Mapping the World-Wide Web

Martin Dodge

Centre for Advanced Spatial Analysis (CASA)

University College London

Gower Street, London, WC1E 6BT

www.casa.ucl.ac.uk | www.cybergeography.org

m.dodge@ucl.ac.uk

14th October 1999

Introduction

Maps of the World-Wide Web will be the focus of this paper. These are maps that usually dispense with many of the conventional notions of real-world cartography, doing away with the framework of latitude and longitude location and the familiar background of countries and cities in favour of arbitrary grids and new visual metaphors. These are maps of the information spaces of Cyberspace that employ graphical techniques and visual metaphors ranging from 2D flat static maps to immersive 3D fly through ones. They are created by academic researchers and companies from a wide variety of disciplines, including computer graphics, information design, user interfaces and human-computer interaction, virtual reality, information retrieval and scientific visualisation to name just a few (see the following for coverage of these, McCormick et al. 1987, Laurel 1990, Tufte 1990, Holtzman 1997, Johnson 1997, Shneiderman 1997, Wurman 1997, Jacobson 1999). In particular, a distinct research field, under the title of information visualisation, has coalesced in the past five years or so, with much of the work focusing on developing new interfaces and interactions for the Internet and the Web (Gershon & Eick 1995, Card et al. 1999, Chen 1999). More recently information visualisation research has been turned into products via commercial ventures such as Visual Insights, Perspecta, Inxight Software and Cartia[1]. By way of a definition, three of the leading practitioners in the field state that:

> "*…information visualization focuses on information, which is often abstract. In many cases information is not automatically mapped to the physical world (e.g. geographical space). This fundamental difference means that many interesting classes of information have no natural and obvious physical representation. A Key research problem is to discover new visual metaphors for representing information and to understand what analytical tasks they support.*" (Gershon et al. 1998:10).

Many of the maps I shall examine, particularly the three-dimensional ones, do not even look like maps in any conventional cartographic sense, although most still serve the fundamental purpose of visual explanation to aid navigation and understanding of a space. But perhaps some can not even make this claim, they are just pictures of the structure of the Web, pictures of Cyberspace, striking and powerful images that give people a unique sense of the space, in similar manner to the Apollo images gave people a new understanding of the Earth (Cosgrove 1994).

Many of the examples I consider are experimental, work in progress, only providing a fragmentary, imperfect view of the Web, just like the Mappae Mundi's gave of the ancient world. However, they are still worth examining because they having a powerful impact on how people are conceiving the shape and form of the Web. Like the Mappae Mundi, today's maps of the Web provide a visual structure for thinking about the world, a world that is now virtual. I now sample the myriad of maps and interactive spatialisation of the Web.

Info Spaces of the Net

The Web is one of a number of information spaces that can be thought of as running on top of the basic Internet infrastructure. These various information spaces have different structures, using different communications protocols and offer users different modes of navigation and interaction with the space. There are information spaces used principally for communications between people, such as asynchronous email and Usenet and synchronous chat and virtual worlds. In this paper I only consider virtual spaces where users search and interact with inanimate information resource, such as Web page, documents, ftp archives, images, database and archives; and particularly those on the Web or accessible through the Web browser. So I focus on mapping what might be termed the 'info-verse' (Card et al. 1999).

These different spaces support different types of information exchange, degrees of synchronicity and levels of user interaction and, therefore, they require differing measurement of their structures and differing forms of spatialisation to appropriately model their true nature. At a fundamental level, the different information spaces are caused by the different network protocols used by software applications to communicate over the Internet, which give rise to the different form and functions apparent to the end-user. **Figure 1** shows a sketch map produced by John December, a leading Internet consultant and writer, showing the principle information spaces and some of the connections between them, as of the end of 1994 (December 1995). The map provides a good way of conceptualising the different information spaces of the Internet as distinct and self-contained domains, but with fluid, complex boundaries and many interconnections and overlaps between them. The map was drawn at the end of 1994 and the nature of the Internet has changed markedly since then, with certain information spaces all but disappearing as they fall out of favour with users (WAIS and Gopher) and the inexorable and exponential growth of Web space. For many end-users the Web, as seen through the browser interface, is the key information space of the Internet today, although email is still the most widely used information space (Clemente 1998). Other important information spaces within the global Internet that have evolved and grown since December drew this map include multi-user chat environments, virtual worlds and instant messaging. Also, the rise of large private networks and Intranets are creating significant, private, information spaces, but they are largely unseen from the outside and so are difficult to quantify and map.

The rise of the Web, since the launch of the graphical browser in 1993, has meant this has become the most powerful space and also a interface to many other information spaces of the Internet. For example, many people access and browse ftp space via the browser, often without even realising they have left the Web and are using a different protocol; one can also use the Web interface for real-time chat, to read and sending email and participate in Usenet news.

Most of the spatialisations of the infoverse are designed to improve navigation through the information space and enable people to find what they are looking for more easily. It is important to distinguish between *browsing* and *searching* for information in a large information space like the Web. They are very different activities which require differing support tools. Browsing is largely an explorative activity, usually with no planning or specific goals, with useful results dependent on serendipity (Marchionini 1997). At present the Web supports two major forms of browsing - link-following and directories. Browsing by link-following uses the fundamental Web function of hyperlinks connecting pages that can be explored using the standard browser application. However, browsing hyperlinks between pages can often be frustrating and unproductive, as it is all too easy to get lost in the complex topologies of links as there is a lack of navigational cues indicating where you started from, where you are at present or where you can go onto. This situation has been termed "lost in hyperspace" (Edwards & Hardman 1989, Cockburn & Jones 1996, Brake 1997). Users waste much time wandering through Web sites without finding anything of interest or gaining any useful insight, "a situation known as the art museum phenomenon" (Chen et al. 1998:584). After a while wandering lost through the Web, users are often forced to go back to the entrance point and start again.

The second popular form of browsing the Web is to use directories which group pages into easily "browseable" categories, often organised in a hierarchical fashion. The categories and the classification of pages is usually done using human judgement. The most well known directory on the Web is Yahoo!, one of the key landmarks of the Web and the most visited site, attracting many millions of visitors a week[2]. There are lots of other directories, many of which focus on particular subject areas, being run by commercial enterprises, academics or the efforts of dedicated individuals. Directories have proved to be a useful, convenient and popular way of organising Web space into a structure that most people find familiar for browsing. However, there are problems with human-constructed directories relating to the granularity of their categories and maintaining currency in the rapidly growing Web (Chen et al. 1998). Perhaps the most important issue is how meaningful are the categories selected by the directory creators and do they match the way the browsing user has conceptualised the structure of that particular information domain.

Maps can ease the navigation burdens by showing the structure of the information space for browsing and also by providing graphical maps of directories and content.

Flat Maps of the Web

The first group of information maps I shall consider are planar, two dimensional maps. They are in some senses flat maps, like looking down on the space from high above, giving an overview of the extent of the information landscape laid out below. The maps use either, arc-node graphs or continuous "land-use" metaphors to provide users with a spatialisation of the content and structure of information, such as a collection of interlinked Web pages. The first two examples maps, ET-Map and NewsMaps, spatialise the contents of Web pages as land-use maps using the thematic and terrain style from conventional cartography.

Hsinchun Chen and the research team in the University of Arizona's Artificial Intelligence (AI) lab are developing tools to aid browsing and searching of information space, with the Web providing a particularly topical and challenging test-beds (Chen et al 1998). The problems of finding relevant information in a timely fashion using the conventional approaches of keyword searching and Web browsing are proving to be real impediments to gaining the maximum benefits from the information resources in cyberspace (Bowman et al. 1994, Brake 1997, Lawrence & Giles 1998, 1999). These problems are being actively researched by the AI lab and one of their most interesting and potentially useful approaches are *category maps* which are visual directories for browsing. (Chen et al. 1996). Their category map prototype called *ET-Map*[3]*,* maps over 110,000 entertainment related web pages listed by the Yahoo! Directory (Chen et al. 1998). It is a hierarchical information map using 2D thematic metaphor to spatialise different categories of a directory. **Figure 2** shows three-layers of maps as one browses for information on jazz music.

Category mapping techniques, like ET-Map, are an innovative graphical aid to browsing activities which may help overcome some of the limits with browsing and hierarchical directories (Chen et al. 1998). Category maps display groupings of associated Web pages as regularly shaped, homogeneous “subject regions”, which can be thought of as virtual 'fields' which all contain the same type of information 'crop'. So the information space is partitioned into meaningful plots of “land” which are designed to be of a manageable size for humans to comprehend and browse.

ET-Map can be interactively browsed, explored and queried, using familiar point-and-click navigation of the Web, to find information resources of interest. Importantly, the map provides the “big picture”, an overview of the whole information space, which users find very useful (Chen et al. 1998). The ability to see the whole expanse of the information space, on a single page, is powerful advantage of this kind of mapping technique over hierarchical text directories.

ET-Map uses two fundamental spatial concepts to represent the structure of this portion of Web space. Firstly, the spatial area of the “subject regions” is directly related to the number of Web pages on that subject. So bigger subjects get larger plots of virtual 'land' on the map. For example the “MUSIC” subject area (bottom right in **figure 2(a)** contains over 11,000 pages and so has a much larger plot on the map than the neighbouring area of “LIVE” which only has 4,300 odd links. (The size of each region is given on the map by the number in brackets). This is intuitive and meaningful, allowing users to instantly gauge which subject areas contain the most Web pages. Clicking on a subject region with less than two hundred pages takes one to a conventional text listing of the page titles from which one can select a page to download. If a region has more than two hundred pages, then a sub-map is created, at a greater resolution, with a finer degree of categorisation. This is repeated and gives rise to a multi-layer map. For example clicking on the “MUSIC” subject region we are presented with a much more detailed category map of music-related Web pages **(figure 2(b))**. In turn, clicking on the "JAZZ" subject region in this map will reveal an even more detailed map for jazz related topics **(figure 2(c))**. In this way Web space is can be visually partitioned into manageable subject areas on a multi-layered map with increasingly finer grained categories.

The second spatial concept is that of neighbourhood, whereby related subject areas in the information space are mapped in close spatial proximity. For example “FILM” and “YEAR’S OSCARS” at the bottom left of **figure 2(a)**. This spatial adjacency of related categories into neighbourhoods matches our common sense notion of the real world in that things that are closer together in geographic space are more likely to be similar than things that are further apart (defined by the geographer, Waldo Tobler, as the "first law of geography", (Tobler 1970)). For example, your home is more likely to be similar in design to those on your street than to houses on a street five miles away. ET-Map also adopts several cartographic conventions. Firstly, the category map is laid out on a regular grid, with definite orientation. Each “subject region” is labelled with a single descriptive term chosen to represent that category. The regions are also shaded with different colour, but interestingly this is not used to convey additional information, it is simply used to aid visual distinction between plots. This differs from real-world thematic maps, such as census or vegetation maps, which use the shading or hatching of the regions to represent data values.

ET-Map was created using a sophisticated AI technique called Kohonen self-organizing map (SOM), which is a neural network approach that has been used for automatic (i.e. no human supervision) analysis and classification of semantic content of text documents like Web pages (Kohonen 1995). Chen et al. believe "*… that Kohonen SOM-based technique (of which our ET-Map is a small working prototype) can be used effectively and scaleably to browse a large information space such as the Internet.*" (Chen et al. 1998:592). However, it is also a challenge to automatically classify pages from a very heterogeneous collection of Web pages and it is not clear that the SOM categories will necessarily match the conceptions of the typical user. From limited usability studies on category maps, like ET-Map, it appears they are good for unstructured, "window shopping" type of browsing, and is less useful for more directed searching.

The next example of a "flat" information map is NewsMaps[4] which has some commonalties with ET-Map, but uses a much more realistic terrain representation. **Figure 3** shows a screen-shot of a typical NewsMaps mapping over nine hundred online news reports of the Kosovan crisis from mid June 1999. It is clearly apparent that the map looks very much like a topographic map of a real landscape, except this is just a map of the content of web news reports aggregated from many sources such as the Washington Post and Fox News. NewsMaps uses a high-powered, sophisticated information mapping system called ThemeScape developed by Cartia, Inc[5].

The visual metaphor of shaded terrain maps, from topographic cartography, uses colour and density of shading, along with contour lines, to give the reader the visual impression of different elevations of the landscape. The colour scheme is conventional for representing elevation, running from low altitude blue and greens, to browns and finally to white for the highest peaks. So despite NewsMaps being flat, planar maps one has the sense of an undulating landscape of hills, snow capped mountain peaks, valleys and plains. NewsMaps is like the view of Web floating high above it; indeed one of Cartia's marketing slogans is that their maps can show "*information from 30,000 feet*". The NewsMaps maps are the result of considerable, off-line, data-crunching that analyses the actual content of the news reports, using proprietary algorithms and techniques, that intelligently summarises the key topics and the relations between them. This is then spatialised visually on the map. Topics which are popular and have many articles about them combine together and kind of "pile-up" to form hills and mountains. The higher the elevation in the terrain, the more articles there are on a given topic. Peaks are indicated with labels that best describe the dominant theme at that point (e.g. paratroopers, Yeltsin & Airport). Using the same neighbourhood idea as ET-Map, similar topics are drawn close together on the map. So the closer together two hills are the more related are the topics. The valleys between hills represent divisions and transitions between topics, where more isolated articles are likely to be found. The actual location of the news articles that form the map is indicated by small black dots.

The maps are not just static pictures, NewsMaps provides an interactive system for users to browse and explore them. Users can interactively browse the map by passing the mouse cursor over areas of interest and then the top five topics within a small radius are displayed in a pop-up window. Clicking once on the terrain will be cause a pop-up list of available articles in the area to be displayed. Clicking on an article title of interest causes the full article to be opened in a new browser window. In this manner it is easy to identify and scan news articles that are of interest from the terrain map. Users can also do a key word search for articles of interest or select articles from a topic list. The results of which are shown prominently by large blue dots on the map, numbered according to their relevancy ranking. It is also possible to stick small red marker flags into the terrain to identify documents of interests for the future reference.

