Sherry Koshman*

University of Pittsburgh Pittsburgh, PA

Amanda Spink

Queensland University of Technology Brisbane, Australia

Bernard J. Jansen

The Pennsylvania State University University Park, PA

Chris Blakely

InfoSpace, Inc. Bellevue, WA

Jonathan Weber

University of Pittsburgh Pittsburgh, PA

Metasearch Result Visualization: an Exploratory Study

Abstract: The Missing Pieces tool visualizes the overlap of search engine results including those generated by the metasearch engine, Dogpile. The major research question is: how well can users interact with and interpret the circular metasearch results display? This study has interesting implications for the use of visualization for presenting information in a web metasearch environment.

Résumé : L'outil "Missing Pieces" visualise le chevauchement des résultats des moteurs de recherche incluant ceux générés par le métamoteur Dogpile. La principale question de recherche est : de quelle manière les utilisateurs peuvent bien utiliser et interpréter l'affichage des résultats des métamoteurs circulaires ? Cette étude met en relief des implications intéressantes pour l'utilisation de la visualisation pour la présentation de l'information dans un environnement de métarecherche web.

1   Introduction

Information visualization techniques have been applied to web information retrieval in an attempt to enhance the user's searching experience. Visualization refers to the application of graphics and symbols to depict information items such as web pages as opposed to the traditional linear search result ranked lists such as those generated by Google and Yahoo.

Theoretical foundations in cognition and perception support the design of visualization interfaces and their application to information retrieval (IR) environments (Ware, 2004). Web IR presents unique challenges since the information is distributed, unstructured, and volatile. Techniques have been developed to visualize relationships between the query and the set of retrieved items, and to visualize the relationship among the search output items so that an abstract document space is made manifest (Koshman, Forthcoming).

Web metasearch engines provide the user with a single solution to entering queries and retrieving results from multiple web search engines. Dogpile (InfoSpace, Inc.), a metasearch engine on the web, created the Missing Pieces visualization tool to depict the overlap of the first page of retrieved search results from Yahoo!, Google, and Ask Jeeves. This paper reports on the first user study with the Missing Pieces visualization.

2   Related Studies

Visualization tools for web information retrieval have focused on the display and navigation of search result sets. Grokker's Map View[1] uses geometric icons and a zoomable clustering

interface to depict the results from Yahoo!, Wikipedia, and/or Amazon Books. The larger the icon, the more items it contains. Users may change the icon type and color scheme of the map as well as export the map to a file. Once an item is identified, the user may click on it and a thumbnail image appears in the split view of the interface.

Kart00[2] offers a metasearch visualization based on a cartographic map metaphor. Query results are depicted as document icons on the screen and a legend helps to decode the icons. The left frame contains a topic list and the right frame displays sponsored sites. User queries may be retrieved in a history function listed under "last queries". Several maps may be generated from one search and the user may browse through the maps of search results. If a single item is selected, then a thumbnail appears and the user may visit the site.

Clusty[3] is a cluster-based metasearch engine that uses the familiar tree metaphor to expand and contract an outline containing clusters which are dynamically generated from a user's search results. The left frame contains the cluster tree and the right frame contains the search results list. Users may cluster their search results according to different categories for web searching (e.g. topic, source, URL).

Spoerri developed a model called MetaCrystal which visualizes the overlap results between search engines (Spoerri, 2004). MetaCrystal uses a layout where the items retrieved by all the search engines are positioned in the center of the display and the use of color, shape, and proximity are used to represent results returned by different search engines. MetaCrystal has a category view that presents the overlap for the top 100 documents found on the search engines and a cluster bulls-eye view which visualizes all the retrieved documents in a concentric ring.

User studies with visualization tools include Heo and Hirtle (2001) who investigated Web visualizations with 80 participants using a large-scale and small web space containing 583 pages and 50 pages respectively. Subjects performed tasks using distortion, zoom, expanding outline, and the control (text) web interfaces. The study showed that performance did not improve with the visualization tools over the control interface, however the expanding outline was shown to be more useful to users and the zoom technique was the most difficult.

Rivadeneira and Bedersen (2003) conducted a user study with an earlier version of Grokker, Grokker text, and the Vivisimo search engine which provides clustered search results. They tested 15 users with factual information queries and showed that the objective measures did not show statistically significant differences and participants preferred the Vivisimo interface.

