

Visualizing Web Search Results

Cyntrica Eaton
ceaton@cs.umd.edu

Haixia Zhao
haixia@cs.umd.edu

(In alphabetical order)

Computer Science Department
University of Maryland at College Park
College Park, Maryland, 20704

ABSTRACT

The explosive growth of the World Wide Web has offered more information and, concurrently, a need for improved navigation. In response to this problem, our aim for this project was to create a query result visualization paradigm that would provide a more comprehensive presentation of returns and better support for result set exploration.

In our particular interface, individual web pages are displayed as points on a two-dimensional plane. An overview of the relevance, popularity, domain, and language of individual web pages returned by the search engine is conveyed through the position, color, and shape of corresponding points. A contextual overview of the result set is provided through a hierarchical topic tree. Exploration of the result set is supported by allowing users to filter out returns based on file size, date published, and other document characteristics. Finally, detailed web page summaries as well as an overview of the types of files supported by a particular webpage (i.e. audio, images, etc.) are available on demand by clicking on points of interest.

Based on the incorporated features and the inherent qualities of information visualization, we believe that the model we have created will be a helpful search tool particularly when the precision of the result set is low, search terms are ambiguous, and/or when users have a difficult time translating search needs to the querying tool. Since we incorporate a topic tree that allows users to dynamically filter out all returns that are not related to topics of choice, users will also have more control over the context of the pages presented to them.

INTRODUCTION

Using computers as tools for archival and dissemination of data has enabled the implementation of the World Wide Web that, at maximum utility, allows expanded access, increased usability, and the establishment of new ways for individuals to interact with information [Hodges 1998]. One of the major issues hindering the realization of maximum utility, however, is the lack of adequate search tools for WWW exploration. While the number of documents available on the web has been increasing by many orders of magnitude, the implementation of tools to help users better navigate this information has not been introduced at a complementary rate. As a result, Web users typically have a difficult time finding information that satisfies their particular search interests.

In a typical search session, users are often burdened with the fact that conventional search engines usually return too many irrelevant documents [Kaindl et al 1998]. While the reasons for this phenomenon involve many faceted issues, user search behavior is definitely a contributor and most current search result presentations do not support an overview of returns that could eventually help users distinguish useful documents from irrelevant ones.

To further illustrate the shortcomings on current search engine paradigms, let's review an example. Suppose, for instance, that a user attempts to conduct a search using the keyword *Java*. Because this particular word assumes a variety of semantic meanings, the query result set would probably contain documents in which

Java means the programming language, an Indonesian island, or a colloquialism for coffee. Because studies on Web search patterns show that users usually issue very short keyword queries, this scenario is not so far off. The problem, however, lies in the fact that short keyword queries make it extremely hard for search engines to accurately predict user information needs. In fact, computers generally have a difficult time determining user intent in cases of ambiguous keywords like *Java*, improper search keywords, or when the search query is vague and covers a wide variety of areas.

Even if the search engine actually returns documents that are in correct alignment with user interests, it is quite possible that those results would be scattered among a multitude of irrelevant returns. Currently, a list of pages typically ranked according to a function of expected relevance and other factors is presented to the user with about 10–20 items per page. Typically, each page is listed with information such as a short summary about the page, the content of the search keywords in the page, and file size. Because the user must process the results using a linear, fixed ordering scheme, the list-based structure of search results offers little assistance in browsing large sets of pages.

In order to address the problems that users face during typical web search sessions, our research efforts have been focused on creating a search result presentation paradigm that would increase efficiency in identifying relevant web documents. In short, our main objective was to create a visualization paradigm that could effectively mediate between the information needs of the user and the information available [Gordon et al 1998]. Realizing the power of human perceptual ability to extract vast amounts of data from graphic representations, our paradigm presents a visualization of search results. By representing individual web pages as points and displaying information like the relevance and popularity of the page through position, the domain (i.e., .edu., .com, etc) through color and the language of the document through shape, we have created a paradigm that employs visual attributes to convey as much information as possible in a constrained space. We have also incorporated direct manipulation ability by incorporating double-clicked filters and sliders that will allow users to specify ranges for the

relevance, popularity, published date, and file size that they would like to review in the result space. In addition, a feature that will allow users to get an understanding of whether a webpage is largely text-based or whether it incorporates audio and image files has also been implemented.

