Visualization of Web Spaces: State of the Art and Future Directions

Ozgur Turetken
Ryerson University

Ramesh Sharda
Oklahoma State University

Abstract

The World Wide Web is a dominant global communication medium and knowledge repository. It is used by a great number of people with a variety of computer skills hence its usability is critical. As with many large information collections, the challenge with web usability is understanding the structure of a collection of information objects (web pages) to find relevant ones for satisfying a specific information need. Web sites are organized in a hyperlinked structure that somewhat addresses this challenge. However, this "connectedness" also causes the now well-known "lost in cyberspace" phenomenon where one may get confused within the complex organization of a web site. Meanwhile, information exploration on the web is not limited to browsing a web site. The problem of finding relevant information applies to a collection of pages that come from various web sites as in the case of the results of a "less than perfectly constructed" search query.

Information visualization has been proposed as a way to cope with these problems by taking advantage of people's innate perceptual skills to support their cognitive skills. Many paradigms have been proposed for the visual presentation of web spaces (i.e. structured or unstructured collection of web pages). This study surveys these paradigms to provide a map of where the research in this field is, and what directions future research and practice can take. For this, we introduce a classification scheme to help in the systematic understanding of web visualization and for providing a framework for the development of future visualizations.

ACM Categories: H.1.2, H.5.2, H.5.4, I.7.2

Keywords: Web Usability, Human-computer Interaction, Information Visualization

Introduction

World Wide Web Usability Issues

The World Wide Web has emerged as the most popular global communication medium and knowledge repository. The almost ubiquitous web is now the platform for a myriad of activities ranging from on-line stock trading to distance-education. As it is, the web is the most dominant element of the information age, and its prominence is likely to continue shaping the nature of information use by virtually every member of the modern society. This trend has already changed many traditional products and services such as newspapers, software, TV, and banking. With the availability of more affordable Internet services and the continuously emerging web-based products and services, the web is no more a

Acknowledgements

The authors are grateful to Jeanette Kelley and Sam Shi for their suggestions in early stages of this project, to Joyce Lucca for her assistance with bibliography, and to Wynne Chin for his continuous support throughout the development of the paper.

privilege for the technology-proficient, but is an IT commodity. Consequently, it is critically important for the web to be easy to use.

A major hurdle against such ease of use is understanding the structure of a collection of web pages and then finding a relevant web page that would satisfy one's information need. The reason for this is that there is usually an overload of irrelevant information one faces both when dealing with a collection that consists of pages that are linked to each other such as those in a web site, or with a collection that is more loosely connected such as the resulting list of a web search. We call such a collection of web pages that are somehow conceived or processed together a "web space". For example, any search on Google or any other search engine typically results in a large web space for further exploration.

Information overload is a major problem in examining web spaces, because it is known that when people are overloaded with information, they begin filtering out portions of it, which then increases the probability that important information will be bypassed during decision-making (Jacoby, 1984). Such behavior in information use leads to lower quality decisions (Keller & Staelin, 1987). For the wealth of web-based information not to lead to such undesirable results (Brandt, 1997), web spaces should be presented in a form that will minimize overload. This is a major challenge. One way to address this challenge is for web designers to provide a structure to the web spaces they are presenting. A traditional method of organizing a large collection of information is to group the information objects in such a way that the amount of information that is presented at a time is not overwhelming. The web is such a domain where most of the information is organized within groups (the typical web site) which are then linked to each other. The biggest problem with presenting this highly linked structure is regarding how the structure is perceived. More often than not, the mismatch between the web designer's organization of the information and the user's perception of this organization causes the user to reach a page that he or she has not planned on (Botafogo et al., 1992; Rivlin et al., 1994). This causes confusion and disorientation problems, a phenomenon usually described as "being lost in cyberspace."

To summarize, the abundance of web-based information combined with its complex organization results in a challenge for web designers to present web spaces in a comprehensive manner so that the end-user is able to locate what they are seeking. In this paper, we discuss how information visualization can help in addressing this problem. To do this, we introduce a classification scheme to help in the systematic understanding of web space visualization and for providing a framework for the development of future visualization systems. Before the discussion of visualization in the specific domain of World Wide Web, we next present a generic description of information visualization

Information Visualization

The challenge of presenting a large collection of information in an understandable fashion is not new. It has long been known that people's cognitive abilities are limited in the simultaneous processing of a high amount of information (see Miller, 1956, for example). In addition to this limit in its capacity, we also know that the cognitive system is slow compared to the perceptual, especially the visual, system. According to Crapo et al. (Crapo et al., 2000), a number of visual features including motion, color, intensity, size, intersection, closure, orientation, lighting direction, and distance from the observer may be extracted "preattentively," without conscious effort and within 100-200 milliseconds. The cognition of the same amount of information without the support of visual clues would take between hundreds of milliseconds and a few minutes (Brautigam, 1996)

Information visualization is the name given to the collection of techniques that take advantage of this fact by using visual clues in information presentation. According to the Model Human Processor, (Card et al., 1986), sensory, i.e. visual and auditory, image buffers are part of the working (short-term) memory, and working memory is part of the long-term memory. Cognitive tasks are performed by taking inputs from working (short-term) memory and long-term memory. If the visual image buffer contains information relevant to the cognitive task at hand, working memory is effectively increased. This way, information overload is reduced since some of the load of processing information is shifted from the cognitive system to the sensory system. This observation has important implications for the design of human interfaces of computerized systems as observed by Shneiderman:

> "Humans can quickly understand the relative position of the different entities and their relationships in a picture. Interface designers can capitalize on this by shifting some of the cognitive load of information retrieval to the perceptual system. By appropriately coding properties by size, position, shape, and color, we can greatly reduce the need for explicit selection, sorting, and scanning operations." (Shneiderman, 1994)

The modern computers' graphical processing capabilities and speed combined with the spread of the development technologies for web-based applications have made various forms of visualization applicable to the web. This paper surveys research on the visualization of web spaces. In doing so, it suggests a classification scheme to discuss how visualization has been, and can be, used to address the problems of overload and disorientation while trying to locate a relevant web page for an information need. This scheme should help in the systematic understanding of research and practice in this field, and should serve as a framework for the development of future visualization systems. It should be noted that the subject of this paper is not those visualization paradigms that use the web as a medium for presenting different forms of content. For example, we are not interested in systems such as WebSpace (from Silicon Graphics) that displays digital versions of visual artwork or 3D models of products in an online catalog using the web simply as a delivery medium (Figure 1). Rather, our interest is in those paradigms where the web is the very content of presentation.

The organization of the rest of the paper is as follows. In the next section we present our classification scheme for visualization of web spaces. Subsequently, we review the state of the art in this field by using our classification method. The last section identifies future research directions and conclusions.

A Classification of Techniques for Visualizing Web Spaces

In this paper, our purpose is to identify important dimensions along which web visualization systems can be classified. While identifying the classification dimensions, we enumerate alternative features a visualization system can have for each dimension. We provide representative examples of systems and/or methods to clarify these dimensions and to provide a snapshot of the state of art in the field. Although we survey more than 60 visualization paradigms, we do not claim (nor is it our aim) that these examples are totally comprehensive. We believe it would not be realistic to expect such comprehensive coverage of an environment like the web that changes almost in real time.

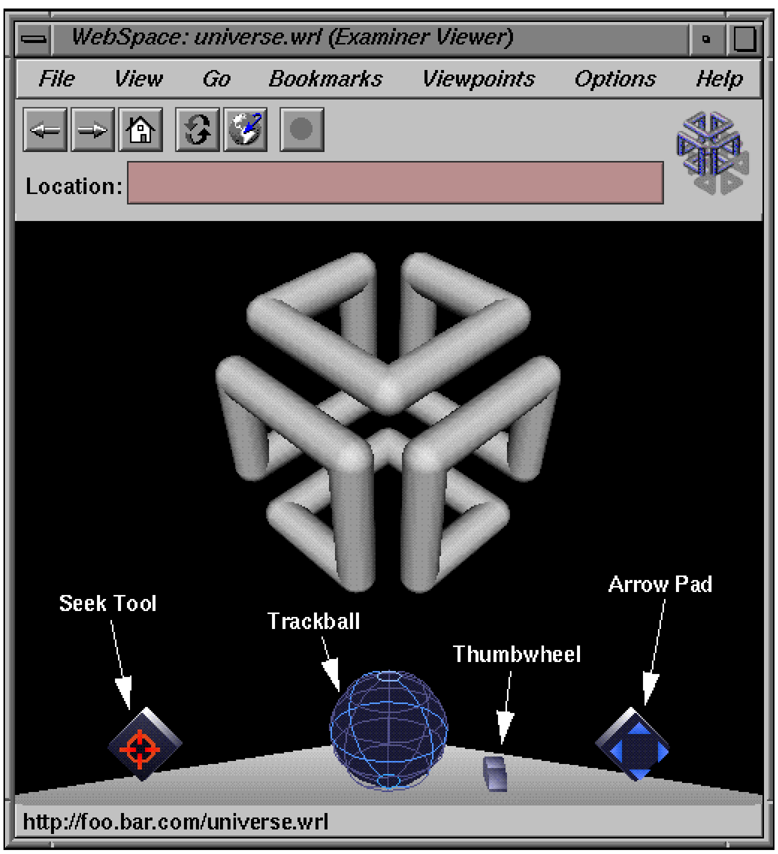

Figure 1. Webspace© Viewer

Through our classification scheme and the following system review, we aim to achieve the following:

1. to provide a systematic way to study the state of art in the field of web visualization,
2. to provide a framework for development of new visualization systems, and
3. to help in recognition patterns for web visualization software, and gaps that can be filled by future research.

For similar purposes, i.e. to sort out visualization systems and guide researchers to new opportunities, Shneiderman (Shneiderman, 1996) proposed a "task by data type taxonomy" for general information visualization. This taxonomy consists of the following seven tasks:

Overview: Gain an overview of the entire collection,
Zoom: Zoom in on items of interest,
Filter: Filter-out uninteresting items,
Details-on-demand: Select an item or group and get details when needed,
Relate: View relationship among items,
History: Keep a history of actions to support undo, replay and progressive refinement,
Extract: Allow extraction of sub-collections and the query parameters.

Meanwhile, the data types mentioned in the taxonomy are the following: 1-dimensional, 2-dimensional, 3-dimensional, temporal, multi-dimensional, tree, and network. The list of the data types presented in this taxonomy is subject to variation and Shneiderman notes that many prototypes use a combination of these data types. The purpose of his taxonomy is to facilitate discussion leading to useful discoveries. This taxonomy defines the object to be visualized as an item with multiple attributes where the basic purpose of visualization is to help in selecting items whose attributes satisfy certain criteria.

In a later study, Card and Mackinlay (Card & Mackinlay, 1997) present another general information visualization framework for purposes similar those of the Shneiderman (Shneiderman, 1996) taxonomy. Card and Mackinlay classify visualizations based on the "Type of the original data to be visualized," the "Function that recodes the original data," the "Recoded data", and "Automatically processed graphical properties" such as position, color, size, and shape, and "Controlled processing graphical properties."

Our classification scheme differs from these mentioned frameworks as follows. The unit object of interest to us is a web page, which is more specific than what is described in the "task by data type taxonomy" as any item as long as it has multiple well-defined attributes, and much more specific than what is described in the Card and Mackinlay framework merely as data to be visualized to communicate some information. Only certain (nominal) data types described in the Card and Mackinlay framework are applicable to web visualization. Meanwhile, partly due to its focused nature, our classification scheme examines the types of transformation (recoding) the original information collection goes through in more detail than either of these other frameworks. Also, the Card and Mackinlay framework admittedly does not include the user tasks a visualization is created to support, and Shnedierman's taxonomy defines the user task as finding objects satisfying certain well-defined criteria. In this dimension, our taxonomy is more similar to Card and Mackinlay's since the purpose of web visualization is more general than just finding objects satisfying certain well-defined criteria. People don't always have such clear purposes while examining web spaces hence the criteria that attributes of the pages are expected to satisfy are often vague and hard-to-define. This leads us to believe that there is value in a classification scheme that is created specifically for the web to fulfill the three objectives identified in the beginning of this section.

In our classification scheme, we take a web page as the unit information object in web space visualization. This approach is more generic than treating a web site (e.g. Terveen et al., 1999) as the unit of interaction, because it is rare that a specific information need will require the examination of a whole web site. Besides, some web sites are not fully connected, i.e. some pages do not link back to the main page and yet may contain what one is looking for.

As discussed in the previous section, information overload is reduced if some form of structure, i.e. organization or summary, is available in the presentation of information. Thus visual presentations of a web space present the collection within a structure. This structure may be the result of a manual or automatic organization. In rare cases, web pages are presented without any high-level organization (Nigay et al., 1998; Wakita & Matsumoto, 2003) or the organization is left to the end-user as in Web Book, Web Forager (Card et al., 1996) (Figures 2 and 3), and VIKI (Shipman et al., 1999) rather than being done automatically. In such cases, the web visualization paradigm simplifies to providing visual representations of web pages with a set of user-controlled attributes so that the user can manipulate these visual objects. Such systems can utilize the ongoing research on end-user support for the manual organization of a collection of objects as in desktop management systems (e.g. Amento et al. 2000).

Figure 2. An example WebBook

Figure 3. The Web Forage

On the other hand, automatic organization of web pages is more attractive to the end-user and the implementation is more computing-intensive and challenging. Therefore, much research has focused on this paradigm. The emphasis on automatic organization techniques is reinforced by the fact that these techniques can be applied manually by the user for smaller size collections; hence they are fairly generic. Accordingly, our emphasis in this paper is on visualization of web spaces that are based on automatic organization of web pages. Figure 4 displays our classification scheme for visualization of web spaces. The next section presents detailed discussion of this scheme.

I. Source of the Web Space
1. Static
2. Dynamic

II. Web Space Organization
1. Connectivity-based
2. Content-based
3. Other

III. Resulting Data (Graph) Structure
1. Generic graph
2. Tree
3. Network

IV. Graph Visualization Paradigm
a. Physical Dimensionality and Structure of the Visual Overview

1. 2D 2. 3D	1. Real life metaphor 2. Map 3. Tree 4. Network

b. Presentation of details
1. “one visualization only”
2. “details only”
3. “details and context”
4. “details embedded in context”

Figure 4. A Classification Scheme for Visualization of Web Spaces

Visualization of Web Spaces – State of the Art

Source of the Web Space

System surveys, including those on web systems, usually study systems according to their purpose. The focus of our classification scheme is a web space, which is a collection of web pages regardless of why they are perceived and processed as a collection. Accordingly for this paper, the purpose of visualization is important not only to understand the systems in general, but also because of the way it affects characteristics of the source of the web space, which, in turn, has direct implications for the design of visualization systems. Following this argument, we discuss the purpose of web space visualization systems simultaneously with the source of the web space they present. Specifically, we observe from a design perspective that it is reasonable to classify visualization paradigms based on whether the source of the web space is dynamic or static, because this characteristic influences the required speed in creating the visualizations. In visualizing web search results, the speed of the visualization process is critical, because the collection of web results are created in real time, and the end user of the visualization expects to use the results immediately. On the other hand, the speed of a system designed to visualize a web site is not that critical since the underlying structure of a web site is static (at least it does not change in real time), and the visualization can be prepared in advance to be displayed immediately on demand. Therefore we present two groups of systems on the source dimension: static versus dynamic.

