Coordinated Views and Tight Coupling to Support Meta Searching

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

This paper addresses how information visualization can support users in the meta search process and some of the design issues that arise. It describes MetaCrystal and its linked tools that enable users to control how to combine and filter the search results by different search engines. A hierarchy of aggregation is employed to provide users with quick insights into the number of documents found by a specific number or combination of search engines. MetaCrystal enables users to visually perform advanced filtering operations, such as complex Boolean constraints. This paper also discusses how to support tight coupling between the filtering controls and tools that operate at different levels of aggregation.

1. Introduction

Information visualization aims to develop visual tools that enable users to explore and gain insight into large and abstract information spaces [7]. It employs *visual coding* and *interaction* principles to create visual abstractions that leverage human perceptual abilities to detect patterns and relationships of interest [7]. This paper addresses how information visualization can support users in the meta search process and discusses some of the design issues that arise. It illustrates some of the visual coding and interaction techniques that can be used to help users combine and filter the top documents returned by different Internet search engines.

Individual search engines only index 20% of the Internet [17] and therefore return different documents for the same query. Meta search engines address this limitation by combining the results returned by different engines. The automatic and effective fusion of different search engine results can be difficult [4]. While meta search engines exist that visually organize the retrieved documents, [9, 12 13, 16, 21] no meta search interface provides users with an overview of the precise overlap between the search engines. Meta searching can benefit from such a visualization, because: a) documents found by multiple search engines are more likely to be relevant [10, 19]; b) it is difficult to predict the quality of coverage for single search engines, because they tend to cover less than 20% of the Internet [17]; c) some engines are more effective than others depending on the search domain [11]; and d) users may prefer or trust some engines more than others. This suggests that active user involvement could make a difference when deciding how to combine the search results returned by different engines.

This paper is organized as follows: section 2 discusses the design requirements for a visual meta search tool. Section 3 briefly reviews related work. In section 4, MetaCrystal and its linked tools are described. Section 5 describes the hierarchy of aggregation used to organize the retrieved documents. Section 6 shows how users can visually perform advanced filtering operations, such as complex Boolean constraints. Section 7 discusses how to support tight coupling between the controls and tools that operate at different levels of aggregation.

2. Design requirements

The hallmark of an effective visualization is that it guides users toward relevant information. 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 documents is visible at any given moment. A meta search interface needs to make it easy for users to decide where to start their visual search and at the same time show as many of the retrieved documents as possible in a structured and compact way. MetaCrystal's design is also guided by the fact that documents found by multiple search methods are more likely to be relevant [10, 19]. It consists of several linked and coordinated tools: the Category View, Cluster Bulls-Eye and RankSpiral. The former groups and aggregates all the documents found by the same combination of search engines. The latter two use related, but complementary organizing principles to display all the individual documents. All these tools incorporate a "bull's eye" layout so that users can expect to find documents found by multiple engines toward the center of the displays. Thus, MetaCrystal shares a key advantage of a ranked list and overcomes its limitations; it guides users toward to potentially relevant documents, while displaying a large number of documents.

While research has shown that documents found by multiple retrieval methods are more likely to be relevant [9, 18], users may also want to examine the documents only found by their preferred engines. MetaCrystal enables users to quickly identify documents found by multiple search engines and at the same time to scan the top documents retrieved by a single engine.

The human visual system can process certain visual properties, such as position, size, orientation, color or shape, pre-attentively or at an early stage of visual processing [24]. The choice of such visual encodings can reduce the cognitive burden for users and help them interpret a visual display more quickly. MetaCrystal uses position, size, orientation, color or shape as visual cues to encode which search engines retrieved a document.

It is critical that visualization tools leverage the perceptual processes that group visual elements based on their proximity, similarity, continuity, connectedness and/or containment relationships. The Category View and Cluster Bulls-Eye employ proximity and similarity principles when positioning documents. Containment is utilized by mapping documents retrieved by the same number of search engines into the same concentric ring. The RankSpiral uses continuity and connectedness principles by displaying the documents along a spiral.

Information visualization supports a variety of interaction techniques to enable users to rapidly explore and progressively filter large data sets. The major interaction methods are: direct manipulation, linked displays, zooming, details-on-demand, dynamic queries, filtering, tight coupling, focus + context and animated transitions [1, 7]. MetaCrystal's direct manipulation interface enables users to iteratively compose and edit meta searches that visualize the precise overlap between up to five search engines. Its linked and coordinated tools support zooming and details-on-demand to give users an immediate sense of a document's content and how the different engines contributed to its total ranking score. Users can perform advanced filtering operations visually. The different filtering controls and linked tools are tightly coupled. Focus + context interaction is supported. Users can visually merge or simplify existing crystals and the transformation is animated to help users track and assimilate the change.

3. Related work

Several meta search engines have been developed that visualize the combined retrieved documents. Vivísimo [23] organizes the retrieved documents using the familiar hierarchical folders metaphor. At the end of each document summary, the search engines are listed that retrieved the document, together with the ranking by each of these engines. Kartoo [16] creates 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. Grokker [12] uses nested circles or rectangles to visualize a hierarchical grouping of the search results. MetaSpider [9] uses a self-organizing 2-D map approach to classify and display the retrieved documents. Sparkler [13] combines a bull's eye layout with star plots, where a document is plotted on each star spoke based on its rankings by the different engines. None of these visual meta search tools provide users with a compact visualization of the precise overlap between the search engines. Instead, they require substantial user interaction to infer the degree of overlap. Further, none of them enable users to control how the search results by the different engines are combined.

