Exploring the Impact of Information Visualization on Medical Information Seeking Using the Web

Ketan K. Mane and Javed Mostafa

School of Library and Information Science (SLIS)
Indiana University, Bloomington, IN 47401
{kmane, jm}@indiana.edu

Abstract

User-friendly client interface is the key component of any search engine. The interface guides the user to acquire relevant information for their queries. The paper discusses different user-interaction interface modules of 'Oncosifter' system – a search engine which provides access to latest information on cancer diagnosis and treatment from Cancer.gov. The system provides support for directory and visualization based queries. A usability test was conducted to evaluate the user-friendliness of two types of interfaces. In the paper, the design rationale and implementation of the two interfaces are presented. The details of the usability test are also discussed in this paper.

1. Introduction

The advent of World Wide Web has revolutionized the way we access information today. More and more information is available in the form of web-pages which makes it globally accessible via any browser in the world. Due to wide-scale availability of data, web-pages are being considered as a medium of communicating a variety of information. Success stories of marketing through web-presence motivated the health industry to use web-pages to disseminate health care information. Different organizations have created healthcare websites such as Cancer.gov[1], Medlineplus[2], Dxplain[3], etc. to provide medical information related to diagnosis and treatment for diseases, information related to clinical trials and statistics. Comprehensive information coverage of different diseases is available through these health care websites. However users often face problem due to information overload. It is likely that a visitor overwhelmed by exposure to a large volume of information may digress from his/her intended purpose and abandon the website. On some websites, disease specific information is embedded in lower levels (far beyond the top levels) and hence it is sometimes difficult to locate such information. Furthermore some websites use medical jargon to represent information which restricts the lay users from taking full advantage of these websites.

This paper describes an information retrieval system named 'Oncosifter'[4] (Mane & Thakur., 2003) which provides direct access to specific cancer treatment and diagnostics information available on Cancer.gov website. Different modes of interactions are available from the Oncosifter system to facilitate search for a specific cancer. Along with the traditional interfaces (keyword and directory-search), an alternate interactive visual interface is also available. All interfaces to the Oncosifter system were designed with the intention to minimize user interaction and provide faster access to information. Using any of the three available interfaces, the user submits his/her query and retrieves information on diagnosis and treatment for a specific type of cancer available on Cancer.gov. All the other information on statistics, clinical trials, etc. is parsed out from the results. Non-technical medical term is used to map cancer information so that information can be accessed by people with little or no medical knowledge. Especially, medical terms suitable for lay user consumption available through Cancer.gov are used to create directory and visualization interfaces covering specific types of cancers. As the Oncosifter system uses the same terms in directory and visualization interfaces, we consider this to be an opportunity to compare and analyze usability of directory and visualization interfaces.

The paper attempts to answer the following questions: Are both interfaces (directory/visualization) equally user-friendly to use? Is there a steeper learning curve, when it comes to adopting a new visual interface to search for

---

[1] Cancer.gov is accessible at URL: http://www.nci.nih.gov/
[2] Medlineplus is accessible at URL: http://medlineplus.gov/
[3] Dxplain is accessible at URL: http://www.openclinical.org/aisp_dxplain.html
[4] Oncosifter is accessible at URL: http://oncosifter.indiana.edu/

information? Is it easy to search for information in visual interface as compared to the directory interface? With prior knowledge of the functionality for both directory and visualization interfaces, do users prefer one interface over the other when it comes to performing a specific task? Are there any major usability concerns that need to be addressed in the directory or visualization interfaces?

The outline of the paper is as follows: section 2 describes some of the related works that reflect user interface study for different systems. Section 3 covers details on the Oncosifter system and its interfaces. Here we describe how the directory and visualization interface can be used to submit query and acquire information on a specific type of cancer. The methodology adopted for the usability study and the analysis of the study are presented in section 4. The paper concludes with a discussion of results from the usability test and future work.