To address the problems associated with ambiguous search terms like *java*, we have incorporated a hierarchical topic tree that will present the user with an overview of the topics covered in a search so that they can make the final decision of the cluster of documents that best fit their search intent. Clustering techniques such as commercial product *Vivisimo* [vivisimo] and research work *Groupier* [Zamir & Etzioni 1999], make use of the fact that some pages in the result set are more similar to each other than the remaining regarding their potential of meeting one aspect of user information needs. Clustering techniques organize these search results into topically hierarchical groups, using short phrases as the group label to describe the common semantic topic of the pages in that group. As a result of this categorized ordering, users are better equipped to quickly narrow the results displayed down to topics that best include their needs.

The use of visualization in interpreting Web search results could be effective for several reasons. First, with graphical presentation, more information about a large number of documents can be conveyed simultaneously in a small space since the characteristics of a graphical object such as color and position can be used to convey document characteristics [Nowell 1996], and quickly direct user attention to areas of interest. Second, the graphical nature of the display allows users to develop a view of the overall trend of the returned result set. More specifically, displaying the documents graphically allows users to navigate the information space as a time instead of a document at a time [Veerasingam 1997]. Overviews are important considering that a problem arises when documents that are relevant to the user's original query are either buried in a barrage of extraneous returns that could take quite a few steps to access if presented in a linear text-based list.

The main motivation of our design lies in following aspects:

- To convey as much fact as possible about the result set and the pages in an

- effective way that can help users to judge the usefulness of the pages.
- To enable quick navigation between overview of the result set and preview of a particular page.
- To provide effective organization of the pages and enable visualized user exploration of the large information set via direct manipulations, such as dynamic query [Shneiderman 1994].

In order to bring our idea to life, we implemented our paradigm based on StarDOM, a general visualization tool that plots information on a two-dimensional scatter plot.

RELATED WORKS

To our knowledge, though there have been significant efforts in visualizing the hypertext space [Andrews 95] [Ayers 95] [Mukherjea 95] [WebTOC] and visualizing the search result from specific document collections [Shneiderman 2000] [SPIRE] there are very few works on visualizing the query results returned by Web search engines to the users.

One type of work that is directly oriented at visualizing Web search results is to convey the contribution of individual search key terms in a query to the overall relevance of a returned document. Because this information is usually hidden in a single rank score in the list-based text presentation, this paradigm presents more information to better assist users in identifying returns that are actually relevant to their queries. [Grewal et al], in particular, proposed and studied two visualization based on the aforementioned ideas. [McCrickard & Kehoe] presented the Search Query Weighted Information Display (SQWID) system that provides an interactive visualization of the search results, allowing users to see the relevance of the results to different key terms as well. In SQWID, the key terms and the result documents are spread out in a two-dimensional space, each as a node. Document nodes are described by their title, and are connected to the key term nodes by colored links with the relevance value as the link label. Although SQWID can show users how the result documents are attracted to the key terms, it does not provide other important information users

need to judge the usefulness of the documents. Problems with this paradigm include the fact that space is wasted by the node-link presentation and occlusion occurs when the number of documents rises above just a dozen.

[Mukherjea & Hara 1997] outlines a system that shows users where pages lie with respect to one another and use one word summaries of a network of Web pages to allow users to not only get an idea of how a particular page connects to other pages based on Web locality, but how the topics that are presented in the pages are related as well. The drawback with this model, however, is that only one result page can be described at a time. Other works in visualizing query returns use spatial metaphors such as space and height to convey result information to users. Spatial Paradigm for Information Retrieval and Exploration (SPIRE) is a suite of visualization tools that support galaxy and terrain based views of return document sets. In the Galaxies model, documents are represented as a galaxy of star cluster. The escapes, on the other hand, incorporate a terrain-based presentation model, and employs features like distance and height to communicate the similarity and overall strength of topics respectively.