4. MetaCrystal

MetaCrystal helps users control and gain insight into how to combine and filter the top documents retrieved by different search engines. Its design is guided by the fact that documents found by multiple search methods are more likely to be relevant [10, 19]. Implemented in Flash, using ActionScript, its linked tools support flexible exploration, enable advanced filtering operations and guide users toward relevant information. The *Category View* displays the number of documents retrieved by different search engine combinations. The *Cluster Bulls-Eye* tool displays *all* the retrieved documents; documents found by multiple engines cluster toward the center and at same time users can easily scan the top documents found by a single engine. The *RankSpiral* tool places *all* the documents sequentially along a spiral based on their total ranking scores to make it easy for users to identify the top documents found by a specific number of engines.

4.1. Category View

Modeled on the InfoCrystal layout [21], the interior consists of *category icons*, whose shapes, colors, positions and orientations encode different search engine combinations. At the periphery, colored and star-shaped *input icons* represent the different search engines, whose top documents are "flowing" into the crystal and are compared to compute the contents of the category icons.

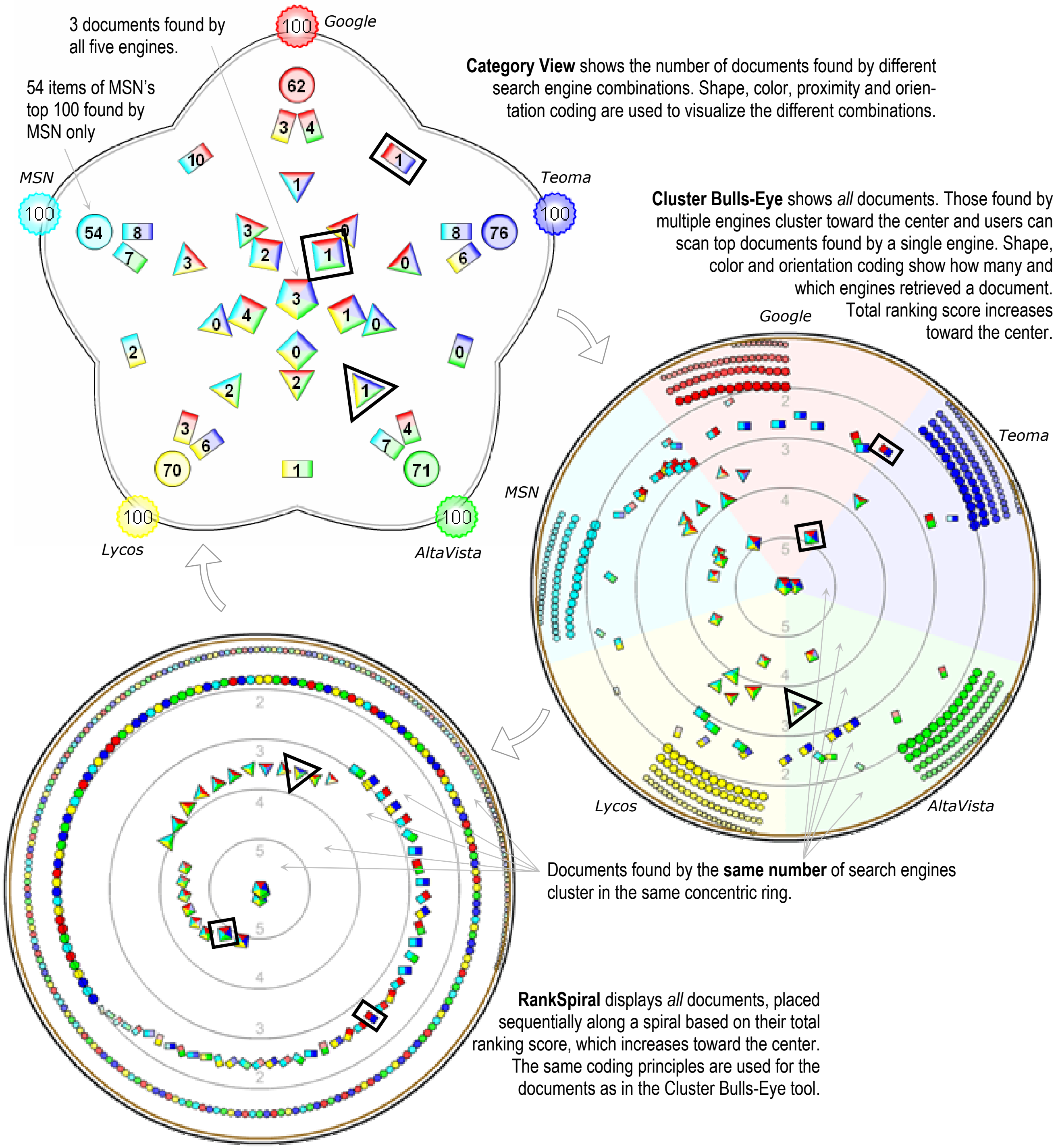

Figure 1: shows MetaCrystal's linked tools: the *Category View*, *Cluster Bulls-Eye*, and *RankSpiral*, which provide users with complementary ways to explore the precise overlap between the top 100 documents found by Google, Teoma, AltaVista, Lycos and MSN, when searching for "information visualization". Several documents are highlighted to show their respective location in the three coordinated tools.