Visualization of Static Web Spaces. The most typical example of a static web space is a web site. Web site visualization is a specific example of visualization of large network structures (Wills 1997) or visualization of hypertext in general (Rivlin et al., 1994). Many systems have been created to visualize a web site such as Hyperspace View (Gershon et al., 1995), Cyberspace Geography Visualization (Girardin, 1995), The Open Text Web Index Visualization (Bray, 1996), WebTOC (Nation et al., 1997), MAPA (Durand & Kahn, 1998), Cybermap (Gloor & Dynes, 1998), DiskTree (Chi et al., 1998), the system described in (Sakairi, 1999), 3D VRML (Mak et al., 2002), 2D DOM (Mak et al., 2002), HDV (Hypermedia Database Visualization) (Owor, 2002), HotSauce (Apple), and Site Manager (Silicon Graphics). Among these, 3D VRML and 2D DOM (Mak et al., 2002) present a site map customized for a requested root page (3D VRML is designed for quick access especially by mobile users) while the other systems provide a single view of a web site regardless of user requests. The system described in (Sakairi, 1999) provides an integrated view of a site's structure along with the collection of keywords extracted from the site to facilitate better interaction whereas The Open Text Web Index Visualization (Bray, 1996) presents visualizations of localities (connectedness) of web sites alongside important web site statistics, and DiskTree (Chi et al., 1998) presents visualization of a web site structure and access statistics (for later use in seeing time trends in analyzing web ecologies in "time tube").

Other static web spaces for which visualizations have been designed are collections of bookmarks as in Data Mountain (Czerwinski et al., 1999) and Lookmark (Breiteneder et al., 2002), a user's database of web pages as in Tecate Web Browser (Kochevar & Wanger, 1995), a generic collection of web pages as in Pathfinder networks with generalized similarity analysis (GSA) (Chen, 1997), WWW access logs as in Webviz (Pitkow & Bharat, 1994), and web directories as in WebBrain (http://www.webbrain.com) and map.net (http://maps.map.net). WebBrain and map.net provide different visualizations of the categories of web sites indexed through the Netscape Open Directory Project. Finally, the VSComp (Visual Site Comparison) system (Liu et al., 2002) visualizes the comparison of different web site contents. There are a number of other web visualization systems such as the self organizing map based system in (Lin et al., 1991) and WAVE (Kent & Neuss, 1994) that are flexible enough to handle dynamic spaces although their original design focus is a static web space.

Visualization of Dynamic Web Spaces. Since one of the most dominant paradigms of web access is the use of a search engine, the most typical examples of dynamic web visualization systems are those for the visualization of search results. Some of these systems search the web based on user-provided query terms as is typical with many commercial search engines today: MultiSurf/Query Graph and MultiSurf/Query Hit List (Hasan et al., 1995a), Relevance Map (Assa et al., 1997), Webquery (Vanish) (Carriere & Kazman, 1997), Adoptive SOM (Roussinov & Ramsey, 1998), VR-VIBE (Benford et al., 1997), Fettucino (formerly WebCutter) (Ben-Shaul et al., 1999), Sentinel-Visualeyes (Fox et al., 1999), Web local search tool (Angelaccio & Buttarazzi, 2000), WTMS (Mukherjea, 2000), CI Spider and Meta Spider (Chau et al., 2001), FISPA (Turetken & Sharda, 2004), Business Intelligence Explorer/Web Community (Chung et al., 2005), Business Intelligence Explorer/Knowledge Map[1] (Chung et al., 2005), Grokker (http://www.groxis.com), anacubis/google enabled visual search (http://www.onlineilink.com/demos/google/), and Kartoo (www.kartoo.com). Other systems follow a user provided URL (rather than a search query) to find similar or linked pages and visualize the resulting web space as in Auditorium Seating Visualization (Terveen et al., 1999), Touchgraph Google Browser (http://www.touchgraph.com/), and WebTracer (NullPointer http://www.nullpointer.co.uk/-/webtracer2.htm). Among these search results visualization systems, MultiSurf/Query Graph and MultiSurf/Query Hit List (Hasan et al., 1995a), VR-VIBE (Benford et al., 1997), and Sentinel-Visualeyes (Fox et al., 1999) present search results along with query terms while Relevance Map (Assa et al., 1997), Webquery (Vanish) (Carriere & Kazman, 1997), Auditorium Seating Visualization (Terveen et al., 1999), Fettucino (Ben-Shaul et al., 1999), and anacubis/google enabled visual search (http://www.onlineilink.com/demos/google/) include structural information extracted from the results collection (such as hyperlinks to and from search results) in the presentation of results. The other search result visualization paradigms present search results and their relationships to each other without query terms or explicit structural information.

The other large category of dynamic web spaces that are commonly visualized is browsing histories: Pad++ (Bederson & Hollan, 1994), WebMap (Doemel, 1994), WebMap (Gaines & Shaw, 1995), MosaicG (Ayers & Stasko, 1995), MultiSurf/Web Structure Graph (Hasan et al., 1995a), Hy+ (Hasan et al., 1995b), CZWeb (Fisher et al., 1997), Aleph/Travel Map (Neves, 1997), WWW3D (Benford et al., 1997), WebPath (Frecon & Smith, 1998), PadPrints (Hightower et al., 1998), WebClass (Greenhill & Venkatesh, 1999), and NESTOR (http://www.gate.cnrs.fr/~zeiliger/nestor.htm). Among these, WebClass (Greenhill & Venkatesh, 1999) can provide visualizations of the browsing history of multiple users while the others are typically for the visualization of one user's browsing pattern in one session. Other examples of dynamic web visualization systems are Ptolomaeus (http://www.dia.uniroma3.it/~ptolemy/) and Webviz/hyperbolic tree (Munzner & Burchard, 1995; Rao et al., 1995; Lamping & Rao, 1996; Pirolli et al., 1996a; Lamping et al., 1995; Pirolli et al., 2001) (also www.inxight.com), which create maps of a web space following a user defined URL and following the links (user-specified number in Ptolomaeus, all the links in Webviz) from there.

Web Space Organization

Whether their source is static or dynamic, the collection of web pages in a web space needs to be

[1] Business Intelligence Explorer works with web search results, however the results collection may be manually processed before visualization

organized in a way that will lead to a data structure that is then visualized. The organization of web pages can be based on numerous kinds of information. Understanding what the organization is based on is essential, because there are fundamentally different approaches in the processing of the web space depending on this basis. Below we discuss the main groups of systems we have identified in this dimension.

Web Space Organization based on Connectivity. A web site offers an already existing organization, i.e. links between pages, and many visualization systems present a web site based on the connectivity structure within or between web sites as in Nicheworks (Wills, 1997), WebTOC (Nation et al., 1997), Auditorium Seating Visualization (Terveen et al.,1999), the system described in (Sakairi, 1999), HDV (Hypermedia Database Visualization) (Owor, 2002), Site Manager (Silicon Graphics), Ptolomaeus (http://www.dia.uniroma3.it/~ptolemy/), HotSauce (Apple), and WebTracer (NullPointer http://www.nullpointer.co.uk/-/webtracer2.htm).

The use of connectivity structure in organizing a web space is not limited to web pages within the same web site. For example, in its visualization of search results, the Fettucino system (Ben-Shaul et al., 1999) displays structural (connectivity) information so as to facilitate dynamic browsing. Similarly, web local search tool (Angelaccio & Buttarazzi, 2000) provides structural (connectivity) information in the visualization of the results of a local search, i.e. search done within a chosen URL.

The use of already existing links in the organization of a web space is straightforward and inexpensive from a design perspective, however, especially for large web spaces, it is a challenge to visualize the connectivity structure of a web space since this structure is often complex. Therefore, organization of the existing connectivity structure into a simplified structure (typically a hierarchy or a graph of web pages with weighted links between pages) is a common approach in web space visualization as in (Noik, 1993), (Rivlin et al., 1994), Cyberspace Geography Visualization (Girardin, 1995), Hyperspace View (Gershon et al., 1995), The Open Text Web Index Visualization (Bray, 1996), DiskTree (Chi et al., 1998), MAPA (Durand & Kahn, 1998), 2D DOM and 3D VRML (Mak et al., Chan 2002), and Webviz/hyperbolic tree (Munzner & Burchard, 1995; Rao et al., 1995; (Lamping & Rao, 1996; Pirolli et al., 2003; Lamping et al., 1995; Pirolli et al., 2001) (also www.inxight.com). For such an approach, the complexity of the transformation of the already existing connectivity structure to a simplified one would directly impact the efficiency of the visualization system. For example, the hyperbolic tree creates a tree structure from a graph simply by replicating nodes that can be reached through multiple paths therefore the transformation is not very computationally complex.

Web Space Organization based on Semantic Content. The visualization of a collection of unconnected web pages is challenging since, unlike a collection of web pages with links between them, such a collection is rather unstructured. The most salient approach for the organization of such web spaces is the use of content similarities. Most material on the web is still textual, and techniques to organize a collection of objects based on textual similarities are relatively well developed (Willet, 1988; Salton, 1989; El-Hamdouchi & Willet, 1989).. The core of these techniques is the representation of each textual document (for our purposes, a web page) as a vector of term frequencies and the consequent processing of these vectors via well-known numeric techniques such as clustering and multi-dimensional scaling. Due to the relative maturity of these techniques, organizing a web space based on textual content of the web pages is common practice as in the system described in (Lin, et al., 1991), MultiSurf/Query Hit List (Hasan et al., 1995a), Aleph/Content View (Neves, 1997), Relevance Map (Assa et al., 1997), VR-VIBE (Benford et al., 1997), Adoptive SOM (Roussinov & Ramsey, 1998), Cybermap (Gloor & Dynes, 1998), Sentinel-Visualeyes (Fox et al., Frieder, 1999), CI Spider and Meta Spider (Chau et al., 2001), FISPA (Turetken & Sharda, 2004), and Kartoo (www.kartoo.com). Among these MultiSurf/Query Hit List (Hasan et al., 1995a), VR-VIBE (Benford et al., 1997) and Sentinel-Visualeyes (Fox, et al., 1999) are designed for the visualization of search results, and they present search results along with query terms showing the semantic similarity between each query term and search result pair. Meanwhile, although the visualizations in Relevance Map (Assa et al., 1997), Adoptive SOM (Roussinov & Ramsey, 1998), CI Spider and Meta Spider (Chau et al., 2001), FISPA (Turetken & Sharda, 2004), and Kartoo (www.kartoo.com) are also designed for presenting search results, the displayed similarities are between different web pages and between web pages and terms (themes) extracted from the web space. The system described in (Lin et al., 1991) and Aleph/Content View (Neves, 1997) have similar visualization principles, but these systems are not limited to working with search results only. Finally, the VSComp (Visual Site Comparison) system (Liu et al., 2002) utilizes textual web content for the clustering and comparison of different web sites.

From a design perspective, the techniques used for the extraction of semantic content from a large collection of text have direct impact on the effectiveness and efficiency of the systems discussed in this section. The research on advanced AI and text mining techniques for automatic summarization and content extraction are directly relevant to web visualization when the web space to be visualized is large and unlinked and hence content based organization is a strong alternative to connectivity. Meanwhile, the amount of non-textual data on the web such as image, audio, and video, is increasing, and there have been efforts to automatically extract attributes of non-textual data for indexing and organization purposes (Chang et al., 1997). This is still a widely open research area, likely to produce interesting results in the future.

Other Web Space Organizations. It has been suggested that, in addition to connectivity and content, various characteristics of web pages including file-system attributes, access statistics, usage statistics (Pirolli et al., 1996a), web sites that web pages come from, author, and time of publication (Baldonado & Winograd, 1997) (Shipman et al., 1999) can be used as bases of web space organization. For example, Data Mountain (Czerwinski et al., 1999) and Lookmark (Breiteneder et al., 2002) use the original structure of a collection of web pages bookmarked by a user as the basis of organization where Tecate Web Browser (Kochevar & Wanger, 1995) organizes web pages based on the geographical location of web servers the pages come from.

As mentioned before, there are many systems designed for presenting browsing histories. These systems use the browsing pattern of a user as the basis of web space organization as in Pad++ (Bederson & Hollan, 1994), WebMap (Doemel, 1994), WebMap (Gaines & Shaw, 1995), MosaicG (Ayers & Stasko, 1995), MultiSurf/Web Structure Graph (Hasan et al., 1995a), Hy+ (Hasan et al., 1995b), CZWeb (Fisher et al., 1997), Aleph/Travel Map (Neves, 1997), WWW3D (Benford et al., 1997), WebPath (Frecon & Smith, 1998), PadPrints (Hightower et al., 1998), WebClass (Greenhill & Venkatesh, 1999), and NESTOR (http://www.gate.cnrs.fr/ ~zeiliger/nestor.htm). Among these, CZWeb uses originating web sites the web pages come from for clustering them in the display of the browsing history. WebViz (Pitkow & Bharat 1994) is based on the same principle as the above systems except for the fact the organization is based on cumulative access log analysis rather than one user's browsing history. On the other hand, MultiSurf/Query Graph is a similar system designed for successive web searches therefore its organization of the web space is based on how different query terms are used in successive queries and how these choices affect the result sets. MultiSurf/Query Graph displays the query terms in a way that suggests the concept space a user is exploring.

Browsing patterns or the pattern of query terms used in web search are natural choices in organizing browsing or search histories where the (browsing and search) data are readily available from system logs and hence the organization is not resource-intensive even when we pages are clustered as in CZWeb. However, such patterns may not be available or appropriate to use for other web spaces, for example a large web directory such as Yahoo or the Netscape Open Directory. The systems designed to visualize these directories (Grokker for Yahoo, WebBrain and map.net for the Netscape Open Directory) organize the web space based on the original organization of the directory. Likewise Touchgraph Google Browser uses relations identified through Google's "similar to" facility to organize the web pages similar to a user-provided page.

There have also been attempts to use a combination of the above mentioned attributes for organizing a web space. The most common combination is connectivity and content as used in Webquery (Vanish) (Carriere & Kazman, 1997), Navigational View Builder (Mukherjea & Foley, 1994), Business Intelligence Explorer/Web Community (Chung et al., 2005), and Business Intelligence Explorer/Knowledge Map (Chung et al., 2005). This combination is potentially powerful in revealing relationship between web pages that is not available through connectivity or content based organizations alone. The design challenge is to find effective techniques to combine information that is separately gathered from connectivity and web page content (such as the algorithm that the Google search engine uses to combine connectivity and content information in determining similarity scores).

Other examples of systems that use a combination of attributes are WAVE (Kent & Neuss, 1994), which makes different combinations of connectivity, content, file-system structure, access statistics, originating web information, time of publication, author, and other user-defined attributes available to provide different organizations of a web space, anacubis/google enabled visual search (http://www.onlineilink.com/demos/google/), which uses links and content similarities (identified by Google) between web pages, and WTMS (Mukherjea, 2000), which bases the web space organization on connectivity and the directory structure of the originating web sites. Finally, Pathfinder networks with generalized similarity analysis (GSA) (Chen, 1997) can use either one of connectivity, content, or browsing patterns as the basis of web space organization.

Resulting Data Structure

Regardless of how the web space is organized, the resulting structure represents web pages and relationships (that are predefined by the site designer or imposed by the system) between them. In the most general sense, such a data structure is a graph with individual web pages or groups of these as its nodes and the relationship between the nodes as its edges. Different kinds of graphs such as a network or a tree are defined based on the characteristics of the edges between the nodes. A tree is a graph structure where there is a hierarchy between nodes and every node has one "parent" where a network implies that different edges of the graph have different strengths (weights). Web visualization systems differ based on the type of data structure that results from the organization of the web space. This structure has implications for the visualization of the web space, because certain visualization paradigms are better suited for certain graph structures as discussed in (Herman et al., 2000).