One can also zoom in and pan around the map to reveal more detail. Zooming-in is done by double clicking on an area of interest which progressively reveals finer grain information structures. A number of different NewsMaps are available including ones for US, world and technology news. NewsMaps is powered by ThemeScape map viewer, part of an expensive "enterprise information mapping" system developed by Cartia, Inc. and aimed squarely at the high-end corporate intelligence market, rather than typical Web users. NewsMaps is in some senses just a marketing exercise to demonstrate its technology to the wider public. ThemeScape technologies are the outcome of several years of research at the Pacific Northwest National Laboratories, in Washington State, USA, to develop visual analysis tools for US intelligence agencies to enable analysts to cope with growing "information overload" from the welter of textual information generated today. From the very beginning they used spatialization techniques and their first prototype using a landscape metaphor was called SPIRE (Spatial Paradigm for Information Retrieval) (Wise et al. 1995). How the ThemeScape system turns documents like web pages into nice browseable maps is not made explicit as the techniques are commercial secrets. However, the end results, as made publicly available at NewsMaps.com, are one of the most useful, usable and engaging maps of junks of cyberspace currently created. The maps and the browsing system is well designed, the metaphor is intuitive and aesthetically pleasing. In many ways NewsMaps is the most literal map of an information space as it borrows so much of the "look and feel" from real world cartography.

There are many other fascinating examples using variations on the flat map metaphors, of ET-Map and NewsMaps, however I do not have time to consider them in any detail. **Figure 4** shows five interesting ones. Firstly, there is category map called Visual SiteMap[6] developed by Xia Lin, school of information and library science, University of Kentucky (Lin 1992, 1997). **Figure 4(a)** shows an example of SiteMap visualising space science related Web pages and has many similarities to ET-Map. The second example, **figure 4(b)**, is a more stylised and abstract terrain map of a region of the Web developed by Luc Girardin, at The Graduate Institute of International Studies, Switzerland, *called "cyberspace geography visualization*"[7]. It is another SOM-based analysis of Web content. Girardin succinctly summarises the aim of his work, which could be applied to many other information visualisation projects, "*These maps provide the equivalent of a bird's-eye view of the world-wide web landscape, such that users can have available a global overview of different areas of interest.*" (Girardin 1995).

Websom maps[8] (**figure 4(c)**) look similar to Girardin's work as they are both based on the same underlying technique of SOM. However, rather than mapping the Web, they map thousands of articles posted on Usenet newsgroups. The WEBSOM maps are multi-layered and can be browsed on the Web. They are being developed by researchers at the Neural Networks Research Centre, Helsinki University of Technology in Finland (Lagus et al. 1996, Honkela et al. 1998). The particular example in **figure 4(c)** is the top-level map of over 12,000 articles posted to the news group comp.ai.neural-nets from June 1995 to March 1997. The next example is a commercial information map called "Map of the Market" developed by SmartMoney.com[9]. It is a web-based interactive map of market performance of the stocks of major US corporations using a land-use spatialisation (**figure 4(d)**). Different companies are represented by different sized plots of virtual land, scaled by their market capitalisation. The colour of the plot is based on recent movements in the company's stock price (red for falls, green for increases) and companies in the same sector (technology, health care, etc) are grouped together into neighbours on the map. The final example is a prototype spatialisation of a digital library[10] (**figure 4(e)**) developed by Sara Fabrikant as part of her graduate research in the Department of Geography, University of Colorado at Boulder (Fabrikant 1999, 2000).

Mapping Individual Web Sites

Another popular theme are maps of individual Web sites designed to provide a visual overview of the key parts of the site to aid users in finding particular information (see Morville 1996, Gloor 1997, Cockburn & Greenberg 1999, Kahn 1999, for research and examples in Web site mapping). A diverse range of visual metaphors and degrees of user interaction are employed to map the structure of a Web site. Although, the quality and usefulness of the maps as navigation aids is very variable, Morville comments:

> "*Since we humans have used maps to navigate our way across oceans, through cities, and around shopping malls for millennia, mapping our Web sites seems natural. The only problem is that we've had thousands of years to perfect the art of drawing maps for navigating the physical world, and we still have a hard time finding the restrooms in our local mall. Given that cyberspace cartography is such*

a new field, it should surprise no one that the Web is being peppered with poorly designed maps in inappropriate places." (Morville 1996)

Many examples are one-off, hand-crafted static maps. There are also several more interactive systems. In addition to site maps for end-users there are also mapping applications aimed at webmasters, the people who build and maintain web sites, this are much more comprehensive in their ability to crawl a site and map the structure in detail.

The simplest Web site maps are those that provide a graphical table of contents of the key pages and sections, often in some form of hierarchical structure. The map is "clickable" in that users simply click on the part of the map that interests them and then the corresponding page is loaded. A wide variety of designs are employed and **figure 5(a)** shows a typical example taken from the Hilton Hotels web site[11] in September 1999. It is simple and functional, a visual "table of contents" providing clear and quick access to key sections of the site, particularly related to destinations and reservations which will be of interest to most visitors to the site.

Another example of a hand-crafted map, which is much more visual and much less like a book "table of contents", is map from the Yell web site[12] (the UK yellow pages) which uses the graphical conventions from the famous London Underground map (Garland 1993). **Figure 5(b)** shows the map which uses four different colour subway lines to represent different sections of the site, and the standard symbol for an inter-change station to represent individual pages. The smoothly curving lines, the stylised river Thames and bold colour scheme are "borrowed" exactly from the Tube map style. The result is a fun and engaging static site map.

A much more interactive and potentially powerful Web site map, called Site Lens, has been developed by Inxight Software[13]. **Figure 6** shows an example of a Site Lens map of the Monterey Bay Aquarium Web site, with hierarchy of the site represented as a graph and the pages are shown by the labelled squares. The graph is displayed using a fish-eye technique which distorts space so elements at the centre of the map are larger than those at the periphery (Sarkar & Brown 1994).

The user is able to grab page rectangle object (by clicking and holding with the cursor) and drag it to the centre of the map and it will be smoothly enlarged, while other elements are pushed aside and shrink. In this manner, the user can explore a large hierarchical graph with local detail in the centre of the map along with a view of the overall context. It is much like an interactive photograph of the real-world taken with fish-eye lens would be like. The Site Lens mapping system is the commercialisation of research into information visualisation by scientists at Xerox PARC (Lamping & Rao 1995). Doubling clicking on a page rectangle on the map will load this page into the main browser window. A small panel of controls at the bottom left of the map allow users to change the size of the text labels for pages, and shrink/stretch of the graph links.

Similar fish-eye techniques are used in another interesting interactive Web site mapping tool called Mapuccino from IBM[14]. It offers various graph styles, such as "wheel-spoke" and horizontal/vertical trees to visualise the structure of a chosen Web site. It is being developed by researchers at IBM's Haifa Research Lab in Israel (Maarek et al. 1997).

There are a number of more specialised Web site mapping applications that are aimed at those who management large sites rather than end-users. They provide tools for detailed graph based visualisation of the structure of pages and hyperlinks. They are not meant as end-user navigational aids. **Figure 7** shows a screen-shot of one particular example, Astra SiteManager[15], from Mercury Interactive, mapping the Web site for Centre for Advanced Spatial Analysis. There are several other products that compete in this market such as CLEARweb, Site Analyst, Visual Web, WebAnalyzer[16].

Astra SiteManager is a stand-alone application that can crawl a complete Web site and show the detailed structure of individual pages, links and images as a large spidery graph and as a detailed list. For large graphs that can not be displayed fully on the screen, a small overview map is also available for panning. SiteManager identifies problems like broken links on the map using colour coding and enables the Webmasters to quickly fix them. It can produce summary reports of the sites structure and performance. One powerful function of Astra SiteManager is its ability to over lay traffic patterns on the structure graph, showing how visitors travel through a site and where the "hot spots" of interest are. Traffic is indicated by colour-coded arrows on the links and nodes of the graph using data from the Web server logs.

Being able to show how a site is being used as well as just its structure is a potentially very interesting aspect of mapping Web sites. It can be instructive for those who design, build and manage Web sites to see what people do within their site, where they start, where they go, how many pages they look at, how deep into the hierarchy do they venture, where do they leave? In larger terms, the ability to combine information structure and user activity in a single map creates a deeper understanding of the geography of cyberspace. One research project that attempts to do this combination for Web browsers is called Footprints[17], developed by Alan Wexelblat as a graduate student at the MIT Media Lab. Web browsing from the point of view of the user is a solitary activity at present, you read pages and click on hyperlinks unaware of other people who are also looking at the same page, and unaware of the many users who passed this way before you. This kind of knowledge could improve peoples navigation, with current users benefiting from effort of those who have explored the space before you. This changes solitary surfing into a more collaborative, social activity like walking around a busy city centre surrounded by others. The Footprints system records the individual traces of users' travels through a Web site, and makes the aggregate patterns visual to other visitors on a map. This concept, in the real-world, is like a well-worn footpaths across a park which shows the routes people follow, providing a visual guide to the shortest path to interesting destinations. Similar work mapping the pathways people take through the Web is being done with the Recer system[18] being developed by Matthew Chalmers at the Department of Computer Science, Glasgow University (Chalmers et al. 1998). Footprints and Recer fit into a broader concept of social navigation, which is an active area of research in navigation of information spaces (see Munro et al. 1999 for a review of current research).

Wexelblat's Footprints system shows the popularity of pathways through a site from the aggregate traces in an interactive visualisation system. Footprints employs "*… a metaphor of navigation - maps, paths and signposts - familiar from the physical world that we have implemented in the digital realm.*" (Wexelblat & Maes 1999). A key component is a fish-eye site map showing traffic through a site, an example of which is shown in **figure 8**. This map may look quite disjointed as it only shows pages and links that have been visited and recorded in the Footprints system. Traffic is aggregated to anonymise it, so it is not possible to identify the activity of any one individual. The traffic is represented graphically by colour-coding the page symbols in the map according to shade of red. The most popular and well-visited pages are darkest red. Taking the Footprints idea further there are browser add-on tools that provide graphical histories of your activities across many Web sites on maps, what has been termed "surf maps". These are like cartographic diaries that can be used to find sites you have visited in the past. I look at 3D examples of surf maps later in this paper.

Landscapes Views of the Web

There are range of spatialisation of the Web that leave the constraining representations of flat, 2D maps and graphs. Instead they use visual metaphors of landscapes to extend the visualisation of information spaces in a higher dimension. This gives one more 'space' to display information, but can often be harder for people to interpret and use. The visualisations are also interactive and many of them run with virtual reality type environments whereby the user can fly over and into the map itself.

The first example is the MAPA system developed by David Durand, Paul Kahn and colleagues at Dynamic Diagrams[19] (Durand & Kahn 1998). It is designed as a mapping tool to improve end-user navigation of large Web sites via an easy to use, interactive, pop-up map to show local structure of pages and options available. **Figure 9** shows a screen-shot of the MAPA Web site map from the IBM site. MAPA is a simple information landscape that stresses the hierarchical structure of a Web site. The visual metaphor looks much like note cards in card index. Dynamic Diagrams call their graphical style the Z-Form diagram, and each vertical card represents one Web page. They use Z-Form map widely in their hand-crafted diagrams for planning the structure of new Web sites or redesigning existing for clients, with thumb-nails images of the actual pages pasted onto the front of cards (see Wurman 1997 for examples)

MAPA can be invoked from any place in the Web site from a standard link on the page, this pop-ups the map window, such as **figure 9**, providing a local map showing your location in relation to the rest of the site. Durand and Kahn believe MAPA is "*Like maps commonly found in public places such as malls, hotels or museums, MAPA aims to answer the question "where am I?'.*" (Durand & Kahn 1998). MAPA uses just simple card projecting vertically from an orthogonal ground plain. The cards are spatially arranged on the ground to reflect the dominant hierarchical structure of the site without cluttering the screen with multiple hyperlinks between pages. Your location is shown at the bottom-left by the card standing on a small red carpet. Above and to the right of your location are further pages in the site hierarchy. Cards are ordered towards the top-right corner of the map, and each layer represents a step deeper into the site. All child pages from your location are shown on the map and they are equally spaces along a horizontal diagonal line so they are spread out without overlap. Further pages below these pages are spaced behind in parallel rows. The number of rows gives instant and intuitive visual indication of the depth of the Web site from your location. Even though some of these cards do overlap to varying degrees, at least a part of them are still visible to the user and can be queried using the cursor. Simple colour coding is also used to further identify hierarchical structure, so immediate child pages are shown in green and further grandchild pages are blue. Below the your location card, toward the bottom-left, ancestor pages are shown by orange coloured cards which delineate the steps back to the site's home page (note, none are shown in **figure 9** as we are at the home page, so there are no ancestor pages). The cards with dark bars across the top are deemed to be most significant one in the site for navigation being section identifiers..

Users can interact with the MAPA map in several ways. Firstly, passing the cursor over a card will cause it to be highlighted and for titles to be for that pages and all its subsequent pages in the hierarchy. This can be seen in **figure 9** where the "IBM Global Product Support Services" page is selected. Titles are displayed in a such a manner that they are all visible with no overlapping. Users can also navigate with in the map, to re-focus the layout centred around a different page. This is achieved by clicking on the cards with a dark bar on them. Finally, MAPA can be used to move the user to a different page in the site by doubling clicking on a card of interest in the map. The MAPA system provides a simple and effective interactive Web site map using an information landscape approach. Unfortunately, it has not been widely implemented and remains an interesting prototype.

I now consider two further examples that extend the information landscape metaphor to a much greater degree by creating immersive spaces that user can 'get inside'. In many respects these look most like the popular imagination of a Gibsonian datascape, which disembodied users effortless glide over, exploring and selecting information. The examples are the *Harmony Information Landscape* and *VR-VIBE* and screenshots of each are shown in **figure 10**. The first example is part of the Harmony browser, the client used to accessing an Internet hypermedia system called Hyperwave[20] (Maurer 1996). The Harmony browser provides an integrated 2D graph map and 3D landscape view of the structure of the hypermedia information space in addition to the conventional page view of the documents. I am concerned with the nature of 3D information landscape component, a typical screen-shot of which is shown in **figure 10(a)**. It was developed by Keith Andrews and colleagues at the Graz University of Technology in Austria (Andrews 1995, Andrews et al. 1996).

In the Harmony Information Landscape users get a three dimensional means of browsing an information space, with resources (such as documents, files, images etc.) represented by blocks and icons laid out across an infinite flat plain. Collections of resources (the Hyperwave equivalent of a Web site) are represented by flat slab blocks onto which the actual resources are placed as iconic glyphs, such as a book to represent a text file and an old fashioned movie camera to identify a video clip. The spatial arrangement of the blocks encodes the hierarchical structure of the Hyperwave information space. The user is able to fly over the landscape and chose objects of interest, which will be display in the conventional browser window. As one browses new collections to are added dynamically to the landscape.