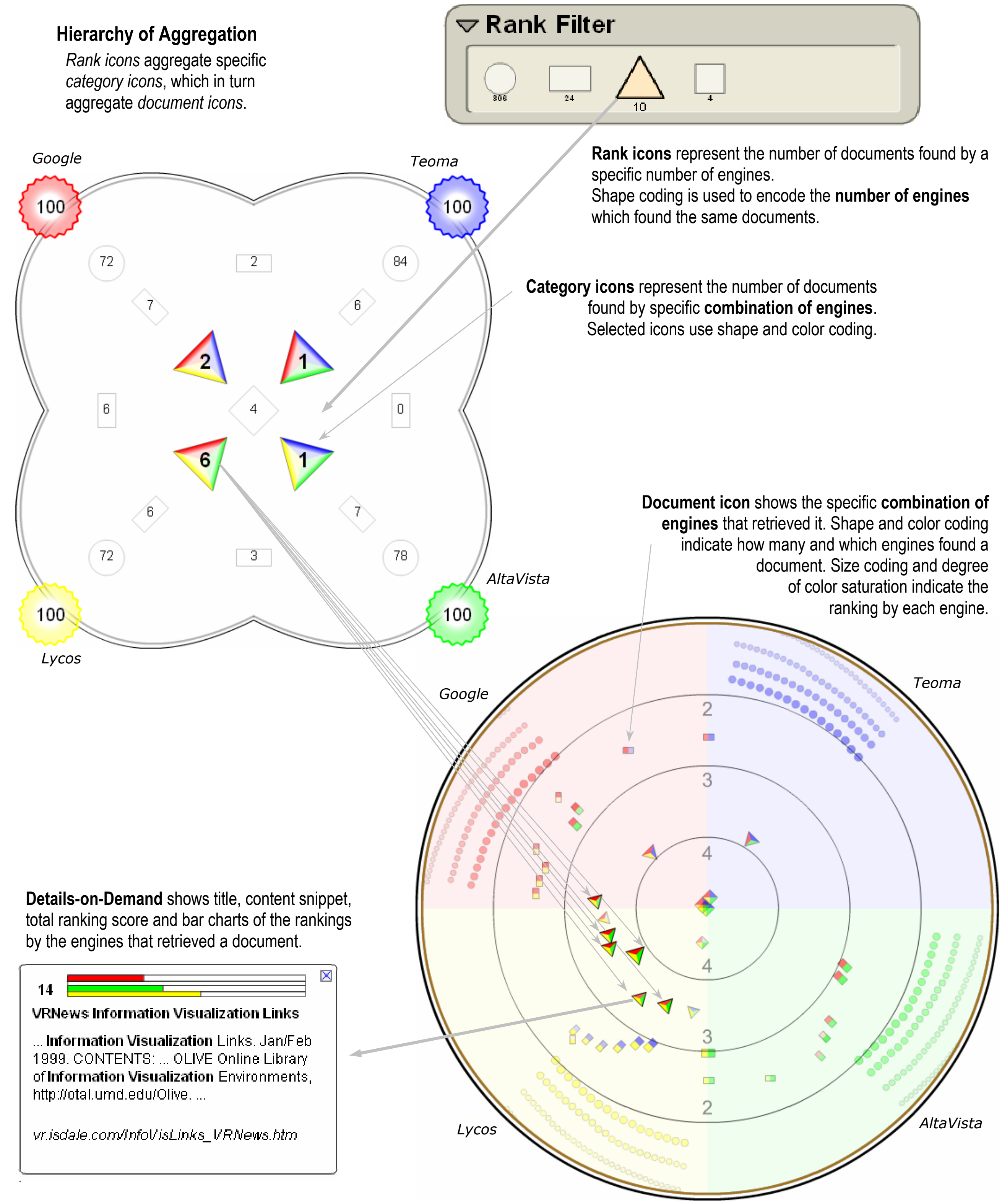

Figure 2: shows the **Hierarchy of Aggregation** used in MetaCrystal. The selected *rank icon* represents all the documents retrieved by three engines. Thus, the selected *category icons* represent all the possible combinations of three engines out of the four being compared. Specifically, a category icon represents all the documents that were retrieved by a specific combination of engines.

The icon in the center of the Category View displays the number of documents retrieved by all engines. The number of engines represented by a category icon decreases toward the periphery. Figure 1 shows the overlap between the top 100 documents found by Google, Teoma, AltaVista, Lycos and MSN, when searching for *information visualization*: three documents are found by all engines; four documents by Google, AltaVista, Lycos and MSN but not Teoma; and most of the documents are retrieved by a single engine. The Category View uses shape (size), color, proximity and orientation coding to visually organize the category icons and show how they are related to the input icons. Each search engine is assigned a unique color code, because color is a good choice for encoding categorical data [24]. Shape coding is used for a category icon if we want to emphasize the number of search engines it represents. Size coding is employed to emphasize the number of documents retrieved by a search engine combination (see Figure 4).

