Appears in 5th Int. Conf. on **Coordinated & Multiple Views in Exploratory Visualization** (CMV 2007), Zurich, Switzerland, July 2, 2007

Coordinating Linear and 2D Displays to Support Exploratory Search

Anselm Spoerri
Department of Library and Information Science
School of Communication, Information and Library Studies
Rutgers University
4 Huntington Street, New Brunswick, NJ 08901, USA
aspoerri@scils.rutgers.edu

Abstract

Linear and 2D displays have complementary strengths and this paper illustrates how they can be used in concert to support users in the search process. Linear lists are the dominant way for presenting search results and users are most familiar with such result displays. However, ranked lists can only display a limited number of results, whereas 2D layouts can display many results. This paper illustrates how several tightly coupled displays and controls can help users to explore large result sets. Further, it shows how linear and 2D displays can be combined to enable users to see structural relationships between the search results, such as being related to the same topics, tags or being stored on the same web host.

1. Introduction

Users are most familiar with linear displays of search results since this is how the major search engines present their results. Users have learned to expect that the results toward the top of the list have the greatest probability of being relevant. An eye tracking study conducted by Eyetools [9] has shown that users pay most attention to the triangle at the top right of the search results page, which includes the top three results. This triangular area, referred to as the "Golden Triangle", was looked at by 100 percent of the participants in their study. If users are searching for only a few relevant results, then focusing primarily on the "Golden Triangle" of the result page is a reasonable strategy. However, if users are engaged in "exploratory search" then a much larger number of search results need to be examined.

Marchionini [16] has identified three types of search activities: "lookup," "learn," and "investigate," where the latter two activities usually require users to engage in exploratory search. If users are performing a "lookup" task, then they may be looking for a fact or website address, for which a single or few results are sufficient to satisfy their information need. Search engines tend to perform well for "lookup" requests, since a short query tends to be sufficient to retrieve the desired information. Users can expect that the correct answer(s) are located toward to the top of the result page, and users do not really need the capability to examine a large number of search results [16].

If users are engaged in a "learning" or "investigative" search activity, which tends to be open-ended and iterative, then they will have to examine more than a few search results. Ranked lists have the advantage that users know where to start their search for potentially relevant documents. However, users have to move sequentially though the list and only a small subset of the results is visible at once. This is why researchers in the field of "*Information Visualization*" [2], and in "*Search Result Visualization*" [12, 15, 18, 20, 29] and "*Exploratory Search Visualization*" [16, 28] in particular, have been developing interfaces to overcome the inherent limits of ranked lists.

This paper describes how searchCrystal, which builds on and extends the functionality of MetaCrystal [21] and InfoCrystal [20], can support users in the "exploratory search" process. This paper is organized as follows: section 2 briefly reviews related work. In section 3, searchCrystal's linked tools and controls are described. A novel "tag cloud" display is presented that enables users to see how the most frequent words are related. Section 4 shows how linear and 2D displays can be combined to visualize multiple structural relationships between the search results.

2. Related Work

There are at least three distinct approaches to enabling users to explore large search result sets: 1) presentation of "tag clouds" that show the most frequent words or topics; 2) application of clustering technology to organize the results into hierarchical topics; 3) creation of 2D or 3D displays to increase the amount of data that can be presented. This article does not attempt to provide an exhaustive discussion of all the different search interfaces that have been developed over the last decade or so [12, 15, 18, 20, 29]. Instead, specific approaches or systems are highlighted to illustrate some of the major designs that have been developed.

Flickr [10] and Delicious [8] have popularized the use of "tag clouds" to display the most frequent tags that have been assigned to photos or bookmarked web pages. The most frequent tags are alphabetically organized, displayed sequentially on multiple lines and the size of a tag indicates how often it has been assigned. Many websites and blogs have adopted the use of tag clouds to visualize the major topics discussed or of special interest to users. Tag clouds are the simplest visualization employed to help users explore large data sets. Since the tags are alphabetically ordered, users can easily determine if a specific tag is present and its relative frequency.

Several "search visualization" systems, especially commercial ones, make use of text clustering technology to organize the retrieved data. For example, Clusty [6] combines the result sets of multiple search engines to organize the unique results into nested topics, which are visualized using a familiar hierarchical folders metaphor. If users select a specific folder, then only the results in the selected subbranch of the topics hierarchy are shown in a ranked list display. Further, Clusty offers a software widget, called Clusty Cloud [7], that can be embedded in a web page and displays the major topics of search as a tag cloud. Grokker [11] uses the commercial version of the Carrot2 clustering framework [4] to identify the hierarchical topics, which can be visualized as a regular nested list or as nested circles. The latter visualization is a special case of a "treemap" visualization [2]. For large result sets, clustering methods tend to produce topics hierarchies that can become quite extensive and contain many subtopics. This requires users to make an additional effort to just understand the topics hierarchy, especially if topic labels are similar or not very informative, before users can begin to examine the actual results and better understand them.

Many search visualization tools present users with a 2D or 3D map of the major topics that have identified by text clustering or multi-dimensional scaling methods [1, 12]. The retrieved documents are placed spatially close to the topics they are related to. Some maps just display the major topics and the individual documents become visible if the user selects a topic or performs a zooming operation. MetaSpider [5] uses a self-organizing 2D map approach to classify and display the retrieved documents. ThemeScape [29] uses a 3D terrain map to visualize the major topics, where the height of a mountain structure reflects how many found documents are related to a specific topic. Other 2D or 3D visual interfaces group related documents by displaying them in close spatial proximity. Kartoo [14] fuses the results from several search engines to create a 2-D map of the highest ranked documents and also displays the key terms that can be added or subtracted from the current query to broaden or narrow it.

Roberts et al. [18] advocate that search visualizations should be used alongside a ranked list. They developed multiform glyphs that encode information collected about the found web pages, such the page size and the number of internal and external links. These glyphs are incorporated in abstract visualizations that are linked to a ranked list display. Roberts et al. [19] also created a search engine similarity tool that displays the textual difference of multiple web searches using a combination of multiple views and visual bracketing. Sparkler [13] compares multiple result lists by combining a bull's eye layout with star plots, where a document is plotted on each star spoke based on its position in the different lists. Finally, several "Points-of-Interest" (POI) visualizations have been developed [3, 17]. The different search terms, which constitute the search query, are placed equidistant along a circle and act as "magnets" so that the retrieved documents are placed inside the circle based their relative relatedness with respect to the individual search terms.