Generic Graph Structure. In our survey of web space visualization, the most commonly observed data structure is a generic graph (graph that assumes no specific relationship between nodes other than absence or presence of a connection). The structure of the WWW is a generic graph (with web pages as nodes and links as edges) therefore systems that visualize this already existing structure are generic graph visualizations as in WebMap (Doemel, 1994), Pad++ (Bederson & Hollan, 1994), Webviz (Pitkow & Bharat, 1994), MultiSurf/Query Graph and MultiSurf/Web Structure Graph (Hasan et al., 1995a), Hy+ (Hasan et al., 1995b), Aleph/Travel Map (Neves, 1997), Nicheworks (Wills, 1997), WWW3D (Benford et al.,1997), WebPath (Frecon & Smith, 1998), the system described in (Sakairi, 1999), WebClass (Greenhill & Venkatesh, 1999), Fettucino (formerly WebCutter) (Ben-Shaul et al., 1999), HDV (Hypermedia Database Visualization) (Owor, 2002), WebTracer (NullPointer http://www.nullpointer.co.uk/-/webtracer2.htm), Ptolomaeus (http://www.dia.uniroma3.it/~ptolemy/), Touchgraph Google Browser (http://www.touchgraph.com/), Site Manager (Silicon Graphics), and NESTOR (http://www.gate.cnrs.fr/~zeiliger/nestor.htm). The approach used in Auditorium Seating Visualization (Terveen et al., 1999) also makes use of the already existing connectivity structure. However this structure is not directly visualized, but is further processed into what the authors call a 'clan graph'.

Other examples of generic graph structures that are visualized are lists (Data Mountain Czerwinski et al., 1999)), (many to many) relationships between search results and query terms (MultiSurf/Query Hit List Hasan et al., 1995a)), search results and clusters (Grokker, http://www.groxis.com)), web pages and categories from a web directory (WebBrain, http://www.webbrain.com), map.net (http://maps.map.net)), and search results and linked pages (anacubis/google enabled visual search, http://www.onlineilink.com/demos/google/),

Tree Structure. The other data structure common in web visualization is a tree. Examples of tree structures are found in WAVE (Kent & Neuss, 1994), Hyperspace View (Gershon et al., 1995), MosaicG (Ayers & Stasko, 1995), WebMap (Gaines & Shaw, 1995), Webviz/hyperbolic tree (Munzner & Burchard, 1995; Rao et al. 1995; Lamping & Rao, 1996; Pirolli et al., 2003; Lamping et al., 1995; Pirolli et al., 2001) (also www.inxight.com), CZWeb (Fisher et al., 1997), Webquery (Vanish) (Carriere & Kazman, 1997), WebTOC (Nation, et al., 1997), Cybermap (Gloor & Dynes, 1998), DiskTree (Chi et al., 1998), PadPrints (Hightower et al., 1998), MAPA (Durand & Kahn, 1998), Web local search tool (Angelaccio & Buttarazzi, 2000), WTMS (Mukherjea, 2000), Lookmark (Breiteneder et al., 2002), VSComp (Visual Site Comparison) (Liu et al., 2002), FISPA (Turetken & Sharda, 2004), Business Intelligence Explorer/Web Community (Chung et al., 2005), Business Intelligence Explorer/Knowledge Map (Chung et al., 2005), and HotSauce (Apple). Among these, some systems simply use the already existing tree structure of the web space such as the hierarchical bookmark structure (Lookmark) and directory/files system structure (WMTS, TopicShop (versions 1 and 2)), or process the generic graph consisting of original links (Webviz/hyperbolic tree, Webquery, WebTOC, MAPA, Web local search, Disktree, Hyperspace View) or browsing patterns (MosaicG) to reduce it to a tree structure. Others apply a (hierarchical) clustering technique to the collection of web pages based on connectivity (Hot Sauce), content (VSComp, FISPA, and CyberMap), usage pattern (WebMap, PadPrints, originating web (CZWeb), a combination of connectivity and content (Business Intelligence Explorer/Web Community and Business Intelligence Explorer/Knowledge Map), or a user chosen attribute (WAVE).

Network Structure. Many web visualization systems do not only depict the presence or absence of a relationship between different objects, but also visualize the strength of these relationships. Examples are Tecate Web Browser (Kochevar & Wanger, 1995) that visualizes the geographic proximity of the web severs that the web pages come from, The Open Text Web Index Visualization (Bray, 1996) that visualizes a web site along with certain statistics, Aleph/Content View (Neves, 1997) that visualizes similarities between web pages and certain terms, Relevance Map (Assa et al., 1997) that visualizes the strength of the relationships between web search results and features extracted form the collection of results, and VR-VIBE (Benford et al., 1997) and Sentinel-Visualeyes (Fox et al., 1999) that visualize the strength of the relationships between web search results and query terms.

The other group of systems in this category are those that visualize a collection of documents with the strengths of the relationship between them and/or between each page and semantic terms extracted from the collection. Examples are Webviz (Pitkow & Bharat, 1994), Pathfinder networks with generalized similarity analysis (GSA) (Chen, 1997), the self organizing map based systems in (Lin et al., 1991), CI Spider and Meta Spider (Chau et al., 2001), Cyberspace Geography Visualization (Girardin, 1995), Adoptive SOM (Roussinov & Ramsey, 1998), 3D VRML and 2D DOM (Mak et al., 2002), and finally Kartoo (www.kartoo.com).

Visual Paradigm

The organization of the web space is not automatically "visible", because none of the data structures discussed above has a unique natural visual metaphor. For example, a network of web page clusters does not necessarily conjure up the same image for everyone. Therefore, a common task in visualization is choosing a paradigm within which the web space organization can be presented in a "visible" (up to three) dimensionality. This task is a major difference between information visualization and (scientific) data visualization in that data visualization deals with representing information that is already two or three-dimensional with an inherent visual structure, e.g. the molecular structure of organic tissue. Yet, information visualization is about the representation of abstract, i.e. non-physical, objects. Naturally, the subject of our review has been information visualization since web spaces are not physical objects.

A comprehensive survey of graph visualization algorithms is presented in (Herman et al., 2000) therefore will not be repeated here. Rather, our aim is to classify, among two major dimensions, alternative paradigms for presenting the graph structure that specifically represents a web space. These major dimensions of interest are 'physical dimensionality and structure of the visual overview', and the 'presentation of the details.'

The visual overview of an information collection is the first (and sometimes the only) visual presentation of that collection that a viewer will be exposed to. As with any information visualization system, visualizations of web spaces are typically 2 or 3-dimensional, and they vary greatly depending on the structure of the visual overview.

In visualization of relatively small web spaces (such as a personal web site) all the details of the web space may be visible in one visualization. In such cases, the details and overview are presented at once. We call this visualization paradigm "one visualization only". More typically, it is not feasible to present all details of the web space simultaneously. Rather a visual overview is presented as a starting point from which details can be further examined on demand. Hence a visualization system should provide means to see "interesting" details while making it possible to de-emphasize the irrelevant ones. Of course, the concept of "interesting" is context and individual specific therefore the presentation of details should be dynamic and flexible. Visual paradigms differ based on how they provide the details and context (overview).

The traditional way of examining a large collection of information is to first look at an overview to identify the portions of interest, and then to look into those portions of in full detail. Doing this, the details that are not of immediate interest are filtered from the view. We call this strict filtering approach "details only" visualization.

The drawback of the "details only" approach is the disorientation problems caused by the required effort to mentally integrate context and details (Furnas, 1986; Furnas, 1997). One alternative to this kind of "strict filtering" is keeping the overview always visible regardless of the details being examined. We call this visualization paradigm "details and context". The major drawbacks with this approach are the wasted space occupied by the overview, and the difficulty (although not as severe as in the "details only" paradigm) to mentally integrate the two (or more) diagrams. Therefore another alternative proposed for presenting details and context is embedding the details into the context with or without the distortion of the context (i.e. the original overview). We call this visualization paradigm "details in context".

Below is a discussion of the major visualization types we have identified along the "visual paradigm" dimension. Because the 'presentation of the details' is closely tied to the 'physical dimensionality and structure of the visual overview' of a visualization, we discuss these dimensions together for each visualization.

Tree and Network Diagrams. The most straightforward visualization of a graph (network, tree, or otherwise) is a display of node images (that represent the nodes of the graph) and lines connecting these node images (that represent the links between the nodes). Such a visualization takes advantage of gestalt theory's *good continuation* principle (Tan & Benbasat, 1993), which states that connecting lines between items emphasize relatedness.

If the underlying data structure is a tree, such a diagram looks like a tree as in WebMap (Gaines & Shaw, 1995), MosaicG (Ayers & Stasko, 1995), Hyperspace View (Gershon et al., 1995), Webviz/ hyperbolic tree (Munzner & Burchard, 1995; Rao et al., 1995; Lamping & Rao, 1996; Pirolli et al., 1996a; Lamping et al., 1995; Pirolli et al., 2001) (also www.inxight.com) , Aleph/Travel Map (Neves, 1997), DiskTree (Chi et al., 1998), PadPrints (Hightower et al., 1998), the system described in (Sakairi, 1999), Fettucino (Ben-Shaul et al., 1999), Web local search tool (Angelaccio & Buttarazzi, 2000), WTMS/tabular view (Mukherjea, 2000), 2D DOM (Mak et al., 2002), VSComp (Liu et al., 2002), and Business Intelligence Explorer/Web Community (Chung et al., 2005), where the tree leaves represent individual web pages. Otherwise, the visualization is a network diagram as in Pad++ (Bederson & Hollan, 1994), WebMap (Doemel, 1994), MultiSurf (Hasan et al., 1995a), Hy+ (Hasan et al., 1995b), VR-VIBE (Benford et al., 1997), Pathfinder networks with generalized similarity analysis (GSA) (Chen, 1997), Webquery (Vanish) (Carriere & Kazman, 1997), WWW3D (Benford et al., 1997), Cybermap (Gloor & Dynes, 1998), WebPath (Frecon & Smith, 1998), WebClass (Greenhill & Venkatesh, 1999), 3D VRML (Mak et al., 2002), Business Intelligence Explorer/Knowledge Map (Chung et al., 2005), anacubis/google enabled visual search (http://www.onlineilink.com/demos/google/), Touchgraph Google Browser (http://www.touchgraph.com/), Site Manager (Silicon Graphics), WebTracer (NullPointer http://www.nullpointer.co.uk/-/webtracer2.htm), WebBrain (http://www.webbrain.com), NESTOR (http://www.gate.cnrs.fr/~zeiliger/nestor.htm), and Ptolomaeus (http://www.dia.uniroma3.it/~ptolemy/).

Figure 5 displays a two-dimensional tree created by the Fettucino (Ben-Shaul et al., 1999) system for the results of the query "Sun Corporation/McLaren deal". The problem with a regular 2-D tree visualization such as this one is that the visualization is not space efficient, i.e. the amount of information that can be displayed at a time is very limited.

Therefore space saving techniques for the display of a tree has been a topic of interest to developers of tree visualizations. One such technique is the use of "hyperbolic space" (Munzner & Burchard, 1995; Rao et al., 1995; Lamping & Rao, 1996; Pirolli et al., 1996a; Lamping et al., 1995; Pirolli et al., 2001) (also www.inxight.com) instead of the "Euclidian space" (that is used in many visualizations such as Fettucino). The use of hyperbolic space renders a 3-D effect and has a better utilization of available space as seen in Figure 6.

This visualization (star tree from inxight.com) displays the first-degree (sometimes second or third, space permitting) links in a web site in a compact form where the viewer is able to quickly see the distribution of the web pages within the site.

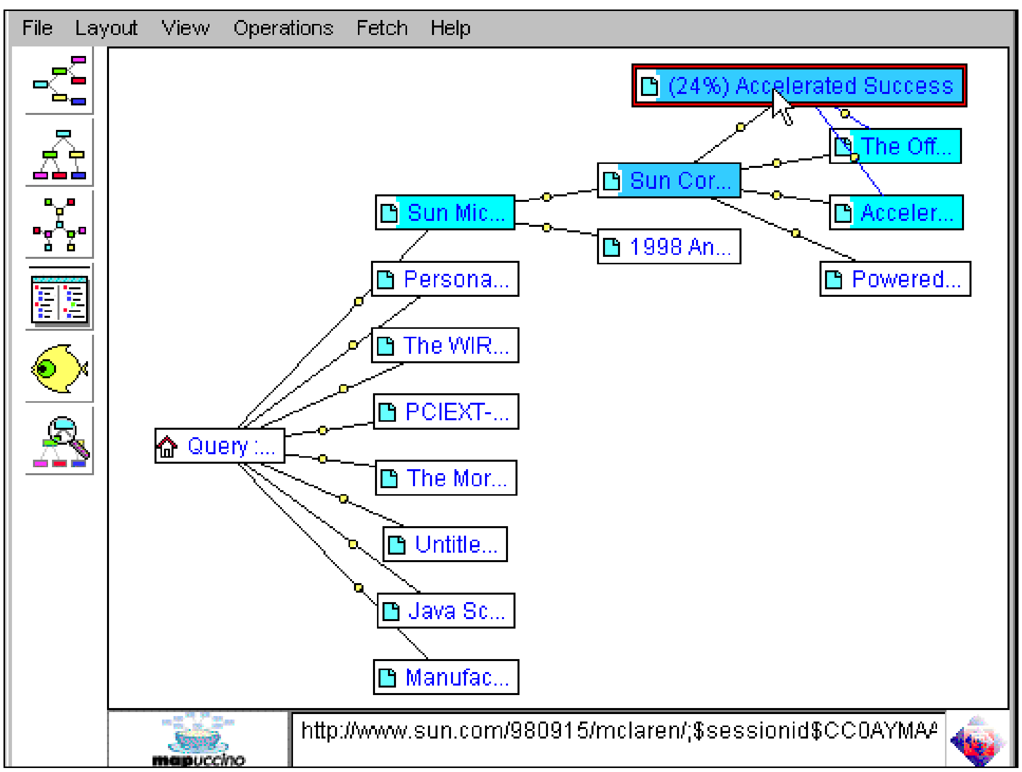

Figure 5. Fettucino Query Results

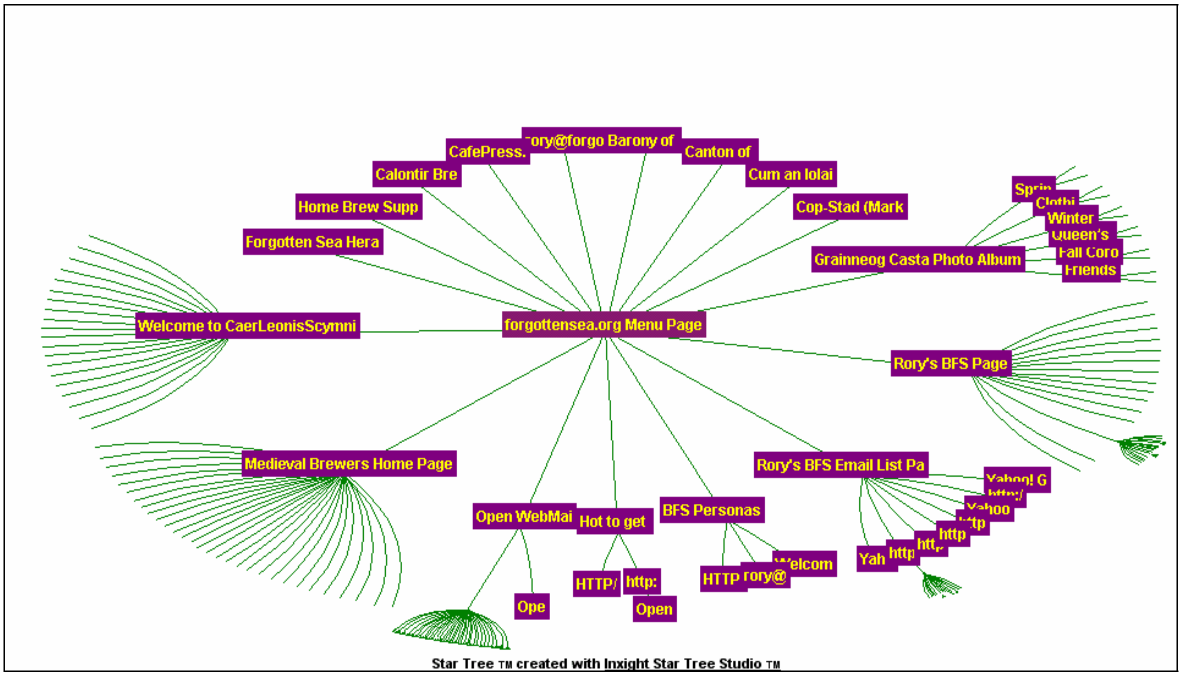

Figure 6. Overview of a Web Site Using Hypertree

This display provides a smoothly varying focus+context view where the display space allocated to a node decreases continuously with the distance from the focus, yet does not disappear abruptly. When the user wants to focus on a specific portion of the visualized web space, she brings that portion to focus as seen in Figure 7, and looks at those details in the context of the rest of the web site. This "details in context presentation" is a specific implementation of the "fisheye view" concept (Furnas, 1986).