4.2. Cluster Bulls-Eye

This tool shows how *all* the retrieved documents are related to the different engines, because a document's position reflects the relative difference between its rankings by the different search engines. Documents with similar rankings by the different engines will be placed in close proximity (see Figure 1). Shape, color and orientation coding indicate which search engines retrieved a document. The Cluster Bulls-Eye tool uses polar coordinates to display the documents: the *radius* value is related to a document's total ranking score so that the score increases toward the center; the *angle* reflects the relative ratio of a document's rankings by the different engines. The *total ranking score* of a document is calculated by adding the number of engines that retrieved it and the average of its different rankings. This causes documents retrieved by the same number of engines to cluster and to be contained in the same concentric ring (see Figure 1). Specifically, documents with high rankings by the different engines cluster in their respective concentric rings so that they are closest to the center of the display and the size of their icons is set to the largest value. Documents with low rankings cluster furthest away from the center in their respective rings and the size of their icons is set to the smallest value. The use of size coding makes it easy for users to identify the top documents found by a specific number of search engines. In addition, a document's position is influenced by the input icons. Although not shown explicitly in this tool, the input icons act as "points of interest" that pull a document toward them based on the document's rankings by the different engines. Several "Points-of-Interest" (POI) visualizations have been developed [4, 9, 15, 19]. A key design goal of MetaCrystal is to map the documents so that users can use the distance from the display's center as a visual cue of a document's potential relevance. It can be shown that a POI visualization does not satisfy this design goal because it can map documents close to the center even though their total ranking scores are low [22]. The Cluster Bulls-Eye combines a POI display with a "bull's eye" mapping to ensure that users will always find documents with high total ranking scores toward its center. It also makes it easy for users to scan the top documents found by a single engine.

4.3. RankSpiral

Search engines tend to display their results as ranked lists, which can only show a limited number of documents in a single screen. The RankSpiral overcomes this limitation by placing *all* documents sequentially along an expanding spiral (see Figure 1). A document's distance from the center is inversely related to its total ranking score. The score increases toward the center, ensuring that users can expect to find relevant documents in the center's vicinity. Consecutive documents are placed adjacent to each other so that they do not overlap, even if they have the same total ranking score. Shape, color and orientation coding are used to visualize which engines retrieved a document. Documents retrieved by the same number of engines are placed consecutively along the spiral and in the same concentric ring as in the Cluster Bulls-Eye. The RankSpiral tool makes it easy for users to identify the "top" documents found by a specific number of engines. Its spiral's structure can be used to solve the labeling problem [22]. For each document, the radial distance to the icon that has the same angle as the document in question can be computed. This distance can be used to display titles so that they not occlude any document icons. Thus, the RankSpiral makes it possible for users to rapidly scan large numbers of documents and their titles in a way that minimizes occlusions and maximizes information density.

5. Hierarchy of aggregation

MetaCrystal's filtering controls and its linked tools aggregate the retrieved documents in multiple ways. The Rank Control groups all the documents found by the same number of search engines. The Category View aggregates all the documents found by the same combination of search engines. The Cluster Bulls-Eye and the RankSpiral tools use related, but complementary organizing principles to display all the individual documents. Figure 2 shows the *hierarchy of aggregation* used to enable users to quickly determine the number of documents that have been found by a specific number or combination of search engines. Shape coding is used to represent the number of engines that retrieved the same document; color coding is used to represent the specific engines that found it.