The second example, *VR-VIBE* (**figure 10(b)**) was developed by researchers in the Communications Research Group at the University of Nottingham, lead by Steve Benford. The focus of their research is into collaborative virtual environments (CVE), which provide virtual spaces that can be shared by many users (Benford et al. 1997). VR-VIBE is an application of CVE technologies to information searching and retrieval to create a three dimensional co-operative system that can be shared by several users at once (Benford et al. 1995, Churchill et al. 1997). Matching documents from keyword queries are displayed as simple blocks floating in patterns above an flat landscape covered with a regular yellow grid. Keywords, represented as octahedrons, are positions across the space and then the documents blocks are displayed in relation to strength of attraction to each keyword. So keywords act as virtual magnets pulling documents towards them with differing strengths depending on their significance to the search. This produces a spatial arrangement of blocks that users can explore and see which are the most relevant documents by their position. **Figure 10(b)** shows a screen-shot of a VR-VIBE session where over 1,500 documents are visualised according to five keyword 'magnets'. The size and colour of a document block encodes the relevance score of that document to the overall query. So large, brighter blocks, that are visually the most prominent, are the best matches to the whole query.

Crucially, the data-space of search results is a shared virtual environment in which more than one person can be present. The screen-shot in **figure 10(b)** is a static screenshot, in first-person perspective, from one user's viewpoint. Other users present in the space are represented by simple avatars, which look like sticks with eyes. Although the avatars are very abstract they are sufficient to indicate the user's presence, the particular location and the direction they are looking. This shared aspects raises interesting possibilities for collaborative searching and exploration of a large information space.

Users are able to dynamically interact with the VR-VIBE dataspace in a number of ways. Most importantly they can freely move throughout the space, in any direction, enabling them to view the block structures from any angle and position. So one could fly in close to examine part of the document space in detail and then quickly fly above to get an overall view of the configuration. Also, users can change the parameters of the search query by adding, deleting and moving the keyword representations. So the user can select a keyword glyphs and move it by dragging it, and the spatial arrangement of documents will dynamically adjust. Users can also dynamic change the thresholds of the query using a 3D scrollbar to limit the number of matching documents. They can also add annotation to blocks, in the form of text "flags", as can be observed in **figure 10(b)**. Finally, they can select documents of interest and these will be fetched and displayed, for example in a conventional Web browser.

I now change scale and look at a fascinating example of an information landscape that maps the wider Web landscape. Rather than mapping individual Web pages and hyperlinks, this spatialisation represents whole Web sites by single graphical objects; thus providing a very generalised, and potentially very powerful, overview of the many millions of Web pages and their interweaving hyperlinks. It is a spatialisation that provides a broader view of the Web than the visualisations I have so far considered.

These fascinating, and unconventional, maps of the Web landscape were produced by Tim Bray in 1996 and **figure 11** shows one view (Bray 1996). Bray is a programmer and all-around general hacker of text, hypertext and xml, based in Vancouver[21]. Before he could map the Web, he needed to take some measurements of it, and he used a large search engine index as the easiest source. He calculated key metrics on the structure of the known Web in 1995, comprising a mere eleven million pages from about 90,000 sites, to answer the fundamental questions - how big it is? how wide is it? where is the centre? how interconnected is it? These questions are still very relevant to academic researchers and commercial pundits today (see for example, Woodruff et al. 1996, Lawrence & Giles 1998 and 1999, Huberman et al. 1998, Albert et al. 1999, Clever 1999). Much of the research into understanding Web morphology focuses on the analysis of the human-built hyperlinks structures. The aim being to improve current Web searching tools and develop new searching algorithms to find the elusive nuggets of gold that lie hidden in the ever growing expanse of Web. The importance of hyperlinks rests on the fundamental assumption that they are conscious recommendations people make to Web sites or page that they think are useful or of interest. In this manner hyperlinks play a similar role to citations in scientific literature.

Bray spent time analysing the hyperlink structures of the Web and found interlinking between sites was surprisingly sparse. Most links were local, within a site and only a few key sites acted as super-connectors tying sites together - Yahoo! being the epitome of this, the master weaver of the Web. He derived two intuitive measures of Web site character based on hyperlinks - *visibility* and *luminosity*. Visibility is a measure of incoming hyperlinks, the number of external Web sites that have a link to them. He found a select few highly visible sites with many thousands of incoming links, they are the bright stars of the Web. In 1995, the most visible Web site was that of the University of Illinois, Urbana-Champaign (UIUC), the home of the Mosaic browser. The vast majority of sites had only a few links to them and nearly five percent had no incoming links at all. Measuring the reverse, the outgoing links, determines a site's luminosity. The top few percent of most luminous sites obviously carry a disproportional amount of navigational workload. Yahoo! was the most luminous site in 1995 and probably still is today.

Bray used these statistical characteristics to map the key landmarks of the Web in 1995, highlighting the largest, most visible and connected Web sites. The visual metaphor employed to represent Web sites was radically different from conventional maps, having more in common with architectural models than planar cartography. He used abstract, 3D models which look like a cross between the pieces on a chess board and totem poles, as can be seen in **figure 11**. He termed his models *ziggurats*, the ancient stepped pyramidal temples.

Bray’s web site totems visualised the degree of luminosity and visibility of a site, along with the size of the site and the broad category it was in (government, education, commercial, etc.). The basic graphic properties of the totem, size, shape and colour, were used to encode these four dimensions of a web site. The overall height of the pole represented visibility, the width of the pole is size of the site, in terms of number of pages. The size of the globe atop the pole maps the site's luminosity and finally colour coding was employed for different categories (green for university, blue for commercial, red for government agencies). The totems were also labelled with the domain name of the site for identification. The spatial layout of the totems on the ground plane were based on the strength of the hyperlink ties between them.

Figure 11 shows a field of totem poles mapping the Web sites at the very core of the Web in 1995. Further from this core region there would be many thousands other, but most would be minuscule in relation to the mighty totems at the heart. The totemic models were constructed automatically from the site statistics using VRML. The example in **figure 11** is simply static screen-grabs of the 3D models. Today, Bray's web site totem poles stand as historic monuments to the early days of the Web and provide a fascinating spatialisation of part of Cyberspace.

3D and Beyond

In the final section I examine a range of abstract spatialisations that use three-dimensional spaces and virtual reality style interaction techniques to create visual environments in which the user becomes immersed in Web space. The first two examples use, what might be termed, virtual molecular models to represent arc-node topological structures. These molecular models are 3D models inside virtual reality environments which allows uses to manipulate and view them from any position and angle, flying into and around them. **Figure 12** shows, (a) the *HyperSpace Visualiser* and (b) a *Semantic Constellation*.

The HyperSpace Visualiser was a prototype spatialisation of the local structure of small portions of the Web that worked as the user browsed to provide a local surf map (Wood et al. 1995). The goal of the system was to provide users with information on the how their current location was connected to the neighbouring Web space to help overcome the symptoms of "lost in hyperspace" by showing where they were, and where they could go next. The system was developed by Andrew Wood, Nick Drew, Russell Beale and Bob Hendley, at the School of Computer Science, the University of Birmingham, UK; and it was based on a general purpose information visualisation system called Narcissus (Hendley et al. 1995). Their spatialisation uses a conventional arc-node metaphor, with solid spheres being individual Web pages and the arcs are the hyperlinks between them, all floating gently in a black void (**figure 12(a)**). The size of the sphere is scaled to the number of hyperlinks on the page. The layout of the arc-node model in three-dimensional space is achieved using a self-organising algorithm based on attraction/repulsion behaviour of individual lines and spheres which acting collectively can produce a complex structure. So the Web page spheres are given a repulsive force, which is counteracted by the hyperlinks which attract each other. Starting from a random placement of a chump of interconnected Web pages in the 3D space, through a series of iterative steps, the spheres and arcs push and pull each other until a stable and coherent spatial arrangement is reached. The result is self-organised equilibrium where, *"… unrelated areas that do not have links between them are pushed apart, whilst highly-interrelated work is pulled together and clustered in the same region of space.*" (Wood et al. 1995:22).

Only pages that have been 'discovered' by the user's browsing are displayed in HyperSpace. The pages at the edge of the explored space are just single nodes, with a single arc back to the parent. This gives the edge spheres a pincushion appearance. The system is synchronised to the users Web browser, so details of pages traversed are automatically transferred to HyperSpace. The actual spatialisation is displayed in a separate window on the desktop. Users can interact with the model as well as selecting what is displayed and nodes can be labelled with page titles and urls.

The second example of 3D virtual molecular model spatialisation is the Semantic Constellation created by Chaomei Chen, in the Department of Information Systems and Computing at Brunel University, UK (Chen 1997). **Figure 12(b)** shows a view of Chen's Constellation which is used within a standard desktop VR environment, while the model itself is made from VRML[22]. It spatialises an information space of over one and fifty conference papers from three years worth of online proceedings. As with HyperSpace, an arc-node 3D graph is employed, with spheres representing the individual papers and the arcs connecting them together based on how related their content is. The spheres are colour code by year.

Unlike HyperSpace, the arcs are not explicit, 'hard-coded' hyperlinks connections. Instead, in the Constellation they are based on a computed measure of semantic similarity between the papers. So the papers that discuss the same or related topics will be semantically linked in the spatialisation. The more closely two papers are related in terms of their content, the nearer they are in the semantic space. The semantic linkages are spatially arranged and connected together into, what is known as, a PathFinder network.

Interaction with the Constellation of papers is typical of desktop VR. The user is able to 'walk' into and around the model, viewing it in close-up to see detail and from a distance to get an overview. Pointing to a particular sphere causes the paper title to be displayed in a pop-up window and clicking on a sphere will display the paper abstract in a linked window in the Web browser.

The next three-dimensional example is WebPath created by Emmanuel Frécon at the Swedish Institute of Computer Science and Gareth Smith at Lancaster University, UK (**Figure 13)**. WebPath is an example of surf map which is linked to the browser to provide a graphical history of routes taken through the Web (Frécon & Smith 1998). There are several interesting examples of surf maps using various graph-like spatialisation. Early examples used two dimensional graphs, for example Peter Dömel's Webmap system from 1994 and Eric Ayers & John Stasko MosaicG system[23] in 1995. More recently, Romain Zeiliger, the Centre National de la Recherche Scientifique, France, has developed a browser extension called Nestor Navigator[24] to provide navigational support through dynamic concept maps and a personal graphical overview (Eklund et al. 1999). In terms surf maps that employ 3D spatialisations, there was notable work from Sougata Mukherjea and his Navigation View Builder system (Mukherjea & Foley 1995), although WebPath presents a much more striking visual appearance. Perhaps the ultimate surf map in some senses was the WWW3D prototype developed by Dave Snowdon and colleagues which did away with the conventional browser page view completely and visualised both the Web pages and their local structure in a multi-user, immersive VR spatialisation (Snowdon et al. 1996).

Figure 13 shows two views of the 3D spatialisation of browsing structures hanging, weightless, in a stylised purple cyber-world. WebPath aim is to provide users with a *"… flexibly tailorable real-time visualisation*" of browsing history within a VR environment, working alongside a conventional browser. The principle advantage of WebPath, the authors contend, will be that visual patterns of previously visited sites will make for easier retrieval. It employs a particularly angular arc-node spatialisation, with many straight lines and ninety-degree angles. Individual Web pages are represented by cubes rather than the spheres of HyperSpace and Semantic Constellations. Cubes were used as their flat surfaces are easier to read from a distance. The page represented by the cubes is indicated by labelling (with the title or url) and texture mapping on the faces. The texture can be the background image of the page, or an image on the page or background colour of the page, depending on the user's choice.

The positioning of the cubes in the space is used to encode data about the Web page and when it was accessed. The three orthogonal dimensions of the space allow one to display three distinct parameters. Firstly, the vertical axis is used exclusively for the time at which the Web page was accessed. So the cubes at the top of the spatialisation are always the most recently visited. The x and y horizontal axes can be used to encode a variety of metrics, such as loading time of the page, page size or number of hyperlinks, which can be selected by the user. The user can change the meaning of the x and y dimensions at any time and the cube positions will be automatically recalculated. WebPath can also position the cube according to approximate real-world geography rather than using an abstract co-ordinate space. A base map is provide on the 'floor' of the information space and the cubes are positioned in the appropriate country based on the domain name of the Web site. Just like many of the other spatialisation I have looked at, clicking on a cube of interest will load that Web page into the browser.

The links between Web page cubes show the paths the user takes via hyperlinks. So when the user goes to a new Web page, a new cube is created, and an arc connects this back to the previous cube. So arcs represent a hyperlink as well as the pathway taken. The colour of the arc is used to indicate the pages are from the same site. Repeat visit to the same Web site at different time are indicated by multiple cubes which are separate vertically but are connected by solid yellow columns. The column clearly delineates the fact that they are the same page. For the most popular Web pages the column turn from yellow to red to indicate repeated accesses.

Distinct session of browsing are also visually separated in the WebPath spatialisation using semi-transparent horizontal planes. This divides the space into obviously separate layers, but still allowing one to see what one did previously. The level of transparency can be set by the user.

Our next example leaves Euclidean space and spatialises the structure of the Web in a different space known as hyperbolic space. **Figure 14** shows two examples of 3D hyperbolic spaces produced by Tamara Munzner, a graduate student in the Computer Graphics Laboratory, Stanford University (Munzner & Burchard 1995, Munzner 1998). Hyperbolic spaces have advantages for visualising the detailed structure of large graphs containing many thousands nodes, such as the Web, as Munzner & Burchard comment, "*The felicitous property that hyperbolic space has 'more room' than Euclidean space allows more information to be seen amid less clutter, and motion by hyperbolic isometries provide for mathematically elegant navigation.*" (1995:33).

The spatialisations provide a novel way for exploratory visual browsing of the page-hyperlink structure of large Web sites. The structure of nodes and links is projected in hyperbolic space inside a ball, known as the "sphere at infinity". The user is able to manipulate the graph, rotation and spinning it inside the sphere, to view it from any angle. Like the fisheye distortion technique, which I looked at in relation to Inxight's Site Lens spatialisation, hyperbolic space gives greater visual presence (in terms of screen-space) to elements at the centre of the space. As objects are moved to the periphery they smoothly shrink in size. At the edge of the sphere the nodes are very small, but the user can easily drag them into the centre to enlarge them and see them in detail. In this manner, the hyperbolic spatialisation can provide a view of the detailed graph structure, whilst still showing the over context.