3. searchCrystal

searchCrystal is a toolset that can be used to visualize diverse data types and sources, such as web, image, video, news, blog and tagging search results. As noted, search-Crystal builds on and is an extension of MetaCrystal [21, 22, 23, 24], which was developed to visualize the overlap between multiple engines searching for the same data type. MetaCrystal in turn is based on InfoCrystal [20], which makes it possible to formulate and visualize Boolean as well as Vector-based queries in the same visualization. These "Crystal" tools have in common that they can visualize the overlap between any fuzzy sets. They have been developed based on the same key design principles: a *radial layout* is employed to ensure that items found by the same number of engines are mapped into the same concentric ring and the greater the number of engines that find an item, the smaller an item's distance from the center. Further, *shape*, *size*, *color*, *proximity* and *orientation coding* are used to indicate how many and which engines retrieved an item and in which rank positions. The higher up an item is placed in the result lists, the greater its visual size. searchCrystal consists of several linked tools: the *Category View*, *RankSpiral*, *Cluster Bulls-Eye* and an enhanced *Ranked List*. The latter two tools will be described in this paper.

The hallmark of an effective visualization is that it guides users toward relevant information. searchCrystal's design is guided by the fact that items found by multiple search engines are more likely to be relevant. Using data from the Text REtrieval Conferences (TREC), Spoerri [25] has shown that the more engines that recommend the same item, the greater the probability that the item is relevant – called the "Authority Effect." Further, the higher up an item is placed in the result lists, the greater its likelihood of being relevant – called the "Ranking Effect." searchCrystal's tools map items found by multiple engines toward the display center, and in the case of the Ranked List, toward the top of the list. This helps to guide users toward potentially relevant information.

searchCrystal is implemented in Flash to make it accessible in a Web browser and can be accessed at www.searchcrystal.com. Its direct manipulation interface enables users to iteratively compose and edit searches as well as perform advanced filtering operations visually [21, 23]. searchCrystal supports "details-on-demand" in all tools and "focus+context" transformations in the Cluster Bulls-Eye or RankSpiral displays. Users can select and drag one of the concentric rings to achieve the desired expansion or contraction of a specific portion of the display (see Fig. 1). Once the mouse is released, the title (fragments) and the size of the visual thumbnails (if available) are recomputed to maximize the information density of the display. searchCrystal has been extended to be able to display both text and images in the same display. This makes it possible to create a visual mashup of web, image, video, news, blog and tagging results [27].

The first goal of this paper is to describe a series of controls that let users explore the search results based on the most frequent words or the major topics. Specifically, these controls are tightly coupled and can visualize how frequent words or major topics are related and correlated. The second goal is to illustrates how linear and 2D displays can be used in concert to visualize multiple structural relationships between the result sets.

3.1. Cluster Bulls-Eye

The Cluster Bulls-Eye tool uses polar coordinates to display the found items: the *radius* value is related to the number of engines that found the item and the average of its rank positions in the lists that contain it; the *angle* reflects the relative ratio of an item's rank positions. This causes items retrieved by the same number of engines to cluster and to be contained in the same concentric ring (see Fig. 1). The use of size coding helps to guide users toward the top items found by a specific number of search engines. In addition, the star-shaped input icons, which are located at the periphery of a crystal, influence an item's position (see Fig. 1). The input icons act as "magnets" that pull an item toward them based on its rank position in the input lists. Items found by multiple engines tend to cluster toward the center and, at same time, users can easily scan the top items found by a single engine.

The Cluster Bulls-Eye tool is extended in searchCrystal to make it possible to visualize relationships between the search results, such as which items are stored on the same web host or are related to the same topic. Further, its layout algorithm is modified to support the display of visual thumbnails of images, videos and websites in a space-efficient way [27].

In Fig. 1, the Cluster Bulls-Eye tool is used to visualize the overlap between the *top 50* Google, Yahoo, MSN, Ask and Exalead search results, respectively, if one searches for "exploratory search result visualization," the results retrieved on March 15, 2007 are shown. This query was chosen since it combines keywords that can be used to describe research efforts that are relevant to the subject of this paper. As shown in Fig. 1, the input icon for Exalead shows that 49 items are found; this is the case since two web pages of the 50 results have identical URLs. A further enhancement made in searchCrystal is the ability to place title (fragments) next to the item icons in a space-filling way. A subtle visual hierarchy is created by varying the typesize and contrast of the titles based on an item's average rank position in the lists that contain it. The algorithm for placing the titles and determining the number of letters to display is quite complex and outside the scope of this paper.