Very often, when a graph structure other than a tree is visualized, the visual paradigm of choice is not limited to a tree, but rather reflects the underlying structure more accurately as a "network diagram".

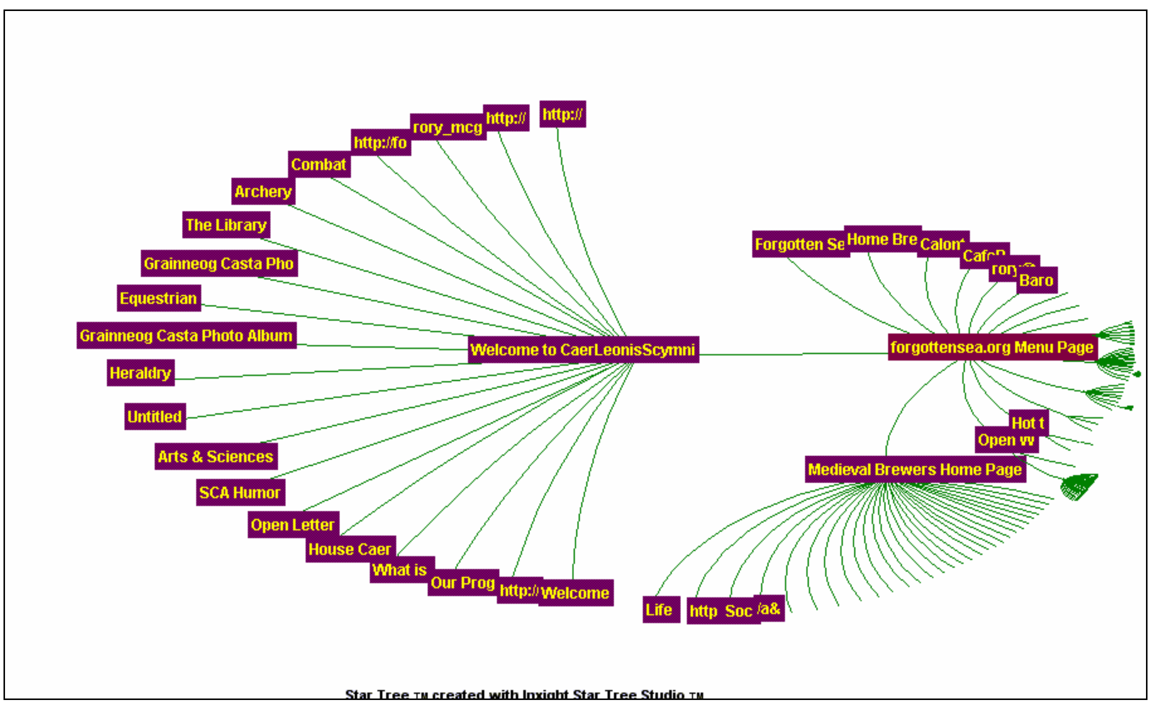

Figure 7. Zooming in on a branch Using Hypertree

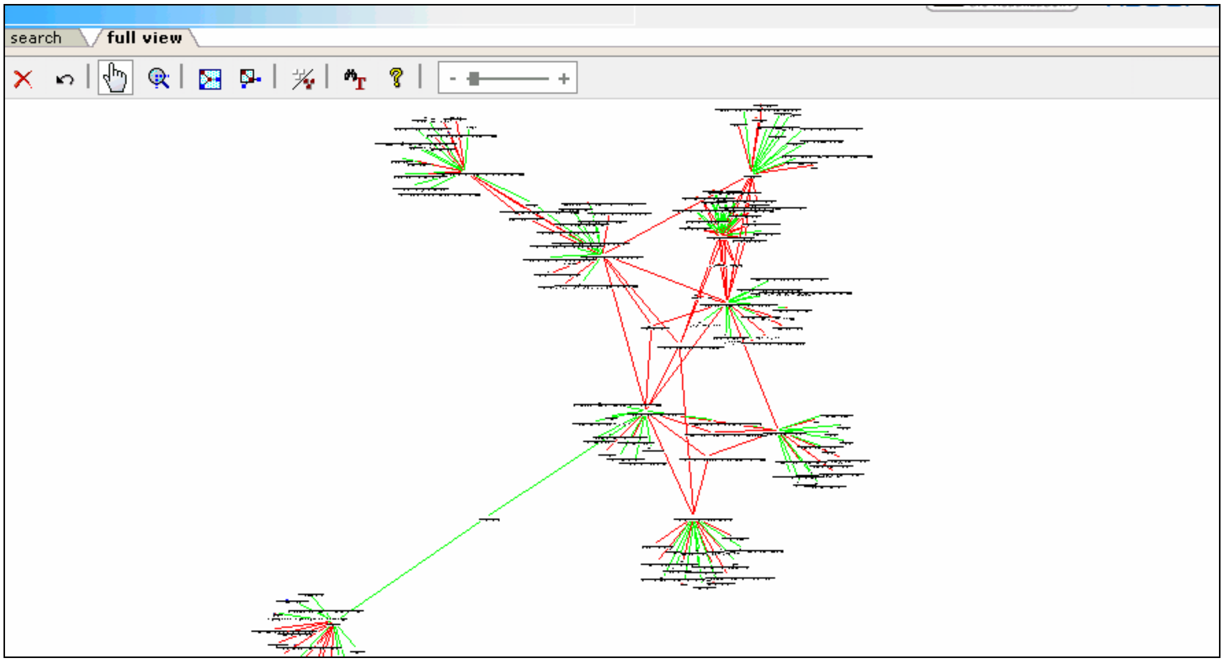

Figure 8. Results to the query "China" from anacubis/Google Enabled Visual Search

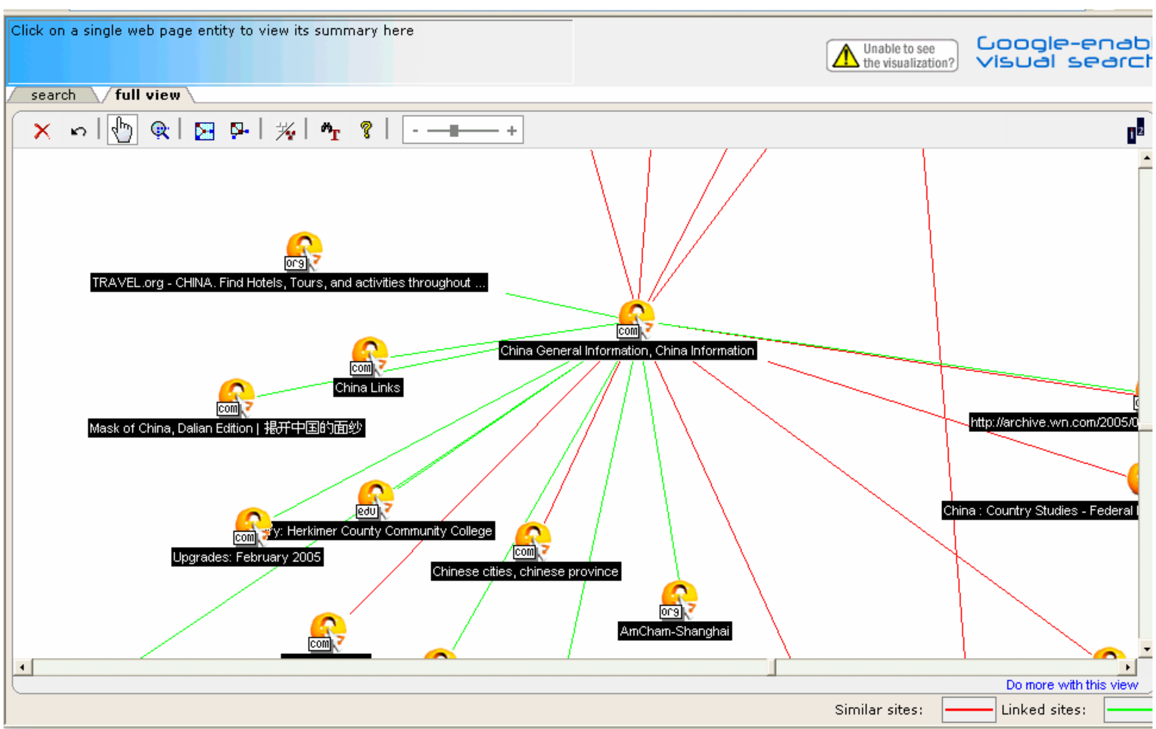

Figure 9. Zooming in on a portion of the network

For example, the Google enabled visual search module of anacubis displays search results with pages they link to along with other pages retrieved with Google's "similar to" facility (Figure 8). It should be noted that our use of the term "network diagram" here simply refers to the appearance of the visual overview rather then the underlying data structure hence it does not mean that the underlying data structure has different weights on the links. In fact, the underlying data structure in this example is a generic graph structure. This visualization is a 2D overview of the selected portion of the results collection. The system provides "details only" zooming into part of this collection as seen in Figure 9.

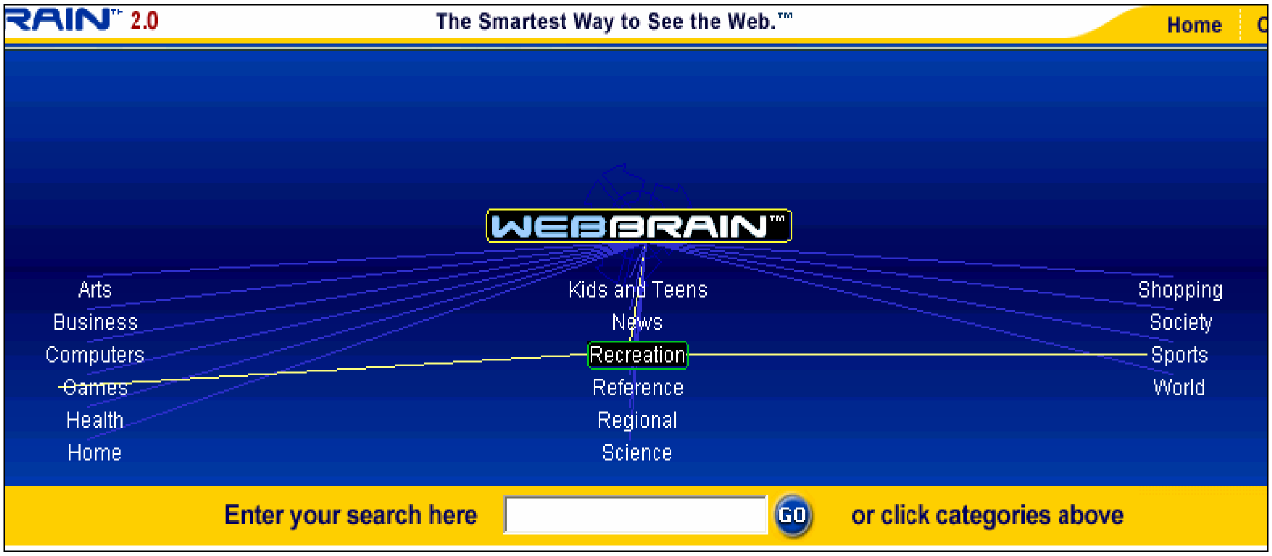

Figure 10. The WebBrain System Displaying Pages from Netscape Open Directory

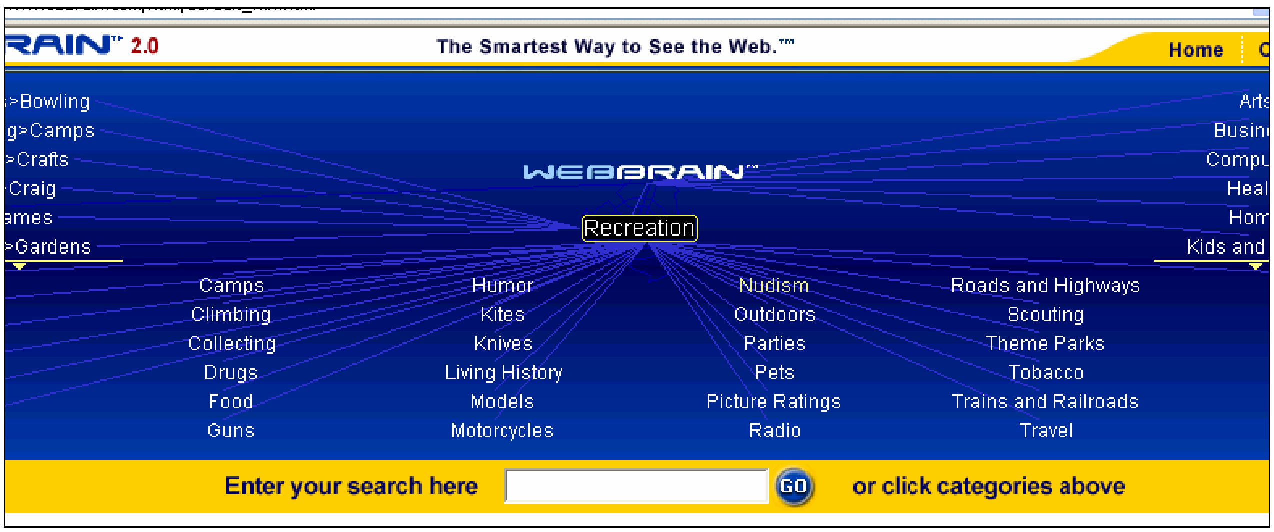

Figure 11. Zooming in on "Recreation"

As with many other examples of the presentation of "details only", the problem with this visualization is the disconnection between the presentation of the overview and the details.

As a possible solution to the cognitive problems due to viewing "details only", the WebBrain system displays a network diagram where the details are viewed in the context of the rest of the directory that the system visualizes. As seen in Figure 11, the details of the "recreation" branch of the directory (seen in Figure 10) can be viewed in their context (Categories Art, Business, etc.). It should be noted that the difference between the "details in context" approach in WebBrain and that in the Hyperbolic Tree example discussed above is that the zoom facilities in WebBrain does not distort the visualization whereas that in Hyperbolic Tree does. The advantage of distortion is the more efficient utilization of space without much clutter while the lack of distortion guarantees a clearer view at all times.

Real Life Metaphors. The world we live in is a visual world. People have no difficulty in understanding the objects that they are used to seeing all the time and can easily interpret the relationships between these objects. Consequently, a natural thought in visual system design would be the representation of abstract objects by familiar physical entities. The systems we describe in this subsection work on this simple principle.

