Visualizing Web Search Results Using Glyphs: Design and Evaluation of a Flower Metaphor

MICHAEL CHAU, The University of Hong Kong

While the Web provides a lot of useful information to managers and decision makers in organizations for decision support, it requires a lot of time and cognitive effort for users to sift through a search result list returned by search engines to find useful information. Previous research in information visualization has shown that visualization techniques can help users comprehend information and accomplish information tasks more efficiently and effectively. However, only a limited number of such techniques have been applied to Web search result visualization with mixed evaluation results. Using a design science approach, this research designed and implemented a glyph (a graphical object that represents the values of multiple dimensions using multiple visual parameters) and a system for visualizing Web search results. A flower metaphor was adopted in the glyph design to represent the characteristics and metadata of Web documents. Following the cognitive fit theory, an experimental study was conducted to evaluate three displays: a numeric display, a glyph display, and a combined display which showed numbers only, glyphs only, and both, respectively. Experimental results showed that the glyph display and the combined display performed better when task complexity was high, and the numeric display and the combined display performed better when task complexity was low. The combined display also received the best perceived usability from the subjects. Based on the findings, the implications of the study to research and practice are discussed and some future research directions are suggested.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Graphical user interfaces (GUI)*; H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*Search process*

General Terms: Design, Experimentation, Human Factors

Additional Key Words and Phrases: Information visualization, Web searching, glyphs, design science

ACM Reference Format:
Chau, M. 2011. Visualizing Web search results using glyphs: Design and evaluation of a flower metaphor. ACM Trans. Manag. Inform. Syst. 2, 1, Article 2 (March 2011), 27 pages.
DOI = 10.1145/1929916.1929918 http://doi.acm.org/10.1145/1929916.1929918

1. INTRODUCTION

The Internet has made a lot of information available to management at various levels. With tens of billions of pages on the World Wide Web, it is possible to find information for a large variety of purposes, such as competitor intelligence, market intelligence, and product analysis, to support decision making [McGonagle and Vella 1999; Chen et al. 2002; Chau et al. 2007]. However, searching for useful ones could be a big problem. Most people rely on Web search engines, such as Google or Yahoo, to search for Web pages

---

This research was supported in part by the University of Hong Kong grant "Applying Information Visualization Techniques to Web Search Results" HKU Seed Funding for Basic Research, March 2008–February 2010 (10208140).
Author's address: M. Chau, School of Business, Faculty of Business and Economics, The University of Hong Kong, Pokfulam, Hong Kong; email: mchau@business.hku.hk.

DOI 10.1145/1929916.1929918 http://doi.acm.org/10.1145/1929916.1929918

that are relevant to their decision making tasks on hand and filter out the irrelevant ones [Spink et al. 2001]. Popular Web search engines, relying on various information technologies, allow users to type in one or several keywords and search for documents containing these words in database systems. This will often result in millions of documents in the result set [Bowman et al. 1994]. Out of this set, the top documents will be returned to users, often as a plain list of page titles and snippets (short summaries of the documents, often with keywords highlighted) and a user has to browse through the list to search for relevant and high-quality results. However, browsing through a long list of documents to locate the information needed could be a mentally exhausting task. This problem of *information overload* is often faced by decision makers in management organizations [O'Reilly, 1980; Farhoomand and Drury 2002].

According to the information search process model of Kuhlthau [1991, 1993], information search consists of six stages, namely *initiation*, *selection*, *exploration*, *formulation*, *collection*, and *presentation*. The formulation and presentation of the search results are the two stages in which the users interact most closely with the search systems in the process. While the formulation stage has received a lot of attention in the computer science and information systems communities (e.g., Storey et al. [2008]), the presentation of search results to users is often overlooked by researchers and commercial search engines. The Web search result display nowadays is still more or less the same as that of the early search engines (like AltaVista and Lycos) developed more than a decade ago: a one-dimensional list showing the top search results. Users have to browse through the list of items one by one to get an overview of the results. Although many new techniques have been proposed and there have been a few commercial efforts (e.g., Kartoo and Grokker), they have not been widely adopted in popular Web search engines like Google or Yahoo, which still return search results as a one-dimensional list.

On the other hand, data visualization and information visualization have drawn a lot of interest in recent years. Studies on two-dimensional and three-dimensional graphs have been widely reported [Tufte 1990; Tan and Benbasat 1990; Bennett and Toms 1993; Cheng et al. 2001; Kumar and Benbasat 2004]. Many techniques also have been proposed for the visualization of multidimensional data. For example, the parallel coordinates display is a widely used method which represents attribute values as points on different coordinates which are parallel lines [Wegman 2003]. Another approach to visualizing multidimensional data is *glyph* representation. Glyphs are graphical objects that represent the values of multiple dimensions by multiple visual parameters such as positions, colors, sizes, and shapes. Glyphs have been studied for visualizing different types of data and different glyphs have been proposed [Chernoff 1973; Chuah and Eick 1997; 1998; Scott 1992; Fanea et al. 2005]. While these techniques have been proven useful for user understanding of information and efficiency in information analysis, only a limited number of visualization techniques such as two-dimensional graph displays [Lin et al. 2000], three-dimensional workspaces [Card et al. 1996], and tree displays [Lamping et al. 1995] have been applied to Web document representation and visualization. These reported representations are often useful for visualizing a large amount of documents (e.g., thousands or millions) and are mostly designed for *browsing* rather than *searching* [Chen et al. 1998; Kohonen et al. 2000]. Previous studies are inconclusive on whether they are good for Web information *search* tasks. Many users are familiar with the linear list-based search results provided by popular search engines like Google, Yahoo, or Bing and find it difficult to use visualization techniques that are significantly different from the one they are familiar with [Xiang et al. 2005]. Commercial efforts like Kartoo that utilized advanced visualization techniques did not gain substantial popularity. When searching the Web, most users concentrate on the keywords they have searched for and the quality of the document [McDonald and Chen 2006], rather than its structure or high-level context. Such structural or semantic information and the novel visualization techniques employed may not fit the mental models of the users [Chen et al. 1998]. In addition, when only looking at a small subset of the Web (e.g., a set of 10 documents in a search result list), the documents may be disconnected and do not form any tree or network.

Complex visualization displays such as two-dimensional or three-dimensional graphs that have been employed suffer two main problems. They are too different from the existing displays that users are familiar with and often do not fit well with users' mental models. On the other hand, list views used by popular search engines are easy to use and widely accepted, but they do not convey as much information as quickly as graphs. It would be desirable to combine the strengths of both graphical visualizations and list-based views. Surprisingly, there is very little research that has attempted to do this. It would be interesting to study how such a combined list-based graphical visualization would be useful in presenting Web search results.

The rest of the article is organized as follows. The next section provides a review of different information visualization techniques, especially those that have been used to address the problem of Web document visualization. The limitations of existing techniques are discussed. Then the research methodology based on a *design science* approach is described. The section that follows presents the design and implementation of the proposed artifacts that aim at addressing the problem in information search on the Web. The article then reports the design and results of an experimental study conducted to evaluate the performance of three different displays. Finally, the findings of the research are discussed and some future directions are suggested.

2. LITERATURE REVIEW

2.1. Searching and Visualization of Web Documents

Although search engines have largely improved their search performances in the past few years, it is still common for irrelevant documents to appear in the search result list, oftentimes even in the top 10 or top 5. To help users quickly gain an overview of the documents in the result list, search engines often provide textual summaries to assist users in judging what documents are potentially relevant and should be clicked for browsing the full content. In addition to query-based summaries where the search keywords are highlighted [McDonald and Chen 2006], additional data such as relevance score, number of keyword matches, document size, or page popularity are often presented as numbers to help users make a better judgment in selecting documents [Brin and Page 1998]. In essence, each document is represented by the values of these multiple textual and numerical attributes. As these attributes are all useful in judging the quality and relevance of a document, one has to process all the information in order to decide what pages out of the result set are worth the time and effort for further exploration. As discussed earlier, this could be a time-consuming and mentally exhausting task. A user relying on only a subset of these attributes to make a judgment may end up wasting time in downloading and reading documents that are not relevant to the search queries, thus wasting even more time and effort.

Numerous previous studies have shown that information visualization techniques can help users comprehend information and perform tasks more effectively and efficiently [Carter 1947; Pinker 1990; Hearst 1995; Xiang et al. 2005]. According to Buja et al. [1996], visualization research consists of two areas, namely *rendering* and *manipulation*. These two areas can be viewed as two distinct phases in information visualization: *graph construction*, which relates to the construction of the graph based on the given information, and *graph manipulation*, which concerns the iterative manipulation of the graphs by users [Kumar and Benbasat 2004].

Shneiderman proposed a framework in which graph construction research can be classified according to the type of data to be visualized. In this framework, data for visualization can be classified into seven different types: *one-dimensional*, *two-dimensional*, *three-dimensional*, *multidimensional*, *temporal*, *tree*, and *network* [Shneiderman 1996]. Turetken and Sharda proposed a framework classifying graph construction techniques designed especially for the Web [Turetken and Sharda 2007]. In their framework, Web visualization techniques can be classified based on four dimensions, namely the source of the Web space, the basis of Web space organization, the resulting data structure, and the visualization paradigm. As the focus of this research is the construction of graphs for the visualization of Web search results, the following subsection will review relevant research in this area.

2.2. Visualization of Web Search Results

Since the advent of the Web, visualization techniques have been applied to Web documents and other Web-related data. The following review focuses on how visualization techniques have been applied to Web search results rather than other types of Web data like particular Web sites or the Web as a whole.

Visualization of search results is different from visualization of the other Web spaces in several ways. First, general Web space visualization often aims to revealing the structure and relationship of the Web sites or Web pages being visualized, while search result visualization should help users with their search tasks. Second, general Web spaces usually contain documents that are already connected to each other, while search results contain documents that may only be semantically similar but not hyperlinked. Third, when visualization techniques are applied to Web search results, the Web data is *dynamic* because every search would return a different set of results, while other Web space visualization research has focused on data that is relatively static.

To visualize Web search results, one first needs to decide what attributes of Web documents would be used for organizing the Web documents for visualization. According to Turetken and Sharda [2007], Web space organization can be based on various attributes, such as semantic content (e.g., title and term frequencies), connectivity (e.g., number of incoming links and outgoing links), and others (e.g., metadata like document size and the Web domain). This dimension is applied to classifying existing literature and each category is reviewed in the following. It should be noted that the current article does not aim to provide an exhaustive review. Readers are referred to Turetken and Sharda [2007] for a more thorough review.