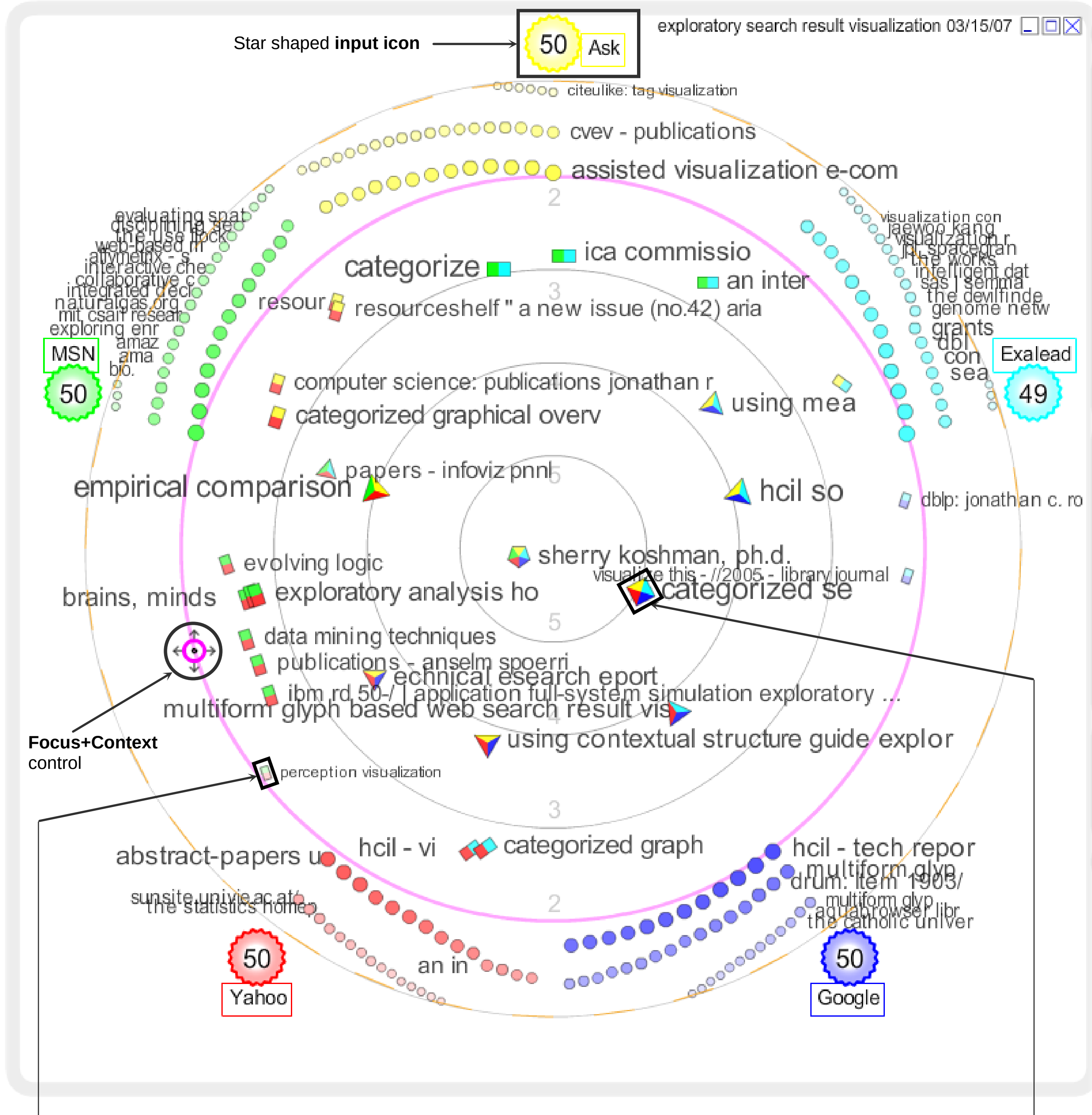
Fig. 1 shows that there is only the web page with the title "Sherry Koshman, Ph.D." that is found by all five engines and this page is a relevant result. The item icon found by four of the five engines has a larger shape since it is more highly placed in the four result lists that contain it; it is also relevant (see Fig. 1). In total, there are 208 unique web pages retrieved by the five engines, where 179 pages are only found by a single engine.

As noted, users can select & drag a circular ring to apply a "focus+context" transformation so that specific pages and their titles can be better examined, since the titles are recomputed once the mouse control is released (see Fig. 1). Users can also place the cursor over an item icon and a "details-on-demand" display will appear that shows the complete title, abstract, URL and a bar chart with the rank positions [21].

Users can also interact with a "list depth" slider to specify how many of the top 50 documents to display (without the need to fetch the results again). Fig. 4 shows the overlap between the result sets if only the *top 20* of the 50 results retrieved per engine are compared. The "Compare Search Home Page" is now the result closest to the center, whereas the "Sherry Koshman, Ph.D." page is only in two of the top 20 lists.

3.2. Enhanced Rank List

searchCrystal offers users an enhanced Ranked List tool so that users can use a familiar display to scroll through a large result set. This tool makes it possible to display detailed information about each of the found items (see Fig. 5). Users can control whether to display the



a) item icon found by the *two* engines Yahoo and MSN, since it has a rectangular shape colored red and green; the item is placed toward the *bottom* of the two results lists, since its size is small and almost transparent.

b) item icon found by *four* engines: Google, Yahoo, Ask and Exalead; item is placed toward the *top* of the results that contain it since its size is large and is almost opaque.

Fig. 1: Cluster Bulls-Eye visualizes the overlap between the **top** 50 search results returned by Google, Yahoo, MSN, Ask and Exalead, respectively, for the search query "**exploratory search result visualization**." The items found by the same number of engines are placed inside the same circular ring. Users can perform a "focus+context" transformation by selecting and dragging a circular ring.