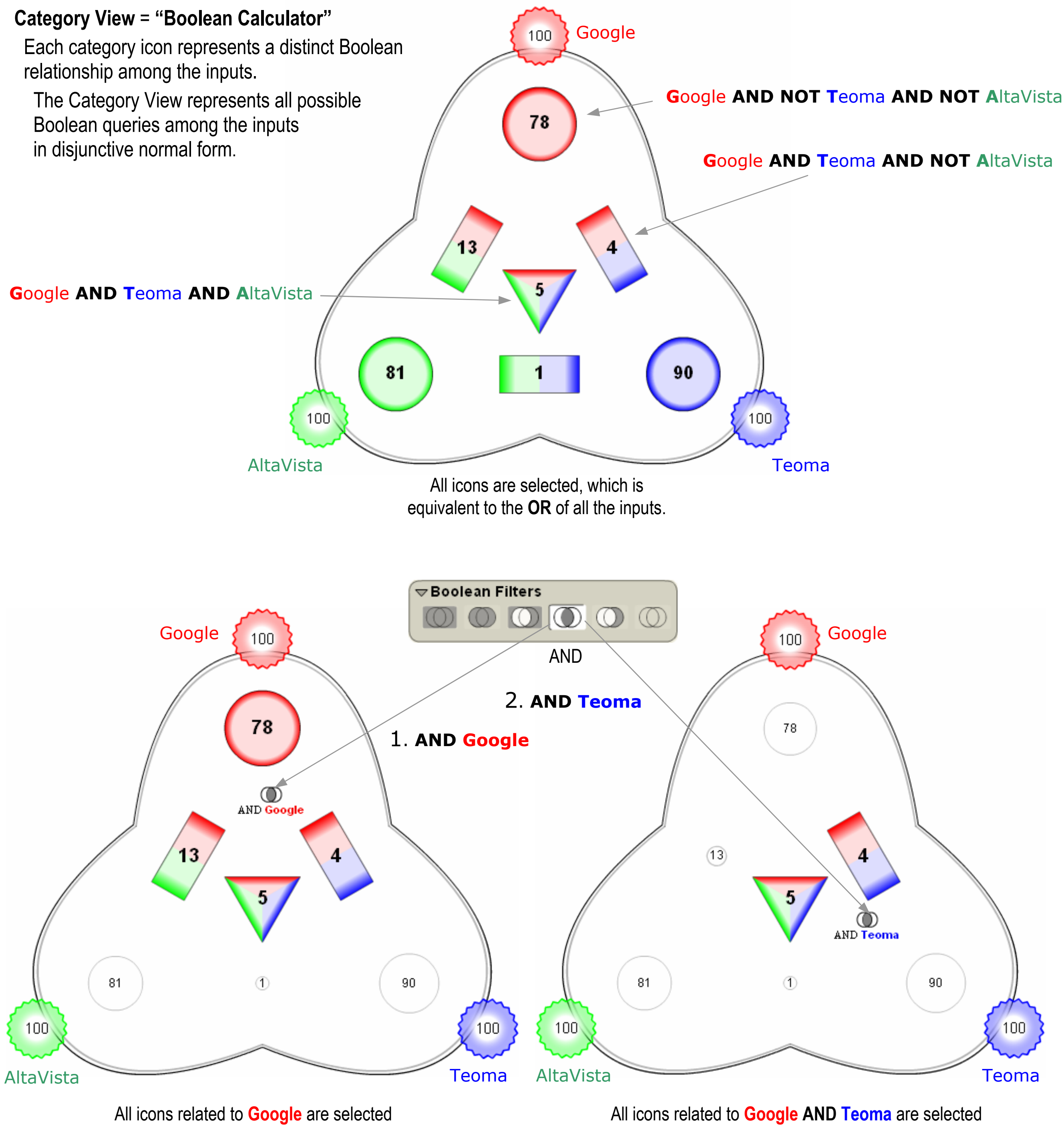

Figure 3: the Category View can be used as a "**Boolean Calculator**", because it represents all possible Boolean queries involving its inputs in disjunctive normal form. The bottom figures show how users can iteratively specify the Boolean query "Google AND Teoma". First, users select the "AND" icon in the "Boolean Filters" control and the cursor changes its appearance. If the cursor is placed in the area associated with a specific input, the cursor shape will reflect this and a mouse click will (de)select the category icons based on the selected Boolean operator and input. Users can sequentially apply Boolean operators to filter the found documents, where category icons not selected are displayed using size, but no color coding.

At the highest level of aggregation, the *rank icon* represents the documents found by a specific number of search engines and it aggregates all the category icons that represent the different combinations of engines with the same rank. Next, the *category icon* represents the documents found by a specific combination of engines. The *document icon* shows which search engine combination retrieved the document, using shape and color coding as well as degree of color saturation to indicate the ranking by each engine.

6. Advanced filtering

MetaCrystal enables users to perform advanced filtering operations visually. Users can filter the found documents to create a short list by: a) requiring category icons and documents to be retrieved by a specific number of search engines; b) specifying Boolean constraints; c) applying a threshold based on the total ranking score; or d) selecting specific category icons or documents by clicking on them.

It is well documented that users find it difficult to formulate Boolean queries [3, 5]. To overcome this problem, visual tools have been developed to help users specify and coordinate Boolean queries [2, 15, 21, 25]. In this section, we illustrate how the Category View enables users to visually specify Boolean constraints. Each category icon represents a distinct Boolean relationship among the inputs and represents a conjunction of the inputs, where some of them are negated (see Figure 3). The Category View represents all the possible Boolean queries involving its inputs in *disjunctive normal form*. If users select several category icons, then the resulting Boolean selection is equal to the disjunction of the Boolean queries associated with the selected icons. In the Category View, the query space defined by its N inputs is partitioned into $2^N - 1$ disjoint subsets, each of which is represented by a category icon. Each category icon can either be selected or not, thus each icon doubles the number of possible queries that can be specified. If the search results returned by five different engines are compared, then over 2 billion possible Boolean constraints can be specified in the Category View.

Users can formulate Boolean constraints in an iterative fashion by successively clicking on specific category icons. Users do not have to use logical operators and parentheses explicitly. Instead, they need to identify the combinations of search engines that are of interest and select them. Thus, users can use the Category View as a *Boolean Calculator*.

Users can also specify Boolean constraints by selecting a Boolean operator , such as "OR", "OR NOT", "AND" or "AND NOT", in the *Boolean Filters* control. The cursor will change its appearance to reflect the selected operator. If the cursor is then placed in the area associated with a specific input, the cursor shape will reflect this and a mouse click will (de)select the category icons based on the selected Boolean operator and input (see Figure 3). In this way, users can specify complex Boolean constraints in a few mouse clicks.

