Metadata Visualization of Scholarly Search Results: Supporting Exploration and Discovery

Taraneh Khazaei
Department of Computer Science
Memorial University of Newfoundland
St. John's, NL, Canada A1B 3X5
taraneh.khazaei@mun.ca

Orland Hoeber
Department of Computer Science
University of Regina
Regina, SK, Canada S4S 0A2
orland.hoeber@uregina.ca

ABSTRACT

Studies of online search behaviour have found that searchers often face difficulties formulating queries and exploring the search results sets. These shortcomings may be especially problematic in digital libraries since library searchers employ a wide variety of information seeking methods (with varying degrees of support), and the corpus to be searched is often more complex than simple textual information. This paper presents Bow Tie Academic Search, an interactive Web-based academic library search interface aimed at supporting the strategic retrieval behaviour of searchers. In this system, a histogram of the most frequently used keywords in the top search results is provided, along with a compact visual encoding that represents document similarities based on the co-use of keywords. In addition, the list-based representation of the search results is enhanced with visual representations of citation information for each search result. A detailed view of this citation information is provided when a particular search result is selected. These tools are designed to provide visual and interactive support for query refinement, search results exploration, and citation navigation, making extensive use of the metadata provided by the underlying academic information retrieval system.

Categories and Subject Descriptors

H.3.7 [**Information Storage and Retrieval**]: Digital libraries—*user issues*; H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*search process*; H.5.2 [**Information Interfaces and Presentation**]: User interfaces—*graphical user interfaces (GUI)*

General Terms

Design, Human Factors

Keywords

Information Retrieval, Information Visualization, Digital Library

1. INTRODUCTION

The ongoing development of automated search technology, along with the rapid growth of available information, has made search a fundamental part of peoples' lives [10]. Search systems are now capable of providing direct access to large information spaces, and are continuing to evolve and grow at a rapid pace. As a result, searching for information is now an integral task undertaken by people daily and is regarded as the second most frequently used online application [5].

In studies on search behaviour, it has become evident that queries crafted by searchers are often poorly formulated and do not reflect their information needs accurately. This may be due to issues such as searchers' tendency to formulate short queries [18, 22], their incomplete knowledge about their information needs [4], or their inability to express their information needs due to a lack of terminology [4]. This problem has long since been documented as one of the main issues in traditional libraries, with reference librarians reporting that few people know how to ask reference questions [15]. In traditional libraries, the active engagement of reference librarians in the search process may enable searchers to subsequently express their information need in a proper way. However, in the context of digital libraries, little assistance is provided for searchers to craft and reformulate their queries.

Similar problems exist with the interfaces for presenting search results, where the simple list-based format is commonplace. Such a representation requires searchers to extensively utilize their cognitive abilities to evaluate and compare result items by reading document surrogates (i.e., titles, snippets or abstracts, and URLs) one-by-one. In addition, there is little support for manipulation and exploration within the search results, and for identification of the overall properties of the retrieved set. This lack of support is even more problematic for large and complex information structures such as content-rich metadata-enhanced digital libraries.

Although an effective ranking method can help searchers for targeted queries, there is still a cognitive burden for exploratory search. Exploratory search tasks are often motivated by a complex information need, a poor understanding of terminology and the information space [27], or a desire to learn [19]. Such conditions are common starting points for library searchers, resulting in their desire to initiate a search process.

With the aim of promoting exploratory search activities in an academic digital library, we have designed and developed a novel search interface called Bow Tie Academic Search.

Information visualization approaches are used to visually encode metadata elements available within digital libraries. Incorporating metadata visualization into the search interface can support searchers in their retrieval tasks by enhancing their abilities to perceive, interpret, and understand features and relationships among the search results and their associated metadata. The goal is to provide a search interface that supports both interactive query refinement and search results exploration activities. By providing simple interaction techniques, an effective combination of searcher control and system retrieval power can be achieved, providing a better search experience.

The remainder of this paper is organized as follows: An overview of the previous research on visual query refinement and search results visualization and exploration is provided in Section 2. In Section 3, an explanation of the design and features of the system is given. An example is provided in Section 4 to demonstrate the benefits the system can provide to searchers via the query refinement and search results exploration features. The conclusion is presented in Section 5, along with an overview of future research activities.

2. RELATED WORK

Many systems have been developed in recent years that incorporate information visualization methods within search interfaces. In these studies, information visualization is used as a cognitive aid, substituting the slow serial process of mentally decoding text with the parallel processing power of the human visual system [28]. We categorize these systems according to whether they attempt to support query refinement or search results exploration.