Semantic content. All popular commercial search engines, such as Google, Yahoo, and Bing, return search results as a one-dimensional list. For each document on the list, the title, URL, a snippet, and other information are shown to users [Brin and Page 1998; McDonald and Chen 2006]. Although it is widely adopted and users are very familiar with it, such a list does not really utilize any visualization techniques except some simple color coding. In recent years, popular search engines have tried to display search results in a more organized way. For example, a search for "Miami" in Bing would return a page with a box containing some information about the city (like population, weather, maps, and attractions) above the regular search results. The same search in Google would result in a list that places a map and some photos of the Miami city on top.

Many previous studies aimed to organize search results by their content similarities in order to come up with a better structure for visualization. Traditional information retrieval research has utilized word tokenization and indexing algorithms to extract words or phrases from documents in order to represent each document as a "bag of words" [Salton 1989]. Represented by a vector of term frequencies (or other metrics based on these frequencies), a document's similarity can be compared with that of another document by using such measures as cosine similarity or Jaccard's score. Clustering algorithms can be used to organize these documents into structure which can be visualized using various paradigms. For example, the Grokker search engine (which is out of business now) presents search results on a two-dimensional, zoomable map. The MultiSurf system shows a network of how the search results relate to the search queries [Hasan et al. 1995]. The NIRVE system presents search results using two-dimensional and three-dimensional displays [Sebrechts et al. 1999]. Another search engine Clusty presents clustered search results using a tree structure. The self-organizing map technique [Roussinov and Ramsey 1998; Chen et al. 1998, 2003] and the fisheye view [Turetken and Sharda 2004, 2005] also have been used to visualize Web search results clustered by semantic similarity. Most of these studies have shown that visualization techniques allow users to perform better in their search tasks.

Connectivity structure. Another way to organize Web documents is to rely on their connectivity structure [Turetken and Sharda 2007]. While this has been widely applied to documents within the same Web sites or a set of related sites [Munzner 1998; Mak et al. 2002; Pirolli et al. 2003], it is seldom applied to Web search results. The reason is that search results may have only a very few or even no links between them. The search results can possibly come from different Web sites that are not directly connected. One system that visualizes the connectivity of search results is the Fetuccino system developed by IBM [Ben-Shaul et al. 1999]. This system visualizes the search results as a tree structure. However, this system uses a local search based on seed URLs, which result in Web documents that are directly connected to each other. Other similar systems that fall into this category include the Google-enabled Visual Search (http://www.onlineilink.com/demos/google/), the CI Spider [Chen et al. 2002], Redips [Chau et al. 2007], and Card-Vis [Mukherjea and Hara 1999]. As mentioned, the results from a general-purpose search engine like Google are often not connected and these techniques may not be applied.

Metadata. Besides semantic and connectivity attributes, a document can be represented by other metadata, such as document length, document type, the domain or site that the document resides in, the number of links, frequency of search keywords, and a variety of other metrics [Chau and Chen 2008; McDonald and Chen 2006]. To visualize these multidimensional attributes, multidimensional visualization such as *glyphs* could be a good fit for Web search result visualization. Glyphs were first proposed for visualizing multidimensional data [Chernoff 1973]. In the well-known Chernoff Faces, different facial features, such as head eccentricity, eye size, nose size, mouth shape, and mouth length, are used to represent the values of different dimensions of the data [Chernoff 1973; Scott 1992]. The Chernoff Faces were later extended to include human body figures [Pflughoeft et al. 2005a, 2005b]. Since the Chernoff Faces were proposed, various types of glyphs have been developed, such as time-wheels, insects, stars, flowers, trees, castles, pies, surfaces, and polygons, among others [Kleiner and Hartigan 1981; Chuah and Eick 1998; Ebert et al. 1997; Xiong and Donath 1999; Roberts et al. 2002; Fanea et al. 2005; Forsell et al. 2006; Wiza et al. 2003; Cellary et al. 2004]. In general, glyph representations map the value of each attribute of an input tuple to a visual dimension of the glyph. The use of glyphs in data visualization has been studied for various types of data, such as economic indicators, weather information, and text documents, among others [Scott 1992; Rohrer et al. 1998; Sangole and Knopf 2002]. Ward provides a detailed review of different examples of glyphs and proposes a taxonomy of strategies for placing glyphs in the display [Ward 2002]. All these glyphs have achieved various levels of success for several reasons. First, glyphs can represent rich information as a small object; a picture is worth a thousand words. Second, many glyphs, such as human faces and flowers, are representations that users are familiar with. Users often find it easy to interact with these glyphs with their innate ability

[Chernoff 1973]. Third, compared with complex visualization structures such as network displays, glyphs are simple and can be easily understood by users. The different visual cues in glyphs can represent the values of different parameters based on their respective scales, allowing users to understand and compare the different attributes [Pinker 1990]. Fourth, glyphs can be easily integrated with the popular list view. This can be used to augment the list view that users are already familiar with, rather than having a new visualization drastically different from the existing paradigm.

Despite the several benefits of glyphs, little research on Web search result visualization using glyphs has been reported. Only two studies were identified [Roberts et al. 2002; Cellary et al. 2004]. Polygon glyphs, including rectangles, triangles, and circles, were used in these studies to visualize the attributes of search results like page size and number of internal links and external links. One shortcoming is that polygons are not as intuitive as other glyphs. Some information conveyed by human faces, for example, cannot be shown by simple polygons. In addition, these displays were not integrated into the traditional search result list, making it difficult for users in Web searching. The polygon glyphs were positioned on a two-dimensional or three-dimensional display which users were less familiar with when compared with the traditional search result list display. This might have made it difficult for users to work with the display. It would be desirable to investigate whether glyphs combined with a list-based display can facilitate users in their Web search process. In addition, no user evaluation was reported in these two studies. The performance of glyphs in visualizing Web search results has not been established and evaluated.

3. A DESIGN SCIENCE APPROACH

Based on the preceding review, the objective of this research is to design, develop, and evaluate an intuitive glyph visualization that can be combined with the traditional list-based visualization in presenting Web search results. In this research, the *design science* approach was adopted as the methodology [March and Smith 1995; Hevner et al. 2004]. Hevner et al. provide a framework for design science research in the field of information systems [Hevner et al. 2004]. They propose seven guidelines for conducting effective and high-quality design science research. While they suggest that these guidelines should be addressed, they also state that there should be some flexibility in applying these guidelines. The following discusses how the current research has followed and addressed these guidelines.

Design as an artifact. Following this guideline, the research aims to design the visualization metaphor and glyph and develop a system. Both the visualization design and system are artifacts that address the problem of information overload faced by many in the information search process on the Web. The design of the artifacts is discussed in detail in the article, and future research and applications can be built on them.

Problem relevance. As discussed in the Introduction, the information overload problem is highly relevant to the decision making process at various managerial levels in organizations in the information age. The Web provides a lot of useful information to management and the effective and efficient use of such information is of high importance [McGonagle and Vella 1999].

Design evaluation. A design must be evaluated in order to show its usefulness and quality. Guided by the cognitive fit theory [Vessey 1991; Vessey and Galletta 1991], an evaluation study was performed using the *defeaturing* approach, which is a widely used evaluation framework for visualization systems [Morse et al. 2000]. Details of the evaluation will be reported in later sections.

Research contributions. This research has two main contributions. First, the research aims to demonstrate the feasibility and usefulness of applying the glyph visualization approach in Web search result visualization. The study fills the research gap due to the lack of application of glyph visualization techniques in Web search result presentation. Future visualization research can be built upon the current study. Second, two artifacts are created, namely the glyph design and the visualization system, and an evaluation study following a theoretical framework is reported.

Research rigor. This research relies on rigorous elements from multiple fields, including information retrieval, human-computer interaction, and system design. Both the construction and evaluation of the artifact were based upon the knowledge base from these disciplines.

Design as a search process. The design of techniques for Web search result visualization is a process that searches for a potential solution to the information overload problem. In early stages of the research, initial feedback is obtained from users and the design is revised a number of times. The design is iteratively revised in order to search for the best display that suits the research purpose.

Communication of research. The research is presented in this article in a way that is accessible to both technology-oriented and management-oriented audiences. Both the artifacts and the evaluation study are clearly presented and discussed in detail in this article, such that both can be easily replicated by researchers and practitioners.

4. THE FLOWER METAPHOR

As discussed earlier, glyphs have been widely used in visualizing multidimensional data. Among the various types of glyphs, the flower glyph is a relatively new metaphor proposed in visualization research. Xiong and Donath proposed the use of the flower metaphor for visualizing user interactions in Web message boards [Xiong and Donath 1999]. The flower metaphor has the advantages of being simple and intuitive. It also has been used to visualize social networks [Lantin and Judelman 2006], DNA sequences [Van Loocke 2004], and other online communication such as newsgroups [Zhu and Chen 2001; Zhu 2002].

During earlier stages in the design, various glyphs were considered as the metaphor, such as human faces, human figures, stars, polygons, books, and flowers. Each of these glyphs was evaluated by collecting different forms of these images, identifying the structure that could be used to represent Web document features, and drawing various forms of the glyph on paper to evaluate suitability. After assessing these glyphs, the flower metaphor was chosen in this research as an example for several reasons. First, as shown in previous research, flowers are intuitive and easy to understand [Xiong and Donath 1999; Lantin and Judelman 2006]. Previous experiments have shown that users have a higher preference towards such visual display [Zhu 2002]. Second, flower glyphs have a rich structure which is sufficient for representing a number of features while not being too complex. It has been shown that the different values represented by the features of a flower glyph can be distinguished easily [Xiong and Donath 1999]. Such features can represent the most important attributes (such as search keywords, page quality, and document type) for Web search result documents. This provides a good cognitive fit with the task of Web information search. Third, the status of a flower, like booming or withering, can represent the status of a document with a positive or a negative meaning, which may not be achieved by other glyphs such as a rectangle. Fourth, flowers have been used for Web forum visualization. It has been shown feasible to apply the flower metaphor to Web-related data and the results reported are encouraging. Previous study reported that in an experiment setting, participants were able to perform information retrieval tasks using flower glyphs more effectively and efficiently [Zhu 2002]. Based on these advantages and review of the literature, it was decided to design and evaluate the flower glyph for its application in Web search result visualization.

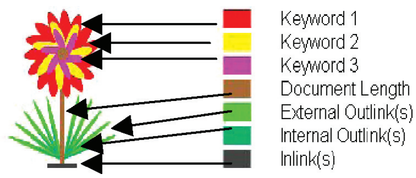

Fig. 1. The flower glyph.