An early example by Munzner and Burchard from 1995 is shown in **figure 14(a)** which spatialises the structure of two layers of their Web site with pyramid glyphs representing pages and the curving lines being the principle hierarchical hyperlinks. This spatialisation was part of a Web mapping system called *webviz* which could gather the structure of specified portion of the Web and then visualise it in a 3D viewer called Geomview (Munzner & Burchard 1995). Munzner undertook further refinements in the underlying hyperbolic spatialisations, developing the H3 layout algorithm and more powerful viewing system (H3Viewer) which enables interactive exploration of graphs of 100,000 or more nodes (Munzner 1998). **Figure 14(b)** shows an example of the H3Viewer incorporated into a product called *Site Manager*[25] from SGI. This is tool is for Webmasters and content creators to give them a fluid and scaleable view of the structure of the Web site they manage. It provides tool to select and manipulate the graph, as well as distinguishing different types of nodes.

The final example in this paper is perhaps the most science-fictional looking cyberspace spatialisation. It was called HotSauce and was a 3D fly through interface to information spaces like the Web[26]. It was developed, largely as a one-man effort, by Ramanathan V. Guha at Apple Research. **Figure 15** shows one screen-grab of a flythrough of the Apple Web site from 1997. The fundamental concept behind HotSauce can is succinctly summarised by the phrase '*why just browse when you can fly?*'.

HotSauce was a specific 3D spatialisation of the Meta Content Framework (MCF) also developed by Guha. MCF was a schema for describing and organising the structure of an information space (Guha n.d.). This is called metadata and is separate from the actual content, for example a library catalogue is vital metadata that enables you to actually find the books you are looking for. (MCF has evolved in the Resource Description Framework (RDF)[27], a proposed standard for the Web's underlying metadata infrastructure that would support advanced cataloguing and searching tools.)

HotSauce was a 3D viewer for MCF and it worked as a plugin to an existing browser. So when you clicked on a hyperlink to an MCF Web site description you were dropped into a first-person perspective view of the information space, with pages floating like brightly coloured asteroids in an infinite space. The view is kind of like that from a starship cockpit and using a combination of the mouse and key you could cruise into and through the pages. The flying is smooth and the pages get larger as they come towards you and then disappear behind. Web pages are represented by rectangular glyphs and are labelled with the page title. Topic areas are indicated by the round cornered rectangles. Different hierarchical level of information are denoted by different colours of glyphs as well as their spatial depth in the 3D display. So in **figure 15**, the top-level is the green "Apple Computer" topic, which is then followed by major sections of the site represented by the bold red glyphs. Further back still are yellow and then purple pages. When you fly towards a clump of pages, more pages deeper in the hierarchy can be revealed. The glyphs are also arranged spatially into distinct groups. The actual MCF description of the site is static and created off-line by the Webmaster.

Flying through the space was done by pointing the mouse cursor in the direction you wanted to go and holding down the mouse button to effortlessly glide forward. It was possible to steer and glide forward at the same time. To fly backwards you held down the control key on the keyboard at the same time. While holding down the shift key would put on in warp speed. One could also instantly zoom into topic area by doubling clicking on its rounded rectangular glyph. Double clicking on a rectangular page glyph caused the page to be open in another browser window. Also clicking on the grid of around the edge panned the 3D window across the space.

The HotSauce spatialisation and the underlying MCF platform was ostensibly the work of Guha and enjoyed a spark of popularity in 1997. Unfortunately, Apple ended its development of HotSauce soon afterwards and Guha moved to Netscape (Andreessen 1999). Since then HotSauce has, unfortunately, become another dead-end prototype, although a different company, Perspecta, has been developing its own 3D information flythrough called PerspectaView, which is based on an underlying metadata structure called SmartContent[28] which has distinct similarities to HotSauce. As a cyberspace spatialisation HotSauce was not terribly usable for serious work. The flying was fun for the first few experiments, but it was not practical means of navigation. It was too easy to become lost and disorientated in the space, once you where in the middle of floating glyphs one lost the overall context. Despite the practical difficulties, HotSauce was important, as Steven Johnson, in his book *Interface Culture*, say,

> "*But a day or two with HotSauce was enough to catch a glimpse of what a genuinely spatial systems might feel like. At a few, enthralling moments, I found myself groping around for a familiar document and thinking: It's back there somewhere, up and to the left a little, about two or three planets deep. For a second or two I was thinking in purely spatial terms, zooming in and out of my own private dataspace. For those few moments, there was a hint of liberation in the air, the promise of things to come.*" (Johnson 1997:80)

References

Albert, R., Jeong, H., & Barabási, A-L., 1999, "Diameter of the World-Wide Web", *Nature*, 9th September 1999, Vol. 401, pages 130-131.

Andreessen, M., 1999, "Innovators of the Net: Ramanathan V. Guha and RDF**",** *Netscape TechVision column*, 8th January 1999. <http://www31.netscape.com/columns/techvision/innovators_rg.html>

Andrews, K., 1995, "Visualising Cyberspace: Information Visualisation in the Harmony Internet Browser", *Proceedings of Info-Vis'95, IEEE Symposium on Information Visualization*, (IEEE Computer Society Press: Los Alamitos, CA.), pages 97-104.
<ftp://ftp.iicm.edu/pub/papers/ivis95.pdf>

Andrews, K., Pichler, M., & Wolf, P., 1996, "Towards Rich Information Landscapes for Visualising Structured Web Space", *Proceedings of Info-Vis'96, IEEE Symposium on Information Visualization*, (IEEE Computer Society Press: Los Alamitos, CA.), pages 62-63. <ftp://ftp.iicm.edu/pub/papers/ivis96.pdf>

Ayers, E.Z. & Stasko, J.T., 1995, "Using Graphic History in Browsing the World-Wide Web", *Proceedings of the Fourth International World-Wide Web Conference*, December 1995, Boston, USA. <http://www.w3.org/Conferences/WWW4/Papers2/270/>

Benford, S., Snowdon D., Greenhalgh, C., Ingram, R., Knox, I., & Brown, C., 1995, "VR-VIBE: A Virtual Environment for Co-operative Information Retrieval", *Proceedings of Eurographics'95*, 30th August - 1st September 1995, Maastricht, The Netherlands, pp 349-360.

Benford S., Snowdon, D., Brown, C., Reynard, G., & Ingram, R., 1997, "Visualising and Populating the Web: Collaborative Virtual Environments for Browsing, Searching and Inhabiting Webspace", *Proceedings of the Eighth Joint European Networking Conference (JENC8)*, 12-15th May 1997, Edinburgh. <http://www.rare.nl/conf/jenc8/papers/123.ps>

Berstein, M., 1991, "The Navigation Problem Reconsidered", in Berk E. & Devlin J., (Eds.), *Hypertext / Hypermedia Handbook*, (McGraw-Hill: New York), pages 285-297.

Bowman, C.M., Danzig, P.B., Manber, U., & Schwartz, M.F., 1994, "Scalable Internet Resource Discovery: Research Problems and Approaches", *Communications of the ACM*, August 1994, Vol. 37, No. 8, pages 98-107.

Brake, D., 1997, "Lost in Cyberspace", *New Scientist*, 28th June 1997, Vol. 154, No. 12, pages 12-13.

Bray, T., 1996, "Measuring the Web", presented at the *Fifth International Conference World Wide Web*, 6-10th May 1996, Paris.
<http://www5conf.inria.fr/fich_html/papers/P9/Overview.html>

Card, S.K., Mackinlay, J.D., & Shneiderman, B., 1999, *Readings in Information Visualization: Using Vision to Think*, (Morgan Kaufmann Publishers, Inc: San Francisco).

Chalmers, M., Rodden K., & Brodbeck, D., 1998, "The Order of Things: Activity-Centered Information Access", *Proceedings of the Seventh International World-Wide Web Conference*, April 1998, Brisbane, Australia, pages 359-367.
<http://www.dcs.gla.ac.uk/~matthew/papers/WWW7/www98.html>

Chen, C., 1997, "Structuring and visualising the WWW with Generalised Similarity Analysis", *Proceedings of the Eighth ACM Conference on Hypertext (Hypertext'97)*, June 1997, Southampton, UK, pages 177-186). <http://www.brunel.ac.uk/~cssrccc2/papers/ht97.pdf>

Chen, C., 1999, *Information Visualisation and Virtual Environments*, (Springer Verlag: Berlin)

Chen, H., Schuffels, C., & Orwig, R., 1996, “Internet Categorization and Search: A Machine Learning Approach”, *Journal of Visual Communications and Image Representation*, Vol. 7, No. 1, pages 88-102.

Chen, H., Houston A.L., Sewell R.R., & Schatz B.R., 1998, “Internet Browsing and Searching: User Evaluations of Category Map and Concept Space Techniques”, *Journal of the American Society for Information Science*, Vol. 49, No. 7, pages 582-603.

Churchill, E.F., Snowdon, D., Benford, S., & Dhanda, P., 1997, "Using VR-VIBE: browsing and searching for documents in 3d-space", paper presented at *HCI International'97: 7th International Conference on Human-Computer Interaction*, 24-29th August 1997, San Francisco, USA.

Clement, P.C., 1998, *The State of the Net*, (McGraw-Hill: New York).

Clever, 1999, “Hypersearching the Web”, *Scientific American*, June 1999, pages 44-52. <http://www.scientificamerican.com/1999/0699issue/0699raghavan.html>

Cockburn, A. & Jones, S., 1996, “Which way now? Analysing and easing inadequacies in WWW navigation.”, *International Journal of Human-Computer Studies*, Vol. 45, pages 105-109.

Cockburn, A. & Greenberg, S., 1999, *Beyond the 'Back' Button: Issues of Page Representation and Organisation in Graphical Web Navigation Tools*, Research Report 99-640-03, Department of Computer Science, University of Calgary, Calgary, Canada. <http://www.cpsc.ucalgary.ca/grouplab/papers/1999/99-WebView/graWeb.pdf>

Cosgrove, D., “Contested Global Visions: One-World, Whole-Earth, and the Apollo Space Photographs”, *Annals of the Association of American Geographer*, Vol. 84, No. 2, 1994, pp. 270-294.

December, J., 1995, "A Cybermap Gazetteer: Maps of the Online World for Browsing and Business", in Staple, G.C., (Eds.), *TeleGeography 1995*, (TeleGeography, Inc.: Washington, DC), pages 74-82.

Dömel, P., 1994, “Webmap – A Graphical Hypertext Navigation Tool”, *Proceedings of the Second International World-Wide Web Conference*, September 1994, Chicago, USA, pages 85-97. <http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings/Searching/doemel/www-fall94.html>

Durand, D. & Kahn, P., 1998, "MAPA: A system for inducing and visualizing hierarchy in websites", *Ninth ACM Conference on Hypertext and Hypermedia (HT'98)*, June 1998, Pittsburgh, USA. <http://www.dynamicdiagrams.com/pdf/papers/mapaht98.pdf>

Edwards, D.W. & Hardman, L., 1989, *Lost in Hyperspace: Cognitive Mapping and Navigation in a Hypertext Environment*, (Intellect Books: Oxford, England).

Eklund, J., Sawers, J., & Zeiliger, R., 1999, “NESTOR Navigator: A tool for the collaborative construction of knowledge through constructive navigation”, in R. Debreceny & A. Ellis (Eds.), *Proceedings of Ausweb99, The Fifth Australian World Wide Web Conference*, (Southern Cross University Press: Lismore), pages 396-408. <http://ausweb.scu.edu.au/aw99/papers/eklund2/>

Fabrikant, S.I., 1999, *Spatial Metaphors for Browsing Large Data Archives*. Unpublished Dissertation, Department of Geography, University of Colorado-Boulder.

Fabrikant, S.I., 1999, "Spatialized Browsing in Large Data Archives", *Transactions in GIS*, January 2000, Vol. 4, No. 1, forthcoming.

Frécon, E. & Smith, G., 1998, "WebPath - A three-dimensional Web History", Proceedings IEEE Symposium on Information Visualization (InfoVis '98), Chapel Hill, NC, USA. <http://www.comp.lancs.ac.uk/computing/users/gbs/webpath/>

Garland, K., 1994, *Mr Beck's Underground Map*, (Capital Transport Publishing: Middlesex, England).

Gershon, N. & Eick, S.G., 1995, "Visualization's New Track: Making sense of information", *IEEE Spectrum*, November 1995, Vol. 32, No. 1, pages 38-56.

Gershon, N. & Eick, S.G., (Eds.), 1995, *Proceedings of Information Visualization'95*, (IEEE Computer Society Press: Los Alamitos, California).

Gershon, N., Eick, S.G., & Card, S., 1998, "Information Visualization", *Interactions*, March/April 1998, pages 9-15.

Girardin, L., 1995, "Mapping the virtual geography of the World-Wide Web", poster presentation at the *Fifth International World-Wide Web Conference*, 6-10th May 1999, Paris.
<http://heiwww.unige.ch/girardin/cgv/www5/index.html>

Gloor, P., 1997, *Elements of Hypermedia Design: Techniques for Navigation & Visualization in Cyberspace*, (Birkhauser: Boston, MA.).

Guha, R.V., n.d., "Meta Content Framework: A Whitepaper", Apple Research.
<http://www.xspace.net/hotsauce/wp.html>

Hendley, R.J., Drew, N.S., Wood, A.M., & Beale, R., 1995, "Narcissus: Visualising Information", *Proceedings of InfoVis'95. IEEE Symposium on Information Visualization*, New York, pages 90-96, 146.

Holtzman, S., 1997, *Digital Mosaics: The Aesthetics of Cyberspace*, (Simon & Schuster: New York).

Honkela, T., Kaski, S., Kohonen, T., & Lagus, K., 1998, "Self-organizing maps of very large document collections: Justification for the WEBSOM method", in Balderjahn, I., Mathar, R., & Schader, M., (Eds.), *Classification, Data Analysis, and Data Highways*, (Springer-Verlag: Berlin), pages 245-252.

Huberman, B.A., Pirolli, P.L.T., Pitkow, J.E., & Lukose, R.M., 1998, "Strong Regularities in World Wide Web Surfing", *Science*, 3rd April 1998, Vol. 280, pages 95-97.

Jacobson, R.E., 1999, *Information Design*, (MIT Press: Cambridge, MA.).

Johnson, S., 1997, *Interface Culture: How New Technology Transforms the Way We Create and Communicate*, (Harper: San Francisco).

Kahn, P., 1999, *Mapping Web Sites: Planning Diagrams to Site Maps*, Dynamic Diagrams, 10th January 1999. <http://www.dynamicdiagrams.com/seminars/mapping/maptoc.htm>

Kohonen, T., 1995, *Self-organizing maps* (Springer-Verlag: Berlin).