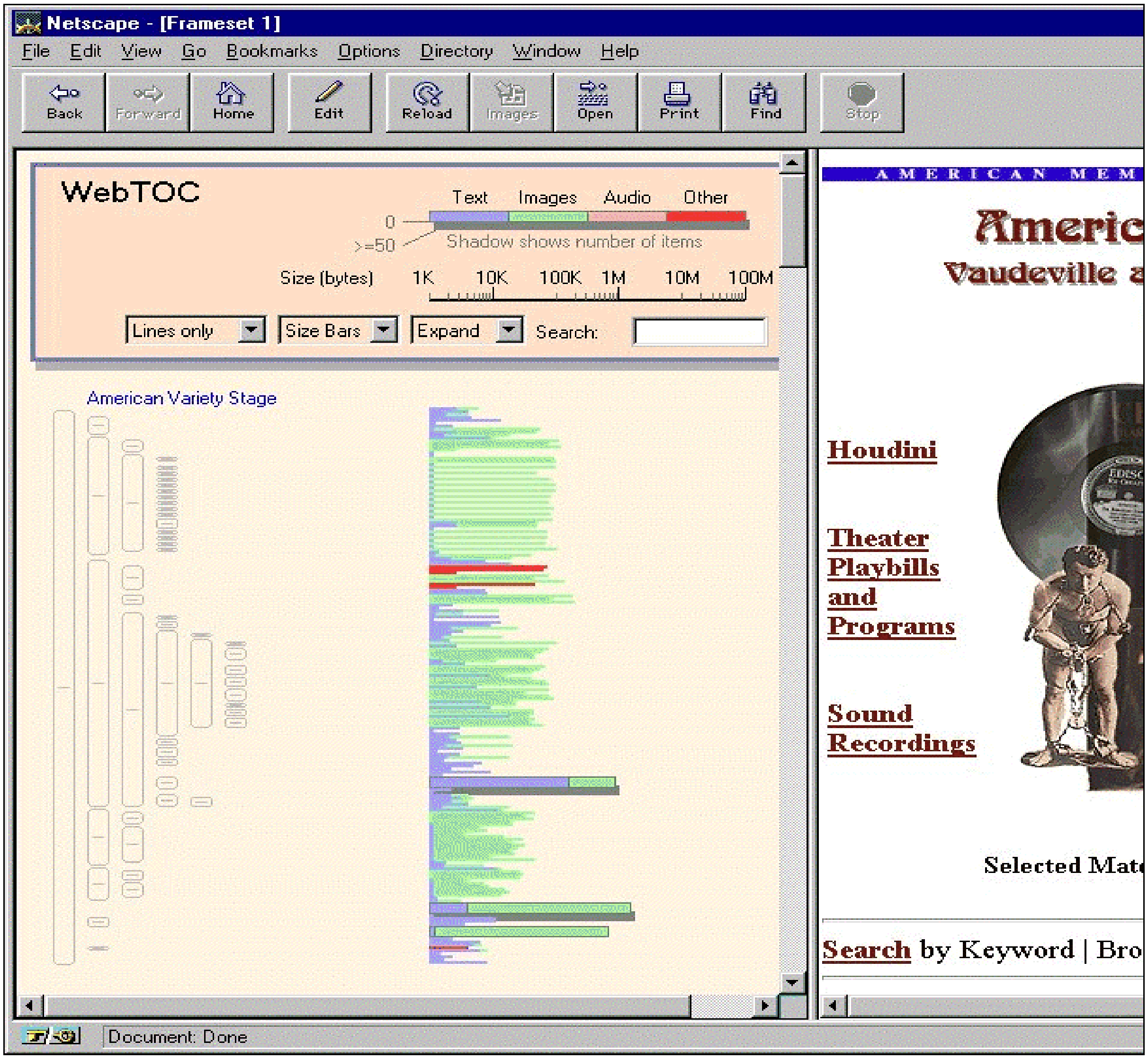

Figure 12. A WebTOC Session

WebTOC (Nation et al., 1997) presents the structure of a web site by a (hierarchical) table of contents metaphor. As seen in Figure 12, the table of contents shows the total size of the files under each branch of the TOC and different file types are represented by different colors. This overview is designed to give viewers a quick idea of the distribution of material in the web site, which, obviously, is not available in the normal view of a web site as seen in the right hand side of Figure 12. The WebTOC system has other similar presentation modes such as using the length of bars to represent the number of documents (as opposed to the total file size) in a certain branch of a hierarchy.

Figure 13 displays the zooming feature of TOC where the details of the "English Playscripts" part of the TOC are viewed within the context of the overall hierarchy. Here the "details in context" zoom has been rendered with a uniform scale down, but without the distortion of the rest of the TOC. Users can adjust the amount of detail they want to see by zooming in further on any part of this hierarchy.

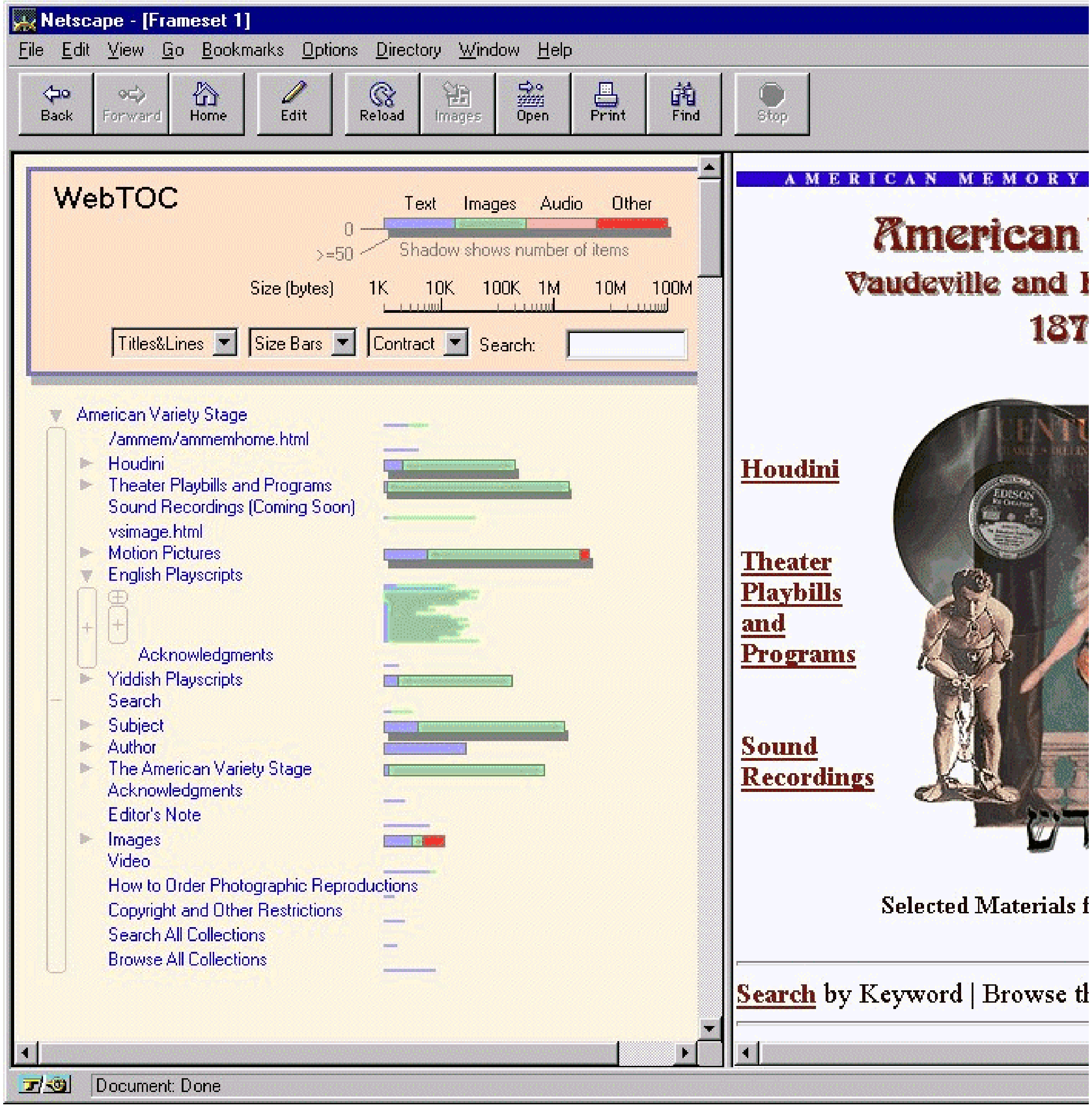

Figure 13. Zooming in on the “English Playscripts” Branch of the TOC

The HDV Information Visualizer system (Owor, 2002) uses the “office space” metaphor to present the web pages on a site as in Figure 14. When a user zooms on part of this overview (in this case a faculty member’s “office, i.e. “web site”) as displayed in Figure 15, the details are viewed separate from the overview hence this is a perfect example of a “details only” zooming approach.

Another system that uses a real life metaphor is the Auditorium Seating Visualization (Terveen et al., 1999) where a graph that represents a collection of web sites is presented through an auditorium metaphor where the sites reside on concentric semi-circles (seats around the stage), which group them into equivalence classes from most to least important on some user-settable property. By default, sites are assigned to rows based on the number of in-links, so the closer to the center a site is, the more of its “peer sites” have linked to it (Terveen et al., 1999). This system also allows dynamic ordering of sites within a row by properties like number and proportion of in and out links, and amount of content (audio files, images, or all types of content). By default, the sites within rows are ordered by the amount of content, so sites with lots of content appear at the top of each row.