A web-based node link visualization, TouchGraph[4], was studied by Koshman (2004). TouchGraph is an open source product available on the web that provides a visualization of items according to Amazon's recommendation feature based on item similarity and Google's "similar to" pages. This research examined how the level of topic knowledge affected similarity-based selections made from the visualization by participants and how participant item selections compared to the system generated output of most similar items to the query. 17 participants were tested with tasks designed for the Amazon and Google implementations. The results showed that there was a high amount of overlap between the participant selected items and the system generated set. The relationship between knowledge level and interface node use was statistically significant for half of the tasks and positive subjective satisfaction data were reported.

Dogpile (InfoSpace, Inc.) in conjunction with academic researchers studied the overlap strength of search engine results from Yahoo!, Google, Ask Jeeves, and MSN for the first page of retrieved results (Spink, et al. Forthcoming). The Missing Pieces tool was developed to visualize the overlap of these search engine results in comparison to the search results generated by Dogpile. The assumption behind the overlap study design is that web searchers typically consult only the first page of web search engine output to make selections. This visualization tool differs from current operational web metasearch visualizations that focus on displaying search results

which can be refined by source, but they do not illustrate the overlap relationships among similar items from different sources. The current investigation tests users with the newly developed Missing Pieces tool to visualize search result overlap.

3   Research Questions

The major research question in this study is: how well can users interact with and interpret the circular metasearch results display?   Related questions include: 1) What is the accuracy with which users selected items from various areas of the display?, 2) What is distribution of participants' selection of items in the rings of the display?, and 3) What is the level of the participants' subjective satisfaction?

4   Study Design

Since the tool was available over the web and printed self-training materials were distributed electronically, a novel methodology used a combination of remote usability testing and on site evaluation. Shneiderman and Plaisant (2004) describe remote usability testing as having advantages for testing web-based applications since a larger number of subjects may be tested. Since participants performed tasks in their own environments with their own machines a more realistic test setting was achieved with remote testing in combination with on-site evaluation. The study was exploratory in nature since it was to test the evaluation methodology and initial responses from users with a new visualization tool. Data collection instruments were pilot tested with 40 online and on-site participants.

4.1   Participants

The thirty-two participants were graduate student volunteers and over half (67.2%) rated themselves as using web resources many times a day. The majority of participants had the study administered online (87.5%) through a web-based course management tool, and the gender distribution was 90.4% female and 9.6% male.

4.2   System

The Missing Pieces tool has three concentric circles used to depict the overlap of the user's query entered in a standard search box (Figure 1). The outer layer represents a single search engine's results, the second layer represents two search engine's results (pairwise) and the center layer represents the results of all three search engines. Colored circular icons represent the retrieved items. Green icons appear in the bull's eye, blue icons in the middle ring, and white icons in the outer ring. Icons containing a paw print represent Dogpile results. The icons may be browsed by mouseover techniques which pop-up a box containing the web site uniform resource locator (URL). The location of items in the circle does not have a direct interpretation. Searchers may highlight each of the three search engine sectors which then reveals the "missing pieces" of information items as red circular icons. Interactivity is facilitated when the user clicks on one search engine and the remainder of the display uses color to indicate the items that would not have been retrieved by that search.
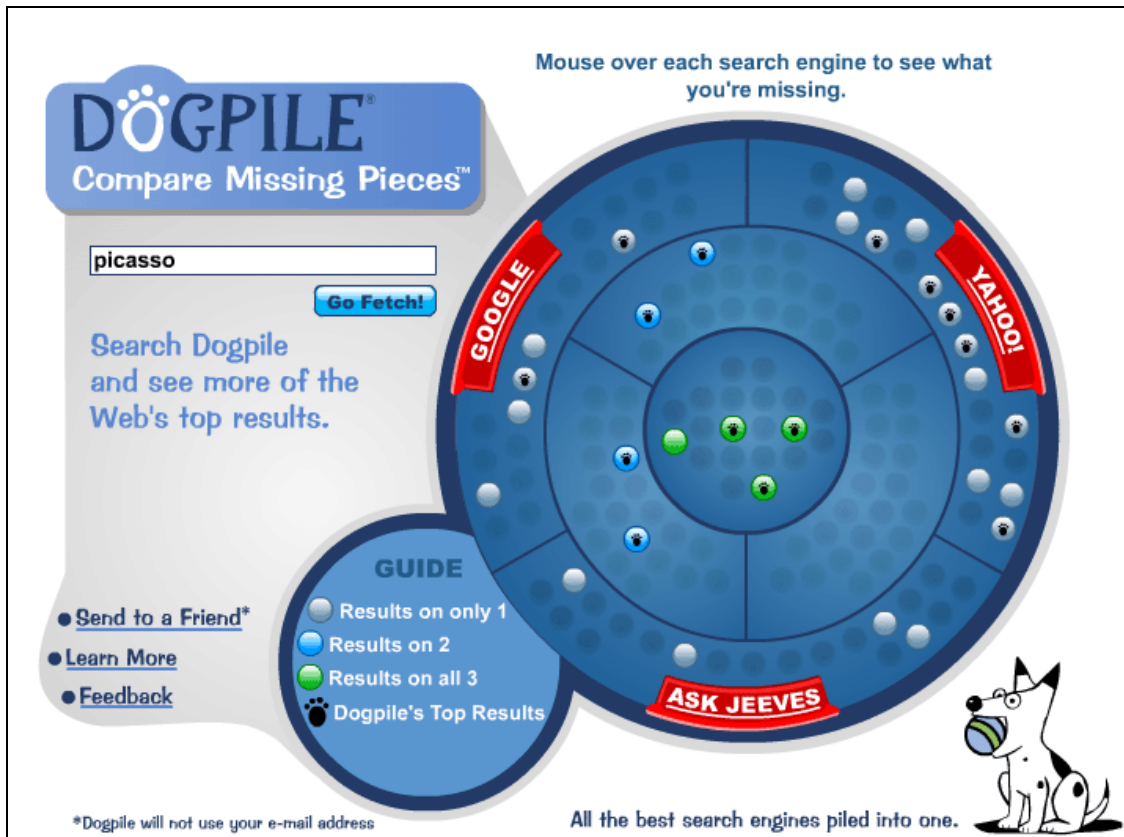
Fig. 1: Missing Pieces Visualization