Lagus, K., Honkela, T., Kaski, S., & Kohonen, T., 1996a, "Self-organizing maps of document collections: A new approach to interactive exploration", in Simoudis, E., Han, J., and Fayyad, U., (Eds.), *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, (AAAI Press, Menlo Park, CA.), pages 238-243.
<http://websom.hut.fi/websom/doc/ps/lagus96kdd.ps>

Lamping, J. & Rao, R., 1995, "The Hyperbolic Browser: A Focus + Context Technique for Visualizing Large Hierarchies", *Journal of Visual Languages and Computing*, Vol. 7, No. 1, pages 33-55.

Lawrence, S. & Giles, C.L., 1998, "Searching the World Wide Web", *Science*, Vol. 280, 3rd April 1998, pages 98-100.

Lawrence S. & Giles C.L., 1999, "Accessibility of information on the web", *Nature*, Vol. 400, 8th July 1999, pages 107-109.

Laurel B., 1990, *Art of Human-Computer Interface Design*, (Addison-Wesley Publishing Company)

Lin, X., 1992, "Visualization for the document space", *Proceedings of IEEE Visualization'92*, pages 274-281. Reproduced in Card, S.K., MacKinlay, J. & Shneiderman, B., (Eds.), 1999, *Readings in Information Visualization: Using Vision to Think*, (Morgan Kaufman Publishers, Inc.: San Francisco), pages 432-440.

Lin, X., 1997, "Map displays for information economy retrieval", *Journal of the American Society for Information Science*, Vol. 48, No. 1, pages 40-54.

Maarek, Y.S., Jacovi, M., Shtalhaim, M., Ur S., Zernik, D., & Ben, I.Z., 1997, "WebCutter: A system for dynamic and tailorable site mapping", paper presented at the *Sixth International World-Wide Web Conference*, April 1997, Santa Clara, USA.
< http://decweb.ethz.ch/WWW6/Technical/Paper040/Paper40.html>

Marchionini, G., 1997, *Information Seeking in Electronic Environments*, (Cambridge University Press).

Maurer, H., 1996, *Hyperwave: The Next Generation Web Solution*, (Addison-Wesley Publishing Company: Reading, MA.).

McCormick, B.H., Defanti, T.A., & Brown, M.D., 1987, "Visualization in Scientific Computing", *Computer Graphics*, November 1987, Vol. 21, No. 6.

Morville, P., 1996, "Mapping Your Site: A Picture's Worth A Thousand Words", *Web Review*, September 1996. <http://webreview.com/wr/pub/96/09/27/arch/index.html>

Mukherjea, S., & Foley, J.D., 1995, "Visualizing the World-Wide Web with the Navigational View Builder", *Proceedings of the Third International World-Wide Web Conference*, April 1995, Darmstadt, Germany.
<http://www.igd.fhg.de/www/www95/proceedings/papers/44/mukh/mukh.html>

Munro, A., Hook, K., Benyon, D., 1999, *Social Navigation: Footprints in the Snow*, (Springer-Verlag: Berlin).

Munzner, T. & Burchard, P., 1995, "Visualizing the Structure of the World Wide Web in 3D Hyperbolic Space", *Proceedings of VRML'95*, December 1995, San Diego, California, pages 33-38. <http://www-graphics.stanford.edu/papers/webviz/>

Munzner, T., 1998, "Exploring Large Graphs in 3D Hyperbolic Space", *IEEE Computer Graphics and Applications*, July/August 1998, Vol. 18, No. 4, pages 18-23.
<http://www-graphics.stanford.edu/papers/h3cga>

Sarkar, M. & Brown, M.H., 1994, "Graphical Fish-eye Views", *Communications of the ACM*, December 1994, Vol. 37, No. 12, pages 73-84.

Shneiderman, B., 1997, *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (3rd edition), (Addison-Wesley Publishing Company).

Snowdon, D., Fahlén, L., & Stenius, M., 1996, "WWW3D: A 3D Multi-User Web Browser", paper presented at *WebNet'96*, October 1996, San Francisco. <http://www.crg.cs.nott.ac.uk/~dns/vr/www3d/webnet96-final.html>

Tobler, W.R., 1970, "A computer movie simulating urban growth in the Detroit region", *Economic Geography (supplement)*, Vol. 46, pages 234–240.

Tufte, E.R., 1990, *Envisioning Information*, (Graphics Press).

Wexelblat A. & Maes, P., 1999, "Footprints: History-Rich Tools for Information Foraging", paper presented at *CHI'99 Conference*, 15-20th May 1999, Pittsburgh, USA.

<http://wex.www.media.mit.edu/people/wex/CHI-99-Footprints.html>

Wise, J.A., Thomas, J.J., Pennock, K., Lantrip, D., Pottier, M., Schur, A., & Crow, V., 1995, "Visualizing the Non-Visual: Spatial Analysis and Interaction with Information from Text Documents", *Proceedings of IEEE Information Visualization'95*, pages 51-58, 140. Reproduced in Card, S.K., MacKinlay, J. & Shneiderman, B., (Eds.), 1999, *Readings in Information Visualization: Using Vision to Think*, (Morgan Kaufman Publishers, Inc.: San Francisco), pages 442-450.

Wood, A.M., Drew, N.S., Beale, R., & Hendley, R.J., 1995, "HyperSpace: Web Browsing with Visualisation", *Proceedings of the Third International World-Wide Web Conference*, April 1995, Darmstadt, Germany, pages 21-25.
<http://www.igd.fhg.de/www/www95/proceedings/posters/35/index.html>

Woodruff, A., Aoki, P.M., Brewer, E., Gauthier, P., Rowe, L.A., 1996, "An investigation of documents from the World-Wide Web", presented at the *Fifth International Conference World Wide Web*, 6-10th May 1996, Paris.
<http://epoch.cs.berkeley.edu:8000/~woodruff/inktomi/index.html>

Wurman, R.S., 1997, *Information Architects*, (Graphis Press Corp.).

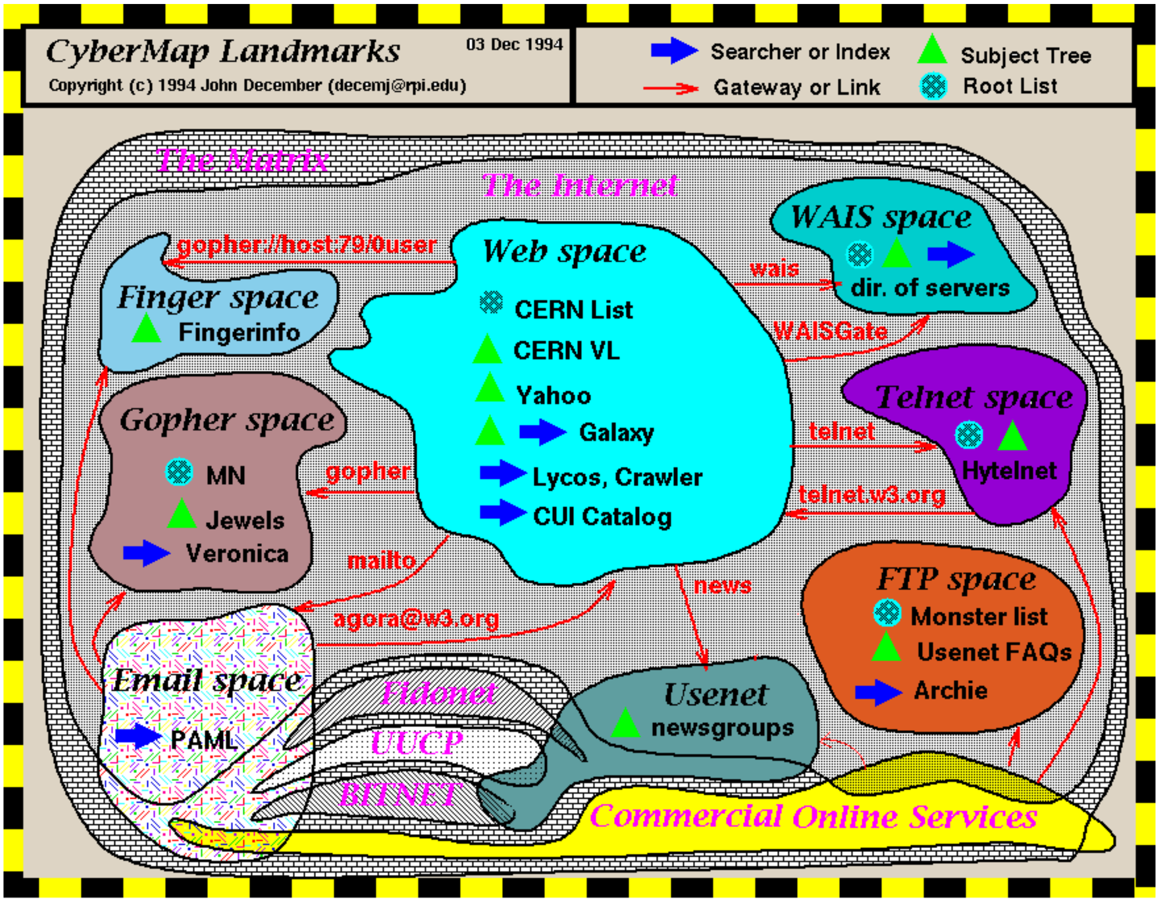


Figure 1: John December's conceptual maps of Internet information spaces.

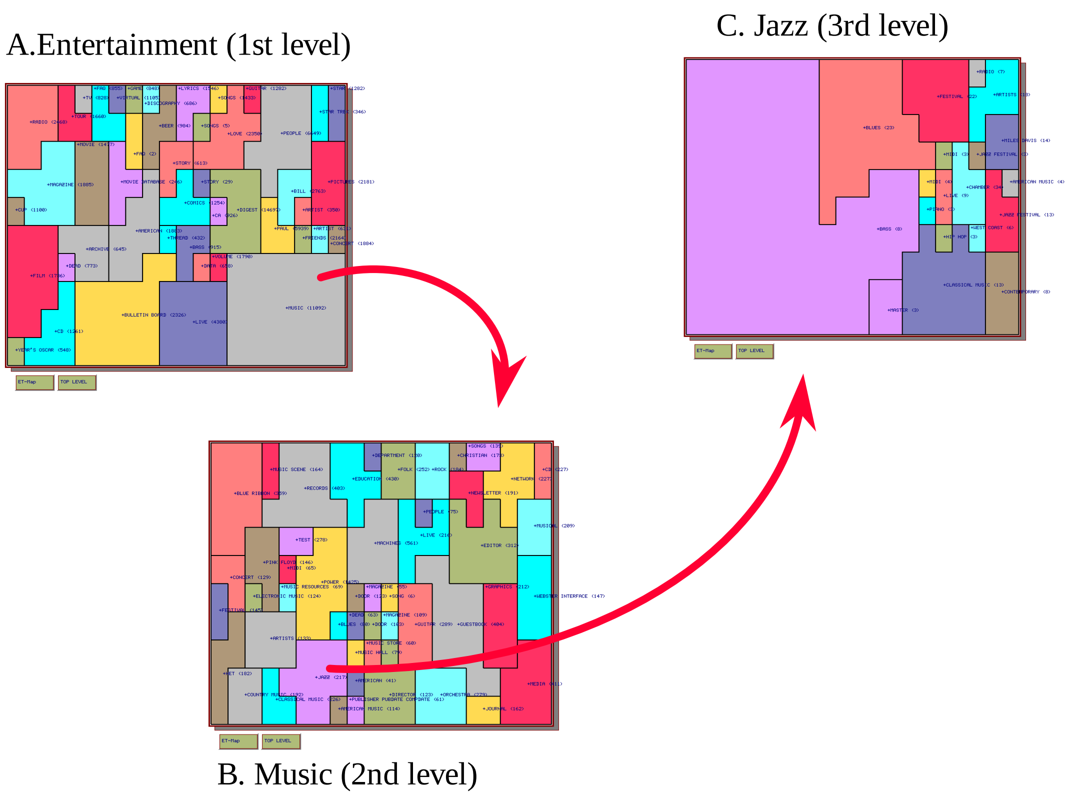


Figure 2: ET-Map category map.

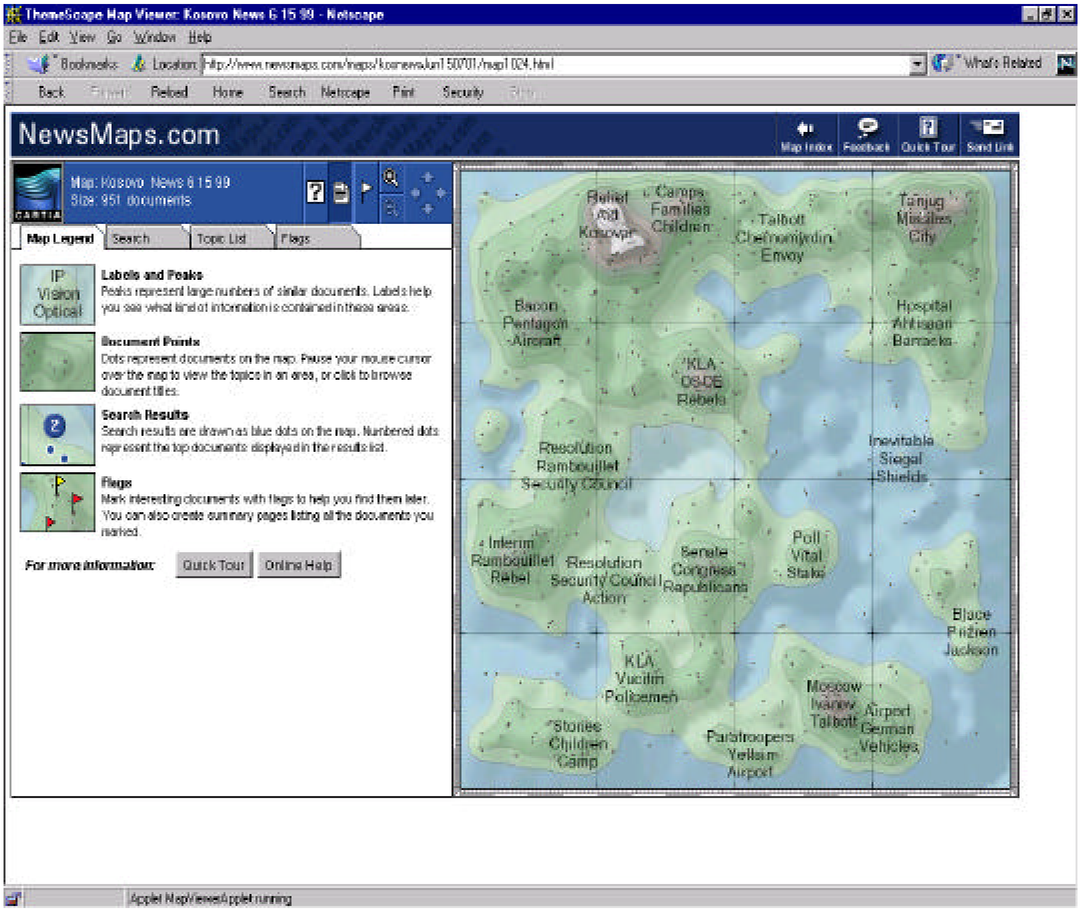


Figure 3: NewsMaps.com information map.