2.1 Visual Query Refinement

Few studies have explored the domain of visual support for query refinement. Those that have mainly provide visual representations of potentially relevant terms and allow searchers to filter the search result set based on these terms, or to add and remove these terms from the initial query. Such research can be divided amongst those that extract terms form the entire collection of documents, the highly ranked search result items, and the external knowledge structures.

In faceted search systems, the metadata are extracted or assigned to documents based on the features of the entire underlying collection. These metadata are offered to searchers, allowing them to filter the search results set and to navigate within the documents by choosing specific facets. There are number of studies focusing on showing faceted navigation visually. Flamenco [29] makes use of hyperlinked Web pages to facilitate navigation within facets. Every hyperlink is displayed with the number of results that will be shown by choosing that facet. Relation Browser [2] is similar to Flamenco, but rather than a number, it uses a graphical bar superimposed on the category label in order to show the relative frequency of each category in the search results.

Extracting terms from the top ranked documents retrieved by the original query is an appealing alternative. Considering the power of current retrieval systems, the highly ranked documents are assumed to be relevant when the initial query is at least partially relevant to searchers' information needs. Extracting information from this collection provides a focused set of terms that may be beneficial to support query refinement. In WordBars [11, 12], a histogram is used to visually depict the most frequent terms found in the highly ranked search results. Using a histogram allows the searcher to identify the potentially relevant terms and interpret their relative frequencies. Searchers are able to remove and add these terms to the query directly from the histogram.

Candidate terms can also be extracted by looking for relevant information in external knowledge structures. For instance, in VisiQ [14], a knowledge base containing relationships between terms and concepts in the domain of computer science topics is used to represent the relationships between the initial query and candidate terms. These relationships are displayed in a graph structure, allowing searchers to quickly understand how their initial query is related to other concepts and terms. Searchers are able to add terms to their query by double-clicking on the node representing that term.

In our work, we have chosen to use the top ranked documents to provide potentially relevant terms. Since such terms are more focused on the given query, we expect this approach to be well suited for supporting the query refinement process.

2.2 Visual Search Results Exploration

Many visual interfaces have been designed and developed to support exploration and manipulation of search results. In general, there are two types of approaches when visualizing search results: visualizing query-document relationships and providing visualizations of metadata extracted from the search results. Even though spatialization of documents [21] and document clusters [25] have been proposed and used in search interfaces, they can make it difficult for searchers to understand the search results set due to the high cognitive load required to decode such visual displays [6].

Visualizing query-document relationships allows searchers to see how the individual terms in a query relate to each of the search results. One of the common methods of representing such information is to augment the search results list by adding small visual representations alongside each document, with each representation visualizing the relation of query terms with that document.

For example, in TileBars [7], a rectangular representation is shown beside each search result item which displays the frequency of the query terms in each document segment by using grey scale encoding. HotMap [8, 13] visualizes the frequency of each query term in each document surrogate by a colour-coded square located alongside the corresponding result item. In addition, HotMap allows searchers to perform nested sorting on the search results based on the query term frequencies, letting them explore within the set of search results. In similar works, a colour-coded pie chart [1] in which segment size indicates the relative frequency of a query term, and a bar chart [24] in which the overall and single keyword relevance is mapped to the length of bars, have been used to depict query-document relationships.

In another set of approaches to support search results exploration, a visualization of metadata extracted from the search results set is offered to the searcher independent of the results list. The search systems augmented with such interface extensions often include interaction tools to let searchers dynamically filter or re-sort search result items.

For instance, PubCloud [16] supports the visual exploration of search results from the PubMed database of biomedical literature by providing tag cloud representations of documents' abstracts. In addition, selecting a term allows the searcher to navigate to the relevant subject matter that otherwise might be hidden far down in the ranked list. WordBars [11, 12], which was discussed in the previous section as a system to support query visualization, also supports interactive search results exploration by using a term frequency histogram. Searchers are able to select a term of interest, which causes the search results set to be re-sorted based on the frequency of the selected term within the titles and snippets of the result items.

In our research, we have included some aspects of both approaches, allowing searchers to choose the one that best suits their current information seeking tasks.

3. BOW TIE ACADEMIC SEARCH

3.1 Data Source

Bow Tie Academic Search is a meta-search engine that uses Microsoft Academic Search (MAS) API [20] data and functions to provide scholarly search results to the searcher. MAS is one the few publicly available digital libraries that provides information on both backward citations (i.e., documents referenced by the original one) and forward citations (i.e., documents that cited the original one).

In response to the searchers' queries, MAS provides various metadata describing different aspects of the documents. The metadata elements primarily used by the system include the title, authors, venues, backward citations, forward citations, publication year, the URL, the first sixty words of the abstract, and the keywords of the document.

All of the supplied metadata is extracted from the author-supplied information for the academic documents. In some cases, however, the metadata are automatically extracted from the contents of the documents. Furthermore, since MAS is still undergoing development, sometimes the metadata is incomplete.