2. Related Work

A large number of studies were conducted to evaluate the advantage of traditional interfaces (keyword search) over visual interfaces. The purpose was to identify if performance on search tasks vary due to presence or absence of visual functions. Information retrieval systems like VIBE, Vineta, Bead, etc. were integrated with a visual interface that displayed results of query-document similarity, as clusters of similar documents (Chalmer & Chitson, 1992; Krohn, 1995; Olsen, Korfhage, Sochats, Spring, & William, 1993). Visual interfaces showing result semantics using Kohonen maps were also studied (Xia, 1992). Tilebars displaying relevance of document results were also studied for their ability to provide more insight into data (Hearst, 1995).

Another study that compared the traditional and visualization interface on TREC data received mixed response on the usage of both interfaces (Veerasamy & N. J. Belkin, 1996). Comparative study of textual, icon, graph and table display revealed that the users favoured graphics as the complexity of the task increased (Morse, Lewis, & Olsen, 2002). Other visual interface based retrieval systems like HomeFinder and FilmFinder rely on the information stored within the database to access information (Ahlberg & Shneiderman, 1994).

To our knowledge, our study is first of its kind that compares traditional directory based interface with a visualization based interface in conducting search tasks for health information. The same medical terms, representing a common set of cancer topics, are used in both interfaces (directory and visualization). This method used to compare two interfaces was not found in any previous studies. A usability test was conducted to give us insight into the usage of different interfaces. Prior to the usability test, we start with a hypothesis that substantial difference exists between the two interfaces which would produce different search performance results. The methodology and results of our analysis are covered in later sections of this paper.

3. Oncosifter System Design

The Oncosifter system is an information retrieval system designed to provide direct access to specific cancer information available on Cancer.gov website. Figure 1 shows the homepage of Oncosifter system. The keyword search is available through the homepage itself. Non-technical medical terms are used to map different cancers so that information can be queried and acquired by the non-technical medical community. For example, eye cancer can be searched using the query terms – 'vision', 'eye' or 'eye cancer'. The directory and visualization interfaces can be accessed by clicking on the respective tabs available on the homepage. Figure 2 shows the directory interface and visualization interface that are configured for cancer search from Cancer.gov website.