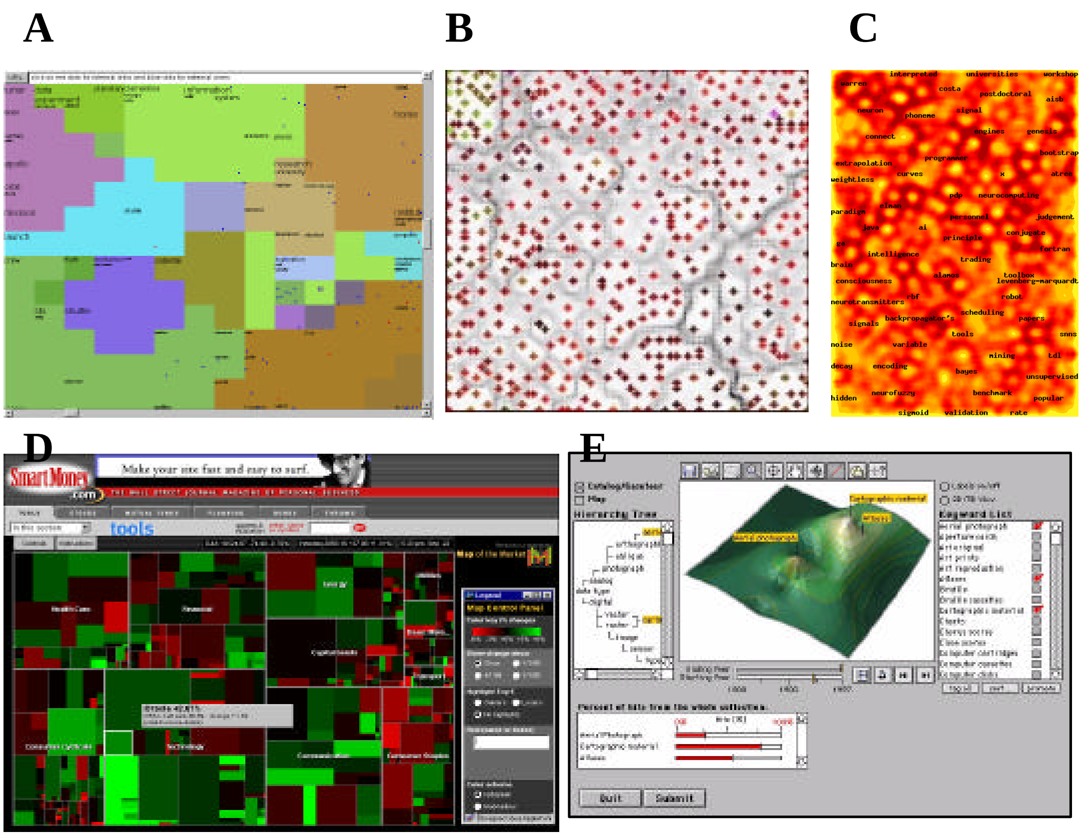
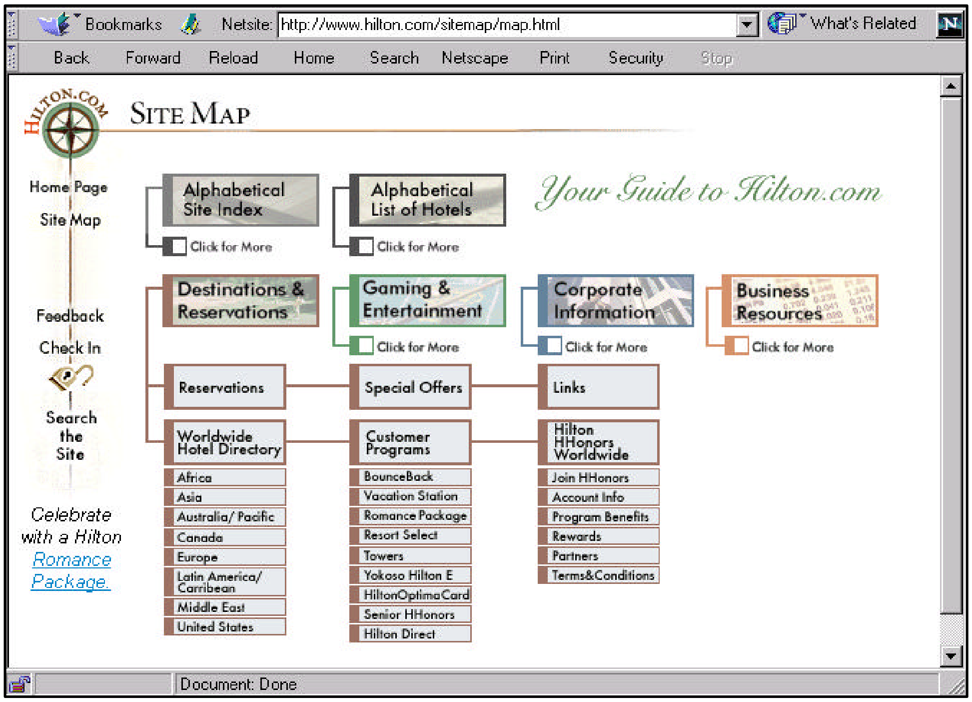


Figure 4: Examples of information maps.

A



B

Figure 5: Hand-crafted, static, Web site maps.

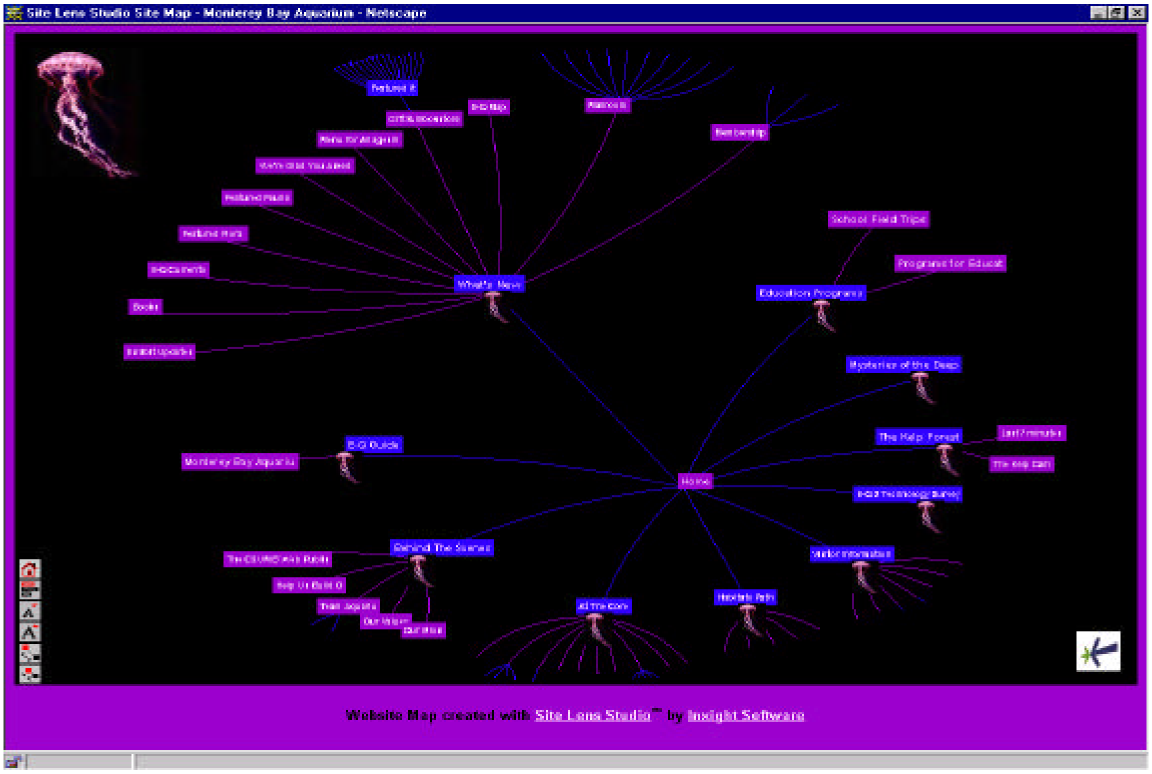


Figure 6: Site Lens interactive Web site map.

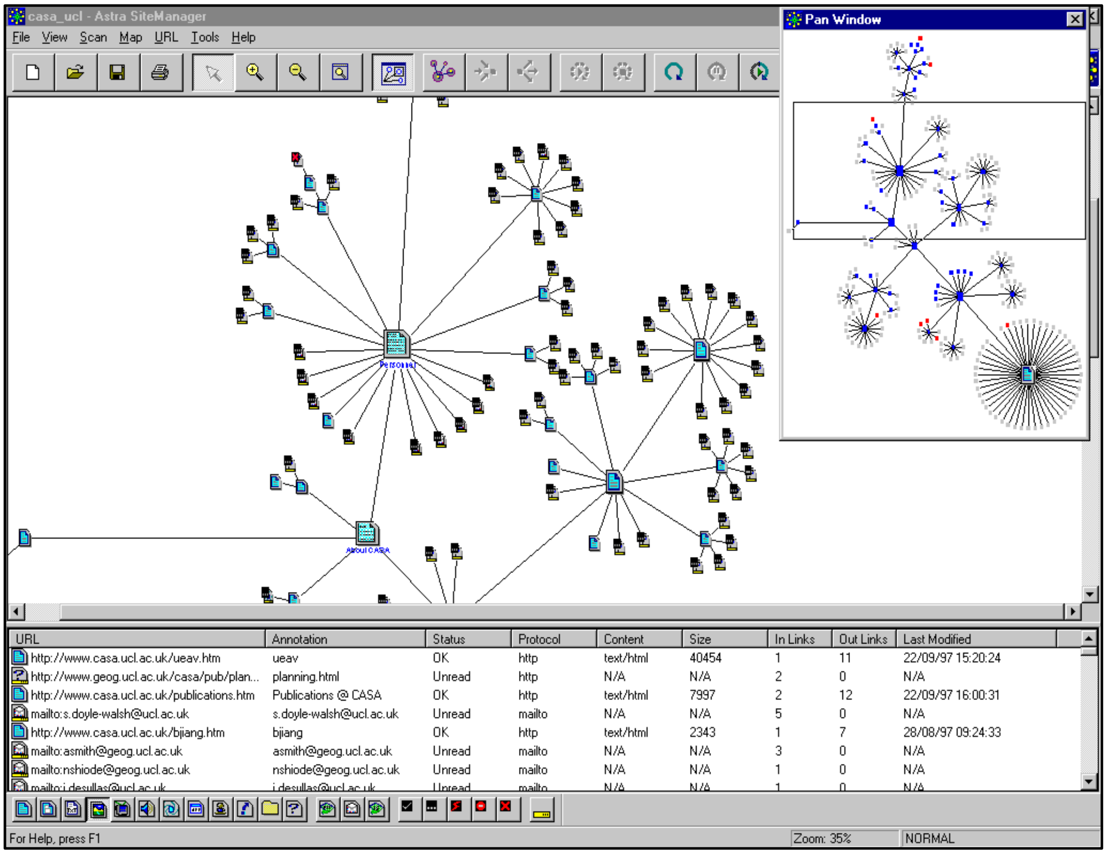


Figure 7: Astra SiteManager visual Web site management tool.

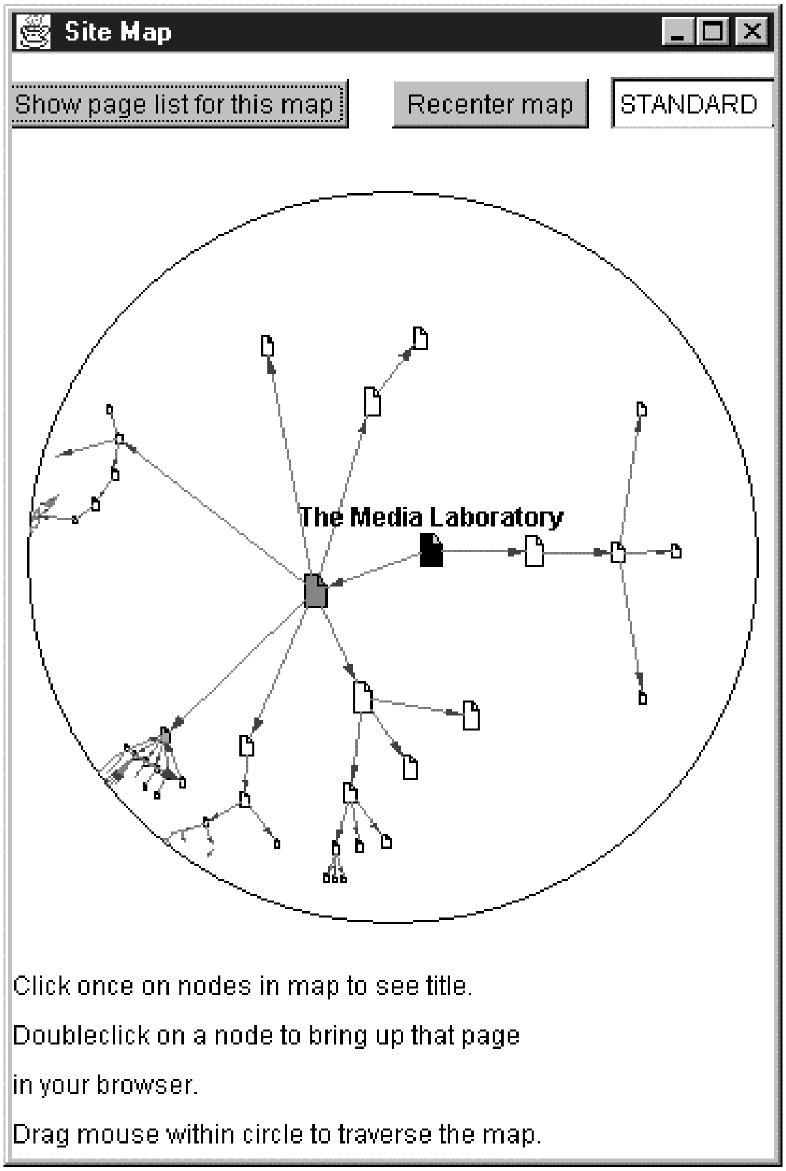


Figure 8: Footprints Web mapping system.