In summary, users can visually formulate Boolean constraints to filter the found documents in these two complementary ways: 1) (de)select specific category icons; 2) select specific Boolean operators and click in the areas associated with a specific input. The former does not require users to explicitly specify which Boolean operators to use, but users may be required to select many category icons. The latter requires fewer user interactions, but users have to specify which Boolean operator to use.

7. Tight coupling

Tight coupling eliminates the strict distinction between query and result displays, because both types of displays can be used to filter the data and to visualize the impact of progressive filtering [1]. In this section, we discuss how to support tight coupling between MetaCrystal's filtering controls and its linked tools that represent the retrieved documents at different levels of aggregation. If users (de)select a category icon, then all of its associated documents are (de)selected. If users select a document icon, then its related category icon is selected, but the icon's appearance also needs to reflect how many of its associated documents are selected. If users deselect a document icon, then its related category icon is deselected only if all of its associated documents are now not selected. In short, a category icon is selected if any of its associated documents are selected and its appearance reflects the number of associated documents that are currently selected. (see Figure 4). In particular, a category icon's area, which is used to display the color coding, is scaled based on the percentage of associated document currently selected.

Users can directly interact with MetaCrystal's linked tools to filter the retrieved documents (as described in the "Advanced filtering" section). Users can also interact with specialized controls to filter the found documents: a) *Rank Filter*: category icons can be (de)selected based on the number of engines that retrieved the associated documents; b) *Boolean Filters*: category icons are (de)selected based on the selected Boolean operators and inputs. The Rank Filter control consists of a series *rank icons* that represent the number of documents found by a specific number of engines (see Figures 2 and 4). The shape of a rank icon encodes the number of engines that found the same documents. Users can (de)select a rank icon to (de)select all the documents found by a specific number of engines. If users apply Boolean operators or interact directly with MetaCrystal's tools, then the Rank Filter control needs to reflect the number of selected documents associated with each of its rank icons.