Table I. Meaning of the Different Parts of the Flower Glyph

Flower part	Representation/Meaning
Petals	Number of keywords
Leaves	Number of outgoing links
Stem	Document length
Supporting ground	Number of incoming links

It is important to note that the choice does not imply that other glyphs are not suitable for this purpose; it only suggests that the flower glyph is a suitable choice and thus is selected as an example metaphor in this study in order to demonstrate the usefulness of glyphs.

4.1. Design of the Glyph

Using the flower metaphor, a glyph was designed for the visualization system. The objective is to design a glyph that can represent several common yet important attributes of Web documents. The attributes include the frequencies of the search keywords, and number of incoming links, the number of outgoing links, and document length. The search keyword frequency shows users how many times a search query term appears in the Web document under consideration in order to evaluate its possible relevance. The number of incoming links is also an important feature of a Web document. It has been suggested that a Web page with more incoming links is usually more *authoritative* such that it is cited by the authors of other Web pages [Brin and Page 1998; Kleinberg 1998]. Another important feature is the number of outgoing links. This provides users with information on whether the page is a *hub* or *directory* page that is likely to contain many links pointing to other pages [Kleinberg 1998]. Lastly, document length is a useful feature that is included in the displays of some commercial search engines (e.g., Yahoo.com). It shows users how much content is included in the document, which allows users to estimate how much time is needed to read the document and how informative it may be.

These parameters are represented by the different parts of the flower glyph. The design of the glyph is shown in Figure 1. Based on the nature of flowers and the characteristics of Web pages, different parts of the flower glyph have been assigned to represent different attributes of a Web page. The representation scheme is shown in Table I.

Petals are the most prominent parts of a flower, and therefore have been chosen to represent the frequencies of keywords because the keywords in a search query are probably the important attributes in judging the relevance of a document. The number of different colors for the petals represents the number of keywords input by the user in the search query. The size of the petals represents the order of the search keyword. The first keyword is represented by the biggest petals in the outer ring, while subsequent keywords are smaller. As the different numbers of petals can visually differentiate the documents from one another [Rodrigues et al. 2007], a user can quickly tell whether a document in the search result list has more occurrences of one of the search keywords than other keywords.

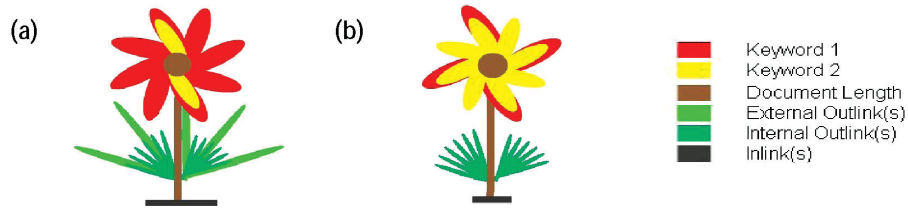

Fig. 2. Two examples of the flower glyphs. After getting familiar with the keys (shown on the right of the figure), one can easily see that the document represented by the glyph in (a) has more occurrences of Keyword 1 (more red petals), fewer occurrences of Keyword 2 (fewer yellow petals), similar document length (similar flower height/stem length), more external outgoing links (more leaves in light green), same number of internal outgoing links (similar number of leaves in dark green), and more incoming links (a larger supporting ground) when compared to the glyph in (b).

If more than five keywords are input by the user, only information of the first five keywords will be included in the flower. Although some information may not be represented in this case, this is a reasonable choice because of the limited space of the glyph and the fact that more than 85% of English Web search queries contain no more than five keywords [Spink et al. 2001]. Otherwise, the display may become too complicated and cause cognitive overload to the user [Ives 1982].

The colors of petals for the first five keywords are red, yellow, pink, orange, and light blue, respectively. These colors have been chosen as they are easily distinguishable from each other, but of course they can be easily revised or customized if needed. Two glyphs that represent two different Web documents are shown in Figure 2 as examples. In the glyphs shown in the figure, it can be easily seen that the flower in glyph (a) contains more red petals, meaning that document (a) has more occurrences of Keyword 1 than document (b). Similarly, one can see that document (a) has only two occurrences of Keyword 2 as it has only two yellow petals. The largest number of petals in a single color which can be displayed in the flower is 45, that is, even if the frequency of a certain keyword is 60, only 45 petals will be displayed.

Leaves have been chosen to represent the number of outgoing links in a page. One reason is that leaves are spanning outwards from the flower and become an intuitive choice for representing outgoing links. More outgoing links, represented by more leaves, mean that the document of interest is more likely to be a hub or directory page and contains many links to other useful resources [Kleinberg 1998]. There are two types of leaves, one in light green color and the other in dark green. The leaves in light green are larger in size and represent the number of external outlinks (hyperlinks that point to documents *outside* the current Web domain) of the corresponding Web page, whereas the other type of leaves are smaller in size and represent the number of internal outlinks (hyperlinks that point to documents *within* the current Web domain) of the page. External outlinks are represented by the bigger leaves because they are often more important; internal outlinks are often used for navigational purpose only. To keep the flower in shape, the maximum number of leaves is set at 12, that is, at most 12 leaves will be displayed for each type of leaves even if the number of the corresponding type of links is greater than 12. In the examples in Figure 2, one can see that document (a) has more external outlinks than document (b), while both documents have approximately the same number of internal outlinks.

The length of the stem is used to represent the length of the document. A longer stem represents a longer document. It is reasonable and intuitive to represent a document with a larger length with a flower with a longer stem, which signals to users that the document is in fact *longer*. For instance, the two glyphs in Figure 2 should represent similar document length. Document length information is shown in the result list in

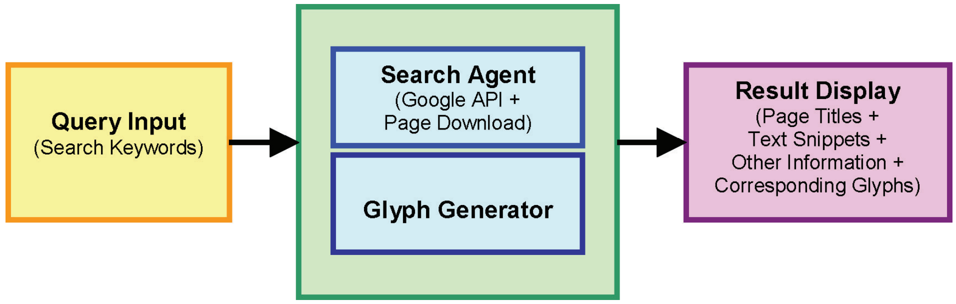

Fig. 3. Overview of system architecture.

many commercial search engines (e.g., Google) and can help users judge whether the document should be explored.

To represent the number of incoming links of a document, the length of the supporting ground has been selected. A large number of incoming links on the Web often means that a page is more authoritative and is well supported by the authors of other Web pages [Brin and Page 1998; Kleinberg 1998]. In other words, a larger number of incoming links often implies larger support and higher authority. Therefore, a longer (stronger) ground of the flower (e.g., glyph (a) in Figure 2) represents a larger number of incoming links and thus larger support from the community.

These representations are different from previous flower glyph designs as they have been specifically designed for Web search result visualization.

4.2. Implementation

The system consists of three main components: query input, processing system, and the result display. The system was implemented in Java Servlets and Java Server Pages (JSP). The overview of the architecture is shown in Figure 3.

Query input. The first part of the system consists of a user interface that accepts search queries from users. The query is checked for syntax errors. If no error is found the query is then forwarded to the processing system.

Processing system. The processing system contains two subcomponents: a *search agent* and a *glyph generator*. The search agent will parse the query from the user interface and search for relevant documents on the Web. Instead of maintaining its own search index, the system uses the Google Web Search API to search for relevant documents [Google 2006]. After getting Google's list of the top 10 search results, the search agent will try to download the full content of these 10 documents in order to perform further analysis. Ten parallel threads will be created to download the pages and calculate the features of the page such as the number of keywords and the number of outgoing links contained in the page. In addition, another call to the Google API is performed for each document in order to get the count of incoming links [Google 2006].

After the search results and the document information have been obtained, the glyph generator, written in Java, will be invoked for generating the flower visualization of the search results. A flower glyph will be produced for each document and the glyph will be saved as an image file.[1]

[1]One might think that such processing could be time consuming. However, in commercial search engines like Google, all such information is already computed during the indexing phase. No extra time is needed for such processing (counting of keywords and links) should these search engines adopt this system. The only time needed is to retrieve and transmit the corresponding glyph images for display, while such images can be easily cached.

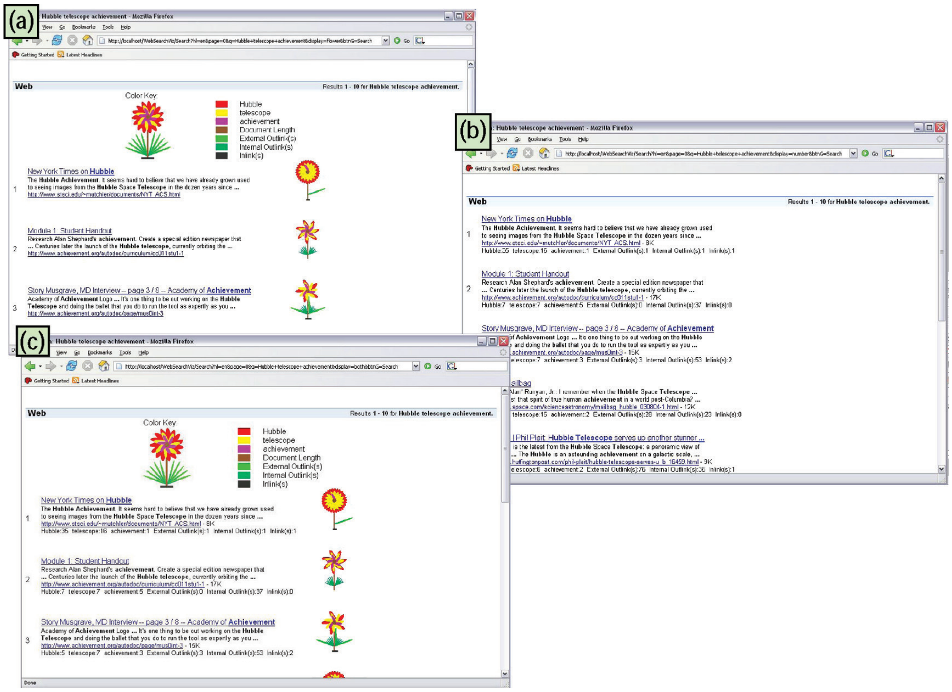

Fig. 4. Sample search result screenshots of the three displays using the query "Hubble telescope achievement": (a) the glyph display: additional information (such as frequencies of keywords, the number of outgoing links, and the number of inlinks) represented as glyphs; (b) the numeric display: additional information displayed by numbers only, representing a traditional display; and (c) the combined display: additional information displayed by both glyphs and numbers.