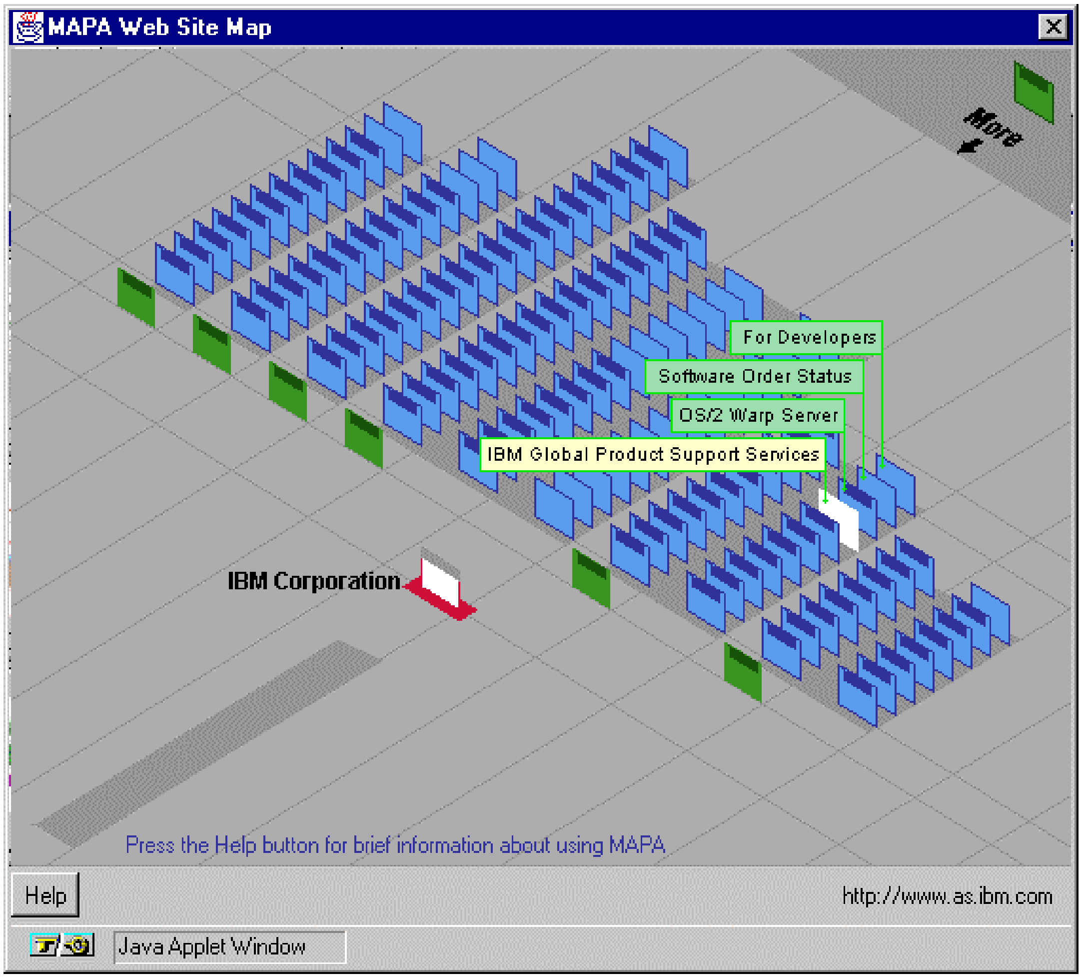
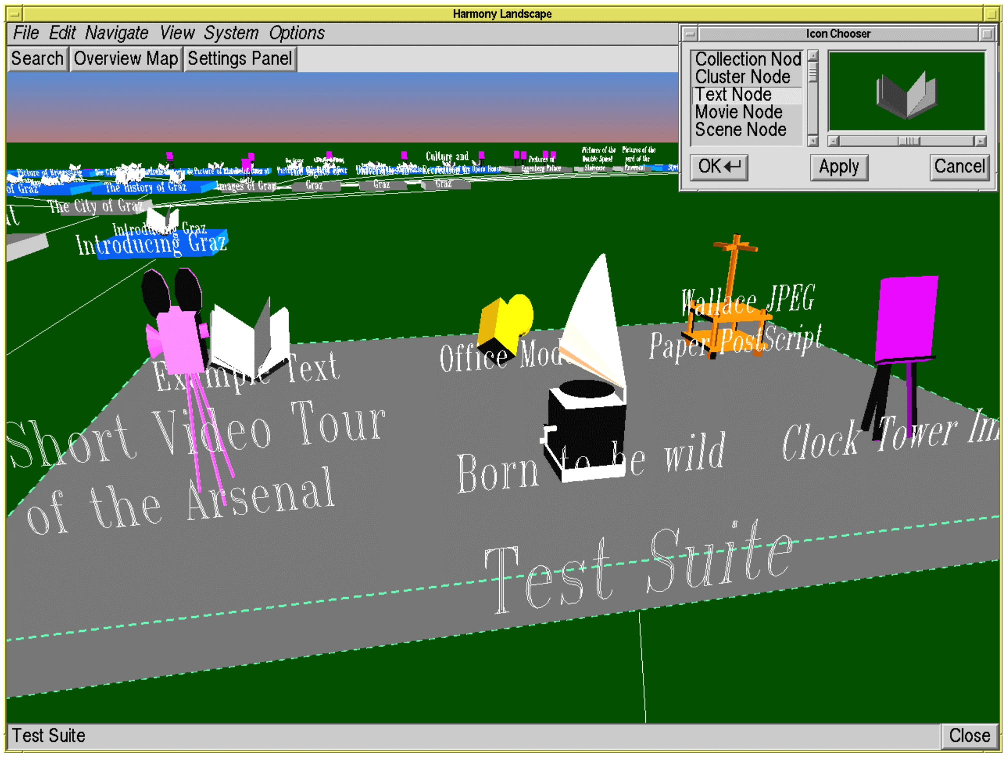


Figure 9: The MAPA Web site map.

A



B

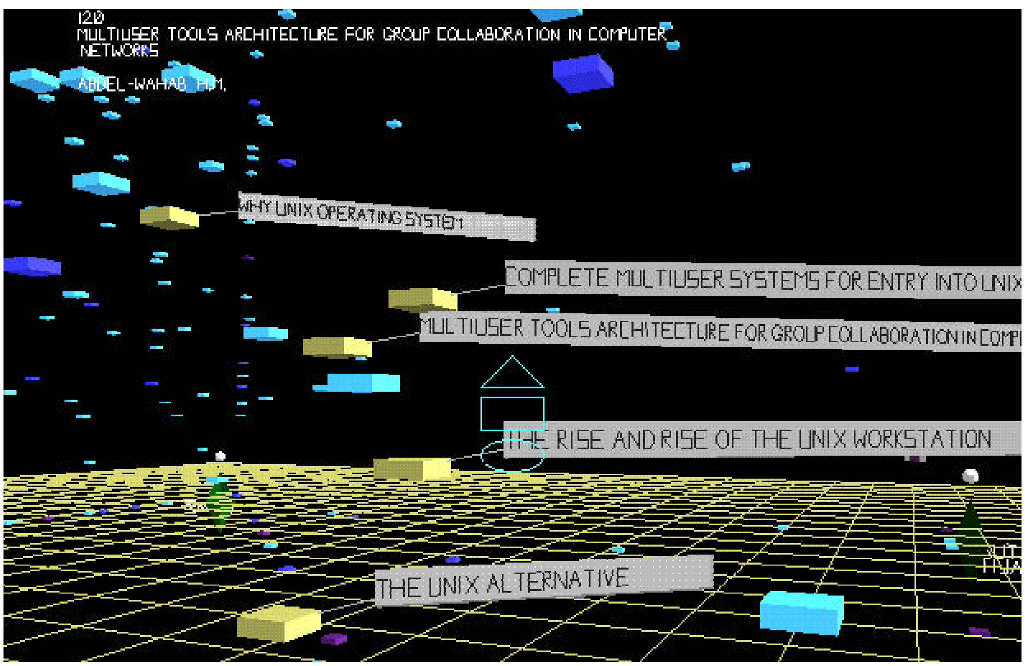


Figure 10: The (a) Harmony Information Landscape; and (b) VR-VIBE.

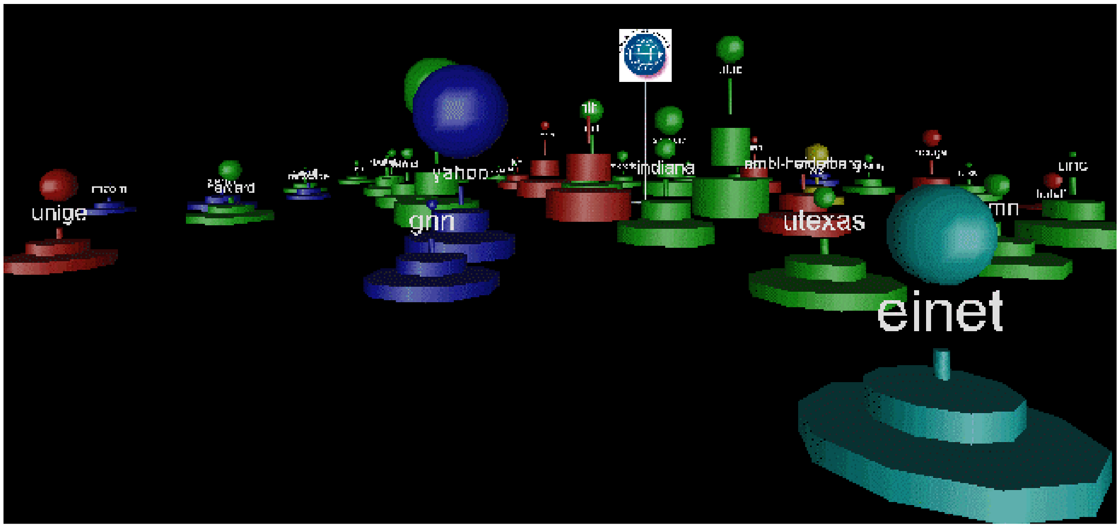
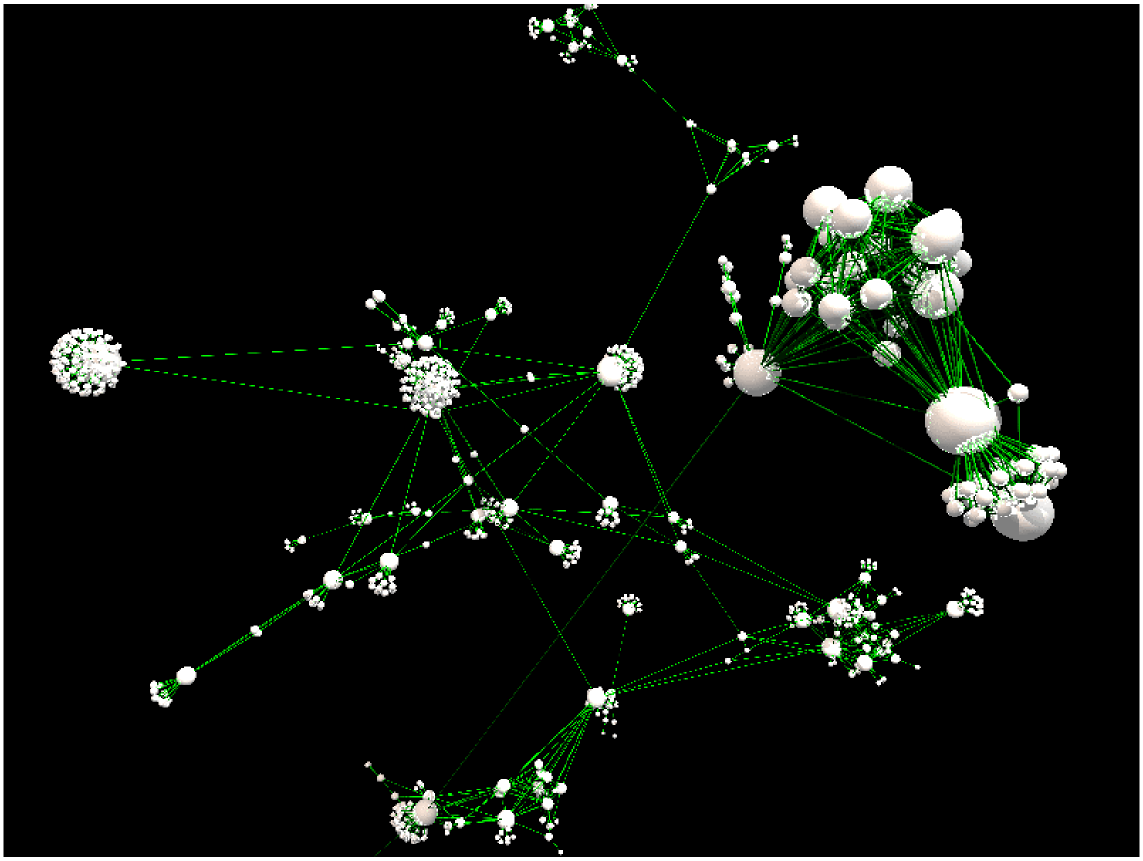


Figure 11: Tim Bray's Web landscape.