3.2 Query Support with Keyword Histogram

3.2.1 *Visual Representation of Keyword Metadata*

In academic libraries, some terms are often assigned to documents in order to represent the core concepts of the documents. Representing keyword information associated with retrieved documents can support identification of important aspects of the search results set. As such, a compact visualization of keyword metadata is offered to the searchers, allowing them to recognize the most frequent keywords used in the top search results, to perceive their relative frequencies, and to understand keyword-document relationships.

As the documents are retrieved from the MAS API, the system counts the keywords associated with the top 500 documents. The keyword frequencies are visually represented using a vertically oriented histogram located at the right side of the search results list, where the size of each bar represents the frequency of the associated keyword. Due to the space limitations, only the first 25 keywords are displayed in the histogram. This approach is similar to what WordBars [11, 12] presents in the Web search context. By browsing this information, searchers are able to understand the general attributes of the search results, and may use this information to evaluate whether the retrieved set is capable of satisfying their information needs.

To represent interrelations of the documents, the histogram is enhanced with a visual depiction of which documents in

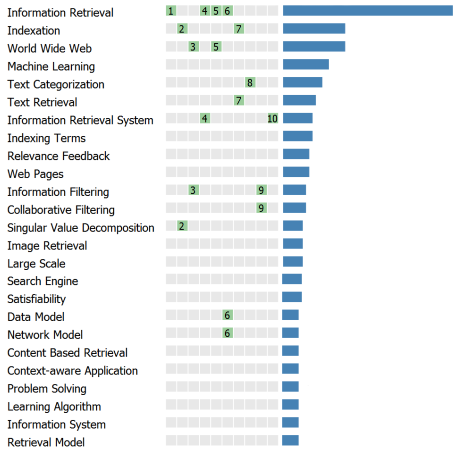


Figure 1: The enhanced histogram of the most frequent keywords allows searchers to identify the most frequent keywords in the search results, their relative frequencies, and document interrelations in terms of their keywords.

the current page of the search results set use each keyword in the histogram. A relation grid is provided, where each row represents a keyword in the histogram and each column represents a document currently being shown to the searcher. Since there are ten documents shown per page of search results, and 25 keywords extracted from the top search results, this grid is of size 10×25 . When a specific keyword is associated with a document of the page, the corresponding cell in the relation grid is highlighted with a colour and tagged with a number, which is the document rank in that page (Figure 1).

Adding the relation grid to the histogram equips this component of the system with the ability to visualize keyword connections across documents, allowing searchers to easily see and understand which documents share the same keyword. Furthermore, scanning down each column reveals the keywords used by each of the search results on the current page. The relation grid can also assist in the document discovery process as searchers are able to see which documents are using the keywords that are interesting to them and by revealing if the keywords of a specific document are among the most frequently used ones. Since the most common keywords may be the most relevant ones to the searchers' information needs, providing this representation can allow searchers to find potentially relevant documents from the search results list.

3.2.2 *Interactive Search Results Exploration*

After identifying the potentially relevant documents by using the relation grid, searchers are able to select the corresponding cell to perform further evaluation of the document. This selection causes the search results list to be scrolled to target the corresponding document. The document is also highlighted, allowing searchers to quickly and easily identify the selected document in the list of search results.

The coordination between the relation grid and the list of search results supports searchers in an interactive exploration of search results, allowing them to both gain an insight into the features of the documents currently being shown as well as investigate the documents in more detail.

3.2.3 *Interactive Query Refinement*

Further interaction techniques are provided to support interactive query refinement by allowing searchers to iteratively convert their initial queries to a well-defined one. When searchers start with a vague and underspecified query, it may be difficult for them to determine alternative ways to improve the query. Here, the histogram assists searchers in identifying useful alternative terms. Providing these terms to the searchers supports recognition of relevant terms instead of requiring searchers to remember them [11, 12]. Assuming that searchers start with an initial query that is at least somewhat relevant to their information needs, the histogram may allow them to find more accurate descriptions of what they are seeking.

Using the checkbox provided at the left side of each keyword, the searcher is able to add or remove terms from the current query. The keywords that consist of multiple terms are added in quotes, allowing searchers to focus on a more specific set of search results. In addition, the searcher is able to change the focus of the search by single clicking on any keyword, which replaces the current query with the selected keyword. After refining the query, searchers can click on the search button to indicate the end of the query refinement process and to retrieve a new set of search results.

3.3 List Augmentation using Bow Ties

3.3.1 *Visual Representation of Citation Metadata*

Representing citation metadata can provide valuable information by drawing attention to important published works [17]. In addition, citations can be used to analyze research trends, identify new or active areas of science, and find out where and how often a particular article is cited [17]. However, current systems do not provide an intuitive way for searchers to evaluate and compare documents based on their citation information or navigate and explore through citations of a document.