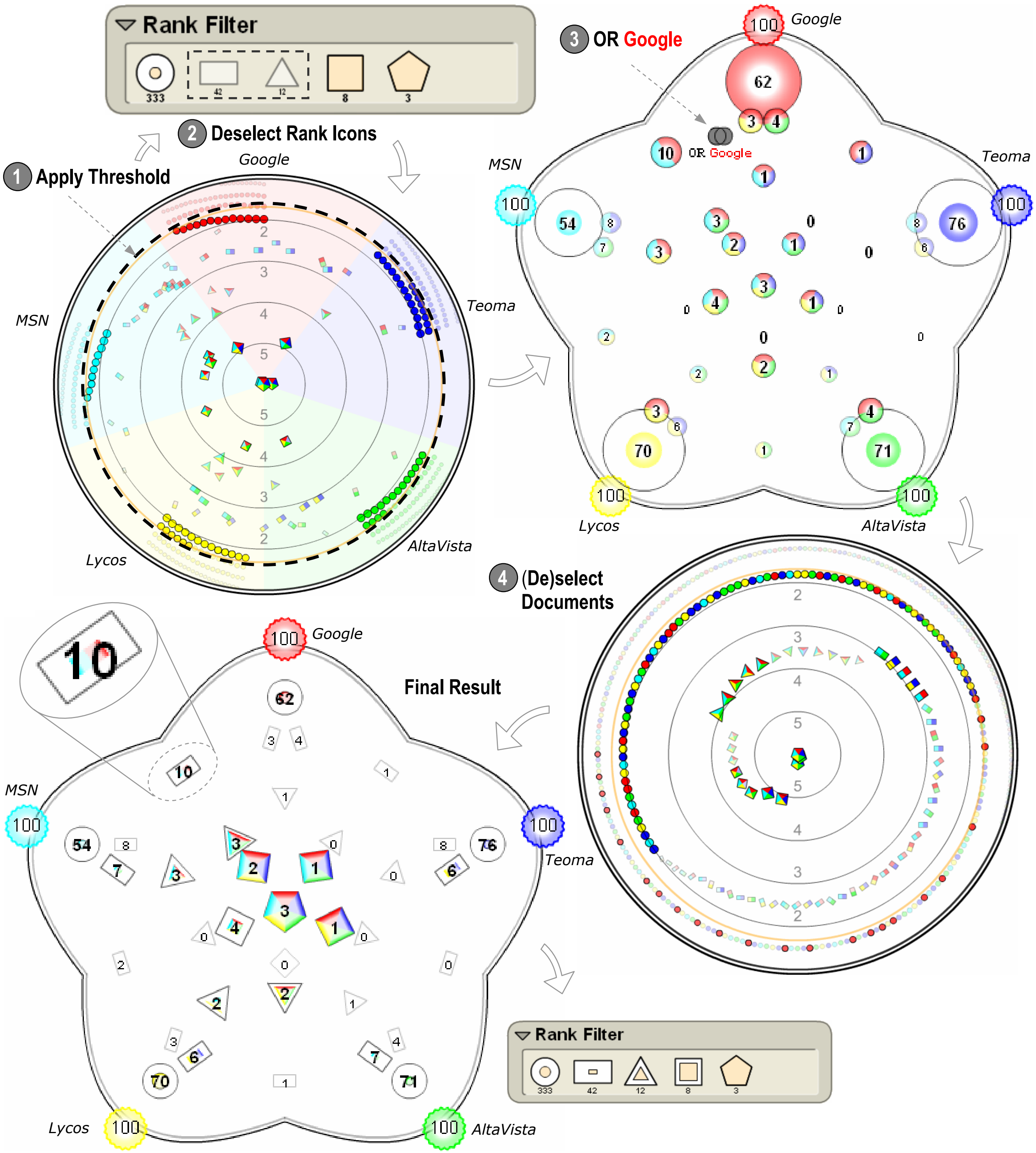

Figure 4: shows the **tight coupling** between MetaCrystal's filtering controls and its linked tools, when the documents are progressively filtered as follows: 1) a threshold is applied to select documents found by multiple engines as well as high-ranking documents found by a single engine; 2) documents retrieved by two or three engines are deselected using the Rank Filter control; 3) all the documents retrieved by Google are selected by choosing the "Boolean OR" operator and clicking in the area associated with the Google input. The selected category icons are displayed using size coding. 4) In the RankSpiral tool, documents are (de)selected so that the top five for each rank are selected and the very low scoring documents only found by Google are deselected. The appearance of a category or rank icon reflects its selection status and how many of its associated documents are selected.

Similar to a category icon, a rank icon is selected if any of its associated documents is selected and its appearance reflects the number of associated documents that are currently selected (see Figure 4). In particular, the colored area of a rank icon is scaled based on the percentage of associated documents currently selected.

Figure 4 illustrates the tight coupling and coordination between MetaCrystal's filtering controls and its linked tools. Specifically, it shows how users can progressively filter the found documents to select the top five documents for each rank as well as all of the documents only retrieved by Google, except for its very low scoring ones. First, users apply a threshold to select documents found by multiple engines as well as high-ranking documents found by a single engine. Second, users interact with the Rank Filter control to deselect all the documents retrieved by two or three engines. Third, all the documents retrieved by Google are selected by choosing the "Boolean OR" operator and clicking in the area associated with the Google input. Fourth, users interact with the RankSpiral tool to make sure that the top five documents for each rank as well as all of the documents only retrieved by Google, except for its very low scoring ones, are selected. Figure 4 shows how the appearance of a category or rank icon reflects its selection status and how many of its associated documents are selected.

8. Discussion and future research

This paper has focused on users conducting meta searches on the Internet, although the MetaCrystal toolset can be used to visualize any ordered set of data and compare the results returned by multiple search methods, such as different query terms or query formulations submitted to the same database or the same query submitted to multiple databases or search engines.

MetaCrystal has been implemented in Flash, using its ActionScript programming language. This has the advantage that it can be deployed using a Web browser and the file size of the application is small. The goal of future versions is to provide users with a versatile search interface that enables them to maintain a search history so that previous searches can be repurposed and combined in an interactive fashion. Future versions will also be better optimized for speed. Currently, users can create multiple MetaCrystals at the same time. This can require the management and animation of thousands of Flash movie clips. It will be explored how to provide increased functionality and speed with fewer movie clips. MetaCrystal currently receives its input data in XML. The current input / output interface will be formalized so that new databases or search engines can be easily added and supported.

Users can create crystals that compare the result sets returned by up to five different search engines. Each crystal can visualize the precise overlap between the engines using the Category, Cluster Bulls-Eye or RankSpiral views. Currently, users don't see these complementary views at the same time. They can rapidly switch between them and an interaction in one is immediately reflected in the ones currently not shown. Thus, it would be easy to modify the current implementation so that all three visualizations are visible simultaneously. However, there is a tradeoff between the number of linked views that are visible at the same time, the amount of screen real estate that can be allocated to each of them and the number of different crystals that can be explored at the same time.

Users can control the number of top documents that are retrieved by a search engine. The maximum number is 100, which is a reasonable choice since search engines make every effort to list the "most relevant" documents as high as possible in their ranked lists and users rarely explore more than 100 documents [20].

As mentioned, MetaCrystal's design is guided by the fact that documents found by multiple retrieval methods are more likely to be relevant [10, 19]. User studies will be conducted to test if guiding users' attention toward documents found by multiple search engines will lead to improved search performance and user satisfaction. It will also be investigated if MetaCrystal and the flexibility it provides lead to greater user satisfaction and help users find relevant information more easily. In particular, it will be tested if the different tools support their indented tasks.

9. Summary

This paper addressed how information visualization can support users in the meta search process. It discussed key design requirements, such as making it easy for users to identify documents found by multiple search engines, because such documents are more likely to be relevant. MetaCrystal's linked tools have been designed so that users can expect to find relevant documents toward the center of the displays, which visualize all the retrieved documents in a compact and structured way. This makes it easy for users to decide where to start their search and they can explore all the documents in a single display. The tools leverage key perceptual grouping processes. The Category View and Cluster Bulls-Eye employ proximity and similarity principles when positioning documents. Containment is utilized in all tools by mapping documents retrieved by the same number of search engines into in the same concentric ring. The RankSpiral uses continuity and connectedness principles by placing the documents along a spiral.