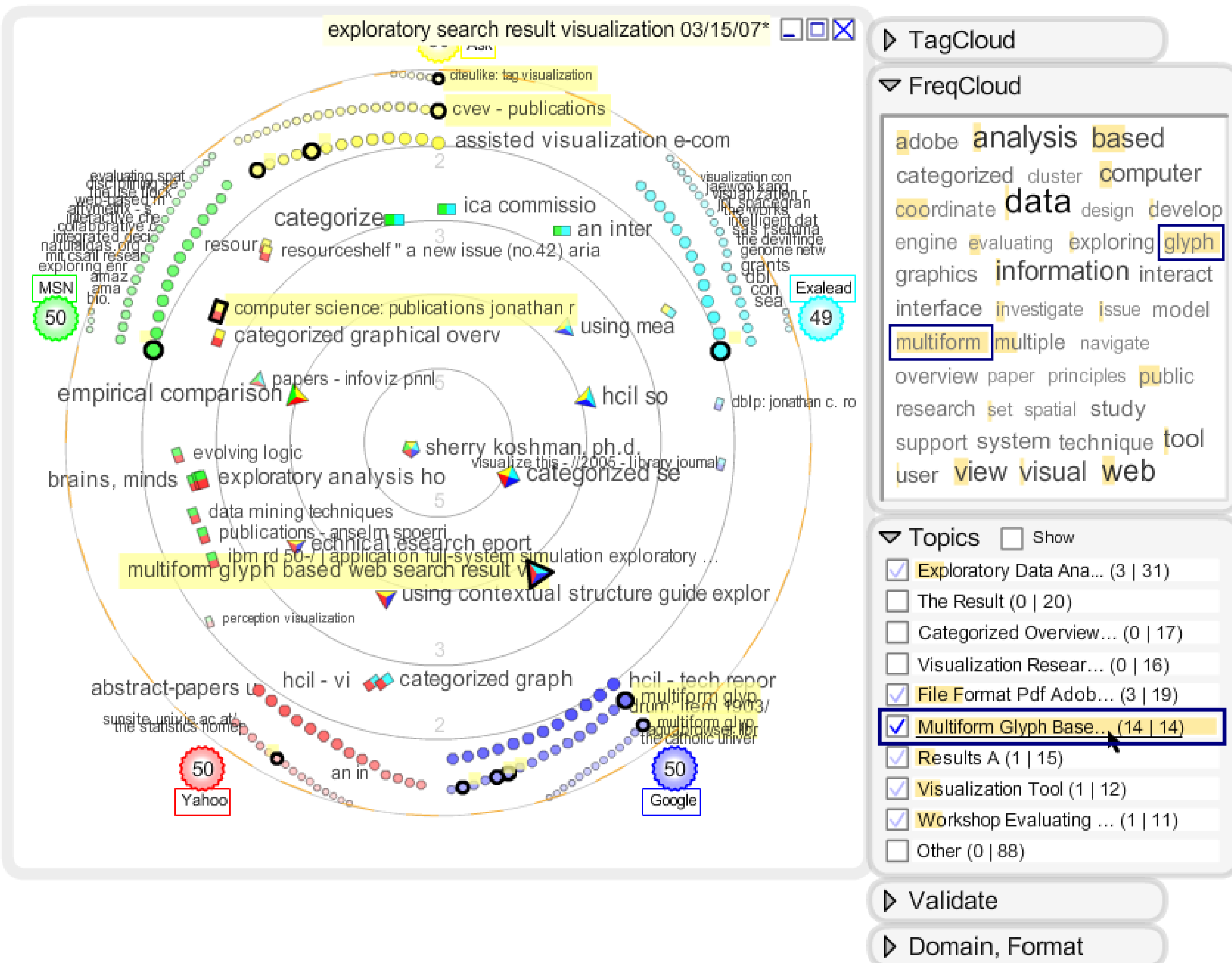


Fig. 2: shows the expanded and magnified **FreqCloud** and **Topics** panels, which are docked on the right side of a crystal. The topic "**Multiform Glyph** ..." is selected and the panels show how this topic is related to other topics and most frequent words. Further, if the mouse is placed over the selected topic, then the items related to this topic are highlighted in the Cluster Bulls-Eye display.

abstract, URL and/or rankings information, where the latter is displayed as a stacked bar chart, whose total length is proportional to the average of the list positions of an item. Fig. 5 shows the Rank List tool and the top four web pages. Below each item icon, a number is displayed that is equal to the overall rank assigned to the found page, which by default is based first on the number of engines that found the page and second on the page's average list positions. Specifically, the found web pages are sorted based the sum of the engines that found a page and the average of its rank positions, where the latter is normalized to a number between 0 and 1. searchCrystal offers users different ways to compute the overall rank of the found items and thus where the item icons are displayed with respect to the display center.

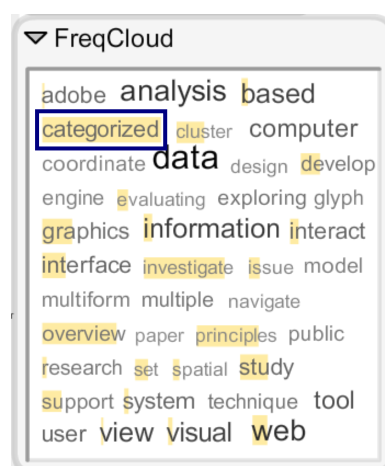


Fig. 3: shows FreqCloud if word "**categorized**" is selected.

3.3. Cloud Panels

searchCrystal supports two types of "cloud" panels that are docked on the right hand side of a crystal (see Fig. 2 right). The *TagCloud* shows the tags that have been added by users to specific search results. Users can add tags by using the Ranked List tool, selecting the "Tags" checkbox and entering tags in the text entry field that becomes visible below the URL of each item. The *FreqCloud* can show the most frequent words that appear in the titles and snippets of the search results (see Fig. 2 right, where the 40 most frequent words are shown). The terms used in the search query are excluded, since they tend to be present in all pages and thus would not provide any new information to help users to understand or filter the search results. Both cloud panels visualize the percentage of items associated with a specific tag or frequent word that are currently selected. This makes it possible to visualize the co-occurrence of specific tags and frequent words. Regular tag clouds do not make this type of information visually explicit.

For the query "exploratory search result visualization," words such as "data," "analysis," "web", "visual" and "tool" have a high frequency (see Fig. 2 right). These words are quite general, whereas "categorized," "cluster," "glyph" and "multiform" are more distinctive and informative. These words suggest that categorization and clustering methods are being considered to enable users to explore search results. If the word "categorized" is selected, then the words "overview," "cluster," and "investigate" co-occur, whereas the word "visual" is weakly correlated (see Fig. 3). This suggests that the categorization efforts do not rely heavily on visualization methods to assist users in their search exploration process. Now, this inference is qualitative in nature. However, this type of qualitative reasoning can help users in their sense-making process [2]. Users can use the FreqCloud to gain quick insights into the dominant concepts that describe the contents of the search results. Depending on the search query, the FreqCloud can "tell a story in keywords" and help users to select better search terms or ways to narrow the search.

3.4. Topics Panel

searchCrystal employs the open source version of Carrot2 clustering framework [4] to identify major topics in the search results. The clusters are allowed to overlap and the open source version of Carrot2 does not support hierarchical clustering. As shown in Fig. 2 (right), the "Topics" panel lists only the ten largest topical clusters in order not to overwhelm users with a long list of topics. If there are more than ten clusters, then the first nine clusters are listed and the remaining clusters are combined and assigned the "Other" label.