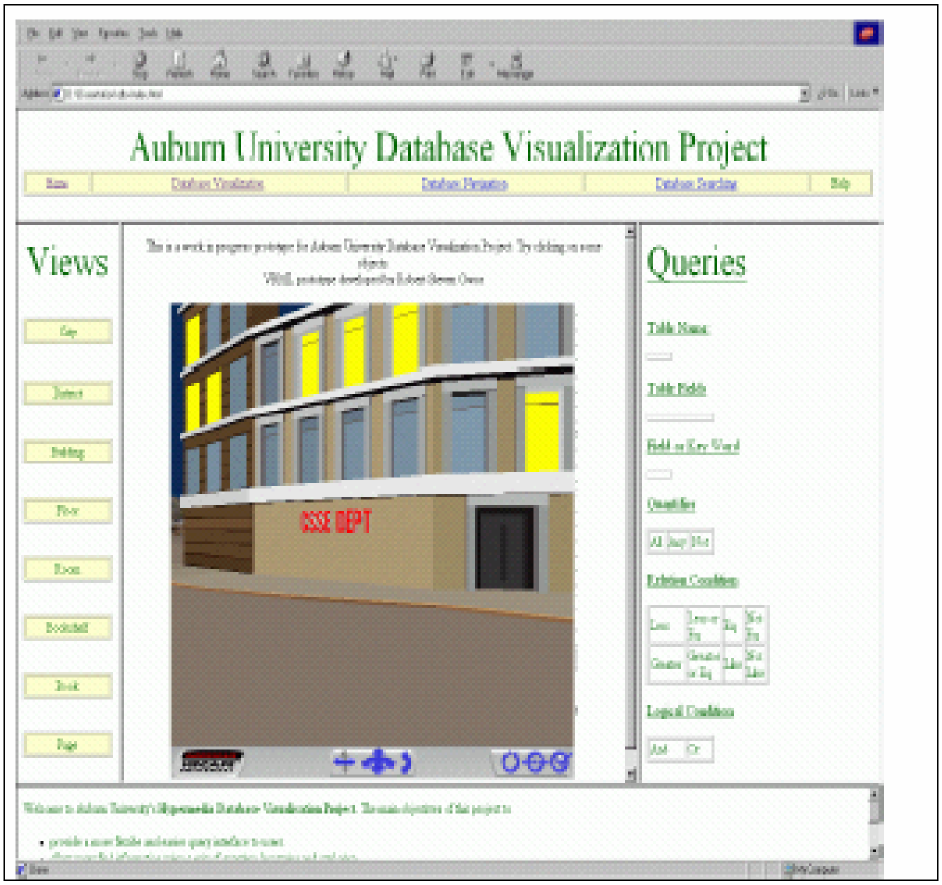

Figure 14. Visualization of a Department Web Site in HDV

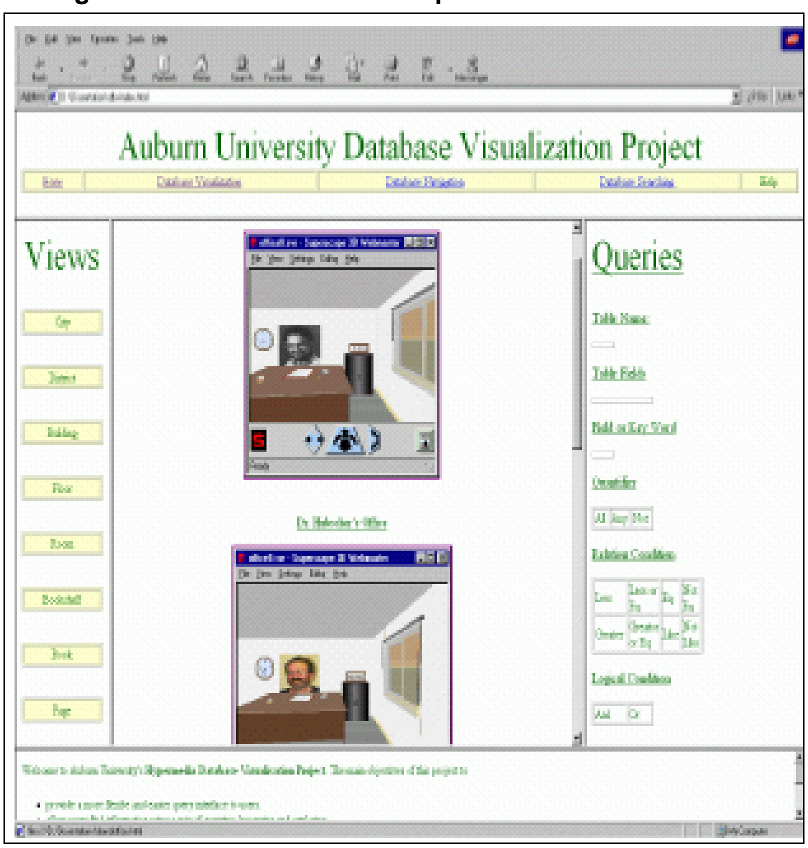

Figure 15. Zooming in on a Faculty Member's Web Site

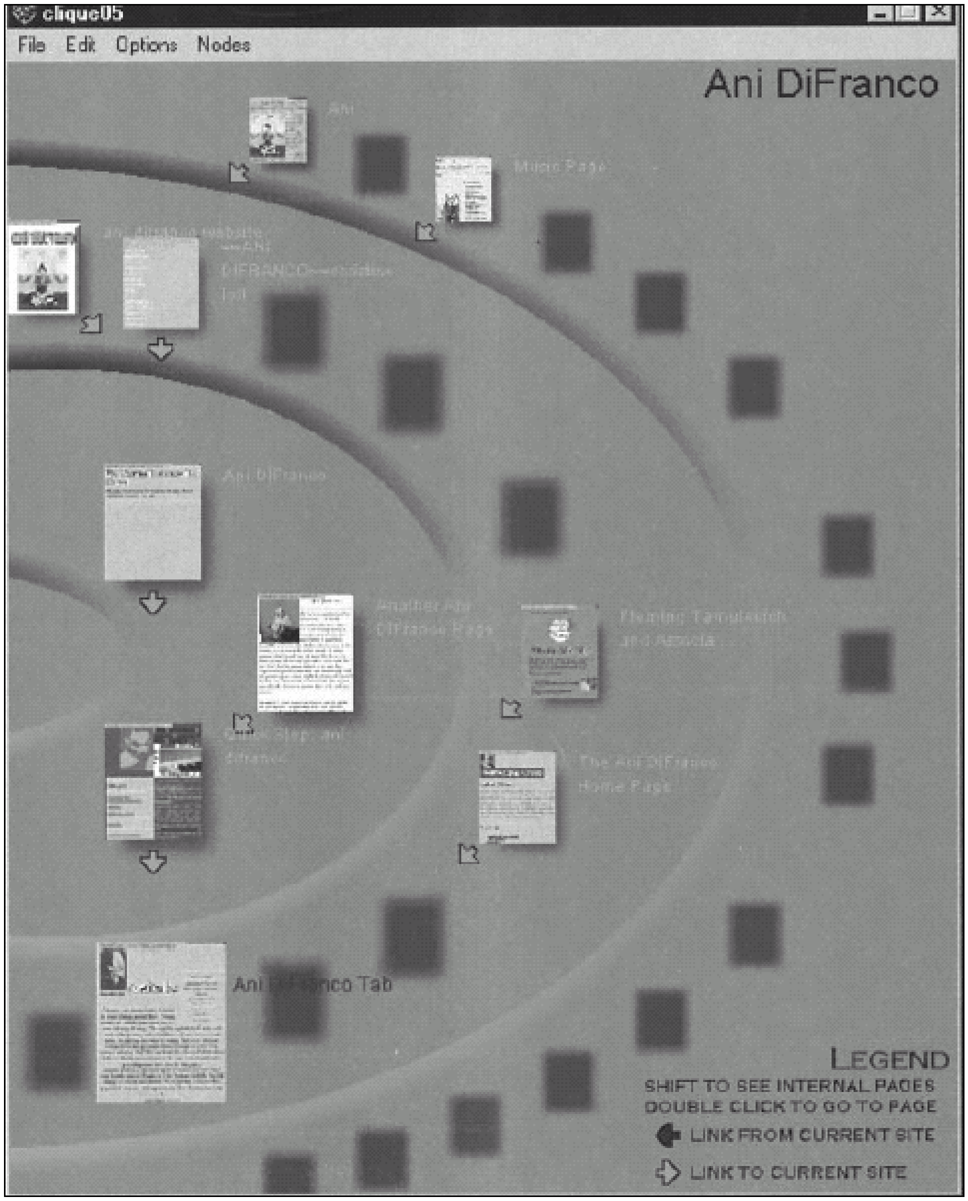

Figure 16. The Auditorium Seating Visualization of Web sites about Ani DiFranco

The zooming approach of this visualization is to display the site in focus larger with its in and out links as seen in Figure 16. Further zooming in on a web site reveals any internal pages of that site that are linked to by other sites. These are pages that the author of the linking site found worthy of special attention. The link text often is more informative in these cases. In addition, by clicking on a site, the user can access a display of the site profile data to find the amount and type of information the site contains and access significant internal pages (Terveen et al., 1999)..

As seen through the examples of this subsection, real life metaphors can be used for the visualization of different data structures (tree, graph, etc.) and with a number of different zooming facilities. The usability advantages combined with this flexibility make it likely for a growing number of real life visualization metaphors to emerge in the near future for the visualization of web spaces.

Maps. A map is a space-filling visualization where relationships between objects are translated to space relationship, i.e. proximity. Just as a network diagram takes advantage *good continuation*, maps are based on the principle that *proximity* between items emphasizes relatedness (Tan & Benbasat, 1993).

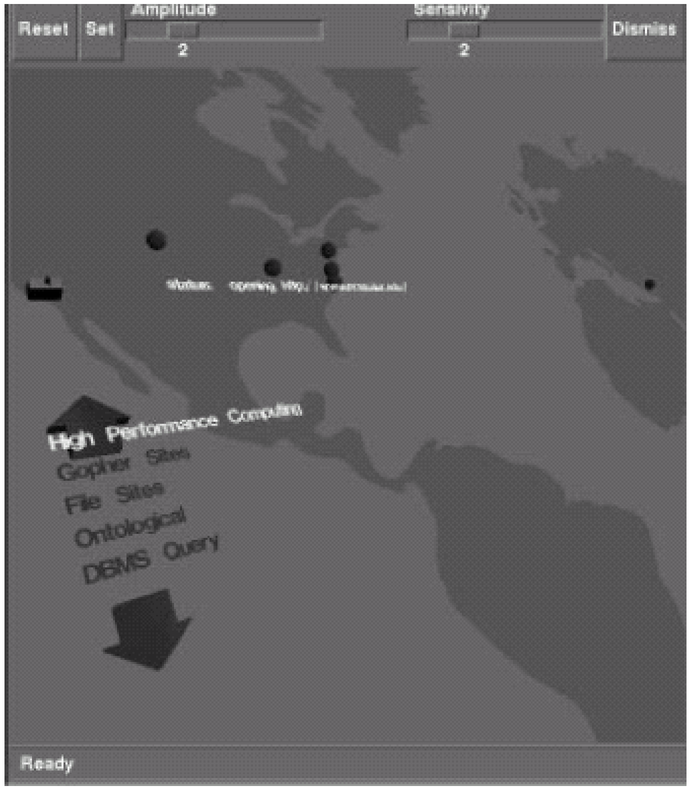

Figure 17. Tecate Browser displaying a web space imposed on the map of the world

The most common maps are geographical where terrains of different geographical regions are displayed along with the distance relationships between these regions. Similarly, maps can be used to present a collection of information objects with their chosen attributes and inter-object relationships. Different regions on the map represent different objects in the collection. The proximity of regions means the underlying objects are similar, while the size of a region is a surrogate for importance of the corresponding object.

This structure is a popular method of presenting web spaces where information objects are web pages or groups of web pages. Due to the ability of a map to represent strength of relationships between objects (nodes in the underlying graph structure), it is a particularly effective way of visualizing network structures as in the self organizing map based system in (Lin et al., 1991), WAVE (Kent & Neuss, 1994), Tecate Web Browser (Kochevar & Wanger, 1995), Cyberspace Geography Visualization (Girardin, 1995), The Open Text Web Index Visualization (Bray, 1996), Aleph/Content View (Neves, 1997), Relevance Map (Assa et al., 1997), Adoptive SOM (Roussinov & Ramsey, 1998), Sentinel-Visualeyes (Fox et al., 1999), CI Spider and Meta Spider (Chau et al., 2001), Lookmark (Breiteneder et al., 2002), and Kartoo (www.kartoo.com).

Figure 17 displays Tecate Web Browser showing a collection of web sites depicted as 3-D icons on a map of the world. The icons are placed on this map based on the geographical location of the web servers the corresponding pages come from. This map presents a single visualization of the overall collection at once therefore is an example of “one visualization only” metaphor of examining visualization details. Another example map of a network structure is Kartoo.com as seen in Figure 18. This search engine presents its results along with themes extracted from the results collection on a map where the relationship between each result and a theme is clearly depicted. The presentation method of Kartoo replaces the “one page at a time” presentation of common search engines with a “one map at a time” metaphor. The overall collection or the relationship between these maps are not displayed therefore this is an example of the “details only” visualization metaphor.

Figure 18. Results to the query "China" from Kartoo.com

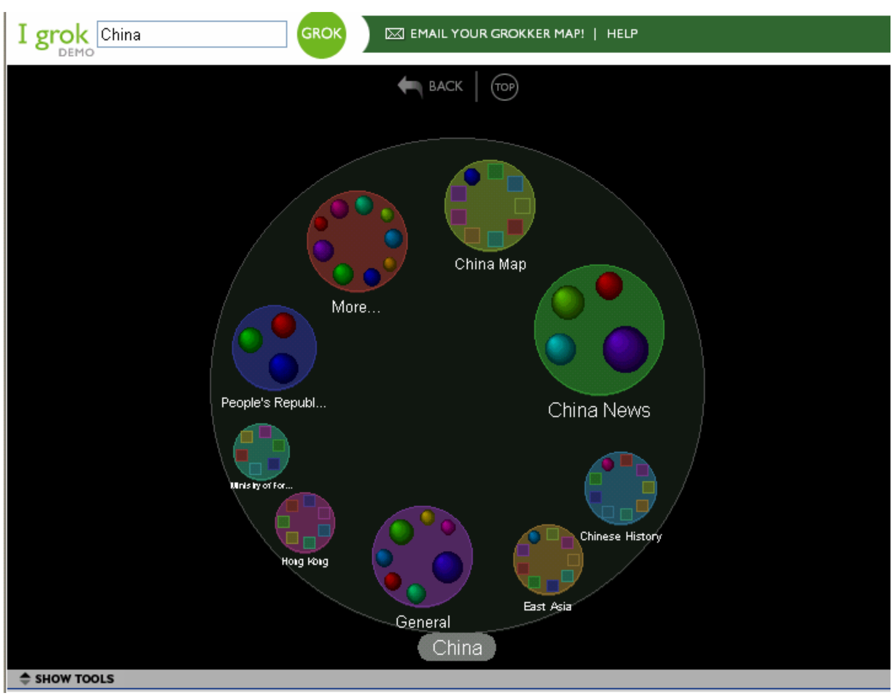

Figure 19. Results to the query "China" from Grokker

Meanwhile, maps are also popular for visualizing tree structures as in MAPA (Durand & Kahn, 1998), FISPA (Turetken & Sharda, 2004), WTMS/scatter plot (Mukherjea, 2000), and HotSauce (Apple) though examples of map visualizations can also be seen for generic graph structures as in Data Mountain (Czerwinski et al., 1999), map.net (http://maps.map.net), and Grokker (http://www.groxis.com).

Figure 19 shows the result set (from Yahoo) of the query "China" visualized by the Grokker interface. We call this overview visualization a map since the size of a region still represents the size of the corresponding group (Yahoo category) although the proximity of regions does not have a specific meaning.

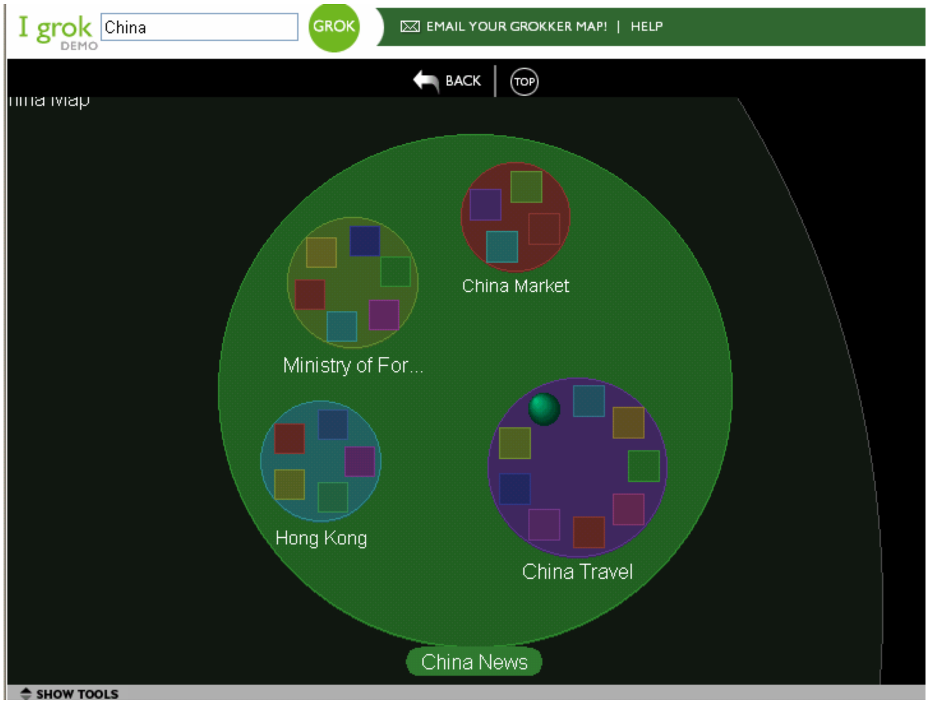

Figure 20. Zooming in on "China News"

In this visualization the documents can belong to multiple categories, and the examination of a category requires zooming into a specific region of the map, which results in a map of that region as seen in Figure 20 for the "China News" region of the map in Figure 19. This is another example of the "details only" zooming paradigm.

Table 1 summarizes the systems discussed in this section by presenting the different attributes (from Figure 4) of each system. In the next section, we discuss our observations from this review to identify future research directions in this field.

System	Source of Web Space	Basis of Web Space Organization	Resulting Data Structure	Visualization Paradigm	
				Overview	Details
Relevance Map (Assa, Cohen-Or and Milo 1997)	Dynamic	Semantic content	Network	2D Map	?
VR-VIBE (Benford, Snowden, Brown, Reynard and Ingram 1997)	Dynamic	Semantic content	Network	3D Network	Details only
The Open Text Web Index Visualization (Bray 1996)	Static	Connectivity	Network	3D Map	One visualization only
Lookmark (Breiteneder, Eidenberger, Fiedler and Raab 2002)	Static	Original bookmark organization	Tree	3D Map	Details and Context
Webquery (VANISH) (Carriere and Kazman 1997)	Dynamic	Combination of content and connectivity	Tree	2D Network	Details only
CI Spider (Chau, Zeng and Chen 2001)	Dynamic	Semantic content	Tree	2D Map	Details and Context
Meta Spider (Chau, Zeng and Chen 2001)	Dynamic	Semantic content	Tree	2D Map	Details and Context

Table 1. Summary of Systems (continued)

Pathfinder networks with generalized similarity analysis (GSA) (Chen 1997)	Static	Could work with connectivity, semantic content, or browsing patterns	Network	2D Network	One visualization only
Data Mountain (Czerwinski et al. 1999)	Static	Combination of content and original bookmark organization	Linear list	3D Map	One visualization only
MAPA (Durand and Kahn 1998)	Static	Connectivity	Tree	3D Map	One visualization only
Hot Sauce (Apple)	Static	Connectivity	Tree	3D Map	One visualization only
webviz/hyperbolic tree (Munzner and Burchard 1995) (Rao et al. 1995) (Lamping and Rao 1996) (Pirolli, Pitkow and Rao 1996a) (Lamping, Rao and Pirolli 1995) (Pirolli, Card and Wege 2001) (also www.inxight.com)	Static	Connectivity	Tree	3D Tree	Details in Context
WebMap (Gaines and Shaw 1995)	Dynamic	Browsing pattern	Tree	2D Tree	One visualization only
WAVE (Kent and Neuss 1994)	Could work with both	Different combinations of connectivity, content, file-system structure, access statistics, originating web information, time of publication, author, and other user-defined attributes	Tree	3D Map	One visualization only
Tecate Web Browser (Kochevar and Wanger 1995)	Static	geographic location of web servers	Network	3D Map	One visualization only
Cybermap (Gloor and Dynes 1998)	Static	Semantic content	Tree	2D Network	Details and Context
CZWeb (Fisher et al. 1997)	Dynamic	Browsing pattern, clusters based on originating web sites, manual organization also facilitated	Tree	2D Map	Details in Context
MosaicG (Ayers and Stasko 1995)	Dynamic	Browsing pattern	Tree	2D Tree	Details in Context
Pad++ (Bederson and Hollan 1994)	Dynamic	Browsing pattern	Graph	2D Map	One visualization only
WebMap (Doemel 1994)	Dynamic	Browsing pattern	Graph	2D Network	Details in Context
Aleph/Travel Map (Neves 1997)	Dynamic	Browsing pattern	Graph	3D Tree	One visualization only
Aleph/Content View (Neves 1997)	Could work with both	Semantic content	Network	3D Map	Details only
Web local search tool (Angelaccio and Buttarazzi 2000)	Dynamic	Connectivity	Tree	2D Tree	One visualization only
Hy+ (Hasan et al. 1995b)	Dynamic	Browsing pattern	Graph	2D Network	One visualization only