MetaCrystal enables users to control how to combine and filter the search results by different search engines. Its direct manipulation interface enables users to iteratively compose and edit meta searches that visualize the precise overlap between up to five search engines. Its linked tools support flexible exploration and guide users toward relevant information. The Category View displays the number of documents retrieved by different search engine combinations and can be used as a Boolean Calculator. The Cluster Bulls-Eye maps documents retrieved by multiple engines toward its center and at the same time helps users scan the top documents found by a single engine. The RankSpiral places all documents sequentially along an expanding spiral based on their decreasing total ranking scores. It makes it easy for users to identify the "top" documents found by a specific number of engines.

This paper also described the hierarchy of aggregation used in MetaCrystal to organize the retrieved documents: rank icons aggregate specific category icons, which in turn aggregate document icons found by the same combination of search engines. MetaCrystal enables users to perform advanced filtering operations visually. Users can interact with specialized controls to filter the found documents. The Rank Filter lets users (de)select category icons based on the number of engines that retrieved the associated documents. The Boolean Filter enables users select category icons based on the selected Boolean operators and inputs. Further, this paper discussed how the Category View can be used to specify Boolean constraints visually, because it represents all the possible Boolean queries involving its inputs in disjunctive normal form. Finally, this paper addressed how to support tight coupling between MetaCrystal's filtering controls and its linked tools. The colored area of a rank or category icon is scaled based on the percentage of its associated documents that are currently selected.

The next step is to conduct formal user studies to investigate the effectiveness of MetaCrystal and its multiple tools as well as further improve the Flash implementation so that it can be made publicly available.

10. References

[1] Ahlberg, C. and Shneiderman, B. (1994) Visual Information Seeking: Tight coupling of dynamic query filters with starfield displays, Proc. of ACM CHI '94.

[2] Anick, P.; Brennan, J.; Flynn, R.; Hanssen, D.; Alvey, B. & Robbins, J. (1990). A Direct Manipulation Interface for Boolean Information Retrieval via Natural Language Query. Proc. ACM SIGIR '90.

[3] Belkin, N. & Croft, B. (1992). Information Filtering and Information Retrieval: Two Sides of the Same Coin. Comm. of the ACM, Dec., 1992.

[4] Benford, S. D., Snowdon, D N., Greenhalgh, C M., Ingram, R J., Knox, I. and Brown, C C., (1995) VR-VIBE: A Virtual Environment for Co-operative Information Retrieval, Computer Graphics Forum, 14, (3), pp. 349-360.

[5] Borgman, C. (1989). All Users of Information Retrieval Systems Are Not Created Equal: An Exploration of Individual Differences. Information Processing & Management, 25 (3).

[6] Callan. J. (2000). Distributed information retrieval. In Croft W.B. (Ed.), Advances in Information Retrieval. (pp. 127-150). Kluwer Academic Publishers.

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

[8] Carey, M., Heesch, D., Rüger, S. (2003). Info Navigator: A visualization tool for document searching and browsing. Proc of Intl Conf on Distributed Multimedia Systems (DMS Sep 2003).

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

[10] Foltz, P. and Dumais, S. (1992) Personalized information delivery: An analysis of information-filtering methods. Comm. of the ACM, 35 (12):51-60.

[11] Gordon, M., & Pathak, P. (1999). Finding information on the World Wide Web: The retrieval effectiveness of search engines. Information Processing and Management, 35 (2), 141–180.

[12] Grokker – www.groxis.com

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

[14] Hemmje, M., C. Kunkel, & A. Willet (1994). LyberWorld - a visualization user interface supporting fulltext retrieval. Proceedings of ACM SIGIR, pp. 254–259.

[15] Jones, S. (1998). Graphical Query Specification and Dynamic Result Previews for a Digital Library. Proceedings of ACM UIST 1998:

[16] Kartoo – www.kartoo.com

[17] Lawrence, S., & Giles, C.L. (1999). Accessibility of information on the Web. Nature, 400, 107–109.

[18] 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.

[19] Saracevic, T. and Kantor, P. (1988). A study of information seeking and retrieving. III. Searchers, searches and overlap. JASIS. 39, 3, 197-216.

[20] Silverstein, C., Henzinger, M., Marais, J. & Moricz, M. (1998). Analysis of a very large Alta Vista query log. Technical Report 1998-014, COMPAQ Systems Research Center, Palo Alto, Ca, USA.

[21] 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). San Francisco: Morgan Kaufmann [also in Proc. of VIS'93].

[22] Spoerri, A. (2004) Cluster Bulls-Eye and RankSpiral: Enhancing Points-of-Interest and Search Results Visualizations. Submitted to InfoVis 2004 - Proc. of IEEE Information Visualization Symposium.

[23] Vivisimo – www.vivisimo.com

[24] Ware, C. (2000). Information Visualization: Perception for Design. San Francisco: Morgan Kaufmann.

[25] Young, D. & Shneiderman, B. (1993). A Graphical Filter / Flow Representation of Boolean Queries: A Prototype Implementation and Evaluation. JASIS, 44 (6).