If users select a topic checkbox or topic name in the Topics panel, then the length of an orange colored bar will indicate that all the items, which are associated with the topic, are selected (see Fig. 2 right). If any selected items are also associated with other topics, then the length of the orange bar for those topics will indicate the percentage of their items that are selected and are thus shared. The numbers at the end of a topic name indicate how many items are selected out of the total number of items associated with the topic. The Topics panel makes it possible to see how the topics are related and which items are members of multiple topics. Further, if the mouse is placed over a topic, then the items related to this topic are highlighted in the main display (see Fig. 2 right).

The Carrot2 clustering framework is used since it aims to compute topic labels that tend to be intelligible and useful for human users [4]. Anecdotal evidence suggests that the created topics labels have room for improvement. As shown in Fig. 2 (right), the topics "The Result" and "Result A" appear on the surface to be highly related, yet they have only one page in common. Further, these two topic labels are not very informative. The commercial version of Carrot 2 supports extensive tuning and ways to improve the topic labels.

The clustering capability in searchCrystal is a "nice to have" feature, whereas for Clusty and Grokker, it is a "must have" feature, since their visual presentation heavily depends on the quality of the identified topical clusters. As searchCrystal, Grokker allows the topical clusters to overlap, but it is difficult to identify which items are contained in multiple clusters in the Grokker interface. This can have the effect of disorienting users, if they encounter items already examined, but they do not understand why this possible. Further, if users select a topic in Clusty and Grokker, then they can only see the items associated with the selected topic, whereas searchCrystal enables users to see these items in the context of all the retrieved items.

4. Visualizing Relationships

The fact that *all* found items are visible in the Cluster Bulls-Eye tool makes it possible to visualize structural relationships between the search results, such as which items are part of the same topical cluster or are stored on the same web host. For example, connecting lines can be drawn between the items that have the same host address. The item that is closest to the center of the display is selected to be the "hub" from which pinkish lines emanate to connect items with the same host (see Fig. 4). A connecting line can help to corroborate the relevance of connected items. The information that two items are stored on same host can be useful to find other relevant items or to identify (close) duplicates.

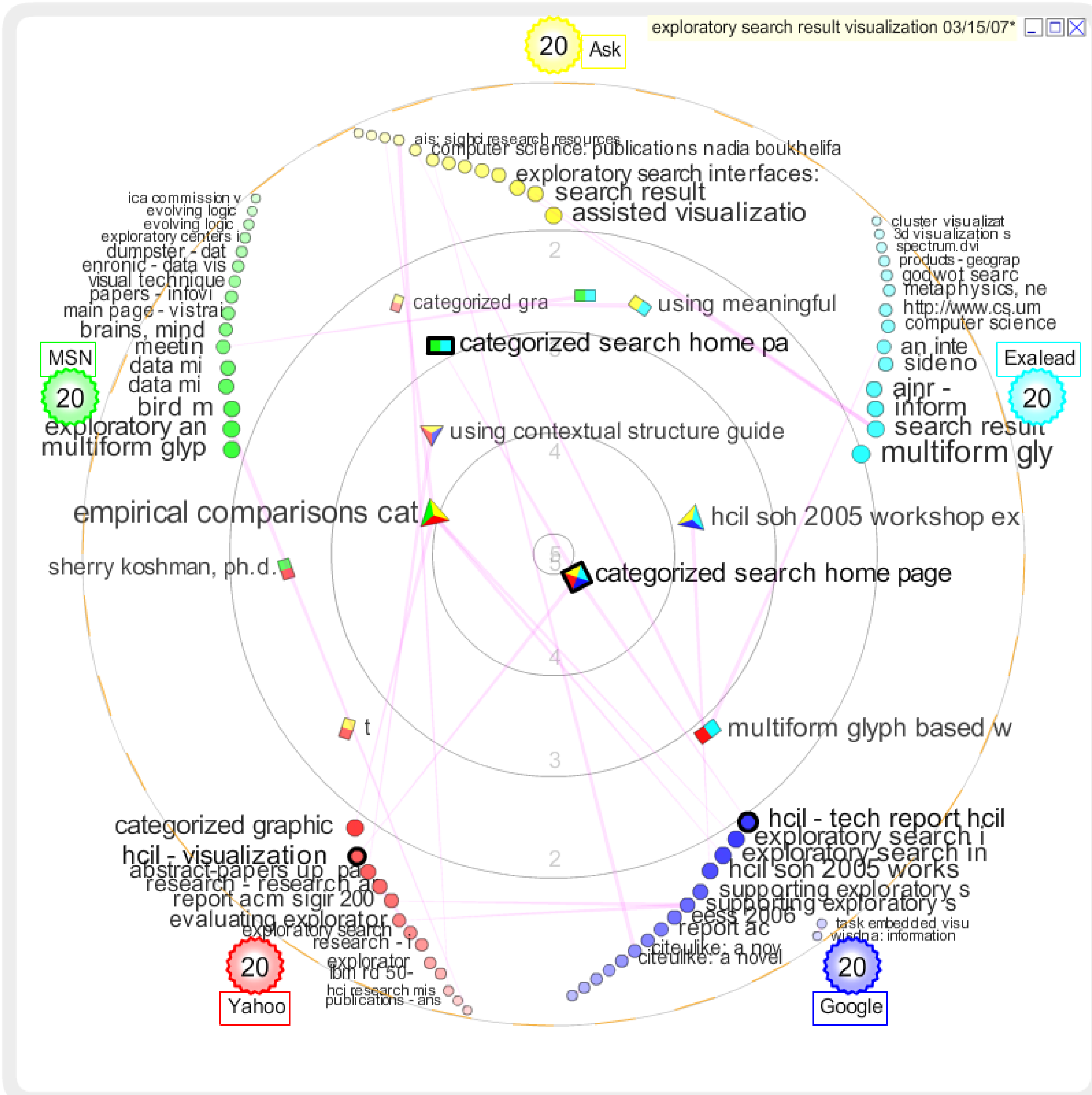


Fig. 4: Cluster Bulls-Eye displays the overlap for the top 20 results for Google, Yahoo, MSN, Ask and Exalead, respectively, for the query “exploratory search result visualization.” Pages stored on the same host are connected by pinkish lines; the four pages stored on “cs.umd.edu” are selected and their item icons have black borders.