Finally, there are also works in which search result visualizations were actually variations on the text list. [Pirolli 1996] described a Scatter/Gather display, where results were clustered by relevancy into five clusters, each of which was shown as a scrollable text box.

OUR APPROACH

Our visualization approach is directly targeted at the query results returned by Web search engines. It includes four key design components:

- An overview of query results in a two-dimensional scatter plot with page characteristics encoded by visual attributes such as color and shape;
- A topic tree that provides a hierarchical overview of the semantic topics covered by the result set;
- Details-on-command features that present page summaries and more in-depth page information;

- Finally, dynamic query and filtering through both the topic tree and Web page attributes.

Overview of Query Results in a 2D Visualization

As addressed before, a one-dimensional text-based ranked list of Web query results cannot provide a rich overview of the result set and page characteristics such as domain type and language. Also, only a weighted ranking can be conveyed, thus, burying some very important ranking factors such as relevance scores and popularity, that can help users to understand why one page precedes another in the result set. We visualize each of the returned pages of a query as a point on a two-dimensional scatter plot. Our paradigm makes it possible to independently view the popularity of each page, the relevance of each page to the search key phrases, and pages that specifically correspond to the contexts provided in the topic tree.

In our prototype, visual attributes such as color, shape, and position of the dot are used to convey the characteristics of the corresponding web page. More specifically, color is used to convey the domain of the document (i.e., .com, .org, etc.), shape indicates whether the page is English or Non-English, and position conveys the relevance and popularity of the page simultaneously.

Topic Tree

To avoid the occlusion problem, the points on the plot are not labeled unless they enter the current focus range in which case the titles of the web pages that correspond to the points automatically show up as excentric labels [Fekete & Plaisant 1999]. Because the points are initially unlabeled, users will not be able to quickly get a feel for the context of the document represented by a particular point. As a result, we have incorporated a topic tree that dually gives a hierarchical overview of the topics covered in the result set and insight as to how each of the individual pages returned fits into the respective categorization shown in the topic hierarchy. This will allow users to identify pages of interest based on the contexts that best fit their information needs.

As pointed out in [Douglass et al 1992], the topic tree can be compared to the current methods for accessing information in a conventional textbook. Because of the vastness of the material that is presented in the textbook, the user will usually consult the index in order to locate passages of interest. Similarly, the semantic organization of the topic tree hierarchy plays an important role in navigation efficiency and also for the effectiveness of the catalogue-oriented browsing user interface as a document set exploration tool [Klement et al 1998]. Cataloging essentially presents a more organized grouping of the results that allows users to focus on categories of interest and provides a solution to conventional linear, rank-based presentations by allowing users to view all of the relevant materials instead of having to browse through all of the results sequentially.

Figure 1 shows a screenshot of our system. The 100 documents represented in Figure 1 are actual Web pages returned by the coupling of commercial search engines Lycos and Vivisimo based on the following search query: *Digital Camera*. Lycos was essentially used to get the web pages that would be returned from the query and Vivisimo was used to populate the topic tree by clustering the documents based on context. The pages returned are repositioned on the 2D scatter plot based on the popularity, or web traffic value, as obtained from [Alexa] and the relevance value of each page which is based on a mimic value from the Lycos rank of the page.

The color of each dot indicates the domain type of the corresponding page. The system distinguishes six types of domains: .com, .net, .edu, .gov, .org, and other. The color legend is shown under the scatter plot. The language of the document is presented through the shape of the point. Rectangles represent English pages and diamonds represent non-English pages.

In the topic tree, the leaf nodes of the tree are the actual web pages, while parent nodes of the tree are labeled with short phrases summarizing the collective context of the pages under it.