While previous research mostly used textual information or node-link diagrams to represent citation information [5], we opted for a simpler approach to enable searchers to compare documents with relative ease and accuracy. In this approach, a visual representation of citation information is located alongside each document as an interface extension to the well-known list-based representation.

In order to provide a compact visualization of citation information, details of the citation tree for each document are abstracted away, generating a representation that takes the form of a bow tie (Figure 2). In this representation, the document node is placed in the middle containing the year that the document was published. The left side represents information about backward citations, while the right side represents forward citation information. This order of backward citations to the left and forward citations to the right follows the common representation of time flowing from left to the right [23]. The number of backward citations in a document is mapped to the width of the left side; and the right side's width indicates how often the article is cited in other works. Although direct decoding of the widths of the bow tie representation into the number of citations will not be possible, the relative differences can be perceived by ensuring that the centre of the document nodes in a collection of bow tie representations are lined up vertically.

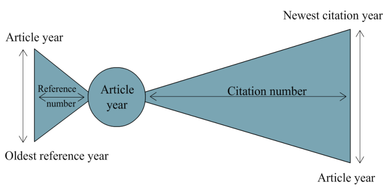


Figure 2: Metadata regarding document citation information is mapped to visual features of the bow tie representation.

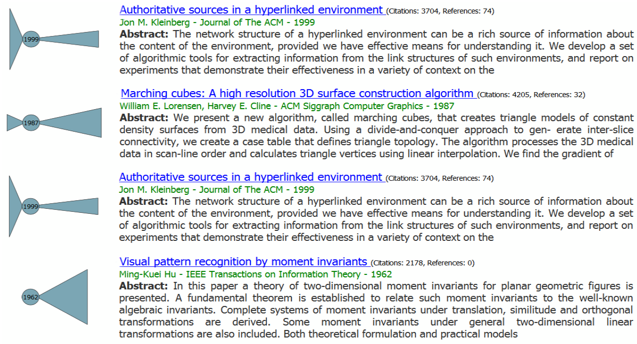


Figure 3: Augmenting the search results list with bow tie representations supports visual identification of potential documents by encoding features of the citation lists, along with the year of publication.

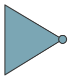
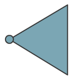
The height of the left side shows the period from the publication year of the oldest backward citation to the year that the document itself was published. This mapping allows searchers to find those documents that provide an extensive coverage of the earlier studies that cover a long period of time. The right side's height represents the period between the document publication year and its most recent forward citation, letting searchers distinguish newly cited documents from potentially obsolete ones. As bow tie representations convey time in two different levels of detail, they consist of two different timelines: A high-level timeline which is mapped to the horizontal dimension, and a low-level one which is mapped along the vertical dimension.

These bow tie representations compactly and simultaneously convey backward and forward citation metadata as well as document year information to the searcher, and can readily be interpreted with little training (see Table 1). Locating these representations alongside the search results list provides searchers with the ability to perceive and interpret this metadata, and then to make fast judgments on the suitability of each document to their given retrieval purpose. For instance, by comparing the widths of bow tie representations of the retrieved documents shown in Figure 3, the last item can be easily identified as a seminal work as it is cited many times. In addition, its height indicates that it is also recently cited and so is not obsolete. On the other hand, if someone is looking for a new document that provides a good coverage of the previous work, the first one is potentially a satisfactory one.

3.3.2 *Document Selection for Detailed Evaluation*

The search results interface augmented with bow tie representations assists searchers in comparing documents and

Table 1: Based on citation information of a document, different forms of bow tie representations can be generated, with each bow tie conveying specific characteristics of the corresponding document.

Backward Citations		Forward Citations	
	Many backward citations that cover a long period of time.		Many forward citations that cover a long period of time.
	Many backward citations, with the oldest one published a few years before document publication.		Many forward citations, with the most recent one published a few years after document publication.
	A few backward citations that cover a long period of time.		A few forward citations that cover a long period of time.
	A few backward citations, with the oldest one published a few years before document publication.		A few forward citations, with the most recent one published a few years after document publication.