Result display. The search results returned from the processing system will be sent to users through their Web browser. The title, snippet, URL, and other relevant information of each document will be presented to users. The flower glyphs also will be displayed along with the search results. Users can browse and scroll through the list of search results just like one would normally do for a regular search engine.

In the design, there are three ways to show the page features to users. The first two displays are the *glyph display* which has the flower glyphs and the *numeric display* which has the data presented as numbers only. The numeric display is the same as a standard search engine display, except that numbers are added in order to allow for a fair comparison. In addition, because the glyphs can be easily integrated with a traditional list with numeric displays, a third display is used: the *combined display* that presents both glyphs and numbers. Other information, such as document titles, snippets, and URLs, would be displayed in all three systems in a way similar to popular commercial search engines. Samples of the three displays are shown in Figure 4.

5. PERFORMANCE MEASURES

5.1. Information Tasks and Task Complexity

Because the visualization has been designed to help users with browsing the search results but not directly related to accuracy of the search results (which are influenced by other factors such as query understanding, document indexing, and result ranking), it is more appropriate to evaluate the system with a focus on how it can help users in browsing the search results. Therefore, a set of low-level visual tasks was used, such that features specific to the domain of the visualization application and search queries could be eliminated. This approach is often known as the *defeaturing* approach [Wehrend and Lewis 1990; Morse et al. 2000; Zhu and Chen 2001]. In this approach, a set of low-level, domain-independent visual tasks are used to examine general and fundamental steps users perform using an interface when trying to retrieve information [Wehrend and Lewis 1990; Morse et al. 2000]. Examples of these low-level tasks include locate, identify, distinguish, emphasize, reveal, categorize, cluster, distribution, rank, compare, associate, and correlate [Wehrend and Lewis 1990; Morse et al. 2000; Zhou and Feiner 1998; Xiang et al. 2005]. Using this approach, it would be possible to evaluate the visualization design independent of the accuracy of the search engines.

These information tasks have different levels of task complexity. In information processing tasks like Web searching, the task complexity is often affected by factors like the amount of information cues, the number of processes involved, and the relationship between information and processes [Wood 1986; Bonner 1994; Campbell 1988; Speier 2006].

The cognitive fit theory [Vessey 1991; Vessey and Galletta 1991] suggests that when the information presentation format matches with characteristics of the task, users will achieve better task performance because there is no need for users to transform the information presented to a different mental model required to solve the task. On the other hand, when the information presentation format does not match with the task, users would need to transform the information into a different mental representation that is required to solve the task. This extra step requires users to spend more time and cognitive effort, leading to possible errors and lower performance.

Following the cognitive fit theory, previous studies have demonstrated the effect of task complexity on visualization system performance [Speier and Morris 2003; Kumar and Benbasat 2004]. For high-complexity tasks, a graphical display often has the advantages of revealing the relationships between multiple attributes and a graphical display is likely to provide a better cognitive fit. For low-complexity tasks, on the other hand, a numeric display often has the advantages of showing clearly the exact values of selected attributes and a numeric display is likely to provide a better cognitive fit. As task complexity is an important variable in evaluating visualization system performance, two sets of tasks were therefore identified based on the characteristics of Web searching. One set has high task complexity and one has low task complexity. The set of tasks with low task complexity includes *identify*, *distinguish*, and *rank*. The set of tasks with high task complexity includes *compare*, *categorize*, and *cluster*.

The three tasks in the first set have relatively low task complexity because they mainly involve a small number of given attributes (usually only one) that are relatively easy to find and compare. These tasks do not require comparing and contrasting a large number of attributes at the same time and require less information processing. In an *identify* task, a user is asked to find an object with a specific property. This task is important in Web searching because users often need to locate documents with some particular characteristics such as high quality (reflected by a large number of inlinks) or high frequencies of a particular search keyword. Similarly, the *distinguish* task, which involves finding an item that is unlike other items for a given attribute, also allows users to quickly identify items that are of particular interest based on certain properties. For instance, a document with a large number of external outgoing links should be given more attention if the user is looking for a hub which can provide pointers to other useful resources. The *rank* task, which involves ranking items based on a certain attribute value, allows users to find the items that are more worthy of further exploration based on particular properties and prioritize their sequence of viewing the search results.

The three tasks with high task complexity are also useful for users when sifting through search results. These three tasks demonstrate high task complexity because they involve much more processing on the different attributes and more actions are needed for comparing and grouping items [Wood 1986; Morse et al. 2000; Speier and Morris 2003]. The *compare* task, which involves comparing the attributes among items, is important in Web information searching because this task allows users to compare the characteristics of different items and decide which items have a higher chance of being useful (e.g., showing a higher frequency of a search keyword). The other two tasks, *categorize* and *cluster*, are also common tasks that users would perform in Web searching. The categorize task allows users to assign documents into groups based on some predefined properties (e.g., with or without a particular keyword) in order to organize the information collected. The cluster task allows users to group documents that are similar to each other based on some patterns that may be previously unknown. Both tasks are useful in allowing users to group similar items together, form a better overview of the result set, and organize their findings in their information collection process.

These low-level tasks are often needed for real Web search. A scenario is described in the following. Suppose a user wants to search for information related to the relationship between radio waves and brain cancer in order to write a simple essay (e.g., as a course assignment). The user first submits the query "radio waves and brain cancer" to a popular search engine like Google and the top ten search results are displayed in the first result page. The user may simply click on the first result page and get some background information, but want to explore more on the topic.

The user glances through the metadata of the top ten results and notices that some Web pages are focused on brain cancer only but do not mention the phrase "radio waves" frequently. The user decides to first *cluster* the search results into two groups: one group with a roughly equal frequency for every keyword and another group with a main focus on brain cancer only. The user may also want to *identify* the search results that contain the word "radio" at least three times. To focus only on search results that are highly relevant, the user may *compare* the results and find the one with the most number of total search keywords. If the user is interested in a detailed Web page, the user may need to *distinguish* the search result with the largest document length from the set of all results.

After reading some Web pages, the user realizes that there are some pages with misleading information. The user then wants to focus only on search results that are more authoritative. The user may *rank* the search results by descending number of incoming links and in order to estimate the authority of each search results. After some exploration on these pages, the user decides to find more Web pages to do further analysis. One way to do this is to *categorize* the search results into two groups: one group with a large number of outgoing links and another group with a few outgoing links only. The user can then visit those Web pages with more outgoing links because these pages are possibly directory pages that provide pointers to other documents.[2]

5.2. Key Performance Variables

The key performance variables are accuracy, time, and perceived usability. Accuracy measures whether users accomplish the information tasks correctly using the display. Time measures the time spent by users to accomplish the information tasks. As discussed earlier, it is more appropriate to measure the time and accuracy for performing

[2]This scenario is artificially designed to contain all the six tasks included in the study. A real search scenario would probably contain some but not all of these tasks. This scenario is therefore just an example to show how all these tasks would be useful in satisfying a user's overall information need.

low-level information tasks rather than complete search tasks. This allows us to separate the performance of the display from the search engine's performance such as indexing accuracy and ranking algorithm. Lastly, users' perceived usability measures how users perceive and feel about the system's usability. All these three measures are commonly used measures for information visualization systems [Morse et al. 2000; Zhu and Chen 2001; Nielsen 2003; Speier and Morris 2003; Ong et al. 2005; Flavian et al. 2006].

6. HYPOTHESES

6.1. High Complexity Tasks

Empirical studies based on the cognitive fit theory have shown that for complex tasks, users perform better on average with graphical (spatial) displays, which are good at revealing relationships between different attributes when there is a fit between the task complexity and the display [Vessey 1991, 1994; Vessey and Galletta 1991; Speier 2006]. In high complexity tasks, users need to examine a large number of attributes and process the information involved. They also have to compare a large number of items and process a large amount of information. In these tasks, graphical displays such as glyphs have been shown helpful in lowering such cognitive demand when compared with working with numbers only [Kosslyn 1989; Pinker 1990; Lohse 1997], by providing a good mapping between the presented information to the reality or the user's mental model and allowing users to perform comparison more easily [Chernoff 1973; Card et al. 1996; Tufte 1990; Speier 2006]. In addition, it has been suggested that glyphs can more easily tap into the access to users' associative, intuitive, and automatic processing of information [Pflughoeft et al. 2005a, 2005b], thus assisting users in performing tasks that involve comparison of multiple objects and attributes. Appropriate visual cues also can help users focus on data that are relevant to accomplishing the task at hand, thus they can process a smaller amount of information and make better judgments [Simon 1979; Speier and Morris 2003]. On the other hand, when using the numeric displays, users would need to put more cognitive effort in memorizing previously seen information for later comparison. They would need to work with more information, requiring more working memory, which has limited capability [Miller 1956]. To deal with cognitive overload, users would spend more time or try to minimize their efforts by trading off accuracy [Johnson and Payne 1995; Speier 2006]. This would result in a lower performance (lower accuracy and longer time) with the numeric display than with the graphical display. Therefore, the following hypotheses are posed.

H1a: Accuracy will be higher with a glyph display than with a numeric display when task complexity is high.

H1b: Time required will be shorter with a glyph display than with a numeric display when task complexity is high.

Both the numeric display and the glyph display could have advantages over the other for different tasks. In order to make use of the advantages of both displays, it is possible to combine both numbers and glyphs by providing both displays to users in a combined display. Previous studies have suggested that such integration of visualizations has the potential to achieve better performance [Xiang et al. 2005]. With the combined display, users can choose to rely on either the glyphs or the numbers, or both of them, in performing their tasks. As users try to reduce their efforts in problem solving because of their limited processing capabilities [Newell and Simon, 1972], they can choose the presentation that matches their tasks best in the combined display. When compared with the numeric display, the combined display offers glyphs which are a better fit for the high complexity tasks. Therefore, we suggest that the combined display would be able to achieve a higher level of performance than the numeric display in the high complexity tasks. The following hypotheses are posed.

H2a: Accuracy will be higher with a combined display than with a numeric display when task complexity is high.
H2b: Time required will be shorter with a combined display than with a numeric display when task complexity is high.