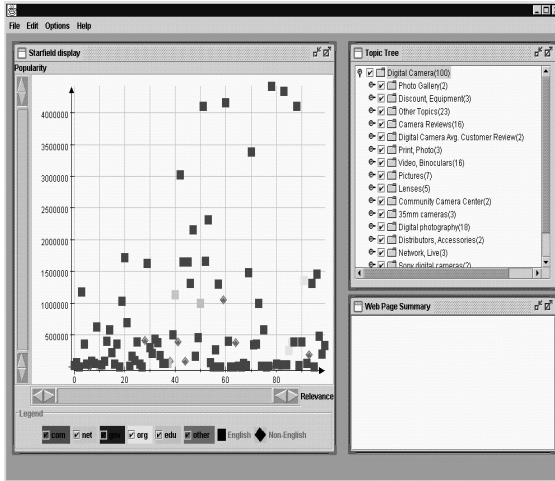


Figure 1: Result Overview by 2-D Scatterplot and Topic Tree

Preview-on-Demand

Insight into more detailed page data is one of the most critical pieces of information users need in order to judge whether a document deserves further examination. Our preview-on-demand feature in the graphical paradigm expands the ability of the interface to effectively convey more information.

Figure 2 shows the preview-on-demand feature of the system. First, automatic display of eccentric labels for the page dots within the current mouse focus range gives users a quick view of webpage titles. Second, when a dot is selected with a mouse click, summary information is displayed in the information panel labeled *Web Page Summary*. While this particular panel is specifically designed to display the page title, a short summary which highlights the context of these search keywords, and the corresponding hyperlink, we have also provided a *File Types View* panel that visualizes statistical data about the distribution of file types (i.e. audio, images) featured on the webpage (Figure 3). The idea of showing the file type distribution was inspired by the work in WebTOC [WebTOC] which presents a website as a table of contents and visualizes information about each entry (such as subsiding file sizes) using colored, shadowed bars with varying thickness and length. File type distribution within a page or even the website where the pager resides can provide users with valuable

clues about potential relevance to search interests. For example, if the user is looking for a song to download, a page with the *File Types View* panel showing a large volume of audio files will probably be a better choice than a page with the metadata showing few audio files.

In order to allow even further exploration of a page of interest outside of the webpage summary, our system even allows the user to click on the hyperlink displayed in the summary to preview the corresponding page (Figure 4).