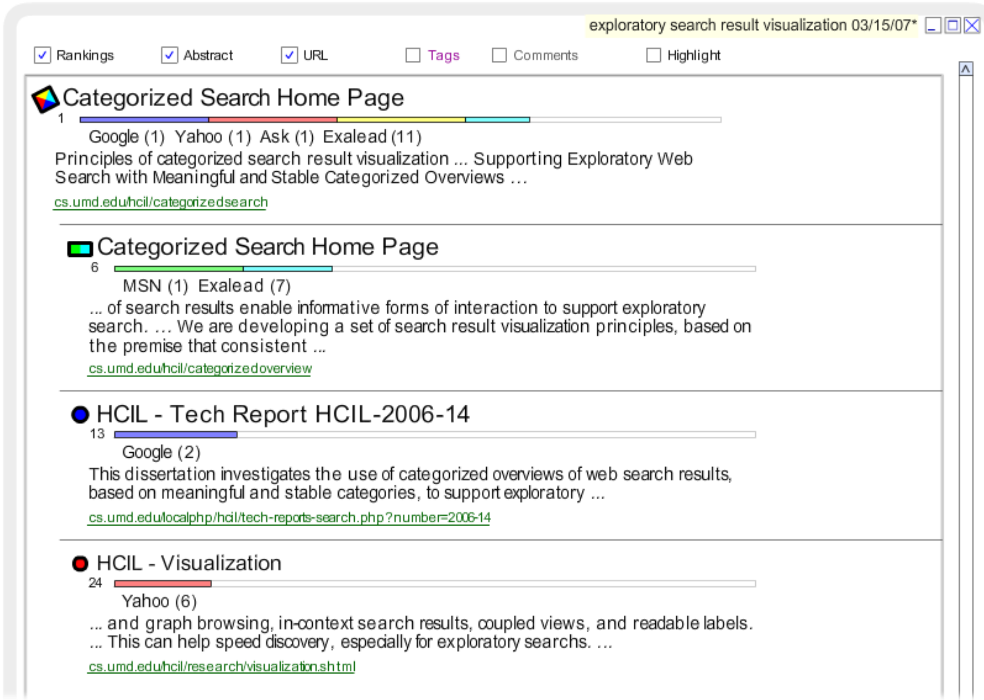


Fig. 5: shows the Ranked List, where the Rankings, Abstract and URL checkboxes are selected. Only the four selected web pages, which are stored on "cs.umd.edu," are shown and the pages, which do not have the highest overall rank, are indented.

In the Ranked List, the "web host structure" can be visualized by sequentially displaying pages stored on the same host and those pages, which do not have the highest overall rank, are indented (see Fig. 5). This makes it easy for users to find the pages that are hosted on the same server. On the one hand, the overall ordering of the pages, which by default is based on the number of engines that have found them and their average rank positions, is lost in the Ranked List. On the other hand, the overall ranking of these pages is preserved in the Cluster Bulls-Eye view. Figures 4, 5 and 6 illustrate how the Cluster Bulls-Eye and Ranked List tools can be used in concert to support users in the search process.

searchCrystal can accommodate the visualization of different types of relationships by using connecting lines that have unique colors to represent different relationship types, such as the web host structure or topical relationships. Blue lines are used to connect pages that are related to the same topic. In Fig. 6, the Cluster Bulls-Eye tool shows the overlap between results for Google, Yahoo, MSN, Ask and Exalead, respectively, if only the top 20 of the 50 pages returned are compared. If a user double clicks on the item icon from which the blue lines emanate for the "Categorized Overviews" topic (see Fig. 6 top right), then all the items, which are related to the same topic and/or stored on the same host, are selected. Fig. 6 (top) only shows the selected pages in the Cluster Bulls-Eye view, since users can choose to show / hide the (non)selected items. Fig. 6 (bottom) shows the same selected items in the Ranked List and the assigned topics are displayed in blue. All the selected pages have the topic "Categorized Overviews" in common and there is one page that is related to these three topics: "Categorized Overviews," "Exploratory Search Interfaces" and "Information Visualization." The Ranked List makes it easy to spot items that are related to multiple topics, whereas the Cluster Bulls-Eye enables users to see the "big picture" as well as the degree of agreement between the search engines and how the search results are related.

5. Discussion and future research

Text clustering methods tend to compute a co-occurrence matrix of the words, where common words are excluded by applying stop word lists. The FreqCloud visualizes a key data structure used by clustering methods to identify major topics. Anecdotal evidence suggests that there can be a high correlation between major words in the FreqCloud and the topics identified by the Carrot2 clustering engine. For example, if the frequent words "glyph" or "multiform" are selected in the FreqCloud (see Fig. 2 right), then all web