When compared with the glyph display, the combined display would show the extra number, which would not cause too much overload to users. In both displays, users can rely on the glyphs to accomplish their tasks. Therefore, we suggest that the combined display would achieve the same level of performance as the glyph display and hypothesized the following.

H3a: There is no significant difference between accuracy with a combined display and that with a glyph display when task complexity is high.
H3b: There is no significant difference between the time required with a combined display and that with a glyph display when task complexity is high.

6.2. Low Complexity Tasks

When task complexity is low, only a few attributes and objects need to be examined and compared [Wood 1986]. Users can easily accomplish their tasks by simply extracting a few pieces of data. According to the cognitive fit theory and previous studies, numeric display provides a better match with this type of task [Vessey 1991; Vessey and Galletta 1991; Speier 2006]. As the task complexity is low, a mental mapping between the display models and the real information is not necessary [Speier and Morris 2003]. On the other hand, in the glyph display, because of the misfit between information presentation and task, users have to map the features of a glyph to the characteristics of the corresponding document before they can make any judgment. In addition, visualization systems often may not show sufficient details for users to make judgments and more effort is required from users for knowing such details [Shneiderman 1996]. These processes for the glyph display would be more error prone and time consuming. Based on these grounds, the following are hypothesized.

H4a: Accuracy will be higher with a numeric display than with a glyph display when task complexity is low.
H4b: Time required will be shorter with a numeric display than with a glyph display when task complexity is low.

Similarly, using a combined display, users can try to use the information presentation that best fits the given task and choose to make use of the numbers when performing low complexity tasks. As the numbers can help users finish their tasks quickly, users can pay less attention to the glyphs despite that they are present. This can save their time in mapping the glyph model to their mental model. Therefore, we hypothesize the following.

H5a: Accuracy will be higher with a combined display than with a glyph display when task complexity is low.
H5b: Time required will be shorter with a combined display than with a glyph display when task complexity is low.

When compared with the numeric display, the combined display would show the extra glyphs, which would not cause too much overload to users. In both displays, users can rely on the numbers to accomplish their tasks. Therefore, we suggest that the combined display would achieve the same level of performance as the numeric display and hypothesize as follows.

H6a: There is no significant difference between the accuracy with a combined display and with a numeric display when task complexity is low.

H6b: There is no significant difference between the time required with a combined display and with a numeric display when task complexity is low.

6.3. Perceived Usability

Perceived usability is a measure often used to evaluate how easy it is to use a system from the user interface perspective [Nielsen 2003; Flavian et al. 2006]. Out of the three displays, a glyph display has the advantages of being more aesthetic and visually appealing. It also has the advantages of showing the overview which allows users to compare between items, and using sharp visual cues which allows users to compare attributes [Chernoff 1973; Chuah and Eick 1998; Zhu 2002]. A numeric display, on the other hand, provides concrete numbers for users to work with and is a model with which users are most familiar.

A combined display brings together the advantages of both glyph and numeric displays. Users are often more satisfied and find a system more usable when they have more information as long as it is not overloading [O'Reilly 1980]. In the combined display, users can choose to use either just the numbers or the glyphs, or both of them when performing a task, based on the task type and user individual preference. Therefore, the following are hypothesized.

H7: Users' perceived usability of the combined display will be higher than that of the numeric display.

H8: Users' perceived usability of the combined display will be higher than that of the glyph display.

In addition, as the numeric display and the glyph display both have their advantages and disadvantages for different tasks, the following are hypothesized.

H9: There is no significant difference between users' perceived usability of the glyph display and that of the numeric display.

7. EVALUATION

This section reports an empirical study conducted to examine the effectiveness and efficiency of the flower glyph in the visualizer. Three different displays were evaluated in the study, as discussed earlier.

The following conditions were satisfied in all the three displays in order to make the evaluation fair: (1) all three displays connected to the same search agent (using the Google Web Search API and user interface); (2) all of them listed the search results based on the ranking from the search engine in the form of a list; and (3) all had the same layout and provided similar functions such as scrolling.

Based on earlier discussion, six tasks were identified for the evaluation. A summary of the task types and some sample tasks are shown in Table II. It should be noted that the task complexity here is manipulated; for instance, one could design a *rank* task with high task complexity or a *cluster* task with low complexity. In the evaluation, the low complexity tasks involving only one to two attributes while the high complexity tasks require comparing and contrasting three or more attributes.

7.1. Experiment Design

A 3×2 design was used in the experiment to test the hypotheses. The two factors were display (glyph, numeric, or combined) and task complexity (high or low). To reduce the effect of individual differences, both factors were used as within-subject variables.

Task complexity was manipulated according to earlier discussion. Based on the six task types identified, four task sets (each with six tasks) were created with different

Table II. Task Types and Examples

Task Complexity	Task Type	Description	Example Task
Low	Identify	Find a document with some specific properties	Find results with equal frequencies of the two keywords found.
	Distinguish	Find a document that is unlike others for a given attribute	Try to find out the result that is unlike others in terms of incoming links.
	Rank	Rank documents based on certain attribute values	Rank the first 3 results in ascending order of total number of keywords found.
High	Compare	Compare the attributes among documents	Find the result with the lowest total keyword frequencies.
	Categorize	Assign documents into group based on some predefined attributes	Categorize the results into 2 groups. One with every attribute value > 0 and the other one with some attribute value = 0
	Cluster	Group documents that are similar	Which subsets of the documents are similar?

search topics. Six of the fifty search topics created by the National Institute of Standards and Technology (NIST) for the TREC-6 ad hoc task were selected and adopted for use in the context of Web searching. The TREC (Text Retrieval Conference) series has been sponsored by the NIST and the Defense Advanced Research Projects Agency (DARPA) to encourage research in information search. The experimental tasks were based on these TREC topics which have been well-studied and many evaluation results of which can be found in the literature [Cormack et al. 1998; Singhal 1998]. This provides a solid foundation and reference framework for the research. The six topics used in the experiment were *Hubble telescope achievement*, *implant dentistry*, *radio waves and brain cancer*, *undersea fiber optic cable*, *new fuel sources*, and *health and computer terminals*. These six topics have been used in previous studies on Web search and visualization systems [Chau et al. 2001; Chen et al. 2002].

One hundred and seventeen undergraduate subjects were recruited in the experiment. These subjects were randomly selected and their participation was voluntary. The average age of the subjects was 21.06. The group consisted of 35.0% female and 65.0% male subjects. The majority of subjects (78.6%) gave a response of 5 or above (out of a 10-point scale with 0 being the lowest and 9 being the highest) when asked about their familiarity with Web searching. Each participant went through four steps in the evaluation: (1) introduction and training, (2) demographic information survey, (3) performing tasks using the systems, and (4) posttest questionnaire survey.

One of the four task sets was first used by the experimenter to help each subject get familiar with the three displays. Afterwards, each subject was asked to fill in a pretest demographic survey. The subjects were told that their performance would be measured by both the accuracy and the time spent for each task. In order to ensure that they would try their best in the experiment, they were also informed that a prize incentive (in the form of cash coupons) would be given to the top three participants with the best performance.

During the main experiment, each subject was asked to perform six tasks (3 with high complexity and 3 with low complexity) using each of the three displays. A different task set was used for each display. As there were four task sets in total, each subject would use each task set exactly once (1 for training and 1 for each of the three displays). Two measures were collected from each task performed: the average accuracy of the answers given by the user and the average time required to complete the tasks. The averages were taken for each task type (high or low complexity). Both accuracy and time have been widely used in the evaluation of visualization systems [Vessey and

Table III. Measures of Accuracy and Time Spent with Each Display

Task Complexity	Display	N	Accuracy Mean (s.d.)	Time (seconds) Mean (s.d.)
High	Glyph	117	0.738 (0.243)	30.04 (15.61)
	Numeric	117	0.678 (0.239)	30.61 (16.22)
	Combined	117	0.769 (0.278)	27.98 (15.01)
Low	Glyph	117	0.738 (0.273)	18.53 (12.23)
	Numeric	117	0.766 (0.271)	16.22 (9.26)
	Combined	117	0.758 (0.268)	15.90 (11.04)
All	Glyph	234	0.738 (0.184)	24.28 (12.53)
	Numeric	234	0.722 (0.195)	23.41 (11.18)
	Combined	234	0.764 (0.222)	21.94 (11.58)

Table IV. Results of Pairwise Comparison Tests for Accuracy and Time

Task Complexity	Comparison	DV	Wilcoxon Signed Rank Test p-value	Hypothesis	Hypothesized difference?	Significant difference found?
High	Glyph vs. Numeric	Accuracy	0.041	H1a	Yes	Yes
		Time	0.332	H1b	Yes	No
	Combined vs. Numeric	Accuracy	0.015	H2a	Yes	Yes
		Time	0.084	H2b	Yes	Yes
	Combined vs. Glyph	Accuracy	0.486	H3a	No	No
		Time	0.302	H3b	No	No
Low	Numeric vs. Glyph	Accuracy	0.183	H4a	Yes	No
		Time	0.021	H4b	Yes	Yes
	Combined vs. Glyph	Accuracy	0.252	H5a	Yes	No
		Time	0.007	H5b	Yes	Yes
	Combined vs. Numeric	Accuracy	0.644	H6a	No	No
		Time	0.377	H6b	No	No

Galletta 1991; Heo and Hirtle 2001; Chau et al. 2001; Xiang et al. 2005; Speier 2006]. The accuracy was judged by two independent judges who went through the answers given by all subjects. When the two judges had a conflict (which rarely happened, as most of the tasks had definite answers), they discussed the answers and came up with an agreed upon judgment.

In order to make sure that the results would not be affected by training effect, primacy effect, and tiring effect, both the assignment of the task sets to the displays and the order of the displays used by subjects were rotated in the experiment.

After performing all the tasks, each subject was asked to fill in a posttest questionnaire. The questionnaire was designed to collect the subjects' demographics information and measure their perceived usability towards the three displays by eight items. The questions for measuring these items were designed based on a modified QUIS instrument, which is a robust and widely used instrument for measuring user satisfaction for computer user interfaces and visualization systems [Chin et al. 1988; Zeng et al. 2003]. For each item, each subject had to choose a value on a 10-point scale, with 0 being least favored and 9 being most favored. In the last part of the questionnaire, the subjects were also asked to give open-ended comments about the experiment and the displays.