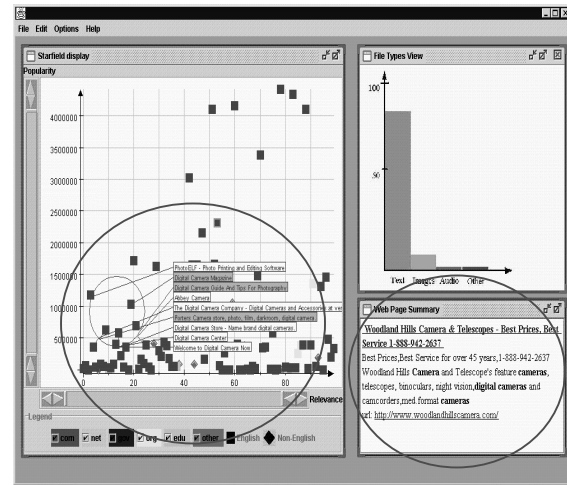


Figure 2: Preview-on-Demand

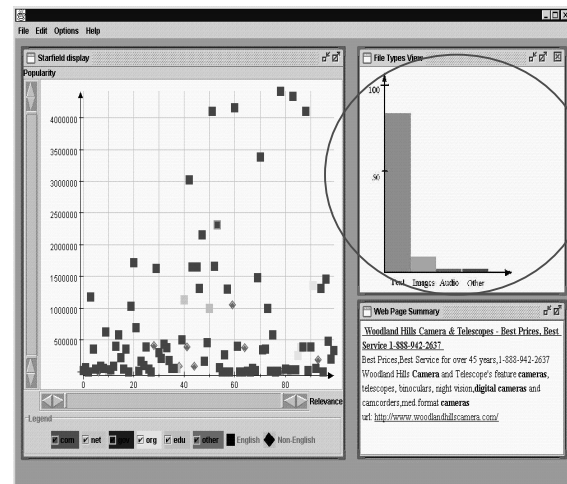


Figure 3: File Type Distribution Feature

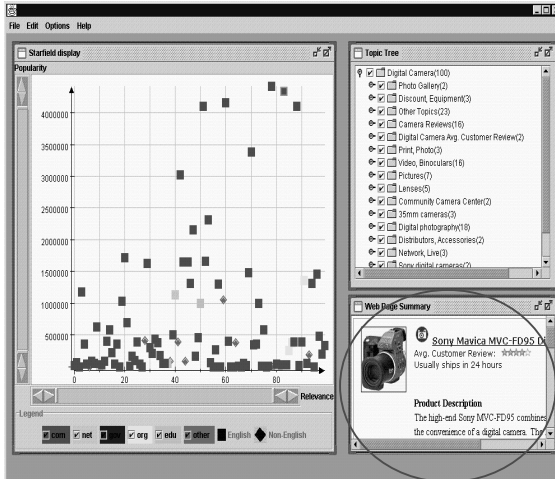


Figure4: WebPagePreview

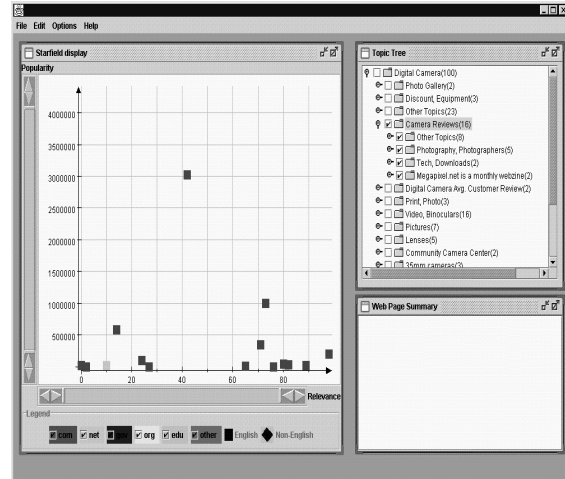


Figure5: Topic -BasedFiltering

DynamicQuery,FilteringandZooming

Realizing the power of direct manipulation as a powerful technique for large dataset exploration [Derthick et al 1997], we implemented such support in our system, including dynamic query, filtering, and zooming. The idea of dynamic filtering is inspired by previous works such as HomeFinder and FilmFinder, as described in [Shneiderman 1999].

Users can perform topical dynamic query by clicking on topics in the topic tree to select or deselect corresponding pages. Dynamic query by topics enables the user to quickly narrow down the displayed results to those which are correctly aligned with the user's needs. Figure 5, for example, shows a screenshot where the user is interested in the topic *Digital Camera Review* in the topic tree. By initially filtering out all of the pages and then selecting the topic *Camera Reviews* in the topic tree, all pages except those under the specified topic are filtered out.

Dynamic filtering is also available in our prototype for page relevance, popularity, publish date, and file size. By manipulating the double-thumbed sliders provided in the *Filters* panel, users can indicate ranges for the aforementioned attributes while the system dynamically filters pages that fall out of the specified range from the display. This feature allows users to focus on high quality and high relevant pages within certain dates and sizes while still keeping an overview of the overall quality of the result set. By selecting or deselecting the domain type checkboxes provided in the legend under the scatter plot, users can show or hide pages from specific domains. Figure 6 shows the *Filters* panel and highlights the checkboxes that can be used to filter based on domain.