The current paper focuses on comparing search performance on directory and visualization interfaces. So the following sub-sections of this paper only cover approaches needed to perform the search. An earlier paper on Oncosifter includes details on system architecture and implementation (Mane & Thakur., 2003).
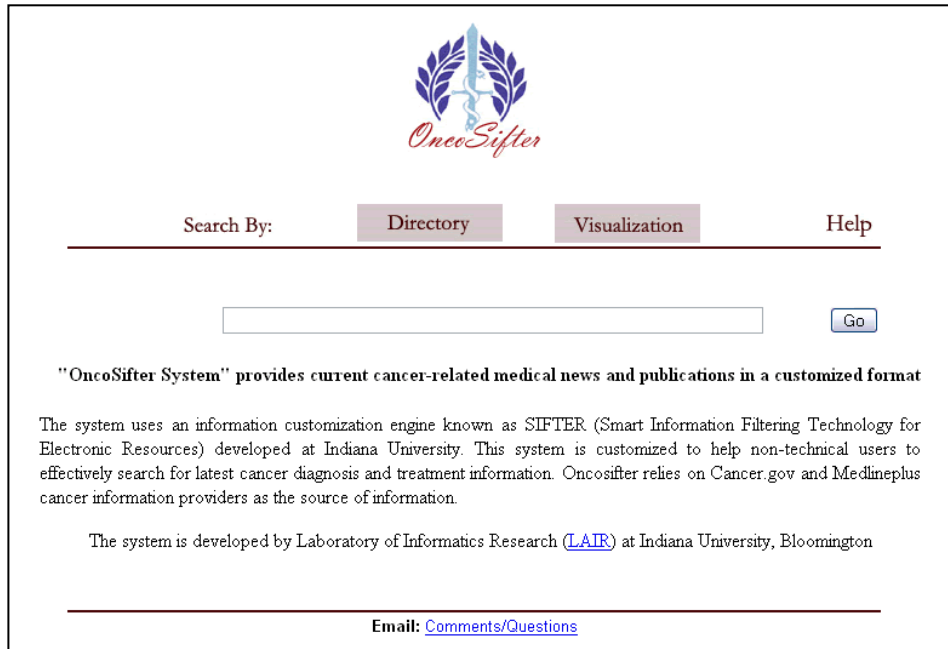
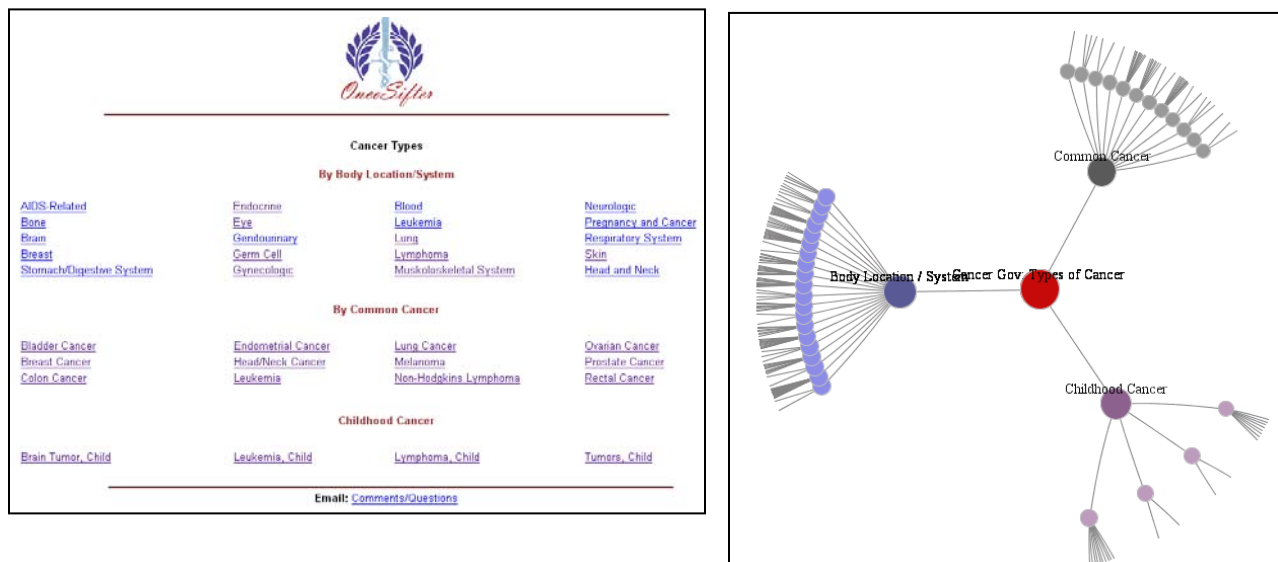**Figure 1: Homepage of Oncosifter system.**

Figure 2: Interface for directory search (left) and visualization search (right).