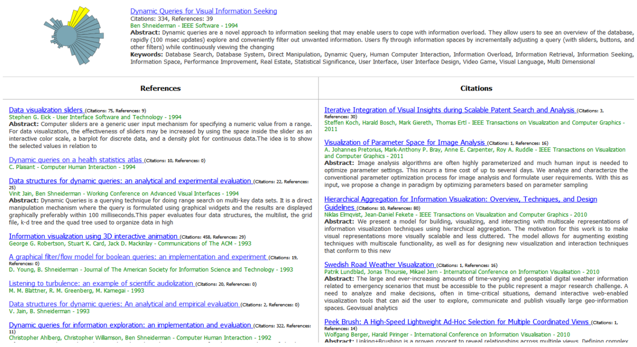


Figure 4: Detailed representation of the document containing the detailed bow tie visualization and lists of forward and backward citations, along with the filtering feature, facilitates further evaluation of individual documents and backward and forward navigation among citations.

finding potentially relevant ones based on their citation information. However, it does not encode the detailed information of backward and forward citations. Therefore, after the recognition of potentially interesting documents in the augmented list of search results, searchers are able to click on the corresponding bow tie representation to perform further detailed evaluation of that document.

When the searcher clicks on a bow tie representation of a document, a detailed bow tie visualization of the document citation information is shown in a new page. The visual and interactive features of the detailed bow tie representations are explained in the following section.

3.4 Document Focus with Detailed Bow Ties

3.4.1 Detailed Visualization of Citation Metadata

The detailed representation of the selected document consists of the detailed bow tie representation of citation metadata and a description of the other metadata elements associated with the document in a textual format. Two columns containing the lists of backward and forward citations of the document are also provided (Figure 4).

While the bow ties added to the search results list represented an abstract view of the citations, the premise here is that the distribution of citations through different years is an important criterion for evaluating individual documents. For example, among two documents that have been cited an equal number of times, the one that has the greater number of recent forward citations may better represent a hot topic. Perceiving this valuable information in textual format is a difficult task that requires careful consideration and evaluation of each document in the citation list. Instead, a visual representation of this information is offered to the searchers, allowing them to evaluate each document in more detail.

In order to visually convey this information to the searcher, the number of backward and forward citations in each year is extracted from the metadata of the document and counted. The height of the bow tie representation of the selected document, which represents the publication period, is subdivided into bars, each representing one year in that period. Further, the number of citations published in each year is mapped to the length of these bars, providing a detailed bow tie representation. Using the bow-tie representation, which the searcher is already familiar with, reduces the learning curve associated with using the system. Furthermore, the radial layout provides a more compact representation in comparison to its alternatives. As the human visual system is capable of comparing and evaluating quantitative information when mapped to the length, searchers are able to easily understand how citations of a document are distributed in different years.

In current search interfaces, if searchers want to view backward or forward citation list of an article, they need to click on a link which shows a new page containing just the requested list, while the original document is left behind. Providing document information in detail on top of the citation lists enables searchers to navigate through these lists without losing focus on the original document.

3.4.2 Exploration and Navigation

Two strategies widely used by library searchers are "backward chaining" and "forward chaining" in which searchers often start with one or more relevant research materials, and then forage for other works through articles that were referenced by this document, or articles that cite this document [3]. To support this strategic behaviour associated with information seeking, the detailed bow tie representation is further augmented with interaction techniques. Searchers are provided with a filtering function to explore within the documents' citations. This feature provides an easy and effective way for searchers to perform backward and forward navigation among documents.

To activate the function, searchers can select different years from the detailed bow tie representation by single clicking on the bar representing that year. This operation causes the backward or forward citation list to be filtered to only show the corresponding data. When a filter is requested, the colour of the year bar is changed as a visual reminder that the filter is in effect. As the colour of the detailed bow tie is blue, yellow has been chosen to label the selected year. Following the opponent process theory of colour, this colour difference between the non-selected and selected years can be pre-attentively processed [26], facilitating easy and fast identification of the selected filters. Re-clicking on the already selected elements deactivates the filter.

Providing this interaction technique promotes searchers' involvement in the search process and facilitates navigation within citations of documents. It also allows searchers to narrow down the scope of their search to a manageable set of documents, enabling them to focus on documents of interest. This is particularly helpful for the documents that have many backward or forward citations. Although these kinds of documents are often good starting points for finding other relevant ones, it is beyond the searchers' tolerance to search within their citations without exploration support.

4. EXAMPLE

Suppose a searcher starts with an initial query "information retrieval" with the goal of finding interesting documents in the field. When the searcher submits the query, the system retrieves documents from MAS API and further offers them in an augmented list to the searcher. Meanwhile, the histogram of the most frequent keywords used in the retrieved documents, along with the relation grid, is generated and presented to the searcher.

By reviewing the top keywords in the histogram, it can be realized that the query "information retrieval" produces documents covering a wide range of subtopics (Figure 5a). Therefore, the searcher may inspect the histogram, looking for terms that better describe the desired information that is being sought. The searcher is able to focus on a more specific set of search results by checking the keyword "relevance feedback", which adds it to the initial query. Then, the searcher can click the search button to submit the new query and review the new set of search results.