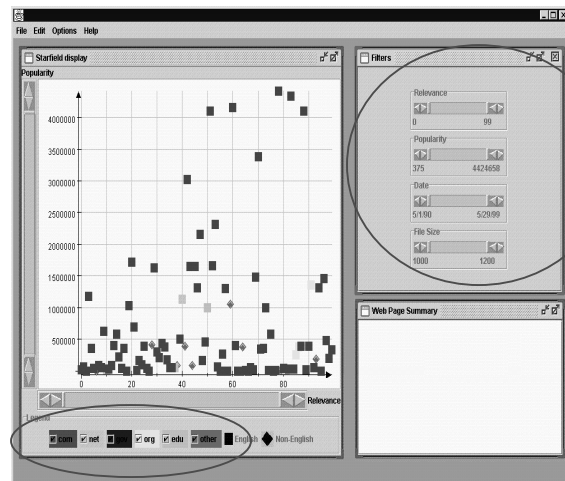


Figure6: Filter Panel and Domain -Specific Filtration

Support for zooming is also available. Users can zero into a specific region of relevance and popularity region by clicking and dragging the arrows along both the relevance and popularity axes of the scatter plot.

ANALYSIS & FUTURE WORK

In contrast to conventional ranked list, text-based query result presentation, our system exploits the synergy between graphic representation in conveying massive information and human perceptual ability to extract vast amount of data from visualizations. Intuitively, our paradigm could be effective in helping users to quickly narrow down search result set to the targets of potential interest, to judge the potential usefulness of the pages returned by the search engines, and to rapidly locate information sources that can meet their information needs, especially under the following scenarios:

- When the precision of the result set returned by the search engine is low.

When the search engine fails to achieve high result set precision, relevant pages are usually scattered and buried among a significant amount of irrelevant documents. If the user fails to quickly find pages of potential interest after linearly scanning the text descriptions, they are likely to abandon the result set. Our paradigm provides an overview of all the documents in the result set along with rich visual cues of the page characteristics. Users can navigate the result space at a time instead of a document at a time, jumping from one document to another quickly and easily, avoiding excessive scrolling.

- When the user information needs are vague, or the user does not know how to correctly express his/her information needs.

Sometimes, users' information needs are vague. Users may know the proximate types of the information they want, but they cannot accurately describe their needs in an effective search query. Statistically, Web queries are typically comprised of short search key phrase. Paradoxically, very short search queries usually cause an overwhelming amount of documents to be

returned. Although these documents may all have very high calculated relevance values based on the appearance of search keywords, some of them may be focused on issues far away from user expectations. Our paradigm will categorize the broad range of pages under different topics in the topic tree, allowing users a clear overview of the outline represented and a quick way to locate pages of interest.

- When the user issues search keys that have ambiguous meaning

When the search keys have ambiguous meaning and the user's information needs are of the less popular meaning, the relevant pages may be washed out by pages which discuss the topic based on the more popular meaning. For example, the user may want to know something about a new drug named DOG that was recently approved by some government drug administration agencies. A search on the word "DOG", or more specifically, the keywords "drug DOG" would probably be conducted. In response to the query, the search engine may retrieve a lot of popular Web pages about the animal dog, pet food and medicines. Since these Web pages may be much more popular than the pages containing the new drug DOG, the pages that are actually useful to the user may be ranked far away from the top of the list and appear after tens or even hundreds of irrelevant documents. The location of the pages in the ranked list may have little chance to get the user's attention since the user's patience probably has been worn out before he reaches the page. Since our paradigm will allow users to view a great deal of pages at a time and use the topic tree to focus on the pages of interest despite their rank in traditional presentations, we feel that it provides an effective solution to this particular problem. Rich visual attributes may give the user extra clues about the potential usefulness of a page. As an example, the user may just focus on the pages from domain types .gov and/or .org by filtering out pages with alternative domains.

- When the user has difficulty understanding why one document comes in front of another.

In a ranked list text-based presentation, retrieved pages are ordered according to a final weighted rank assigned by the search engine. Details about how the final ranks are generated are hidden from the users. Consequently, users may have difficulty understanding why one document comes in front of another in the pure ranked list. In our paradigm, those details can be shown to the users, thus allowing them to have more insight into the relevance and popularity of the web pages returned. The users can even set thresholds of the ranks of the pages to be displayed in order to ensure that only pages of the specified relevance and popularity range are shown.