3.1. Directory Search Interface

The directory interface replicates the different cancer categories available on Cancer.gov website. Information on all known cancers is categorized under three different sub-sections – body locations/systems, common cancers and childhood cancers, as shown in Figure 3. The terms used on the directory interface are usable as query terms. If the user is interested in acquiring diagnostic information for a particular type of cancer, he or she clicks on the term. The system then presents the user with a list of known related cancer. If the user finds a match for the cancer name, the details of which he or she is interested in, then more information can be acquired by clicking on details.
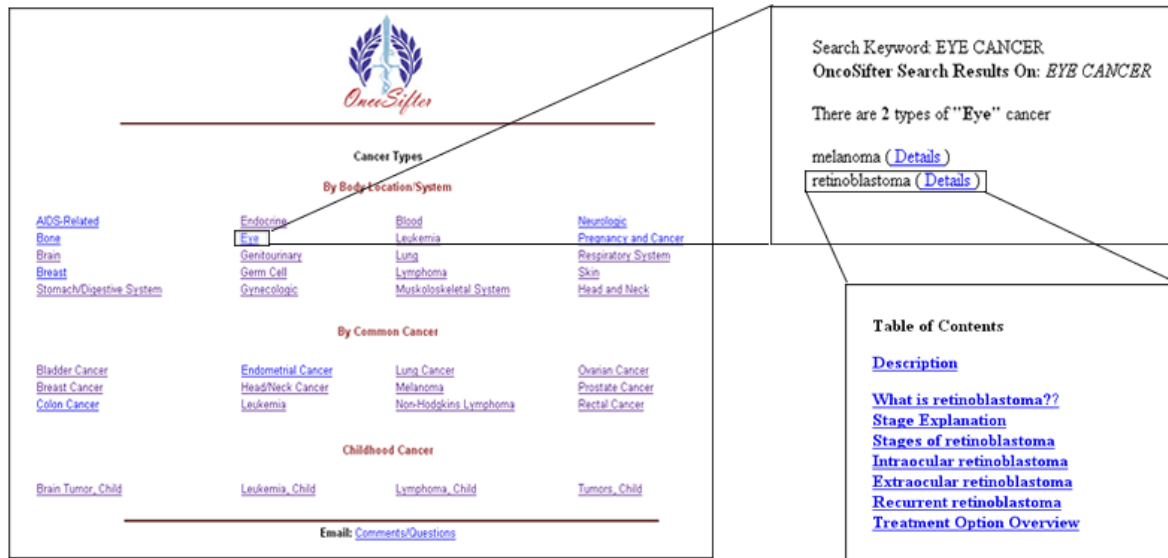
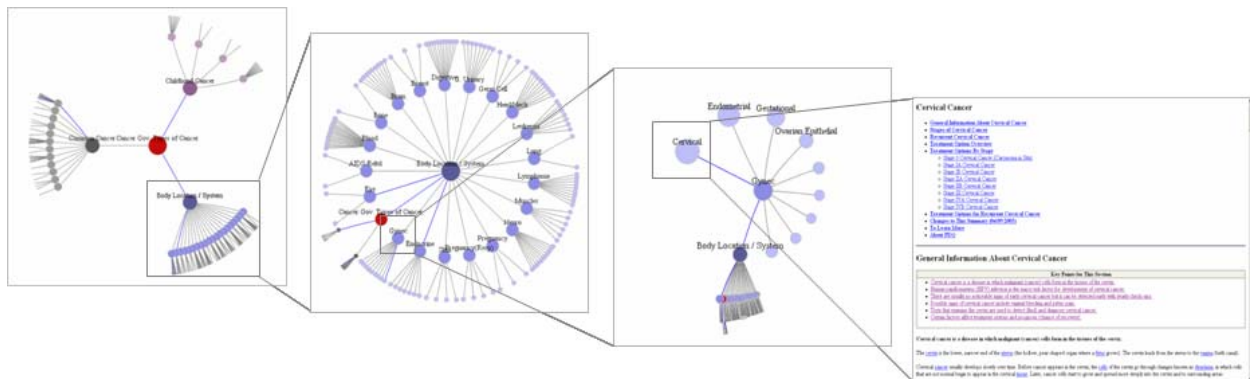**Figure 3: Flow of search through the directory interface.**

Figure 4: Flow of search through the visualization interface

Figure 3 shows the search approach that one would take to find more information about an eye cancer called 'retinoblastoma'. The final result comprising only of diagnosis and treatment information is made available to the user in a new browser window. The concept of different results popping up in different windows was adopted for efficient user interaction. If two cancer treatments are to be compared then existence of multiple browser helps to facilitate side by side comparison. In addition, consistency in information layout and interaction is maintained to reduce the cognitive load on the user (Card, Robertson, & Mackinlay, 1993).

3.2. Hierarchical Visualization Interface

Graphical visualization is a powerful mechanism to communicate and support exploration of information (Tufte, 1997). Visualization help to convey structural relationships that exists within the data (Card et al., 1993). An effective visualization complies with the information seeking mantra – Overview, Zoom and filter, and Details on Demand (Shneiderman, 1997).

Similar to the directory structure, the visualization interface also replicates the cancer categories available on Cancer.gov website. Hyperbolic tree visualization is used to visualize the cancer hierarchical structure. Figure 2 shows the initial visualization interface that is available to the user. Within the visualization, the body

location/systems, common cancers and childhood cancers act as parent nodes for their individual categories. In one level below, the group terms serve as parents to the different types of cancers within the group. Our visualization also has explicit color-coded nodes to provide navigational cues. The final level of node comprises of different cancer types. These nodes are click-able hyperlinks to specific cancer information. The final results page contains diagnosis and treatment information which is made available to the user in a new browser window. Figure 4 shows the search approach that one would take in the visual interface to find more information about cervical cancer.