7.2. Experiment Results

A summary of the accuracy and time measures collected is shown in Table III. First, hypotheses H1 to H6 were tested by performing data analysis on accuracy and time. The Kolmogorov-Smirnov test was first conducted on the data and showed that the accuracy and time data did not follow normal distribution ($p < 0.05$). Because both accuracy and time did not satisfy the assumption of pairwise t-tests, the nonparametric Wilcoxon signed rank tests were performed for pairwise comparisons among the three displays. The results are shown in Table IV and the charts are shown in Figure 5.

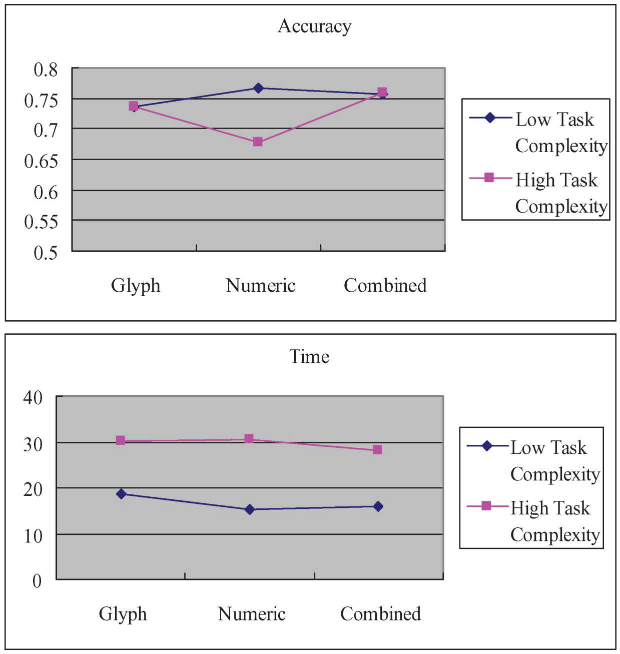

Fig. 5. Experimental results for accuracy and time.

The experimental results are consistent with hypotheses H1a, H2a, H2b, H3a, H3b, H4b, H5b, H6a, and H6b. When task complexity was high, the accuracies achieved by both the combined display (0.769) and the glyph display (0.738) were significantly better than that of the numeric display (0.678). This supported H1a and H2a. There was no significant difference between the time spent for the glyph display (30.04 seconds) and the numeric display (30.61 seconds), but the combined display (27.98 seconds) performed better than numeric display. The results did not support H1b but supported H2b.

There was no significant difference between the accuracies and time spent achieved by the glyph display and the combined display. The results did not reject hypotheses H3a and H3b.

When task complexity was low, there is no significant difference between the accuracies of the combined display (0.758), the numeric display (0.766), and the glyph display (0.738). H4a and H5a were not supported. One possible reason is that the low complexity tasks were fairly easy and most participants, regardless of what display they used, were able to achieve the same level of accuracy.

The average times spent by subjects for the low complexity tasks with the numeric display (15.22 seconds) and the combined display (15.90 seconds) were significantly better than that of the glyph display (18.53 seconds). Both H4b and H5b were supported. For low complexity tasks, the numbers available in both the numeric display and the combined display match with the characteristics of the task and therefore subjects finish the task more efficiently as predicted by cognitive fit theory. When using the glyph display, subjects had to perform the mental mapping of the visualization to the real-world data and to examine the attributes using visual cues. This resulted in more time spent in accomplishing the tasks. There was no significant difference between the accuracy and the time spent with combined display and the numeric display, and the result was consistent with H6a and H6b.

Table V. Perceived Usability

Display	N	Perceived Usability Mean (s.d.)
Glyph	117	4.10 (1.57)
Numeric	117	5.71 (1.22)
Combined	117	6.54 (1.36)

Table VI. Results of t-Tests for Perceived Usability

DV	Comparison	t-test p-value	Hypothesis	Hypothesized difference?	Significant difference found?
Perceived Usability	Combined vs. Glyph	<0.001	H7	Yes	Yes
	Combined vs. Numeric	<0.001	H8	Yes	Yes
	Numeric vs. Glyph	<0.001	H9	No	Yes

Table VII. Summary of Hypothesis Testing Result

Hypothesis	Task Complexity	Hypothesis (> denotes perform better than; = denotes no difference in performance)	Measures	Pairwise Comparison p-value	Hypothesized difference?	Significant difference found?
H1a	High	Glyph > Numeric	Accuracy	0.041	Yes	Yes
H1b			Time	0.332	Yes	No
H2a		Combined > Numeric	Accuracy	0.015	Yes	Yes
H2b			Time	0.084	Yes	Yes
H3a		Combined = Glyph	Accuracy	0.486	No	No
H3b			Time	0.302	No	No
H4a	Low	Numeric > Glyph	Accuracy	0.183	Yes	No
H4b			Time	0.021	Yes	Yes
H5a		Combined > Glyph	Accuracy	0.252	Yes	No
H5b			Time	0.007	Yes	Yes
H6a		Combined = Numeric	Accuracy	0.644	No	No
H6b			Time	0.377	No	No
H7	All	Combined > Numeric	Perceived Usability	<0.001	Yes	Yes
H8		Combined > Glyph	Perceived Usability	<0.001	Yes	Yes
H9		Numeric = Glyph	Perceived Usability	<0.001	No	Yes

The data collected from the posttest questionnaires were analyzed to test hypotheses H7 to H9.[3] The reliability of the instrument was first validated by calculating Cronbach's alpha. The alpha value of perceived usability is 0.905, showing that the construct is highly reliable. The questionnaire data on perceived usability are summarized in Table V. The Kolmogorov-Smirnov test was conducted on the perceived usability and showed that the data do not depart from normal distribution ($p > 0.650$). Pairwise t-tests were performed and the results are shown in Table VI.

H7 and H8 were both supported. In other words, the subjects demonstrated a higher perceived usability (6.54/9.00) towards the combined display when compared with the other two displays. One possible reason for the preference towards the combined display was that the subjects liked to have access to both numbers and glyphs and felt that they were in more control in this case. The data also showed a higher perceived usability towards the numeric display (5.71) and the glyph display (4.10) and did not support H9. A summary of all hypothesis testing results is shown in Table VII.

[3]In the experiment, each participant filled in one usability questionnaire for the three designs, without a separation on high complexity and low complexity tasks. Therefore, it was not possible to analyze the results on perceived usability separately according to different levels of task complexity due to the experiment design.

8. DISCUSSIONS

8.1. Summary of Findings

The experimental results were encouraging. Overall, it was found that both the glyph display and the numeric display have their own strengths and weaknesses. In general, the glyph display performed better for high complexity tasks and the numeric display performed better for low complexity tasks. These results were in accordance with the findings of previous studies that glyphs are especially effective at tasks which involve examining and comparing a large number of items and attributes [Chernoff 1973; Chuah and Eick 1997, 1998].

The results also showed that the performance of the combined display was better than the numeric display for high complexity tasks, and was better than the glyph display for low complexity tasks. In addition, users preferred the combined display to the other two displays in terms of perceived usability (H7 and H8). Combining all factors, the combined display appeared to be the best. One may intuitively think that the combined display would always perform better than the other two displays as more information is available. However, it was noticed that the combined display sometimes performed worse. The only difference between the combined display and the numeric display is that the former contains more information: the flower glyphs. For high complexity tasks, the glyphs were useful and helped the combined display to achieve a better accuracy than the numeric display. However, for low complexity tasks, the glyphs provided extra visualization for the users to process, and sometimes distracted users.

Because Web search often involves both high complexity and low complexity tasks and it is impractical to separate them during a user's information search session, it is difficult to decide when one display should be strictly preferred to another. One possible way to incorporate glyphs into Web search system is to let users have a choice on whether they would like to see numbers only (a traditional display), glyphs only, or both, that is, the three displays tested in the study. This would allow users to select what they are most comfortable with. These suggestions could possibly be applied to other visualization systems in which multiple displays can be used [Chen et al. 2003; Speier and Morris 2003; Xiang et al. 2005], but further testing would be needed.

Another suggestion is that the metaphor used for the glyph should be customizable by users. Although a flower glyph was designed and used in the study, it is also possible to design and implement other glyphs. As stated by one subject in the open-ended comments in the posttest questionnaire, it would be desirable to use various different types of glyphs and allow users to select their glyphs according to their preferences. For example, one user may choose the flower glyphs while another may choose human faces or other metaphors.

The findings also have important practical implications. The results have demonstrated that visualization techniques such as glyphs can help users in satisfying their information needs for certain tasks. However, as discussed, the adoption of visualization in commercial search engines such as Google or Yahoo is still wanting. Web search engine companies should put more resources in designing and evaluating different visualization techniques and deploy the ones that are helpful. As these search engines have already computed these metadata in their indexing process, the overhead cost of using such a display is very small. As users perceive better user satisfaction and usability with these techniques, the search engine may be able to draw more customers and traffic to their sites and thus more revenue. Based on the findings, a combined display would be helpful for users. It is also desirable to investigate under what circumstances such display should be used. It would be good to make it as an optional rather than a compulsory setting, such that users who do not like such display can use the simple text display. The improved user interface for search result display will also help managers and decision makers in their information search tasks.

8.2. Limitations of the Study

While the results of the study are encouraging, a few limitations of the study should be noted when interpreting the findings. First, only the flower glyphs were used in the study. Although we believe the findings are generalizable, caution should be taken when applying the results to other glyphs, such as faces, stars, and polygons, in the Web searching context. More studies would be needed.

Some of the design choices, such as the maximum number of links and the maximum frequencies of each keyword that can be represented by the glyphs, were decided based on design and display constraints and the characteristics of Web pages used for testing. This may not be applicable to Web pages with different characteristics, for example, a set of Web pages all having a large number of outgoing links. It is possible that different design choices may affect the performance of the proposed display. A mechanism to dynamically adjust the maximum numbers of links or keyword frequency will be desirable. Further testing on the optimal numbers of links will also help improve the display.

Another limitation of this study is that a small sample of student subjects in a single city was used in the experiment. These undergraduate subjects could be possibly more familiar with computer graphic displays (e.g., video games) than other populations. The average age of the subjects was only slightly above 21, which could be different from the average age of Web users. The gender ratio is not perfectly balanced. Cultural differences may also affect the generalizability of the findings, as people with different cultures may have different perceptions towards visualization systems and user interface design [Marcus and Gould 2000]. The results of this study may not generalize well to other populations. The findings should be interpreted with caution when applying to a larger and broader population, and more studies would be needed in this aspect.