4.3 Tasks

Participants completed a pre-search and post-search questionnaire along with 21 tasks on three topics: "Picasso", "Beethoven" and "gardening" using the Missing Pieces tool. Tasks were deliberately designed to be of general interest without requiring specialized domain knowledge. The query terms were provided and were counted as the initial task for each set.

Tasks were designed according the visualization's functionality of categorizing, highlighting, and distinguishing search results. Each set of tasks contained questions relating to the display such as identifying the search engine that has the most results, listing a URL that is common to two of the search engines, and counting results that were found on all three search engines. Participants were also asked to list URLs that they would have missed if only using one search engine and also to list five URLs that they thought were most useful to the query topic.

5   Results

5.1   Accuracy of Selected Items

An interesting pattern emerged among the three sets of tasks for answering how many results are found on all three search engines. The answer was 5 results for each task scenario. The percentage of participants who correctly answered the question was 68.8% (gardening task), 56.3% (Beethoven task), and a much lower 28.1% for the Picasso task.

The next question asked participants to identify which search engine had the most results. The majority of participants identified the correct search engine in each task and they were 98.3% (Beethoven), 90.6% (gardening), and 84.4% (Picasso).

4

Participants were asked to estimate the percentage of Dogpile search results on the visualization, which were indicated by a dog paw symbol. The categories included "most" items, "some", "few", "none", and "can't determine". Over half of the participants (62.5%) identified the "some" category defined by a percentage range of items on the Beethoven task and 21.9% of the participants identified the "most" category range which reflects a more accurate range of results. For the gardening task, 56% of the participants correctly identified the "some" category. The Picasso task was inconclusive.

Participants were asked to identify two URLs of items that would have been missed had the search been conducted on only one search engine. The responses were complex in that varying percentages of participants selected a range of items on the display that met this criterion. The interesting pattern in these results showed that in general 94-97% of the participants responded correctly to this question and were able to select items that would have been missed.

The last task asked participants to identify no more than five URLs that they thought were most useful to the query. For each task set three URLs were selected by 40% or more of the participants and eight out of these nine URLs were Dogpile designated results.

5.2   Distribution of Items in the Display

The following three figures show that the highest number of participant responses were not always in the bull's-eye; however, they indicate a general trend of more responses gravitating form the center to the outer circle. The range of items along the x-axis varied according to the items selected by the participants.