4. Usability Test

A formal usability test was conducted to test our hypothesis that 'substantial differences exist in performance when search is conducted using directory and visualization search interfaces'. In addition to finding answers to some of the questions raised in the initial section of the paper, the search goal was to gather feedback from users about their opinion on system speed, accuracy and ways to improve interfaces. No advanced search knowledge is required to use the system. For the usability test we recruited ten subjects from a large mid-west university to test our system. The subjects received $15 each for participating in the study.

4.1. Methodology

In order to evaluate the performance of individual interfaces, the total participants were split into two groups of five participants each. To avoid confusion, the group that evaluated the directory interface is referred to as 'directory group'. The other group that evaluated the visualization interface is called the 'visualization group'.

At the start of the test, each participant was asked to fill out a pre-test questionnaire to gather background information related to their profession and experience in using the information visualization tool. In addition, they were also asked, if they were familiar with any website that provided information on cancer diagnostic and treatment.

Further, participants were given a short demonstration to familiarize them with various features available within the Oncosifter system. For just one time, users were shown different ways of interacting with the system to acquire information on a specific cancer. Subsequently, each participant was asked to find answers for a task list.

The task list comprised of nine questions. The top five questions were common to both the groups. All participants were told to use their own discretion to choose any of the two interfaces to find answers to the questions. The remaining four questions from the task list could only be answered using either the directory or visualization interface. The directory group was asked to solve the last four questions using the directory interface. The visualization group was restricted to answer the last four questions using the visualization interface only. There were two purposes to perform this kind of an exercise. The first purpose was to see, if a trend exists in the usage of a particular interface over the other, when it comes to answering top five questions. This would help us determine the preferences of the user in answering a particular kind of task. The second purpose was to establish the user-friendliness of each of the interface by restricting each group to the choice of the interface. Error in locating the correct answer from either of the interfaces would negate the user-friendliness of that interface against the other one.

In addition, special emphasis was placed on topics in developing the questions. All questions could be classified into two categories: 1) Topical – where direct instructions were provided to the user on the location of cancer information in the interface (example Q-2) and 2) Non-topical – where ambiguity was maintained in providing information on the location to find the cancer (example Q-3). The purpose was to identify, if the non-technical medical terms used to represent information in either interface were easy to understand and could be used to perform related searches using the interfaces.

The common questions applicable to both groups are listed below:
1. Among cancers categorized under body location/system, how many cancers are listed under germ-cell cancer?
2. Colon cancer is a type of common cancer. Go to the webpage showing information on colon cancer. Reading through the information, please list the body parts that are likely to be affected by recurrent colon cancer.
3. Under common cancer category, can you find the category name to which oropharyngeal cancer belongs? Please write the category name here.

4. Find the common cancer category with highest number of sub-categories? Please write the category name and the sub-category count here.
5. 'Melanoma cancer' is categorized as skin cancer as well as an eye cancer. Is the information shown on the webpage different or same?

The questions targeted to the directory group are listed below:
6. Use directory search approach to find how many categories of cancer belong to common cancer?
7 'Ovarian germ cell cancer' is a common type of ovarian cancer. Using directory search approach, Can you locate   the body location/system cancer category which has the same information?
8. 'Wilms cancer' belongs to which category of childhood cancer? Please use directory search approach.
9. 'Hodgkin cancer' is a cancer related to Lymph glands. Using directory search approach, first find 'Adult's Hodgkin Lymphoma' under body   locations/systems. Further find the child related 'Hodgkin cancer' under 'Child Lymphoma'. Do these two pages share the same information? Answer saying yes/no.

The questions targeted to the visualization group are listed below:
6. Use the visualization approach to locate the common cancer category with information on 'Mycosis'. Please write the category name here.
7. 'Osteosarcoma' is categorized as a bone cancer under body-location/systems. Using visualization based approach, go to the webpage with osteosacoma information. From the page locate and write the process name that indicates the spread of cancer to other body parts.
8. Use visualization approach to locate the body location/system category for 'small cell lung cancer. Please write the category name here.
9. 'Child Ependymoma' is categorized as childhood cancer. Use visualization approach to locate the category name and list it here.