While we feel that our paradigm will be a helpful tool in WWW navigation and have provided intuitive scenarios where it would be particularly useful, the proper HCI studies must be conducted before any real conclusions could be drawn about the true effectiveness and usefulness of this model. Before detailed user study plans are made, there are several issues that should be considered as part of the user study design guidelines.

Compared to conventional text-based ranked list presentation, our paradigm trades the availability of individual context preview at the very first glance for an overview of the entire document set. Although the rich visual attribute of the graphical objects convey a lot of extra information, and more data about a page can be displayed through the preview-on-command feature, user studies have to carefully examine the effect of losing preview at the very first glance, and how well the other features compensate for the lack of initial detail.

In addition, user studies should examine performance and satisfaction of users with different backgrounds and experience, especially regarding their skills and experiences with Web searching.

Before a final conclusion can be drawn, however, richer data sources are needed. Although our paradigm allows great potential to simultaneously display a large amount of information about the result set, the information in our prototype is limited because of the lack of good data sources. We created a parser that can extract the topic tree and Web pages returned by commercial search engines and obtain popularity

values, but a great deal of the data is mimicked. An ideal resource of data would be the information stored and used by actual search engines to perform the retrieval and ranking upon user search requests. In the specific terms of running our prototype, however, we are able to quickly generate a data file from the parser based on actual search results.

The system we show here is just a prototype to illustrate our basic idea. When richer, more accurate data is available, some other features could be added to polish up the system. E.g., in the prototype, the scatter plot is fixed as overall relevance vs. page popularity. It could be expanded so that the user can choose relevance values to an individual search keyword as an axis. Also, our prototype uses two different shapes for English and non-English pages. The shape-coding could be expanded so that more languages can be conveyed, and the language can serve as an attribute for dynamic filtering, just as the domain type.

CONCLUSION

By creating a web query paradigm that allows users to visually extract relevance, popularity, domain, and language of the pages in a result set, we believe that our model will effectively allow users to navigate through query results by shifting mental load from the slower, thought-intensive process of reading to the faster perceptual process of digesting information through visual cues [Auetal2000]. Also, the incorporation of tools for dynamic query will allow users to narrow the results returned to ones that satisfy criteria that they have specified as opposed to a linear representation of static results that they have no control over post-generation.

Our main intuition is that the combined features of this tool will enhance browsing capabilities and allow users a better opportunity of finding relevant documents in larger result sets, when search terms are ambiguous, when too many irrelevant documents are returned, and/or when the field that the keywords cover is extremely broad. In order to confirm the validity of our model, however, user studies must be conducted to determine the cost of presenting result set overviews as opposed to immediate page-specific details.

ACKNOWLEDGEMENT

We thank our professor Dr. Jean -Daniel Fekete, Assistant Professor at the Ecole des Mines de Nantes, Invited Professor at the University of Maryland at College Park, for working closely with us during the forming and developing of our project.

REFERENCES

[Alexa] www.alexa.com

[Andrews95] Andrews, K. Visualizing Cyberspace: Information Visualization in the Harmony Internet Browser, *InfoVis'95*, IEEE Press, Atlanta (1995) p.97 -104.

[Au et al 2000] Au, P., Carey, M., Sewraz, S., Guo, Y., and Ruger, S. M. New Paradigms, in *Information Visualization. Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (2000).

[Ayers95] Ayers, E. Z. and Stasko, J. T. Using Graphic History in Browsing the World Wide Web, *Proceedings of the Fourth International World Wide Web Conference*, Boston (1995).

[Fekete & Plaisant 1999] Jean -Daniel Fekete and Catherine Plaisant, *Excentric Labeling: Dynamic Neighborhood Labelling for Data Visualization*, *ACM CHI'99*.