After the submission of the new query, the updated keyword histogram and the bow tie representations can be used to explore the search results list. As can be seen in Figure 5b, by scrolling down the list, the searcher can easily realize that the fifth and the sixth documents of the page are relatively old documents that are cited many times. In addition, by looking down the document columns in the relation grid, the searcher can quickly understand that the eighth and ninth documents of the page have most of their keywords among the most frequent and potentially prominent ones. So, even though they are new documents that have not been cited considerably, they can be still identified as potentially relevant documents by the searcher.

In addition, when a document of interest is found, the searcher may again inspect the keyword histogram to find other relevant documents. For instance, if the ninth document is identified as an interesting one, the searcher may check for other documents having mutual keywords with the ninth one by looking down its corresponding column in the relation grid. Clicking on a cell in the grid will scroll the search results list to this document and will highlight it, allowing the searcher to quickly determine its relevance.

The searcher may decide that the fifth document is worthy of further examination, and may wish to perform further evaluation and exploration with the goal of finding other interesting but more recent documents. Clicking on the bow tie for this document loads a new page with the detailed bow tie representation. The searcher can filter the forward citations of the document to the most recent ones by selecting the upper bars in the right side of the detailed bow tie (Figure 5c). If needed, the filtering can also be performed in the backward citation list to find older documents.

Conducting this search task with the assistance of Bow Tie Academic Search has not only guided the searcher towards crafting a better query, but also in exploring the search results set, inspecting a particular search result in detail, and performing further exploration of the citation information. Performing similar activities with current search interfaces would have been time consuming and requiring significant cognitive effort.

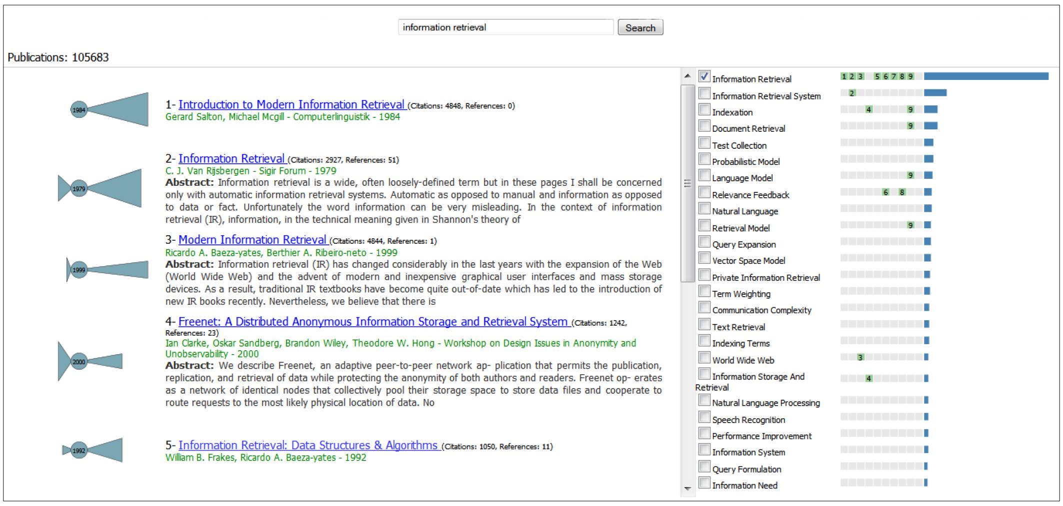
5. CONCLUSION

While the Web facilitates fast and direct access to large collections of academic resources, the ways people obtain information from these collections are based on Web search technologies and interfaces. Bow Tie Academic Search takes advantage of the rich metadata associated with academic documents and employs information visualization techniques to support searchers during their search process. This is an example of next-generation information retrieval systems to support the search process beyond the simple query box and search results list [9, 10].