After completion of the questions, each participant was asked to fill out a post-test questionnaire. The questions were designed to acquire information on the efficiency, accuracy and speed of the Oncosifter system in acquiring relevant information. Additional interest was to obtain subjective feedback about the positive and negative features of the Oncosifter system.

The post-test questionnaire are listed below:
1. The Oncosifter functionality is easy to determine
2. The interface is easy to use
3. Oncosifter provides fast results
4. Oncosifter provides accurate results
5. It was easy to find the information that I needed to complete the task
6. What did you like most about the Oncosifter?
7. What did you like least about the Oncosifter?
8. Do you have any recommendations to improve the Oncosifter?

Participants were asked to respond to the above questions 1-5, using a Likert scale of 1-5. On the scale 1 represents strongly disagree while 5 represents strongly agree.

4.2. Results

Analysis of pre-test questionnaire revealed the following: 6 men and 4 women participated in the usability test. All participants belonged to the age group 21-30 years except for one who was in the 31-40 age group. All the participants had advanced degree (M.S. and Ph.D.) but none of them specialized in medical or health area. Out of the ten participants, one person also worked as an instructor at the university. One of the questions was: Have you used any tool that involves information visualization? It was answered with 'Yes' by all the participants. The information visualization tools listed were Information Visualization CyberInfrastucture (IVC) and other network visualization software packages like Pajek, R, UCINet, Timesearcher and GraphViz. A second question was: Have you used any website that provides access to diagnosis and treatment information on cancer? It was answered with 'No' by eight of the participants and 'Yes' by the remaining two participants.

All the participants completed nine task questions in an average time of 20 minutes. Since all participants showed familiarity with information visualization tool and answer each question was retrievable in two mouse-clicks, we anticipated that there would not be substantial difference in the time taken to answer individual questions.

A typical trend was noticed in the choice of the interface for completing Q4 by all participants. To successfully complete Q4 users had to name a common cancer category with the highest number of sub-categories. For this question all users employed the visual interface. Hence, to get a global view of the dataset and ease the decision making process, the visual interface was a popular choice. For other four common questions, a mixed response was observed in the choice of the interface. Among the top five questions, only two participants answered Q5 incorrectly.

Apart from the top five common questions, the remaining four questions were answered correctly by the directory group. Except for one participant from the visualization group, who failed to answer Q6 correctly, all other group members successfully answered all the questions.

The post-test questionnaire was analyzed to evaluate the subjective feedback on interfaces for the Oncosifter system. These results are presented in Table 1. Average of feedback scores obtained for post-test questions 1-5 is shown in Figure 5.

User #	Q1	Q2	Q3	Q4	Q5
1	4	4	4	3	4
2	3	4	5	5	3
3	4	4	4	4	4
4	5	4	4	5	5
5	2	2	4	4	1
6	4	4	5	4	3
7	4	4	4	4	3
8	5	4	5	5	3
9	4	4	5	5	4
10	4	3	3	4	2

Table 1: Feedback to post-test questionnaire Q1-5 using a Likert scale of 1-5. On the scale 1 represents strongly disagree while 5 is used to represent strongly agree.

Feedback to Post-test Questions (1-5)

	Q1	Q2	Q3	Q4	Q5
Avg. Score	3.9	3.7	4.3	4.3	3.2

Question
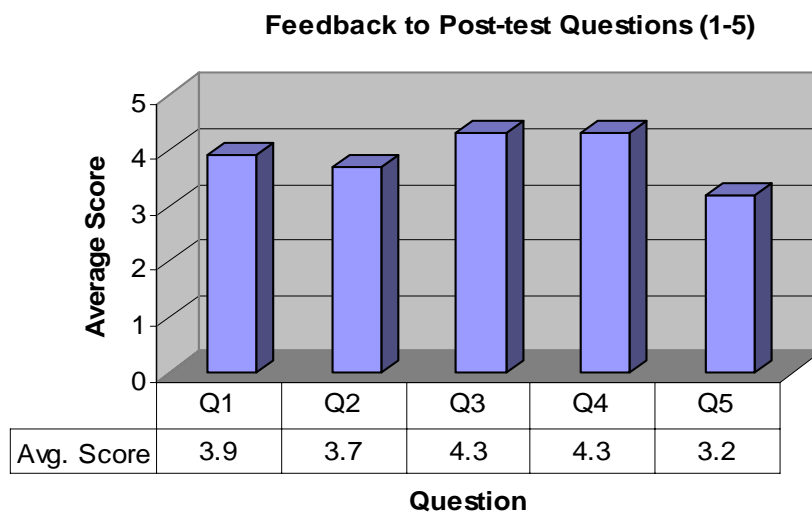
Figure 5: Average score of feedback for Q1-5 from 10 participant

The decrease in the average score for Q5 is observed because two participants reported low scores. A low score was reported due to the absence of search box to search for relevant information.

