaughter L. Helping healthcare consumers understand: an "interpretive layer" for finding and making sense of medical information. Stud Health Technol Inform 2004;107:931–5.
[55] Mu X, Lu K. Towards effective genomic information retrieval: the impact of query complexity and expansion strategies. J Inform Sci 2010;36(2):194–208.
[56] Keselman A, Browne AC, Kaufman DR. Consumer health information seeking as hypothesis testing. J Am Med Inform Assoc 2008;15(4):484–95.
[57] Wildemuth BM. The effects of domain knowledge on search tactic formulation. J Am Soc Inform Sci Technol 2004;55(3):246–58.
[58] Toms EG, Latter C. How consumers search for health information. Health Inform J 2007;13(3):223–35.
[59] Akbarzadeh-T M, Moshtagh-Khorasani M. A hierarchical fuzzy rule-based approach to aphasia diagnosis. J Biomed Inform 2007;40(5):465–75.
[60] Klöckner K, Wirschum N, Jameson A. Depth- and breadth-first processing of search result lists. In: CHI '04: CHI '04 extended abstracts on human factors in computing systems; Vienna, Austria. New York (NY, USA): ACM; 2004. p. 1539.
[61] Information Pirolli P. Information foraging theory: adaptive interaction with information. New York: Oxford University Press; 2007.