Figure 2 shows that the Beethoven task item selections were primarily made from the bull's eye and the outer ring of the visualization which primarily contained items listed by single search engines.
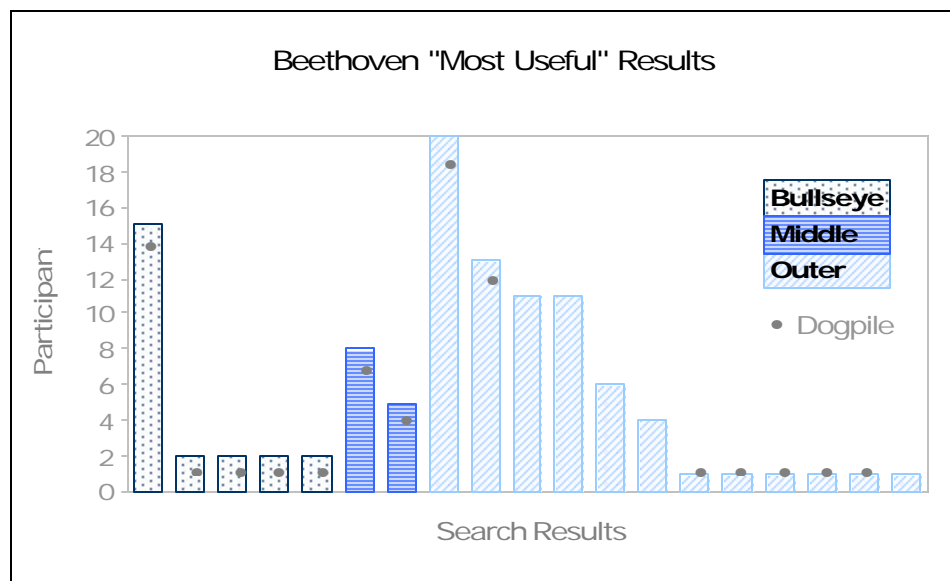

Fig. 2 Beethoven Visualization Participant Response Distribution

Figure 3 shows that the distribution of item selections in the gardening task by participants from the visualization primarily fell in the bull's eye of the display. These items were common to all three search engines and were all Dogpile results and the descending slope of the distribution indicates that items selected radiated from the bull's eye for this task.
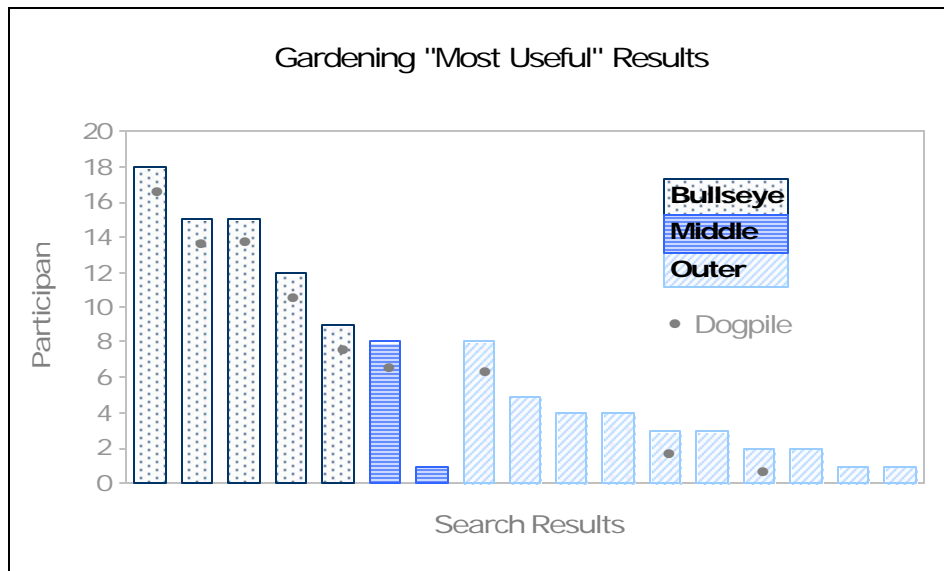
Fig. 3 Gardening Visualization Participant Response Distribution