Table 1. Summary of Systems (continued)

Hyperspace View (Gershon et al. 1995)	Static	Connectivity	Tree	2D Tree	Details in Context
Cyberspace Geography Visualization (Girardin 1995)	Static	Connectivity	Network	2D Map	One visualization only
VSComp (Liu, Zhao and Yi 2002)	Static	Semantic content	Tree	2D Tree	Details only
3D VRML (Mak, Leong and Chan 2002)	Static	Connectivity	Network	3D Network	One visualization only
2D DOM (Mak, Leong and Chan 2002)	Static	Connectivity	Network	2D Tree	One visualization only
WTMS/Scatterplot View (Mukherjea 2000)	Dynamic	Connectivity	Tree	2D Map	One visualization only
WTMS/Tabular View (Mukherjea 2000)	Dynamic	Directory structure of a web site	Tree	2D Tree	Details only
Adoptive SOM (Roussinov and Ramsey 1998)	Dynamic	Semantic content	Network	2D Map	Details only
Auditorium Seating Visualization (Terveen, Hill and Amento 1999)	Dynamic	Connectivity	Graph	2D Real Life Metaphor	Details in Context
Fettucino (Ben-Shaul et al. 1999)	Dynamic	Connectivity	Graph	2D Tree	One visualization only
MultiSurf/Query Graph (Hasan et al. 1995a)	Dynamic	How query terms are used in successive queries	Graph	2D Network	One visualization only
HDV (Hypermedia Database Visualization) (Owor 2002)	Static for visualizing a web site Dynamic for search results visualization	Connectivity	Graph	3D Real Life Metaphor	Details only
MultiSurf/Query Hit List Graph (Hasan et al. 1995a)	Dynamic	Semantic content	Graph	2D Network	One visualization only
MultiSurf/Web Structure Graph (Hasan et al. 1995a)	Dynamic	Browsing pattern	Graph	2D Network	One visualization only
FISPA (Turetken and Sharda 2004)	Dynamic	Semantic content	Tree	2D Map	Details in context
Site Manager (Silicon Graphics)	Static	Connectivity	Graph	3D Network	Details only
Ptolomaeus (http://www.dia.uniroma3.it/~ptolemy/)	Dynamic	Connectivity	Graph	2D Network	One visualization only
WebBrain (http://www.webbrain.com)	Static	Organization of Netscape Open Directory	Graph	2D Network	Details in context
map.net and map.net (http://maps.map.net)	Static	Organization of Netscape Open Directory	Graph	2D Map	Details only
WebTracer (from NullPointer http://www.nullpointer.co.uk/-/webtracer2.htm)	Static when working with existing maps Dynamic for creation of new map	Connectivity	Graph	3D Network	One visualization only
Nicheworks (Wills 1997)	Static	Connectivity	Graph	2D Network	

Table 1. Summary of Systems (continued)

Touchgraph Google Browser (http://www.touchgraph.com/)	Dynamic	relations identified through Google's "similar to" facility	Graph	2D Network	One visualization only
SENTINEL-VisualEyes (Fox, Frieder, Knepper and snowberg 1999)	Dynamic	Semantic content	Network	3D Map	One visualization only
WebClass (Greenhill and Venkatesh 1999)	Dynamic	Browsing pattern	Graph	2D Network	One visualization only
WebPath (Frecon and Smith 1998)	Dynamic	Browsing pattern	Graph	3D Network	One visualization only
PadPrints (Hightower, Ring, Helfman, Bederson and Hollan 1998)	Dynamic	Browsing pattern	Tree	2D Tree	Details only
NESTOR (http://www.gate.cnrs.fr/~zeiliger/nestor .htm)	Dynamic	Browsing pattern	Graph	2D Network	One visualization only
WWW3D (Benford, Snowden, Brown, Reynard and Ingram 1997)	Dynamic	Browsing pattern	Graph	3D Network	One visualization only
DiskTree (Chi et al. 1998)	Static	Connectivity	Tree	2D Tree	One visualization only
KartOO (www.kartoo.com)	Dynamic	Semantic content	Network	2D Map	Details only
Business Intelligence Explorer/Web Community (Chung, Chen and Jr. 2005)	Dynamic	Both connectivity and semantic content	Tree	2D Tree	Details only
Business Intelligence Explorer/Knowledge Map (Chung, Chen and Jr. 2005)	Dynamic	Both connectivity and semantic content	Tree	2D Network	Details only
WebTOC (Nation, Plaisant, Marchionini and Komlodi 1997)	Static	Connectivity	Tree	2D Real Life Metaphor	Details in context
the system described in (Sakairi 1999)	Static	Connectivity	Graph	2D Tree	Details only
the system described in (Lin, Soergel and Marchionini 1991)	Could work with both	Semantic content	Network	2D Map	Details only

Table 1. Summary of Systems

Future Research Directions and Conclusions

In this paper, the application of information visualization to the web domain is surveyed. Visualization technologies are far from maturation, but are promising for the future of knowledge management especially for web-based knowledge. We believe the classification scheme introduced in this paper will be insightful to designers of web space visualization systems and visualization researchers. Based on our review of the visualizations through our classification scheme we next discuss a number of directions for future research.

One observation from this review is that although web spaces have many attributes that have been identified by previous research as viable bases of organization, current state of the art in web space visualization mainly organizes a web space based on connectivity, semantic content, or browsing patterns. These organizations assume that the end user has a unique predefined purpose for interacting with the web space. For example, visualizations of search results are designed to help the searcher find one or a few relevant pages to the search query. For that reason, they are very similar to content-based visualizations of search results from more structured text databases (e.g. scatter/gather (Cutting et al., 1993; Pirolli et al., 1996b; Rao et al., 1995), butterfly (Mackinley et al., 1995; Rao et al., 1995; Robertson et al., 1993), and envision (Heath et al. 1995; Nowell & Hix, 1993)..

However, it has been proposed that (Shneiderman, 1997) there are fundamentally different tasks an end-user may be pursuing while exploring an information collection such as a web space. Shneiderman (Shneiderman, 1997) identified these tasks as specific fact-finding (searching directly for a readily identifiable outcome), extended fact-finding (searching indirectly for relatively uncertain but replicable outcomes), open-ended browsing (gaining an understanding of a general subject area), and exploration of availability (self explanatory).

According to the "fit" theories (task-technology fit (Zigurs & Buckland, 1991) and cognitive fit (Vessey & Galletta, 1991)), certain technologies would better fit certain tasks. Therefore, one design implication of this is that the basis of web space organization should be flexible in pursuit of meeting the demands of the task that an end-user is pursuing. Similar flexibility would be desirable for the visualization paradigm adopted by a system. Alternatively, a good fit between a system and a task may be achieved through detailed analysis of user tasks and consequent design of system features. The above mentioned list of tasks from Shneiderman is a good start for this; however this list needs to be validated for the web.

What complicates web usability even further is the fact that the web environment is shared by almost every member of the information society. Since there are known differences in people's cognitive styles and information processing patterns, the kind of flexibility we identified above is particularly relevant for the customization of visualizations for different individuals. Most of the systems studied in this paper have their own way of pre-visualization information processing (steps I through III in Figure 4) depending on the visual paradigm they have chosen. However, for a multi-view presentation system to be feasible, the language and expressions used to describe the graph structure should be somewhat standardized so that different visual paradigms are readily applicable. Therefore another research direction would be the identification of features for such standard data structure descriptions. Once this is done, visualizations that are designed to work with already existing standard data structures in other domains such as Perspective Wall (Mackinlay et al., 1991; Rao et al., 1995; Robertson et al., 1993), CHEOPS (Beaudoin et al., 1996), Cone tree (Rao et al., 1995), and Information cube (Rekimoto, 1993) can easily be adopted to the visualization of web spaces. The 3-D FSN-file System Navigator© (from Silicon Graphics) displayed in Figure 21 is such an example. It visualizes a 3D presentation of a hierarchy to support managing large collections of hierarchically structured data such as computer file systems.

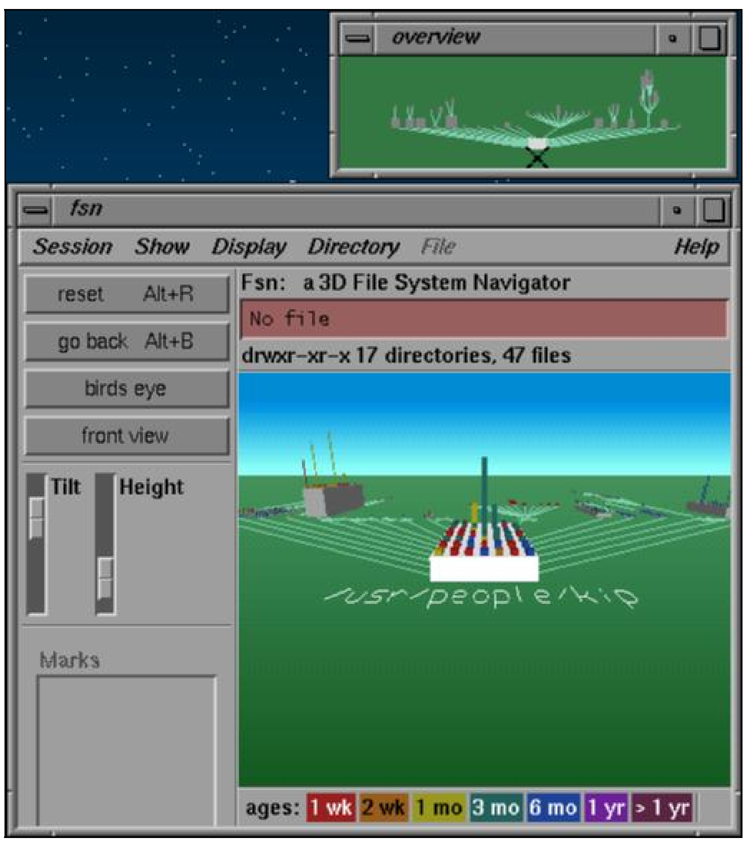

Figure 21. 3-D FSN-file System Navigator©

Our next observation is on the interactivity of web space visualizations. A great majority of the systems surveyed in this paper do not use animations in spite of the technical feasibility of these visual aids. Although conceptual models that lead systems development and testing are relatively abundant for static visualizations, the same cannot be said for highly interactive visualizations. Therefore there is need for theory to specifically address the specifics of highly interactive animated interfaces. On a more technical note, such interactive systems should be studied from an implementation point of view paying close attention to the advantages and disadvantages of server side versus client side technologies for the technical feasibility of basic visualizations and animations.

Meanwhile, very few (e.g. Grokker, Kartoo, WebBrain) of the visualization systems surveyed in this study have gained much commercial success. This is in spite of the fact that many have passed rigorous performance and usability tests. This observation requires an understanding of technology adoption and diffusion of innovation concepts in this domain. What is likely is that certain criteria that support diffusion of these innovations (Rogers, 1995) have not been fulfilled. One of these criteria might be that (most) people do not perceive a "crisis" -- that is, using the web can be *a bit* inconvenient, but most people settle for a use strategy that is adequate instead of finding the best one. It is likely that for most tasks that require information use from the web, such "satisficing" is reasonable, while for others it is not. Therefore an important research question is, "Can the cost of inferior presentations, i.e. the difference in one's performance due to this satisficing be quantified?" and if so "for what kinds of tasks?" This questions again point to the importance of web task analysis, and it is very likely that if system developers cannot do a better job in showing a cost due to ineffective use of web-based information, at least for certain tasks, they may not be able to generate a market for their visualization systems.

To conclude, as observed by Card and colleagues, "While the term 'information visualization' is coming into use, the goal is really 'information perceptualization'. The latter implies a richer use of many senses, including sound and touch, to increase the rate at which people can assimilate and understand information." (Card et al., 1996). The information perceptualization idea is at a very early stage of development, but understanding the very important visualization dimension of information perceptualization is a major step in the right direction. This paper has aimed to provide such understanding in the crucial web domain.

This paper has covered a wide range of web space visualization systems developed over time. The long development/review cycle experienced by this paper has the potential that some systems are no longer accessible at their original sites. But we chose to include them for illustrating the diversity of research approaches.

These systems have contributed to the richness of research in web space visualization, something that commercial search engines are exploring now. We also note that the copyrights of all the figures reproduced from other publications are explicitly recognized as resting with the original copyright holders. We include the figures in this paper to only illustrate the concepts and use them under the fair use approach of the copyright law.

References

Amento, B., Terveen, L., Hill, W., and Hix, D. (2000). "TopicShop: Enhanced Support for Evaluating and Organizing Collections of Web Sites," *UIST '00,* San Diego, CA USA. pp. 201-209.

Angelaccio, M., and Buttarazzi, B. (2000). "A Visualization System for Web Local Search," *International Conference on Information Visualisation (IV2000),* London, England. pp.

Assa, J., Cohen-Or, D., and Milo, T. (1997). "Displaying Data in Multidimensional Relevance Spacewith 2D Visualization Maps," *8th IEEE Visualization '97 Conference*, pp. 127-134.

Ayers, E., and Stasko, J. (1995). "Using Graphic History in Browsing the World Wide Web," *Fourth International World Wide Web Conference,* Boston, MA. pp.

Baldonado, M.Q.W., and Winograd, T. (1997). "SenseMaker: an Information-exploration Interface Supporting the Contextual Evolution of a User's Interests," *Proceedings of the SIGCHI conference on Human factors in computing systems,* Atlanta, Georgia, United States. pp. 11--18.

Beaudoin, L., Parent, M.-A., and Vroomen, L. C. (1996). "Cheops: a Compact Explorer for Complex Hierarchies," *Proceedings of the 7th conference on Visualization*, pp. 87--ff.

Bederson, B. B., and Hollan, J. D. (1994). "Pad++: A Zooming Graphical Interface for Exploring Alternate Interface Physics," *7th annual ACM symposium on User interface software and technology*, pp. 17--26.

Ben-Shaul, I., Herscovici, M., Jacovi, M., Maarek, Y.S., Pelleg, D., Shtalhaim, M., and Soroka, V. (1999). "Adding Support for Dynamic and Focused Search with Fetuccino.," *Computer Networks* Vol.31, No. 11-16,pp. 1653-1665.

Benford, S.D., Snowden, D.N., Brown, C.C., Reynard, G.T., and Ingram, R.J. (1997). "Visualising and Populating the Web: Collaborative Virtual Environments for Browsing, Searching and Inhabiting Webspace," *Computer Networks and ISDN Systems* Vol. 29, No.,pp. 1751-1761.

Botafogo, R.A., Rivlin, E., and Shneiderman, B. (1992). "Structural Analysis of Hypertexts: Identifying Hierarchies and Useful Metrics," *ACM Transactions on Information Systems* Vol.10, No. 2,pp. 142-180.

Brandt, D.S. (1997). "Constructivism: Teaching for Understanding of the Internet," *Communications Of The ACM* Vol.40, No. 10, pp. 112-117.

Brautigam, M. (1996). "Applying Information Visualization Techniques to Web Navigation," Dissertation Proposal. University of California - Santa Cruz, Santa Cruz, CA.

Bray, T. (1996). "Measuring the Web," *Fifth International World Wide Web Conference,* Paris, France. pp. 993-1005.

Breiteneder, C., Eidenberger, H., Fiedler, G., and Raab, M. (2002). "Lookmark: A 2.5D Web Information Visualization System," *Proceedings of Eurasia - ICT,* Tehran, Iran. pp.