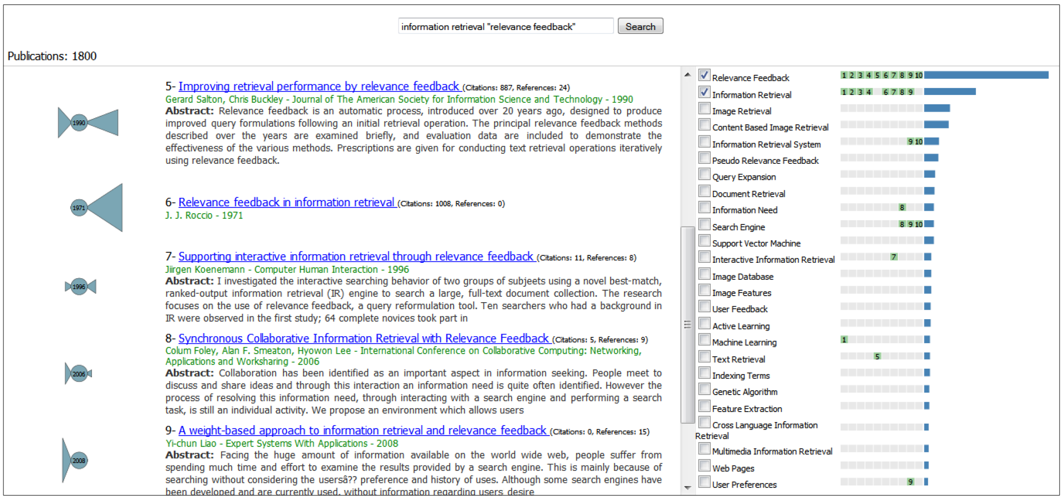
The enhanced keyword frequency histogram provides a visual overview of the search results, allowing searchers to evaluate the overall relevance of the retrieved set, to interpret the relevance of each keyword, and to perceive documents similarities based on the co-use of keywords. Query refinement is supported through the interaction methods of the histogram. In addition, visual representations of citation metadata provides searchers with the ability to make quick judgments about the potential importance and relevance of the documents. The interactive features of these representations allow searchers to navigate within the citations of a document, aiding them in employing the common backward and forward chaining strategies.

Since the system is a meta search engine, the constraints and limitations of the underlying search engine may affect the quality of the output. For example, since the current prototype implementation is based on Microsoft Academic Search, when that search engine provides incomplete information (e.g., missing abstracts, document years, and/or citations), our system must also exclude this information. However, as the underlying search engine improves so will the Bow Tie Academic Search.

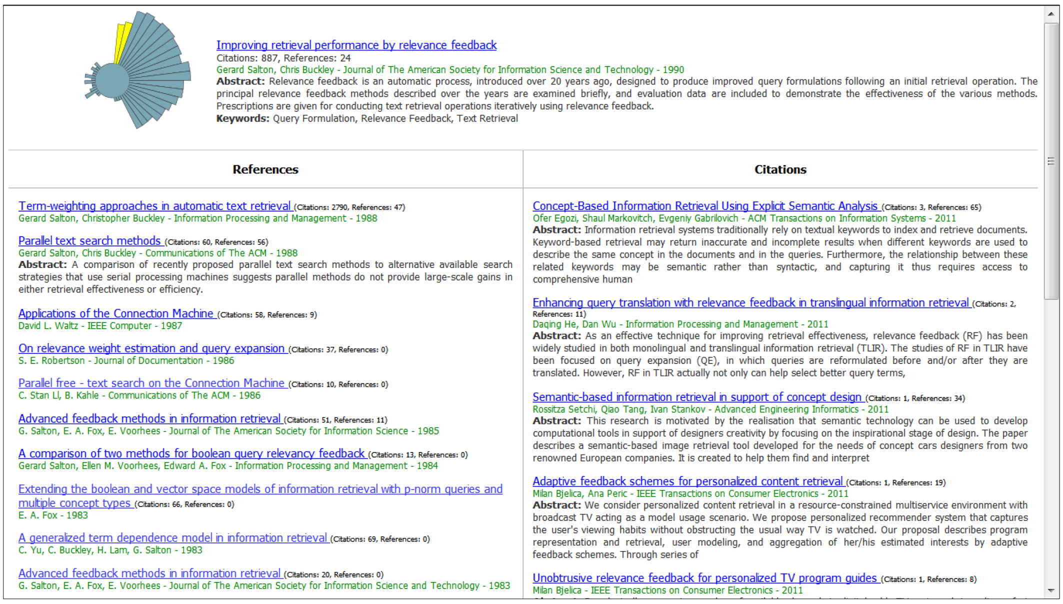
Future work includes further refinement and enhancement of the prototype, adding more visual features to support the query refinement and document discovery process, and exploring alternative methods for supporting the interactive forward and backward navigation within the entire retrieved



(a) The enhanced histogram and the augmented list of search results are generated by the searcher's initial query.



(b) After the refinement of the query by the searcher, the visual representations of keyword and citation metadata are updated based on the characteristics of the new set of search results.



(c) The searcher is able to explore and manipulate citation lists for a document of interest.

Figure 5: An example of query refinement and search results exploration using Bow Ties Academic Search.

set of documents. User studies are currently in the planning stages, the goal of which are to measure the benefits of the specific design choices, in comparison to digital library search interfaces that follow the traditional query box and list-based search results paradigms.

6. REFERENCES

[1] T. J. Anderson, A. Hussam, B. Plummer, and N. Jacobs. Pie charts for visualizing query term frequency in search results. In *Proceedings of the International Conference on Asian Digital Libraries*, pages 440–451, 2002.