Figure 4 shows the item selection distribution where the largest number of participants selected items from the bull's eye and middle rings of the display. Dogpile results were scattered throughout the distribution for the Picasso task.
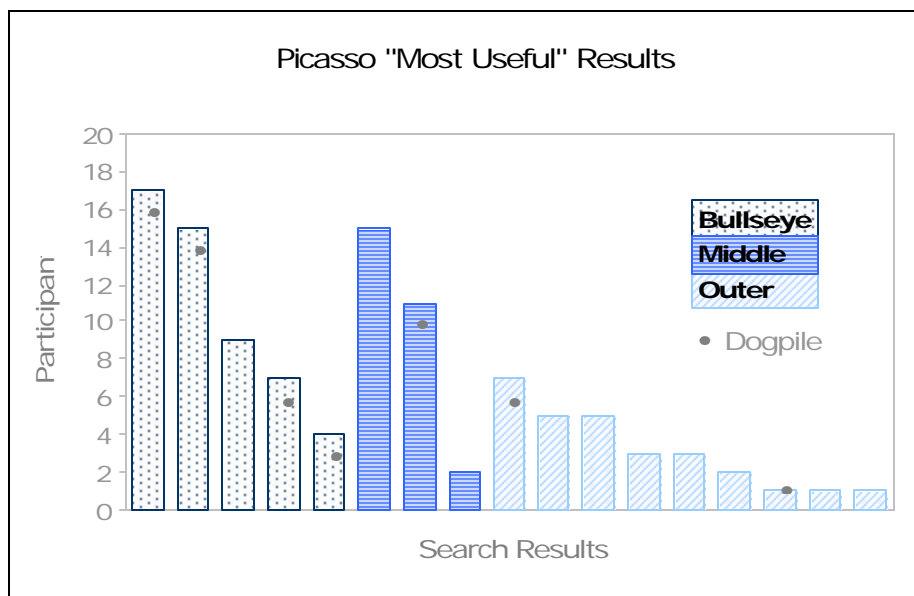

Fig. 4 Picasso Visualization Participant Response Distribution

5.3   Level of Subjective Satisfaction

In the post-task questionnaire the majority of the online participants (82%) rated the tasks as "easy" or "moderately easy". The respondents rated their primary criterion for selecting search results as URL (42.8%), proximity to center/location on the display (21.4%), paw print (17.9%), and color of the dot (7.1%). The majority of participants (72.9%) rated the Missing Pieces display as "intuitively easy" or "moderately easy" to understand. Participants were asked if the tool enabled them to identify items that would have been missed if they had searched with only one search engine. 43.6% of them responded in the "all of the time" category and 43.6% responded in the "frequently" category. The majority of participants categorized the rate which

6

the Missing Pieces tool as providing a sufficient number of results on the screen, and 64.3% selected "frequently" and 25% selected the "all of the time" category. Over half of the participants (53.6%) rated the Missing Pieces tool as presenting search results "somewhat better" than a standard text-based list and 25% rated it as "much better".

The open-ended responses asked participants to describe the Missing Pieces tool to a friend and to list what they liked most and liked least about the visualization. The majority of participants understood the primary function of the visualization as showing the results from the top web search engines, comparing results, and identifying what was missed. Five responses specifically cited metasearching or Dogpile in their description (Table 1).

Table 1: Examples of Open-ended Responses that Cited Dogpile

A meta-search engine that combines results chosen from google, askjeeves and yahoo search engines. Results are divided into those chosen by all, 2 or 1 of the search engines, then further divided into dogpile top results.
The dogpile tool is a better way to search the web because it searches with at least three engines at a time. So you may possibly get three times the results than if you just used one.
Dogpile Missing Pieces Tool is a visual "hits" tool that compares three search engines, Google, Yahoo! and Ask Jeeves. The Dogpile's search engine is incorporated within the results and are show by the pawprint dots.
The tool shows the searcher what they are missing searching just one to three search engines when not using Dogpile. Therefore, Dogpile is a robust search engine that captures up to 40% more relevant hits compared to the others. I will recommend it to my family, friends and co-workers.
Visual graph of the results of a metasearch engine using Google, Yahoo!, and AskJeeves.

Responses to what participants like most about the Missing Pieces tool fell into two major descriptive categories, the tool was easy to use and that the visualization allowed them *to see* "web sites", what they "would normally miss", and "other results". The primary concern reported by 30% of the participants was the inability to interact with the visualization by clicking on the dots to follow the hyperlink. Other participants noted that the URLs only appear when the mouse pointer is hovering so it is difficult to compare and remember the location of specific URLs. Some participants did cite the lack of context for examining search results such as text snippets found in a standard ranked list of linear results.

6   Discussion

This study, while exploratory in nature, has interesting implications for the use of visualization for presenting information in a metasearch environment and the use of a circular visualization. The study demonstrated that in general users understood the basic principles of metasearching and the comparison of results among the search engines. In instances where participants were asked to identify items found on all search engines, there appeared to be uncertainty in one topic area (Picasso) which may suggest that there was some confusion over what it meant to be a Dogpile result.