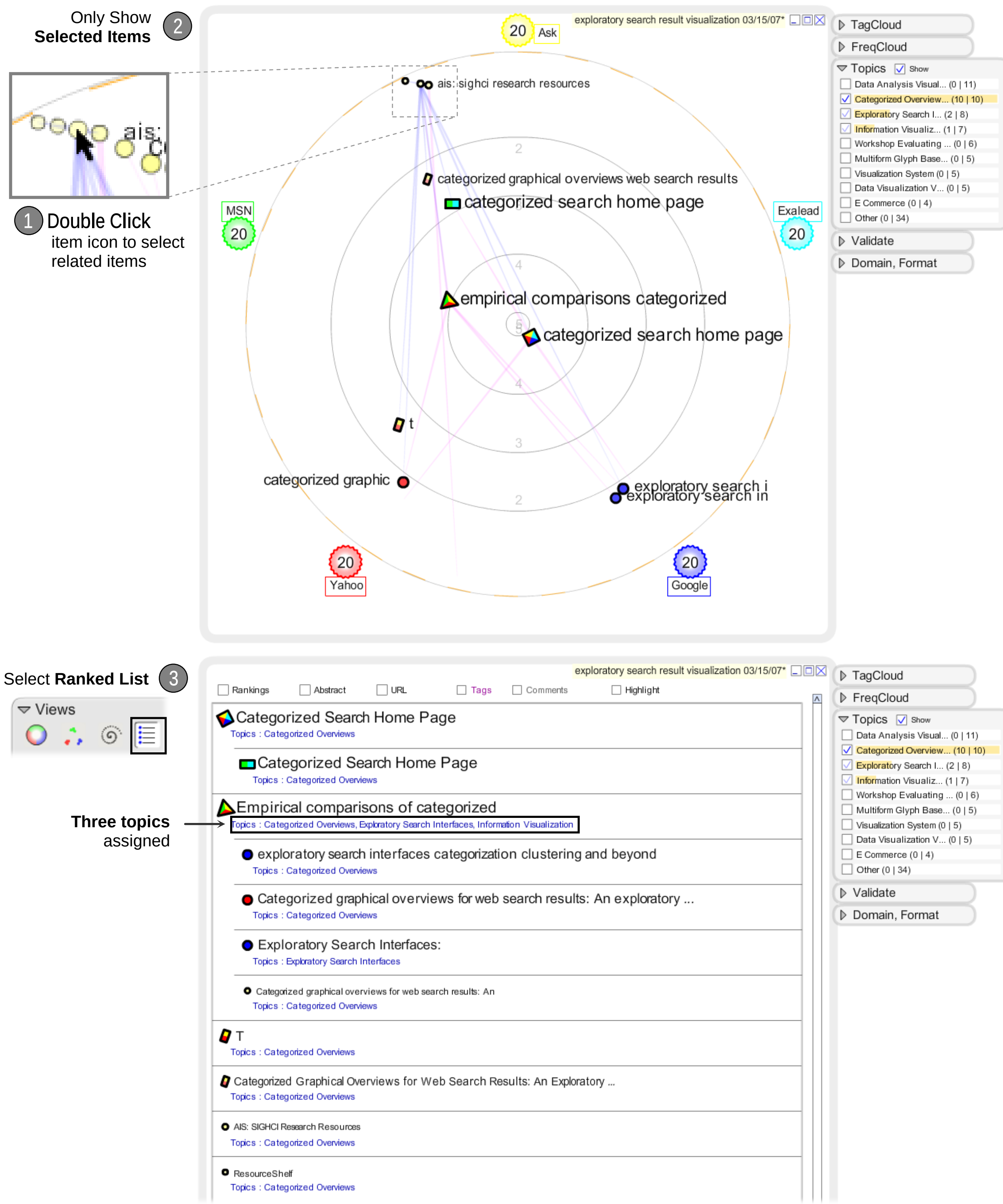


Fig. 6: (1) if users **double click** on an item icon in Cluster Bulls-Eye display, then all its related items are selected. (2) shows only the selected items in the Cluster Bulls-Eye view. Blue lines connect items related to the same topic; pinkish lines connect items stored on the same web host. (3) shows the selected pages in the Ranked List and the assigned topics are displayed.

pages associated with the topic "Multiform Glyph …" are selected and visa versa. Now, "glyph" and "multiform" are specific and specialized words. If "research," another frequent word, is selected, then 15 of the 16 pages associated with the topic "Visualization Research …" are selected. This raises the question if the FreqCloud captures much of the value provided by a clustering method. The FreqCloud suggests "fuzzy" labels that are less brittle since topic labels do not have to be assigned in a binary way. Instead, multiple words can be used to describe in varying degrees the found items. Future research will investigate how the FreqCloud computation can be improved to take better advantage of text retrieval methods [1].

As noted, for the "exploratory search result visualization" query, 179 pages are only found by a single engine when the top 50 results returned by Google, Yahoo, MSN, Ask and Exalead are compared. This implies that on average at least 70 percent of an engine's results are only found by it and are not corroborated by other engines. Further, only one page is recommended by all five engines; only one page by four of the five engines. This suggests that there is some, but not a great deal of agreement between the search engines. The Cluster Bulls-Eye and the Category View [21] in particular provide users with a quick insight into the degree of agreement between the "experts" being consulted – the more items that are located toward the display center, the greater the agreement. The fact that there is not a great deal of agreement for the "exploratory search result visualization" query can indicate that this is still an emerging field of research. If the top 50 results are compared when searching for "information visualization," then the agreement between the search engines is much greater and on average 50 percent of an engine's results are also shared with another engine. Finally, it is important to keep in mind that the results returned by search engines can change on a daily basis. searchCrystal makes it possible for users to compare search results across time to detect how the results have changed.

As shown in Fig 1, the search engine inputs are mapped to specific locations along the periphery of the Cluster Bulls-Eye. For example, the Google results are displayed in the bottom right since the items only found by Google can be organized in such a way that a greater number of title (fragments) can be shown. Further, the titles can be right aligned next to the item icons and read top to bottom to examine items in the ranked order in the Google result list. For the other inputs, the titles need to be left aligned next to the item icons or read bottom to top (see Fig. 1).

A user study [26] has been conducted to determine if novice users can use the RankSpiral and Cluster Bulls-Eye displays to find the documents that are most likely to be relevant. Specifically, it was shown that novice users can use the provided visual cues, such as the icon's shape and position, to decide which icons to explore first to find highly relevant documents. This ability is a prerequisite for users being able to make effective use of searchCrystal's full functionality. The user study also showed that novice users could identify highly relevant documents more rapidly and accurately by using the RankSpiral than the Cluster Bulls-Eye tool. Further, it was shown that users tend to use the distance from the center as the primary visual cue when deciding which icons to select and in which order. This raises the question if it would be advisable to relax some of searchCrystal's design principles, such as mapping items found by the same number of engines into the same concentric ring. The specific data fusion strategy used to sort the found items affects how the item icons are visually organized. searchCrystal offers users different ways to fuse and sort the different results lists so that users can select to have both an item icon's size and its distance from the center can be used to encode the item's probability of being relevant. This is equivalent to only using an item's average rank positions, but ignoring how many engines have found the item. Further user studies are in the planning stages to test searchCrystal's extensive functionality. In particular, a user study is being prepared to test if users can identify a sufficient number of known relevant documents and the Cluster Bulls-Eye tool will be compared with a standard ranked list display. It will be tested if the searchCrystal interface will lead to improved search performance, better user satisfaction and greater confidence in the relevance of the selected documents.