When asked about what they liked most about the Oncosifter system, most participants had a positive response about the layout and ease with which information was accessible. Most users reported that the information was clearly presented. When asked about what they liked least about the Oncosifter system, the participants suggested incorporating the search function to facilitate the search of cancer term more easily. Recommendations for visual interface include: avoid node label overlap, avoid changing the position of nodes, and disable opening of new browser window to display information. Some users reported that feature to switch directly to initial visual display was not available. This information is not explicitly stated but is available when one clicks on a blank area within the visualization.

5. Discussion and Future Work

The usability test analysis revealed a trend of using the visual interface to get a global insight of the data set. However this observation must be understood in the context of the hyperbolic tree visualization: a rather transparent hierarchical visualization and it's structure makes it easy to count sub-categories. This factor alone may explain the popularity of the visual interface over the directory interface.

All participants were able to effectively use both interfaces to perform the search tasks. Perhaps the single instance of training was sufficient to familiarize users with the main functions of the interfaces and lower the learning curve significantly. No participant had any medical background, yet they were generally able to complete the search tasks effecttively.

In our opinion, this phase of our study is best characterised as a "Pilot" phase and a more thorough study is the next logical step. In the next phase, we plan to use a larger dataset and a more complex set of tasks. We also intend to measure other performance variables such as number of steps to conduct tasks and different classes of interaction errors.

Acknowledgments

References

Ahlberg, C., & Shneiderman, B. (1994). *Visual information seeking: Tight coupling of dynamic query filter with starfield displays.* Paper presented at the Boston, MA, In Proceedings of CHI'94.

Card, S. K., Robertson, G. G., & Mackinlay, J. D. (1993). Information visualization using 3d interactive animation. *Communications, SCM, 36*(4), 51-60.

Chalmer, M., & Chitson, P. (1992, June 27-July1). *Bead: Explorations in information visualizations.* Paper presented at the Proceeding of SIGIR' 92, Copenhagen, Denmark.

Hearst, M. (1995). *Tilebars: Visualization of term distribution information in full text information access.* Paper presented at the Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), Denver, CO, May.

Krohn, U. (1995). *Visualization of navigational retrieval in virtual information spaces.* Paper presented at the In Proceedings of the Workshop on New Paradigms in Information Visualization and Manipulation, Baltimore, MD.

Mane, K. K., & Thakur., S. (2003). *Oncosifter: A customized approach to cancer information.* Paper presented at the Information Visualization Interfaces for Retrieval and Analysis (IVIRA), JCDL, Houston, TX.

Morse, E., Lewis, M., & Olsen, K. A. (2002). Testing visual information retrieval methodologies case study: Comparative analysis if textual, icon, graphical and spring displays. *Journal of American Society and Information Science, 18*(1), 28-40.

Olsen, K. A., Korfhage, R. R., Sochats, K. M., Spring, M. B., & William, J. G. (1993). Visualization of document collection: The vibe system. *Information Processing and Management, 29*(1), 69-81.

Shneiderman, B. (1997). Human factors of interactive software. In *Designing the user interface: Strategies for effective human computer interaction* (pp. 1-37): Addison-Weasly.

Tufte, E. (1997). *The visual display of quantitative information.*Cheshire, CT: Graphics Press.

Veerasamy, A., & N. J. Belkin. (1996). *Evaluation of a tool for visualization of information retrieval results.* Paper presented at the Annual ACM Conference on Research and Development in Information Retrieval, Zurich, Switzerland.

Xia, L. (1992, Oct 19-23). *Visualization for the document space.* Paper presented at the In Proceedings of Visualization, Boston, MA.