URLs were cited as one of the top reasons as to why participants chose a particular result, and matching this with the lack of text snippets may indicate that domain name may play a role in the identification of a search result as being useful to the query. This is supported by the finding that not all of the selections were located within the visualization's bull's eye although metaphorically the "most important" results may be selected by the user from the center of the display. The majority of the highly listed URLs were Dogpile results so the visualization may enhance the user's understanding of the effectiveness of a metasearch engine. This was reflected in the open-ended responses where the majority of the participants successfully described the tool's functionality in regard to the display of multiple search engine results.

Subjective satisfaction in regard to the visual interface was rated positively by a majority of the participants. Interestingly, the tool was reported as easy to use or easy to interpret and this lends

support to the bull's eye circular visualization as one that may offer a successful representation in regard to decoding the display.

The limitations of the study include elements of remote testing such as control over the participant actions in the study and the ability to observe participants directly. Since the visualization was limited to visualizing the overlap of one page of search results, it is not known if the users' responses would change given a greater range of comparative overlap data. Further, there may be some inherent limitation in the wording of the questions which draws upon conventional set theory that could potentially confuse users. This was not observed in the pilot study, but may play a role in participant responses.

7   Future Work

The application of visualization techniques to mapping an information space of overlap data holds potential to aiding the user in comparing and selecting retrieved search results. This research has implications for improving the interactivity with Missing Pieces visualization by rendering the items in the display as active hyperlinks so that users can view the pages directly from the display. Also, a method of tracking, selecting, or listing URLs from the interface would aid users in navigating the overlap search results. If context is to be embedded in a visualization to enhance the user interaction with the display, then the use of pop-up windows with text snippets to describe web pages may prove to be beneficial.

The implications for enhanced interface design stems from a fundamental tenet of visualization and human-computer interface design which is to reduce the cognitive load of users when using the interface (Card, Mackinlay, & Shneiderman, 1999). The Missing Pieces tool appears to support this principle in communicating the basic notion of overlap data and significance of a metasearch engine, however the visualization interface requires additional features to enhance its functionality as a tool to examine search results.

Future research could include the comparison of the Missing Pieces tool and Dogpile's text-based metasearch results. Investigation into directly observing user interaction would enhance the knowledge of how a circular visualization is parsed by users and further post-task questions could be developed to determine the utility of the visualization for solving web information seeking tasks. Additional features could be incorporated into the interface design and tested with users to determine the appropriate implementation. The visualization may be extended to display the overlap of more than the first page of search results.

The results of the user study with the Missing Pieces tool point toward positive results in using visualization techniques to model overlap data in a web metasearching environment and the findings offer potential to incorporate interactive functionality to enhance the tool for web users.

8   References

Card, S. K., Mackinlay, J.D. & Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think*. San Francisco: Morgan Kaufmann Publishers.

Heo, M., & Hirtle, S. (2001). An empirical comparison of visualization tools to assist information retrieval on the Web. *Journal of the American Society for Information Science and Technology, 52*(8), 666-675.

Koshman, S. (In Press).Visualization-based information retrieval on the web. *Library & Information Science Research*.

Koshman, S. (2004). Web-based visualization interface testing: similarity judgments. *Journal of Web Engineering, 3*(3/4), 281-296.

Rivadeneira, W., & Bederson, B. (2003). *A study of search result clustering interfaces: comparing textual and zoomable user interfaces*: Human-Computer Interaction Lab, University of Maryland.

Shneiderman, B., & Plaisant, C. (2004). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (4 ed.). Reading, Mass: Addison-Wesley Longman.

Spink, A., Jansen, B.J., Koshman, S. , & Blakely, C. (In Press). A study of results overlap and uniqueness among major web search engines. *Information Processing & Management.*

Spoerri, A. (2004). Metacrystal: visualizing the degree of overlap between different search engines, *13th international World Wide Web conference on Alternate track papers \& posters*. New York, NY, USA: ACM Press.

Ware, C. (2004). *Information Visualization: Perception for Design, Second Edition*. San Francisco: Morgan-Kaufmann.

9   Acknowledgements

---

[1] http://www.grokker.com
[2] http://www.kartoo.com
[3] http://www.clusty.com
[4] http://www.touchgraph.com