6. Summary

This paper has addressed the question of how visual tools can support users in exploratory search. The searchCrystal toolset has presented and the Cluster Bulls-Eye and Ranked List tools were described in more detail. It was shown that the linear Ranked List and the 2D Cluster Bulls-Eye have complementary strengths and can be used in concert to support users in the exploratory search process. Several tightly coupled controls were introduced that let users explore large results sets based on the most frequent words, assigned tags or the major topics. These controls enable users to see how the different topics, tags and frequent words are related by visualizing the percentage of items associated with a specific topic or frequent word that are currently selected. In particular, the FreqCloud control visualizes how the most frequent words co-occur and are correlated, which something a regular tag cloud display does not do. It was also shown how the Cluster Bulls-Eye tool can visualize structural relationships between the search results, such as the web host structure or topical relationships.

7. Acknowledgments

The author would like to thank Alexander Stanton for his development help with searchCrystal and its data acquisition and server components in particular. Further, the author would like thank Marija Dalbello for her continual support and encouragement as searchCrystal continues to be developed.

8. References

[1] Baeza-Yates, R. and Ribeiro-Neto B. (eds.) Modern Information Retrieval. Addison-Wesley, 1999

[2] Card S., Mackinlay J. and B. Shneiderman (Eds.), Readings in Information Visualization: Using Vision to Think. San Francisco: Morgan Kaufmann.

[3] Carey, M., Heesch, D., Rüger, S. (2003). Info Navigator: A visualization tool for document searching and browsing. Proceedings of Int. Confernce on Distributed Multimedia Systems (DMS 2003).

[4] Carrot2 Clustering Framework (2007). Comparison of the Commercial and Open Source version of the Carrot2 Clustering Engine. Accessed 14 March 2007 at http://company.carrot-search.com/lingo-3g-vs-classic.html

[5] Chen H., Fan H., Chau M. and Zeng D. MetaSpider: (2001). Meta-Searching and Categorization on the Web. JASIST, Volume 52 (13), 1134 - 1147.

[6] Clusty – www.clusty.com

[7] Clusty Cloud - http://cloud.clusty.com/

[8] Delicious – http://del.icio.us/

[9] Eyetools Inc., 2005. "Eyetools, Enquiro, and Did-it uncover Search's Golden Triangle," accessed 20 February 2007 at http://www.eyetools.com/inpage/research_google_eyetracking_heatmap.htm.

[10] Flickr – www.flickr.com

[11] Grokker – www.groxis.com

[12] Hearst M. (1999) User interfaces and visualization. Modern Information Retrieval. R. Baeza-Yates and B. Ribeiro-Neto (eds.). Addison-Wesley, 257-323.

[13] Havre, S., Hetzler, E., Perrine K., Jurrus E., and Miller N. (2001). Interactive Visualization of Multiple Query Results. Proc. IEEE Information Visualization Symposium (InfoVis 2001).

[14] Kartoo – www.kartoo.com

[15] Mann T. (2002) Visualization of Search Results from the WWW. Ph.D. Thesis, University of Konstanz.

[16] Marchionini, G. (2006). Exploratory Search: From Finding to Understanding, Communications of the ACM, v.49 n.4, April 2006.

[17] Olsen, K. A., Korfhage, R. R., Sochats, K. M., Spring, M. B., & Williams, J. G. (1993). "Visualization of a Document Collection: the VIBE System", Information Processing & Management, 29(1), 69-81.

[18] Roberts, J., Boukhelifa, N., Rodgers, P. (2002). Multiform glyph based web search result visualization, Proceedings of Sixth International Conference on Information Visualisation (IV 2002).

[19] Roberts, J. Suvanaphen, E. (2003). Visual Bracketing for Web Search Result Visualization, Proceedings of Seventh International Conference on Information Visualization (IV 2003).

[20] Spoerri, A. (1999). InfoCrystal: A Visual Tool for Information Retrieval. In Card S., Mackinlay J. and B. Shneiderman (Eds.), Readings in Information Visualization: Using Vision to Think, pp. 140-147.

[21] Spoerri, A. (2004). Visual Editor for Composing Meta Searches, Proceedings of the 67th Annual Meeting of the American Society for Information Science and Technology (ASIST 2004), volume 41, issue 1, pp. 373-382.

[22] Spoerri, A. (2004). RankSpiral: Toward Enhancing Search Result Visualizations, Proceedings IEEE Information Visualization Symposium (InfoVis 2004), pp 18.

[23] Spoerri, A. (2004). "Coordinated Views and Tight Coupling to Support Meta Searching," Proceedings of the 2nd International Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV 2004), pp 39-48.

[24] Spoerri, A. (2004). Toward Enabling Users to Visually Evaluate the Effectiveness of Different Queries or Engines. Journal of Web Engineering. Vol. 3, No.3&4 (2004) 297-313.

[25] Spoerri, A. (2005). How the Overlap Between the Search Results of Different Retrieval Systems Correlates with Document Relevance. Proceedings of the 68th Annual Meeting of the American Society for Information Science and Technology (ASIST 2005).

[26] Spoerri, A. (2006). Visualizing Meta Search Results: Evaluating the MetaCrystal toolset. Proceedings of the 69th Annual Meeting of the American Society for Information Science and Technology (ASIST 2006).

[27] Spoerri, A. (2007). Visual Mashup of Text and Media Search Results. Proceedings of Eleventh International Conference on Information Visualisation (IV 2007).

[28] White, R., Kules, B., Drucker, S., Schraefel MC (2006). Supporting exploratory search. Communications of the ACM 49:44, 36-39.

[29] Wise, J.A., Thomas, J.J., Pennock, K., Lantrip, D., Pottier, M., Schur, A., Crow, V. (1995). Visualizing the non-visual: spatial analysis and interaction with information from text documents," Proceedings IEEE Information Visualization Symposium (InfoVis 1995).