9. CONCLUSIONS

In this article, we have reported research on designing, implementing, and evaluating a flower glyph for information search on the Web. This study has several important contributions. Using a design science approach, two artifacts were created, namely a flower glyph design and a Web search visualization system based on glyphs. This instantiation is a new design which is different from previous research using flower glyphs. The study has addressed the lack of research and practical systems in the area and demonstrated the feasibility and strengths of the design. In addition, a systematic user evaluation on the artifacts was reported. The results suggest that a display that combines both numbers and glyphs would be the most appropriate, while users should also be given choices on what kind of display and what kind of glyph are to be used based on their preferences. With the continuous growth of the Web, this study sheds light on how Web search engines and visualization systems could be further improved in order to help managers and other decision makers in their information search processes.

One direction for future work would be to expand the evaluation study in order to address some of the limitations outlined earlier. Besides having a larger sample size, some potential variables, such as individual differences, cultural differences, user mathematical ability, and user spatial ability [Speier and Morris 2003], should be investigated. This could provide insights on how glyphs should be designed for individuals with different abilities and backgrounds.

It would also be interesting to customize the display for different types of queries. For example, for a search query that contains a phrase (i.e., multiple words enclosed by double quotations that require the words to appear consecutively), the glyph can be customized such that each petal represents a phrase instead of a single word. In many situations, it would also be desirable to remove stop-terms (such as “to” and “of”) from the display unless the user has explicitly specified that these terms are needed. The incorporation of these options should be considered when adopting the display.

Another research direction is to explore the design of glyphs and its effect on user performance. The number of features included in the current design could be increased or decreased. More evaluation will be needed to find out the desirable balance between information richness and simplicity. Moreover, while the flower metaphor is used in the design, other metaphors may perform equally well or even better. For example, Chernoff’s faces have the advantage of being easily processed by humans [Chernoff 1973], while a book metaphor has the advantage of being a closer mapping with user’s mental models of documents and reading [Card et al. 1996]. What would be the effect of glyph design on user performance? This would be an interesting question to investigate.

Another possible direction for future research is to study how glyphs should be positioned in a search result display. In the current study, the glyphs are listed on the right side of the search results in the display. One limitation is that users have to scroll up and down in order to see all the glyphs. This may require more effort from users to make comparisons between different items. In addition, when the number of top search results is not 10 (e.g., 5 or 20), some of the findings of this article may be affected. Previous research has tried to design displays, called *garden*, where a larger number of flower glyphs are visible without any scrolling [Xiong and Donath 1999; Zhu and Chen 2001]. It would be interesting to study the performance of a *Web document garden* by displaying different numbers of search results and combining glyph visualization and document clustering techniques.

In recent years, a quickly increasing number of Web searches are being conducted on devices with small screens, such as mobile phones. The display of search results on these displays has become an interesting research problem as well. One possible extension of the present research is to investigate how visualization techniques can be used to display search results on small screens, for example, by summarizing and aggregating search results into glyphs.

ACKNOWLEDGMENTS

I would like to thank F. C. W. Law and K. S. C. Lee of the University of Hong Kong for their work on coding the visualizer system and the evaluation study and their constructive ideas during the design of the system; J. Tam, A. Lee, and C.-W. Chan of the University of Hong Kong for their help in the evaluation study; B. Zhu of Boston University for the inspiration from her flower design for visualizing newsgroups; the editors and the anonymous reviewers of ACM TMIS for their constructive and insightful suggestions; the audience at Bentley University and the City University of Hong Kong for their feedback on presentations of this study; and Google for making its services and API freely available.

REFERENCES

BENNETT, K. AND TOMS, M. 1993. Emergent features and graphical elements: Designing more effective configural displays. *Hum. Factors 35*, 1, 71–97.

BEN-SHAUL, I., HERSCOVICI, M., JACOVI, M., MAAREK, Y. S., PELLEG, D., SHTALHAIMB, M., SOROKA, V., AND UR, S. 1999. Adding support for dynamic and focused search with Fetuccino. In *Proceedings of the 8th International. Word Wide Web Conference (WWW8).* 575–587.

BONNER, S. 1994. A model of the effects of task complexity. *Account. Organiz. Soc. 19*, 213–234.

BOWMAN, C. M., DANZIG, P. B., MANBER, U., AND SCHWARTZ, F. 1994. Scalable Internet resource discovery: Research problems and approaches. *Comm. ACM 37*, 8, 98–107.

BRIN, S. AND PAGE, L. 1998. The anatomy of a large-scale hypertextual Web search engine. In *Proceedings of the 7th World Wide Web Conference (WWW).*

BUJA, A., COOK, D., AND SWAYNE, D. 1996. Interactive high-dimensional data visualization. *J. Comput. Graph. Statist. 5,* 1, 78–99.

CARD, S. K., ROBERTSON, G. G., AND YORK, W. 1996. The WebBook and the Web forager: An information workspace for the World-Wide Web. In *Proceedings of the ACM SIGCHI Conference*.

CARTER, L. F. 1947. An experiment on the design of tables and graphs used for presenting numerical data. *J. Appl. Psychol. 31,* 640–650.

CELLARY, W., WIZA, W., AND WALCZAK, K. 2004. Visualizing Web search results in 3D. *IEEE Comput. 37*, 5, 87–89.

CHAU, M. AND CHEN, H. 2008. A machine learning approach to Web page filtering using content and structure analysis. *Decis. Supp. Syst. 44*, 2, 482–494.

CHAU, M., SHIU, B., CHAN, I., AND CHEN, H. 2007. Redips: Backlink search and analysis on the Web for business intelligence analysis. *J. Amer. Soc. Inf. Sci. Technol. 58*, 3, 351–365.

CHAU, M., ZENG, D., AND CHEN, H. 2001. Personalized spiders for Web search and analysis. In *Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL'01).* 79–87.

CHEN, H., CHAU, M., AND ZENG, D. 2002. CI spider: A tool for competitive intelligence on the Web. *Decis. Supp. Syst. 34*, 1, 1–17.

CHEN, H., HOUSTON, A. L., SEWELL, R. R., AND SCHATZ, B. R. 1998. Internet browsing and searching: User evaluation of category map and concept space techniques. *J. Amer. Soc. Inf. Sci. 49,* 7, 582–603.

CHEN, H., LALLY, A. M., ZHU, B., AND CHAU, M. 2003. HelpfulMed: Intelligent searching for medical information over the Internet. *J. Amer. Soc. Inf. Sci. Technol. 54*, 7, 683–694.

CHENG, P. C., CUPIT, J., AND SHADBOLT, N. R. 2001. Supporting diagrammatic knowledge acquisition: An ontological analysis of Cartesian graphs. *Int. J. Hum.-Comput. Studies 54*, 4, 457–494.

CHERNOFF, H. 1973. The use of faces to represent points in k-dimensional space graphically. *J. Amer. Statist. Assoc. 68*, 361–368.

CHIN, J. P., DIEHL, V. A., AND NORMAN, K. L. 1988. Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'88).* 213–218.

CHUAH, M. C. AND EICK, S. G. 1997. Glyphs for software visualization. In *Proceedings of the International Workshop on Program Comprehension*. 183–191.

CHUAH, M. C. AND EICK, S. G. 1998. Information rich glyphs for software management data. *IEEE Comput. Graph. Appl. 18*, 4, 24–29.

CHUNG, W., CHEN, H., AND NUNAMAKER, J. F. 2005. A visual knowledge map framework for the discovery of business intelligence on the Web. *J. Manag. Inf. Syst. 21*, 4, 57–84.

CORMACK, G. V., PALMER, C., AND CLARKE, C. 1998. Efficient construction of large test collections. In *Proceedings of the Annual ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'98).*

EBERT, D. S., KUKLA, J. M., SHAW, C. D., ZWA, A., SOBOROFF, I., AND ROBERTS, D. A. 1997. Automatic shape interpolation for glyph-based information visualization. In *Proceedings of the IEEE Visualization Conference.*

EBERT, D. S., ROHRER, R. M., SHAW, C. D., PANDA, P., KUKLA, J. M., AND ROBERTS, D. A. 2000. Procedural shape generation for multi-dimensional data visualization. *Comput. Graph. 24*, 375–384.

FANEA, E., CARPENDALE, S., AND ISENBERG, T. 2005. An interactive 3D integration of parallel coordinates and star glyphs. In *Proceedings of the IEEE Symposium on Information Visualization (INFOVIS).* 149–156.

FARHOOMAND, A. AND DRURY, D. 2002. Managerial information overload. *Comm. ACM 45,* 10, 127–131.

FLAVIAN, C., GUINALIU, M., AND GURREA, R. 2006. The role played by perceived usability, satisfaction and consumer trust on website loyalty. *Inf. Manag. 43,* 1, 1–14.

FORSELL, C., SEIPEL, S., AND LIND, M. 2006. Surface glyphs for efficient visualization of spatial multivariate data. *Inf. Visualiz. 5*, 112–124.

GOOGLE. 2006. Google SOAP search API (beta), 2006. http://www.google.com/apis/

HASAN, M., GOLOVCHINSKY, G., NOIK, E., CHAROENKITKARN, N., CHIGNELL, M., MENDELZON, A., AND MODJESKA, D. 1995. Browsing local and global information. In *Proceedings of the Conference of the Centre for Advanced Studies on Collaborative Research*. 228–240.

HEARST, M. 1995. TileBars: Visualization of term distribution information in full text information access. In *Proceedings of the ACM SIGCHI Conference*. 59–66.

HEO, M. AND HIRTLE, S. C. 2001. An empirical comparison of visualization tools to assist information retrieval on the Web. *J. Amer. Soc. Inf. Sci. Technol. 52*, 8, 666–675.

HEVNER, A. R., MARCH, S. T., PARK, J., AND RAM, S. 2004. Design science in information systems research. *MIS Quart. 28*, 1, 75–105.

IVES, B. 1982. Graphical user interfaces for business information systems. *MIS Quart. 6*, 15–47.

KLEINBERG, J. 1998. Authoritative sources in a hyperlinked environment. In *Proceedings of the 9th ACM-SIAM Symposium on Discrete Algorithms (SODA'98).* 668–677.

KLEINER, B. AND HARTIGAN, J. A. 1981. Representing points in many dimensions by trees and castles. *J. Amer. Statist. Assoc. 76,* 260–269.

KOHONEN, T., KASKI, S., LAGUS, K., SALOJÄRVI, J., HONKELA, J., PAATERO, V., AND SAARELA, A. 2000. Self organization of a massive document collection. *IEEE Trans. Neural Netw. 11*, 3, 574–585.