A



B

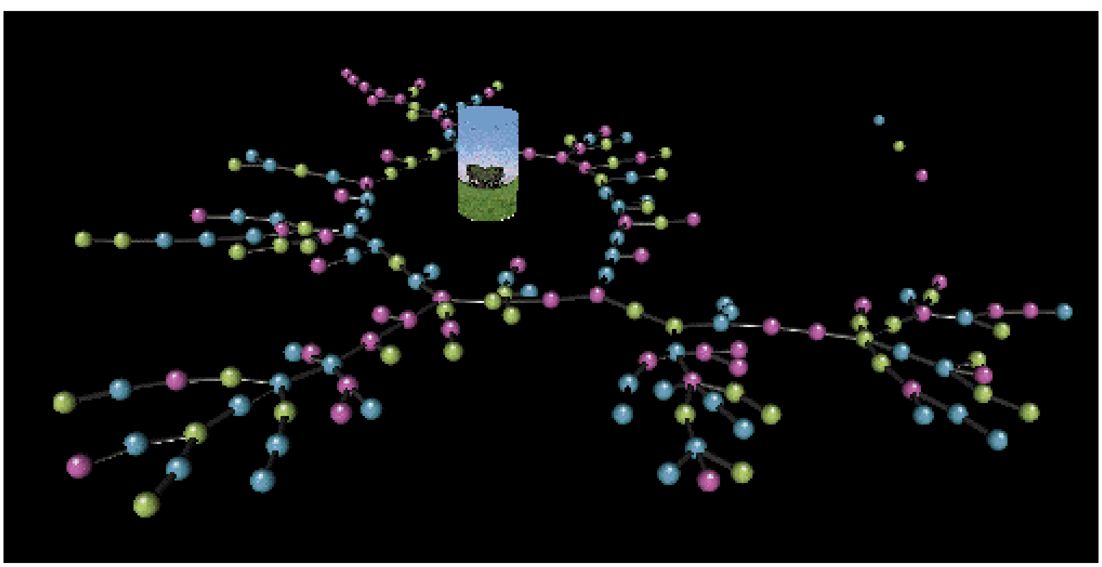


Figure 12: (a) HyperSpace Visualiser; and (b) Semantic Constellation.

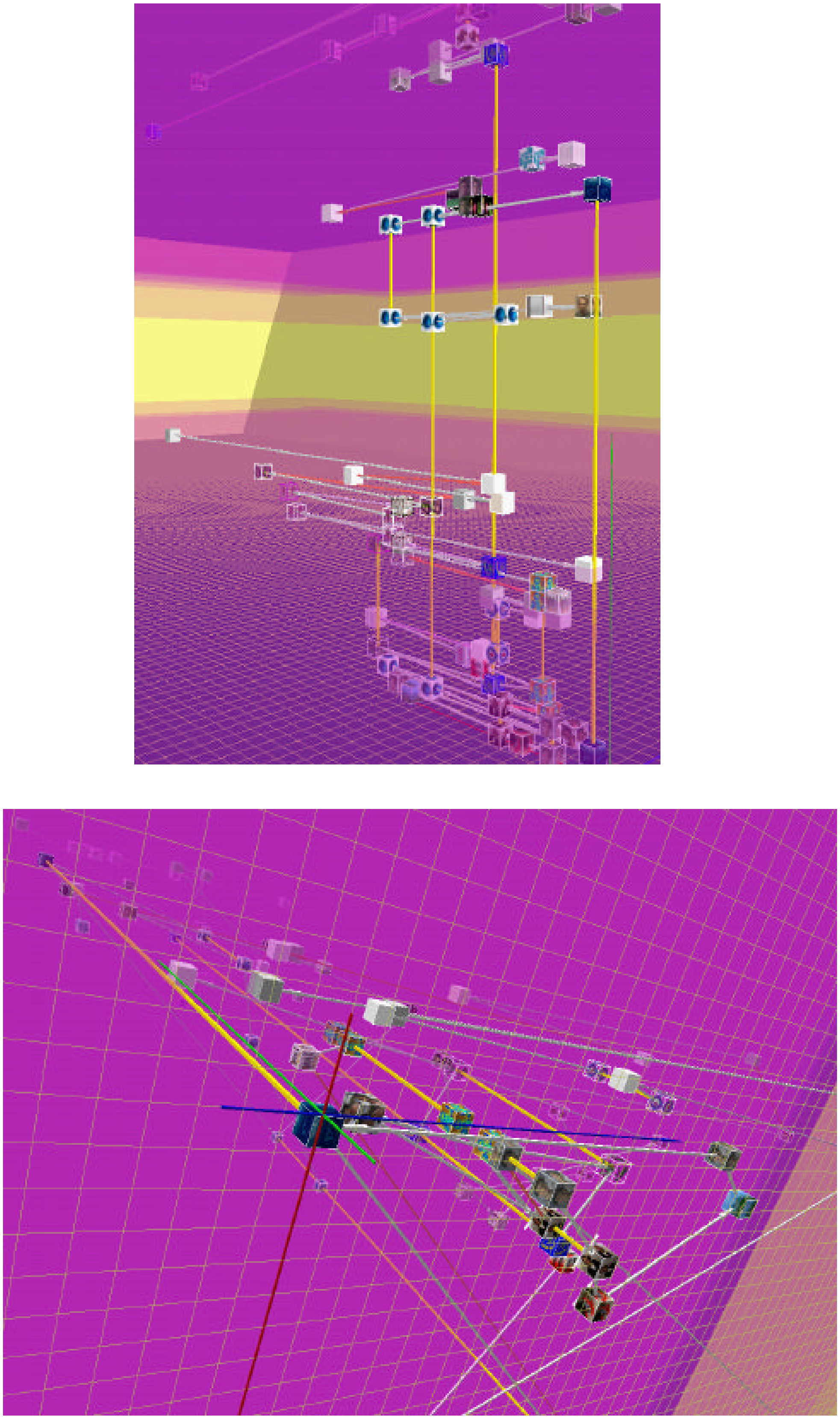


Figure 13: WebPath surf map.

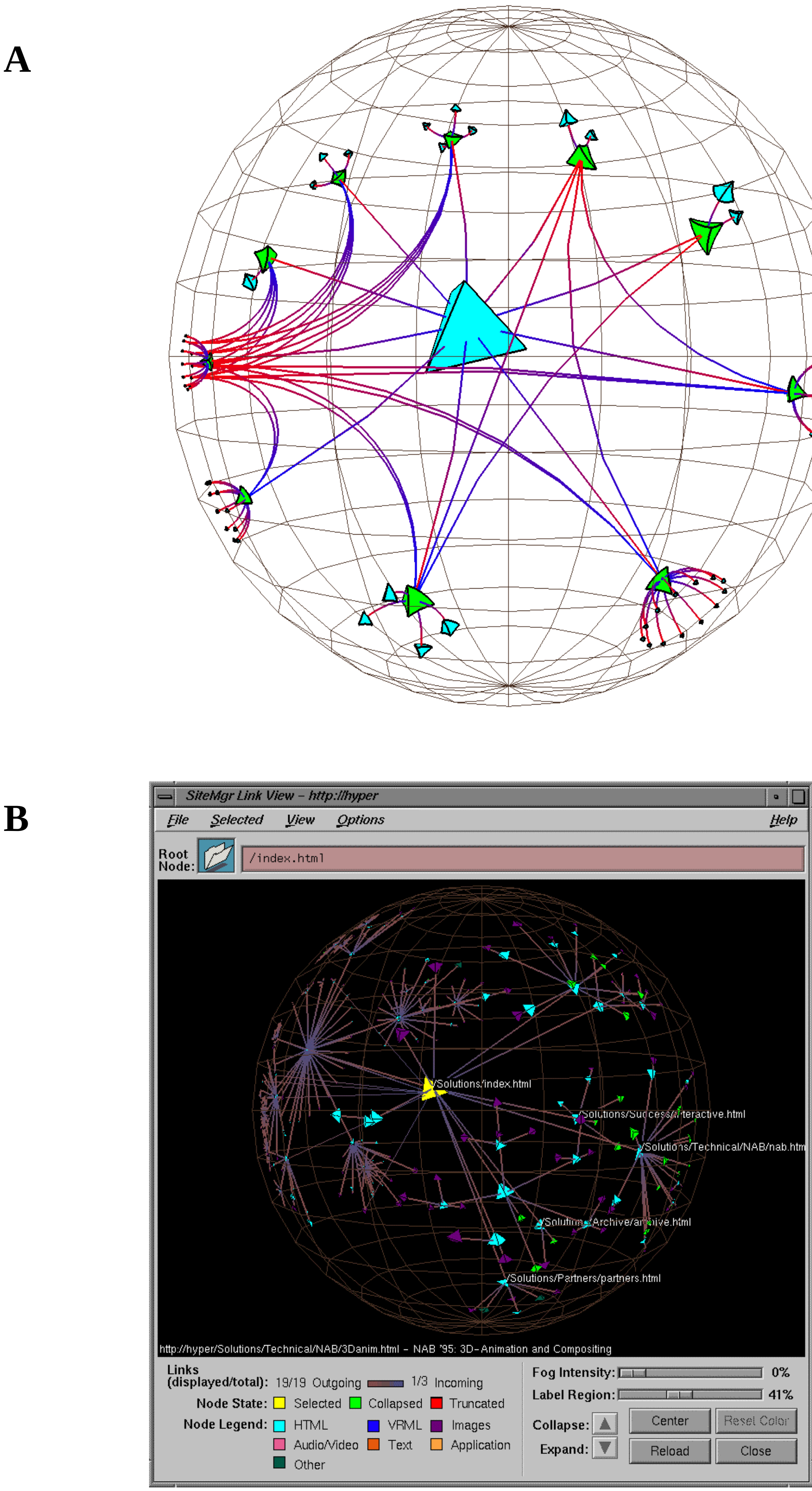


Figure 14: Hyperbolic spatialisation by Tamara Munzner.

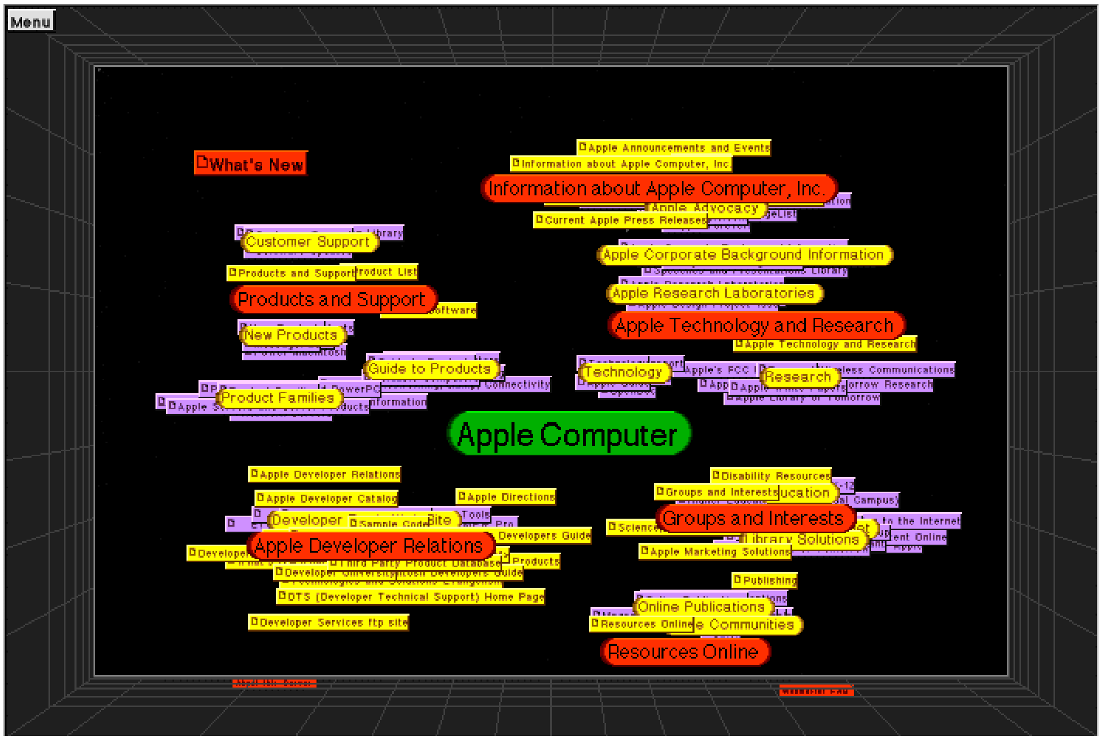


Figure 15: HotSauce Web site fly thru.

---

[1] These four pioneering companies are spins-off from cutting-edge research at Bell Labs-Lucent Technologies, MIT Media Lab, Xerox PARC, and Pacific Northwest National Laboratories. For further information consult their web sites at, <http://www.visualinsights.com/>, <http://www.perspecta.com/>, <http://www.inxight.com/> < http://www.cartia.com/>.

[2] According to Media Metrix's web site popularity measurements, Yahoo! and its associated online properties received 38 million unique visitors in July 1999. <http://www.mediametrix.com/TopRankings/TopRankings.html>

[3] <http://ai2.bpa.Arizona.edu/ent/>

[4] <http://www.newsmaps.com/>

[5] <http://www.cartia.com/>

[6] <http://lislin.gws.uky.edu/Sitemap/>

[7] <http://heiwww.unige.ch/girardin/cgv/>

[8] <http://websom.hut.fi/websom/>

[9] <http://www.smartmoney.com/marketmap/>

[10] <http://ucsu.colorado.edu/~fabrikas/research/spatialization.html>

[11] <http://www.hilton.com/sitemap/map.html>

[12] The map was in use on the Yell web site <http://www.yell.co.uk/> in 1997 and, unfortunately, it is no longer available.

[13] <http://www.inxight.com/Products/Web/SLS.html>

[14] <http://www.ibm.com/java/mapuccino/>

[15] <http://www.merc-int.com/products/astrasitemanager/>

[16] < http://www.clearweb.com/>, <http://www.microsoft.com/siteserver/site/>, <http://world.isg.de/World/2_Internet/Visual_Web/index.html>, <http://www.incontext.com/WAinfo.html>

[17] <http://footprints.media.mit.edu/>

[18] <http://www.dcs.gla.ac.uk/~matthew/>

[19] <http://www.dynamicdiagrams.com/>

[20] <http://www.hyperwave.com/>

[21] See his homepage at <http://www.textuality.com/>.

[22] Semantic Constellation can be access from Chen's homepage at <http://www.brunel.ac.uk/~cssrccc2/>.

[23] <http://www.cc.gatech.edu/gvu/internet/MosaicG/MosaicG_1.0_about.html>

[24] <http://www.gate.cnrs.fr/~zeiliger/nestor.htm>

[25] <http://www.sgi.com/software/sitemgr.html>

[26] Apple ended research into HotSauce and stopped supporting it in 1997. However, the plug-in and documentation are still available from the Lightbulb Factory, see <http://www.xspace.net/hotsauce/>.

[27] See <http://www.w3c.org/RDF/> for more information on RDF.

[28] <http://www.perspecta.com/>