[Google] Sergey Brin and Lawrence Page, *The Anatomy of a Large -Scale Hypertextual Web Search Engine*. Also available at: <http://www.db.stanford.edu/~backrub/google.html>

[Gordon et al 1998] Gordon, A., Domeshek, E. *Déjà Vu: A Knowledge -Rich Interface for Retrieval in Digital Libraries*. *Proceedings of the 1998 International Conference on Intelligent User Interfaces* (1998) p.127 -134

Also available at: www.acm.org/pubs/articles/proceedings/uist/268389/p127-gordon/p127-gordon.pdf

[Grewa et al] Ratvinder Grewal, Mike Jackson, Jon Wallis, Peter Burden, *Using Visualization to Interpret Search Engine Results*.

[Hodges 1998] Hodges, R. *Digital Libraries: Impact on Science, Medicine, and Libraries*

(1998). Available at: http://muspin.gsfc.nasa.gov/conferences/previous_conf/1998/98proceedings/hodges.html
[Kaindl et al 1998] Kaindl, H., Kramer, S., Afonso, L. M. *Combining Structure Search and Content Search for the World -Wide Web. Proceedings of The Ninth ACM Conference On Hypertext And Hypermedia: Links, Objects, Time And Space -Structure In Hypermedia Systems* (1998) p.217 -224

[McCrickard & Kehoe] D. Scott McCrickard & Colleen M. Kehoe, *Visualizing Search Results using SQWID In Proceedings of the Sixth International World Wide Web Conference*, (1997).

[Miller] Miller, T. *Federated Web Search Visualization* available at: <http://www.cc.gatech.edu/~tomiller/7001/project1.html>

[Mukherjea 95] Mukherjea, S. and Foley, J. *Visualizing the World Wide Web with the Navigational View Builder*. *Proceedings of the Third International World Wide Web Conference*, Darmstadt, Germany (1995).

[Nowell 1996] Nowell, L. T., France, R. K., Hix, D., Heath, L. S., Fox, E. A. *Visualizing Search Results: Some Alternatives To Query -Document Similarity*. *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (1996) p.67 -75

[Pirolli 1996] Pirolli, P., Schank, P., Hearst, Diehl, M. *Scatter/Gather Browsing Communicates the Topic Structure of a Very Large Text Collection*. *Conference Proceedings on Human Factors in Computing Systems* (1996) p.213 -220

[Shneiderman 1994] Shneiderman, Ben, *Dynamic queries for visual information seeking*, *IEEE Software* 11,6(1994),70 -77.

[Shneiderman 1999] Shneiderman, B., *Dynamic queries, starfield displays, and the path to spotfire* (1999). Also available at: <http://www.cs.umd.edu/hcil/spotfire/>

[Shneiderman 2000] Shneiderman, B., Feldman, D., Rose, A., FerreFrau, X. *Visualizing Digital Library Results with Categorical and*

HierarchicalAxes.DL'00:Proceedingsofthe
5thACMInternationalConferenceonDigital
Libraries(2000)p.57 -66

[SPIRE]SPIRE -SpatialParadigmfor
InformationRetrievalandExploration,
<http://www.pnl.gov/infoviz/spire/spire.html>

[Veerasamy1997]Veerasamy,A.,Heikes,R.
EffectivenessofaGraphicalDisplayofRetrieval
Results.ProceedingsOfThe20thAnnual
InternationalACMSIGIRConferenceOn
ResearchAndDevelopmentInInformation
Retrieval(1997)p.236 -244

[Vivisimo]<http://www.vivisimo.com>

[WebTOC]Visualizingwebsitesusinga
hierarchicaltableofcontentsbrowser:WebTOC,
<http://www.cs.umd.edu/hcil/webtoc/>
[Zamir&Etzioni1999]OrenZamir,Oren
Etzioni,Groupier:Adynamic Clustering
InterfacetoWebSearchResults.InProceedings
ofthe21stAnnualInternationalACMSIGIR
Conference(1998)p.46 -54