KOSSLYN, S. M. 1989. Understanding charts and graphs. *Appl. Cogn. Psychol. 3*, 2, 185–225.

KUHLTHAU, C. C. 1991. Inside the search process: Information seeking from the user's perspective. *J. Amer. Soc. Inf. Sci. 42*, 5, 361–371.

KUHLTHAU, C. C. 1993. *Seeking Meaning: A Process Approach to Library and Information Services*. Ablex, Norwood, NJ.

KUMAR, N. AND BENBASAT, I. 2004. The effect of relationship encoding, task, type, and complexity on information representation: An empirical evaluation of 2D and 3D line graphs. *MIS Quart. 28*, 2, 255–281.

LAMPING, J., RAO, R., AND PIROLLI, P. 1995. A focus+context technique based on hyperbolic geometry for visualizing large hierarchies. In *Proceedings of the ACM CHI '95 Conference on Human Factors in Computing Systems.*

LANTIN, M. AND JUDELMAN, G. 2006. FlowerGarden: An interactive visualization of concept-sharing. In *Proceedings of the CHI'06 Workshop on About Face: Interface Creative Engagement in New Media Arts and HCI.*

LIN, C., CHEN, H., AND NUNAMAKER, J. F. 2000. Verifying the proximity hypothesis for self-organizing maps. *J. Manag. Inf. Syst. 16*, 3, 57–70.

LOHSE, G. L. 1997. The role of working memory on graphical information processing. *Behav. Inf. Technol. 16*, 6, 297–308.

MAK, J. Y. S., LEONG, H. V., AND CHAN, A. T. S. 2002. Dynamic structuring of web information for access visualization. In *Proceedings of the ACM Symposium on Applied Computing*. 778–784.

MARCH, S. T. AND SMITH, G. 1995. Design and natural science research on information technology. *Decis. Supp. Syst. 15*, 4, 251–266.

MARCUS, A. AND GOULD, E. W. 2000. Cultural dimensions and global web user-interface design: What? So what? Now what? In *Proceedings of the 6th Conference on Human Factors and the Web.*

MCDONALD, D. M. AND CHEN, H. 2006. Summary in context: Searching versus browsing. *ACM Trans. Inf. Syst. 24,* 1, 111–141.

MCGONAGLE, J. J. AND VELLA, C. M. 1999. *The Internet Age of Competitive Intelligence.* Quorum Books, London.

MILLER, G. A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.,* 81–97.

MORSE, E., LEWIS, M., AND OLSEN, K. A. 2000. Evaluating visualizations: Using a taxonomic guide. *Int. J. Hum.-Comput. Studies 53,* 637–662.

MUKHERJEA, S. AND HARA, Y. 1999. Visualizing World-Wide Web search engine results. In *Proceedings of the International Conference on Information Visualisation*. 400–405.

MUNZNER, T. 1998. Exploring large graphs in 3D hyperbolic space. *IEEE Comput. Graph. Appl. 18*, 4, 18–23.

NEWELL, A. AND SIMON, H. 1972. *Human Problem Solving*. Prentice Hall.

NIELSEN, J. 2003. Usability 101. http://www.useit.com/alertbox/20030825.html

ONG, T. H., CHEN, H., SUNG, W., AND ZHU, B. 2005. Newsmap: A knowledge map for online news. *Decis. Supp. Syst. 39*, 4, 583–597.

O'REILLY, C. A. 1980. Individuals and information overload in organizations: Is more necessarily better? *Acad. Manag. J. 23*, 684–696.

PFLUGHOEFT, K., ZAHEDI, M., AND SOOFI, E. 2005a. Data visualization using figural animation. In *Proceedings of the 11th Americas Conference on Information Systems.*

PFLUGHOEFT, K., ZAHEDI, M., AND SOOFI, E. 2005b. Figural animation visualization: The system and its application. In *Proceedings of the AIS SIGDSS Pre-ICIS Workshop.*

PINKER, S. 1990. A theory of graph comprehension. In *Artificial Intelligence and the Future of Testing*, R. Freedle Ed., Lawrence Erlbaum Associates, Hillsdale, NJ, 73–126.

PIROLLI, P., CARD, S. K., AND VAN DER WEGE, M. M. 2003. The effects of information scent on visual search in the hyperbolic tree browser. *ACM Trans. Comput.-Hum. Interact. 10*, 1, 20–52.

ROBERTS, J., BOUKHELIFA, N., AND RODGERS, P. 2002. Multiform glyph based web search result visualization. In *Proceedings of the 6th International Conference on Information Visualisation.*

RODRIGUES, J. F., TRAINA, A. J. M., DE OLIVEIRA, M. C. F., AND TRAINA, C. 2007. The spatial-perceptual design space: A new comprehension for data visualization. *Inf. Visualiz. 6*, 261–279.

ROHRER, R., EBERT, D., AND SIBERT, J. 1998. The shape of Shakespeare: Visualizing text using implicit surfaces. In *Proceedings of the IEEE Symposium on Information Visualization.* 121–129.

ROUSSINOV, D. AND RAMSEY, M. 1998. Information forage through adaptive visualization. In *Proceedings of the 3rd ACM Conference on Digital Libraries.* 303–304.

SALTON, G. 1989. *Automatic Text Processing*. Addison-Wesley, Reading, MA.

SANGOLE, A. AND KNOPF, G. K. 2002. Representing high-dimensional data sets as closed surfaces. *Inf. Visualiz. 1*, 111–119.

SEBRECHTS, M. M., CUGINI, J. V., LASKOWSKI, S. J., VASILAKIS, J., AND MILLER, M. S. 1999. Visualization of search results: A comparative evaluation of text, 2D, and 3D interfaces. In *Proceedings of the 22nd Annual ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'99).* 3–10.

SCOTT, D. W. 1992. *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley and Sons, New York.

SHNEIDERMAN, B. 1996. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the IEEE Symposium on Visual Languages.* 336–343.

SIMON, H. A. 1979. Information processing models of cognition. *Annual Rev. Psychol. 30*, 363–396.

SINGHAL, A. 1998. AT&T at TREC-6. In *Proceedings of the 6th Text Retrieval Conference (TREC-6).*

SPEIER, C. 2006. The influence of information presentation formats on complex task decision-making performance. *Int. J. Hum.-Comput. Studies 64,* 1115–1131.

SPEIER, C. AND MORRIS M. 2003. The influence of query interface design on decision-making performance. *MIS Quart. 27*, 3, 397–423.

SPINK, A., WOLFRAM, D., JANSEN, B. J., AND SARACEVIC, T. 2001. Searching the Web: The public and their queries. *J. Amer. Soc. Inf. Sci. Technol. 52*, 3, 226–234.

STOREY, V., BURTON-JONES, A., SUGUMARAN, V., AND PURAO, S. 2008. CONQUER: A methodology for context-aware query processing on the World Wide Web. *Inf. Syst. Res. 19*, 1, 3–25.

TAN, J. K. H. AND BENBASAT, I. 1990. Processing of graphical information: A decomposition taxonomy to match data extraction tasks and graphical representations. *Inf. Syst. Res. 1*, 4, 416–439.

TURETKEN, O. AND SHARDA, R. 2004. Development of a fisheye-based information search processing aid (FISPA) for managing information overload in the web environment. *Decis. Supp. Syst. 37*, 3, 415–434.

TURETKEN, O. AND SHARDA, R. 2005. Clustering-Based visual interfaces for presentation of web search results: An empirical investigation. *Inf. Syst. Front. 7*, 3, 273–297.

TURETKEN, O. AND SHARDA, R. 2007. Visualization of Web spaces: State of the art and future directions. *Datab. Adv. Inf. Syst. 38*, 3, 51–81.

TUFTE, E. R. 1990. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT.

VAN LOOCKE, P. 2004. Generative flowers and flower canopies as visualizations of binary information and of DNA sequences. In *Proceedings of the Generative Art Conference (GA'04).*

VESSEY, I. 1991. Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decis. Sci. 22*, 219–240.

VESSEY, I. 1994. The effect of information presentation on decision making: A cost-benefit analysis. *Inf. Manag. 27*, 103–119.

VESSEY, I. AND GALLETTA, D. 1991. Cognitive fit: An empirical study of information acquisition. *Inf. Syst. Res. 2*, 63–84.

WANIEK, J., LANGNER, H., AND SCHMIDSBERGER, F. 2005. MemoSpace: A visualization tool for Web navigation. In *Proceedings of the World Wide Web Conference*. 900–901.

WARD, M. O. 2002. A taxonomy of glyph placement strategies for multidimensional data visualization. *Inf. Visualiz. 1*, 3-4, 194–210.

WEGMAN, E. J. 2003. Visual data mining. *Statist. Med. 22*, 1383–1397.

WEHREND, S. AND LEWIS, C. 1990. A problem-oriented classification of visualization techniques. In *Proceedings of the 1st IEEE Conference on Visualization*. 139–143.

WIZA, W., WALCZAK, K., AND CELLARY, W. 2003. AVE - Method for 3D visualization of search results. In *Proceedings of the International Conference on Web Engineering.* 204–207.

WOOD, R. E. 1986. Task complexity: Definition of the construct. *Organiz. Behav. Hum. Decis. Process. 37*, 1, 60–82.

XIANG, Y., CHAU, M., ATABAKHSH, H., AND CHEN, H. 2005. Visualizing criminal relationships: Comparison of a hyperbolic tree and a hierarchical list. *Decis. Supp. Syst. 41*, 69–83.

XIONG, R. AND DONATH, J. 1999. PeopleGarden: Creating data portraits for users. In *Proceedings of the 12th Annual ACM Symposium on User Interface Software and Technology*. 37–44.

ZENG, D., CHEN, H., DASPIT, D., SHAN, F., NANDIRAJU, S., CHAU, M., AND LIN, C. 2003. COPLINK agent: An architecture for information monitoring and sharing in law enforcement. In *Proceedings of the NSF/NIJ Symposium on Intelligence and Security Informatics (ISI).* 281–295.

ZHOU, M. X. AND FEINER, S. K. 1998. Visual task characterization for automated visual discourse synthesis. In *Proceedings of the ACM CHI '98 Conference on Human Factors in Computing Systems*. 392–399.

ZHU, B. 2002. Information visualization for knowledge repositories: Applications and impacts. Ph.D. dissertation, The University of Arizona.

ZHU, B. AND CHEN, H. 2001. Social visualization for computer-mediated communication: A knowledge management perspective. In *Proceedings of the Workshop on Information Technology and Systems (WITS'01).*

Received January 2010; revised July 2010, December 2010; accepted January 2011