[2] R. G. Capra and G. Marchionini. The relation browser tool for faceted exploratory search. In *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, page 420, 2008.

[3] D. Ellis. A behavioural model for information retrieval system design. *Journal of Information Science*, 15(4-5):237–247, 1989.

[4] G. Furnas, T. Landauer, L. Gomez, and S. Dumais. The vocabulary problem in human- system communication. *Communications of the ACM*, 30(11):964–971, 1987.

[5] M. Hearst. *Search User Interfaces*. Cambridge University Press, Cambridge, 2009.

[6] M. Hearst. User interfaces for search. In R. Baeza-Yates and B. Ribeiro-Neto, editors, *Modern Information Retrieval: The Concepts and Technology Behind Search*, pages 21–55. Addison-Wesley, Harlow, England, second edition, 2011.

[7] M. A. Hearst. TileBars: Visualization of term distribution information in full text information access. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 59–66, 1995.

[8] O. Hoeber. The visual exploration of Web search results using HotMap. In *Proceedings of the International Conference on Information Visualization*, pages 157–165, 2006.

[9] O. Hoeber. Web information retrieval support systems: The future of Web search. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence - Workshops (International Workshop on Web Information Retrieval Support Systems)*, pages 29–32, 2008.

[10] O. Hoeber. Human-centred Web search. In C. Jouis, I. Biskri, J.-G. Ganascia, and M. Roux, editors, *Next Generation Search Engines: Advanced Models for Information Retrieval*, pages 217–238. IGI Global, Hershey, PA, 2012.

[11] O. Hoeber and X. D. Yang. Interactive Web information retrieval using WordBars. In *Proceedings of the IEEE/WIC/ACM Conference on Web Intelligence*, pages 875–882, 2006.

[12] O. Hoeber and X. D. Yang. Evaluating WordBars in exploratory Web search scenarios. *Information Processing and Management*, 44(2):485–510, 2008.

[13] O. Hoeber and X. D. Yang. HotMap: Supporting visual explorations of Web search results. *Journal of the American Society for Information Science and Technology*, 60(1):90–110, 2009.

[14] O. Hoeber, X. D. Yang, and Y. Yao. VisiQ: Supporting visual and interactive query refinement. *Web Intelligence and Agent Systems: An International Journal*, 5(3):311–329, 2007.

[15] W. Katz. *Introduction to Reference Work: Reference Services and Reference Processes*. McGraw-Hill, NY, 1978.

[16] B. Y.-L. Kuo, T. Hentrich, B. M. Good, and M. D. Wilkinson. Tag clouds for summarizing Web search results. In *Proceedings of the International Conference on World Wide Web*, pages 1203–1204, 2007.

[17] S. Lawrence, C. Lee Giles, and K. Bollacker. Digital libraries and autonomous citation indexing. *Journal of Computer*, 32(6):67–71, 1999.

[18] M. Mahoui and S. J. Cunningham. Search behavior in a research-oriented digital library. In *Proceedings of the Conference on Research and Advanced Technology for Digital Libraries*, pages 13–24, 2001.

[19] G. Marchionini. Exploratory search: From finding to understanding. *Communications of the ACM*, 49(4):41–46, 2006.

[20] Microsoft. Microsoft Academic Search. `http://academic.research.microsoft.com/About/Help.htm`.

[21] B. Shneiderman, D. Feldman, A. Rose, and X. F. Grau. Visualizing digital library search results with categorical and hierarchical axes. In *Proceedings of the ACM Conference on Digital Libraries*, pages 57–66, 2000.

[22] A. Spink, D. Wolfram, B. J. Jansen, and T. Saracevic. Searching the Web: The public and their queries. *Journal of the American Society for Information Science and Technology*, 52(3):226–234, 2001.

[23] E. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, Conn, 1983.

[24] A. Veerasamy and N. Belkin. Evaluation of a tool for visualization of information retrieval results. In *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 85–92, 1996.

[25] J. Wang, A. Agrawal, A. Bazaza, S. Angle, E. A. Fox, and C. North. Enhancing the ENVISION interface for digital libraries. In *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, pages 275–276, 2002.

[26] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann, San Fransico, second edition, 2004.

[27] R. W. White, B. Kules, S. Drucker, and m. c. schraefel. Supporting exploratory search: Introduction to special section. *Communications of the ACM*, 49(4):36–39, 2006.

[28] J. A. Wise, J. J. Thomas, K. Pennock, D. Lantrip, M. Pottier, A. Schur, and V. Crow. Visualizing the non-visual: Spatial analysis and interaction with information from text documents. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 51–58, 1995.

[29] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image search and browsing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 401–408, 2003.