Card, S.K., and Mackinlay, J.D. (1997). "The Structure of the Information Visualization Design Space," *Information Visualization '97*, pp.

Card, S.K., Moran, T.P., and A., N. (1986). "The Model Human Processor: An Engineering Model of Human Performance," In *Handbook of Perception and Human Performance.* Ed. J. W. a. Sons. New York, New York, 1-35.

Card, S.K., Robertson, G.G., and York, W. (1996). "The WebBook and the Web Forager: an information workspace for the World-Wide Web," *Conference on Human Factors and Computing Systems,* Vancouver, British Columbia. pp. 111-117.

Carriere, J., and Kazman, R. (1997). "WebQuery: Searching and Visualizing the Web through Connectivity," *WWW6,* Santa Clara CA. pp.

Chang, S.-F., Smith, J.R., Beigi, M., and Benitez, A. (1997). "Visual Information Retrieval from Large Distributed Online Repositories," *COMMUNICATIONS OF THE ACM* Vol.40, No. 12, pp. 63-71.

Chau, M., Zeng, D., and Chen, H. (2001). "Personalized Spiders for Web Search and Analysis," *JCDL'01,* Roanoke, Virginia, USA. pp. 79-86.

Chen, C. (1997). "Structuring and Visualising the WWW by Generalised Similarity Analysis," *Hypertext 97,* Southampton UK. pp.

Chi, E.H., Pitkow, J., Mackinlay, J., Pirolli, P., Gossweiler, R., and Card, S.K. (1998). "Visualizing the Evolution of Web Ecologies," *the SIGCHI conference on Human factors in computing systems,* Los Angeles, CA. pp. 400 - 407.

Chung, W., Chen, H., and Jr., J.F.N. (2005). "A Visual Framework for Knowledge Discovery on the Web: An Empirical Study of Business Intelligence Exploration," *Journal of Management Information Systems* Vol. 21, No. 4, pp. 57-84.

Crapo, A.W., Waisel, L.B., Wallace, W.A., and Willemain, T.R. (2000). "Visualization and The Process of Modeling: A Cognitive-theoretic View," *KDD 2000,* Boston, MA USA. pp. 218-226.

Cutting, D.R., Karger, D.R., Pedersen, J.O., and Tukey, J.W. (1993). "Scatter/gather: a cluster-based approach to browsing large document collections," *1992 Conference on Research and Development in Information Retrieval,* Copenhagen, Denmark. pp. 318 - 329.

Czerwinski, M., Dumais, S., Robertson, G., Dziadosz, S., Tiernan, S., and Dantzich, M.v. (1999). "Visualizing Implicit Queries For Information Management and Retrieval," *CHI '99,* Pittsburgh PA USA. pp. 560-567.

Doemel, P. (1994). "WebMap - A Graphical Hypertext Navigation Tool," *Second International World Wide Web Conference,* Chicago, IL. pp.

Durand, D., and Kahn, P. (1998). "MAPAIM: A System for Inducing and Visualizing Hierarchy in Websites," *HyperText 98,* Pittsburgh PA USA. pp. 66-76.

El-Hamdouchi, A., and Willet, P. (1989). "Comparison of Hierarchic Agglomerative Methods for Document Retrieval," *The Computer Journal* Vol.32, No. 3, pp. 220-227.

Fisher, B., Agelidis, M., Dill, J., Tan, P., Collaud, G., and Jones, C. (1997). "CZWeb: Fish-Eye Views for Visualizing the World-Wide Web.," *HCI International*, pp. 719-722.

Fox, K.L., Frieder, O., Knepper, M.M., and Snowberg, E. J. (1999). "SENTINEL: A Multiple Engine Information Retrieval and Visualization System," *Journal of the American Society for Information Science* Vol.50, No.7, pp. 616-625.

Frecon, E., and Smith, G. (1998). "WebPath - A Three-Dimensional Web History," *IEEE Symposium on Information Visualization*, pp. 3.

Furnas, G. W. (1986). "Generalized Fisheye Views," *CH 1'86 Proceedings*, pp. 16-23.

Furnas, G.W. (1997). "Effective View Navigation," *CHI 97,* Atlanta GA LISA. pp.

Gaines, B., and Shaw, M. (1995). "WebMap: Concept mapping on the Web.," *Fourth International World Wide Web Conference.,* Boston, MA. pp.

Gershon, N., Winstead, J., Levasseur, J., Croll, J., Pernick, A., and Ruh, W. 1995. "Visualizing World Wide Web Information Resources."

Girardin, L. 1995. "Cyberspace Geography Visualization - Mapping the World-Wide Web to help people find their way in cyberspace."

Gloor, P., and Dynes, S. (1998). "Cybermap: Visually Navigating the Web," *J. Visual Lang. Comput.* Vol.9, No. 3,pp. 319–336.

Greenhill, S., and Venkatesh, S. (1999). "Constructing and Navigating Personalised Views of the Web," *Information Processing and Management* Vol.35, No.,pp. 679-689.

Hasan, M., Golovchinsky, G., Noik, E., Charoenkitkarn, N., Chignell, M., Mendelzon, A., and Modjeska, D. (1995a). "Browsing Local and Global Information," *Proc. CASCON'95,* Toronto, Canada. pp. 228-240.

Hasan, M.Z., Golovchinsky, G., Noik, E.G., Charoenkitkam, N., Chignell, M., Mendelzon, A.O., and Modjeska, D. (1995b). "Visual Web surfing with Hy+," *CASCON'95,* Toronto. pp.

Heath, L.S., Dalal, K., Fox, E.A., Hix, D., Nowell, L.T., Wake, W.C., Averboch, G.A., Labow, E., Guyer, S.A., Brueni, D. J., and France, R. K. (1995). "Envision: A User-centered Database of Computer Science Literature," *Communication of the ACM* Vol.38, No. 4,pp. 52 - 53.

Herman, I., Melancon, G., and Marshall, M.S. (2000). "Graph Visualization and Navigation in Information Visualization: A Survey," *IEEE Transactions on Visualization and Computer Graphics* Vol.6, No. 1,pp. 24-43.

Hightower, R., Ring, L., Helfman, J., Bederson, B., and Hollan, J. (1998). "Graphical multiscale web histories: A study of PadPrints," *Hypertext '98*, pp. 58-65.

Jacoby, J. (1984). "Perspectives on Information Overload," *Journal of Consumer Research*, No. 10,pp. 432-435.

Keller, K. L., and Staelin, R. (1987). "Effects of Quality and Quantity of Information on Decision Effectiveness," *Journal of Consumer Research*, No. 14,pp. 200-213.

Kent, R.E., and Neuss, C. (1994). "Creating a Web Analysis and Visualization Environment," *Second International WWW Conference*, pp.

Kochevar, P., and Wanger, L. 1995. "Prototype System Provides Three-dimensional Interface to the World Wide Web."

Lamping, J., and Rao, R. (1996). "Visualizing Large Trees Using the Hyperbolic Browser," *CHI '96*, pp. 388-389.

Lamping, J., Rao, R., and Pirolli, P. (1995). "A Focus+Context Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies," *CHI '95*, pp.

Lin, X., Soergel, D., and Marchionini, G. (1991). "A Self-organizing Semantic Map for Information Retrieval," *Proceedings of the 14th annual international ACM SIGIR conference on Research and development in information retrieval table of contents,* Chicago, Illinois. pp. 262-269.

Liu, B., Zhao, K., and Yi, L. (2002). "Visualizing Web Site Comparisons," *WWW 2002, May 7-11, 2002, Honolulu, Hawaii, USA.*, No.

Mackinlay, J. D., Robertson, G. G., and Card, S. K. (1991). "The Perspective Wall: Detail and Context Smoothly Integrated," *Conference on Human Factors in Computing Systems,* New Orleans, Louisiana. pp. 173-176.

Mackinley, J. D., Rao, R., and Card, S. D. (1995). "An Organic User Interface for Searching Citation Links," *CHI'95,* Denver, CO. pp. 67-73.

Mak, J. Y. S., Leong, H.V., and Chan, A.T.S. (2002). "Dynamic Structuring of Web Information for Access Visualization," *SAC 2002, Madrid, Spain.*, No.,pp. 778-784.

Miller, G. A. (1956). "The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information,," *Psychological Review* Vol.63, No. 2,pp. 81-97.

Mukherjea, S. (2000). "Organizing Topic-Specific Web Information," *The Eleventh ACM Conference on Hypertext and Hypermedia*, pp. 133-141.

Mukherjea, S., and Foley, J. D. (1994). "Navigational View Builder: A Tool for Building Navigational Views of Information Spaces," *CHI'94,* Boston, Massachusetts USA. pp. 289-290.

Munzner, T., and Burchard, P. (1995). "Visualizing the Structure of the World Wide Web in 3D Hyperbolic Space," *VRML '95,* San Diego CA USA. pp. 33-39.

Nation, D. A., Plaisant, C., Marchionini, G., and Komlodi, A. (1997). "Visualizing Websites Using a Hierarchical Table of Contents Browser: WebTOC," *3rd Conference on Human Factors and the Web,* Denver, CO. pp.

Neves, F. D. (1997). "The Aleph: A Tool to Spatially Represent User Knowledge about the www Docuvers," *ACM Hypertext '97*, pp.

Nigay, L., Vernier, F., and the, P. p. f. (1998). "Navigational Interaction Techniques in the Search Results Space," *CHI98 workshop on Innovation and Evaluation in Information Exploration Interfaces,* Berkeley, CA., pp.

Noik, E. G. (1993). "Exploring Large Hyperdocuments: Fisheye Views of Nested Networks," *the 5th ACM Conference on HyperText*, pp. 192-205.

Nowell, L.T., and Hix, D. (1993). "Visualizing Search Results: User Interface Development for the Project Envision Database of Computer Science Literature," *Human-Computer Interaction Conference International 1993*, pp. 56-61.

Owor, R. S. (2002). “A Data Model and Architecture for Hypermedia Database
Visualization,” *Web3D’02*, No.,pp. 121-126.

Pirolli, P., Card, S. K., and Wege, M. M. V. D. (2001). “Visual Information Foraging in a Focus + Context Visualization,” *CHI 2001* Vol.3, No. 1,pp. 506-513.

Pirolli, P., Card, S. K., and Wege, M. M. V. D. (2003). “The Effects of Information Scent on Visual Search in the Hyperbolic Tree Browser,” *ACM TransactIons on Computer-Human Interaction* Vol.10, No. 1,pp. 20-52.

Pirolli, P., Pitkow, J., and Rao, R. (1996a). “Silk from a Sow’s Ear: Extracting Usable Structures from the Web,” *Conference on Human Factors and Computing Systems,* Vancouver, British Columbia, Canada. pp. Pages: 118 - 125.

Pirolli, P., Schank , P., Hearst, M., and Diehl, C. (1996b). “Scatter/gather Browsing Communicates the Topic Structure of a Very Large Text Collection,” *Conference on Human Factors in Computing Systems,* Vancouver, British Columbia, Canada. pp. 213-220.

Pitkow, J.E., and Bharat, K.A. (1994). “WebViz: A Tool for World Wide Web Access Log Analysis,” *the First International World Wide Web Conference,,* Geneva, Switzerland. pp.

Rao, R., Pedersen, J. O., Hearst, M. A., Mackinlay, J. D., Card, S. K., Masinter, L., Halvorsen, P.-K., and Robertson, G.G. (1995). “Rich Interaction in the Digital Library,” *Communication of the ACM* Vol.38, No. 4,pp. 29-39.

Rekimoto, J., Green, M. (1993). “The Information Cube: Using Transparency in 3d Information Visualization,” *Proc. 3rd Annual Workshop on Information Technologies & Systems (WITS ‘93)*, pp. 125-132.

Rivlin, E., Botafogo, R., and Shneiderman, B. (1994). “Navigating in Hyperspace: Designing a Structure-based Toolbox,” *Communication of the ACM* Vol.37, No. 2,pp. 87-96.

Robertson, G.G., Card, S.K., and Mackinlay, J.D. (1993). “Information Visualization Using 3D Interactive Animation,” *COMMUNICATIONS OF THE ACM* Vol.36, No. 4,pp. 57-71.

Rogers, E., M. (1995). *Diffusion of Innovations, (Fourth ed.) The Free Press, New York, 1995.*, New York, New York: The Free Press.

Roussinov, D., and Ramsey, M. (1998). “Information Forage through Adaptive Visualization,” *Digital Libraries 98,* Pittsburgh PA USA. pp. 303-304.

Sakairi, T. (1999). “A Site Map for Visualizing Both a Web Site’s Structure and Keywords,” *Proceedings of the IEEE Conference on System, Man, and Cybernetics*, pp.

Salton, G. (1989). *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*, Reading, MA: Addison-Wesley.

Shipman, F. M., Marshall, C. C., and LeMere, M. (1999). “Beyond Location:Hypertext Workspaces and Non-Linear Views,” *Hypertex 99,* Darmstadt Germany. pp. 121-130.

Shneiderman, B. (1994). “Dynamic Queries for Visual Information Seeing,” *IEEE Software* Vol.11, No. 6,pp. 70.

Shneiderman, B. (1996). “The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations,” *IEEE Visual Languages,* Boulder, CO. pp. 336-343.

Shneiderman, B. (1997). *Designing The User Interface, Strategies for Effective Human-Computer Interaction*, Reading, MA: Addison Wesley.

Tan, J. K. H., and Benbasat, I. (1993). “The Effectiveness of Graphical Presentation for Information Extraction: A Cumulative Experimental Approach,” *Decision Sciences* Vol.24, No. 1,pp. 167-191.

Terveen, L., Hill, W., and Amento, B. (1999). “Constructing, Organizing, and Visualizing Collections of Topically Related Web Resources,” *ACM TransactIons on Computer-Human Interaction* Vol.Vol. 6, No. 1, No.,pp. 67–94.

Turetken, O., and Sharda, R. (2004). “Development Of A Fisheye-Based Information Search Processing Aid (FISPA) For Managing Information Overload In The Web Environment,” *Decision Support Systems* Vol.37, No. 3,pp. 415-434.

Vessey, I., and Galletta, D.F. (1991). “Cognitive Fit: An Empirical Study of Information Acquisition,,” *Information Systems Research* Vol.2, No. 1,pp. 63-84.

Wakita, A., and Matsumoto, F. (2003). “Information Visualization with Web3D Spatial Visualization of Human Activity Area and Its Condition,” *Computer Graphics*, No.,pp. 29-33.

Willet, P. (1988). “Recent Trends in Hierarchical Document Clustering: A Critical Review,” *Information Processing and Management* Vol.24, No. 5,pp. 577-597.

Wills, G.J. (1997). “Nicheworks -- Interactive Visualization of Very Large Graphs.,” *Graph Drawing ‘97 Conference*, pp.

Zigurs, I., and Buckland, B. K. (1991). “A Theory of Task/Technology Fit and Group Support Systems Effectiveness,” *MIS Quarterly* Vol.22, No. 3, pp. 313-334.

About the Authors

Ozgur Turetken is Associate Professor at the Ted Rogers School of IT Management at Ryerson University. He holds a BS in EE, an MBA (both from Middle East Technical University – Ankara, Turkey), and a PhD (Oklahoma State University). His general research areas are human computer interaction and decision support. His work has appeared in *Communications of the ACM, Decision Support Systems, IEEE Computer, Information & Management*, and *Information Systems Frontiers.*

Ramesh Sharda is Director of the Institute for Research in Information Systems (IRIS) , and a Regents Professor of Management Science and Information Systems in Spears School of Business at Oklahoma State University. His research has been published in major journals including *Management Science, Information Systems Research, Decision Support Systems, INFORMS Journal on Computing,* and many others. His research interests are in decision support systems, NN applications, and technologies for managing information overload.