Visual elements in search and information retrieval systems

Mari Carmen Marcos

Citación recomendada: Mari Carmen Marcos. *Visual elements in search and information retrieval systems* [en linea]. "Hipertext.net", num. 3, 2005. <http://www.hipertext.net> [Consulted: 12 feb. 2007]. .

"The next generation of animated GUIs and visual data mining tools can provide users with remarkable capabilities if designers follow the Visual Information-Seeking Mantra:

"Overview first, zoom and filter, then details-on-demand.

Overview first, zoom and filter, then details-on-demand.

Overview first, zoom and filter, then details-on-demand.

Overview first, zoom and filter, then details-on-demand..."

(Ben Shneiderman)

1. Introduction

Using this mantra Ben Shneiderman prophesises in many of his articles and conferences what the next generation of data recovery interfaces will be like : a general vision first , followed by a closer approach and a filter, and finally details about the part of interest . Shneiderman thus synthesises four stages of the process of information search, and to each of these he adds a visual component:

The user comes to the system and sees the overall information he can find there, what subjects are contained within, and how it works.

Once viewed in a general way, the user zooms in, that is to say, he centres in on one part that is of major interest.

In order to refine his search better, he applies a filter, so that the results obtained better match his information needs.

To end the process, the user asks for more detail from some of the results in order to determine if it will be of interest.

The mantra refers to "animated GUIs and visual data mining tools", and we will emphasise the visual in the rest of this article. The area of information visualization attempts to offer visual representations that communicate information in a speedy and effective way. Diverse techniques have been developed to achieve this objective and to make the process of data recovery a more comprehensible job and on occasions, more interactive. We present the most used of these:

- use of icons to represent concepts

- use of colours and textures to highlight or differentiate elements
- graphics with hierarchical presentations that are easy on the eye
- maps that present information arranged according to similitude, using grouping techniques
- zoom effects to show detailed information
- animation and perspective in three dimensions

Visualization when applied to information helps people form a mental image of the information space. If the visualization takes place in an interface whose objective is information rettrieval, the expression used for this type of system is *Visual Information Retrieval Interfaces (VIRIs).*

The four processes named previously are susceptible to being presented through visual components. In fact, in the sections that follow we are going to see some systems of information presentation - mainly projects that have not been commercially implemented - that will serve as examples tshowing these elements of visualization. They have been classified into three groups: overall vision of the collection, visualization of results, and visualization of the attributes of each document.

What remains is to tackle the final phase of the process of search and recovery of information, that which links the moment of presentation of results obtained by the system with the query. This deals with the reformulation of the query, which at the present time does not use visualization, for which reason this item will not be dealt with.

2. The collection visualization (overview)

In the case of the complete collection visualization (overview), the objective is to prevent the user from being confronted with a "query box" in which he must express his information need in the best possible way, using descriptors.

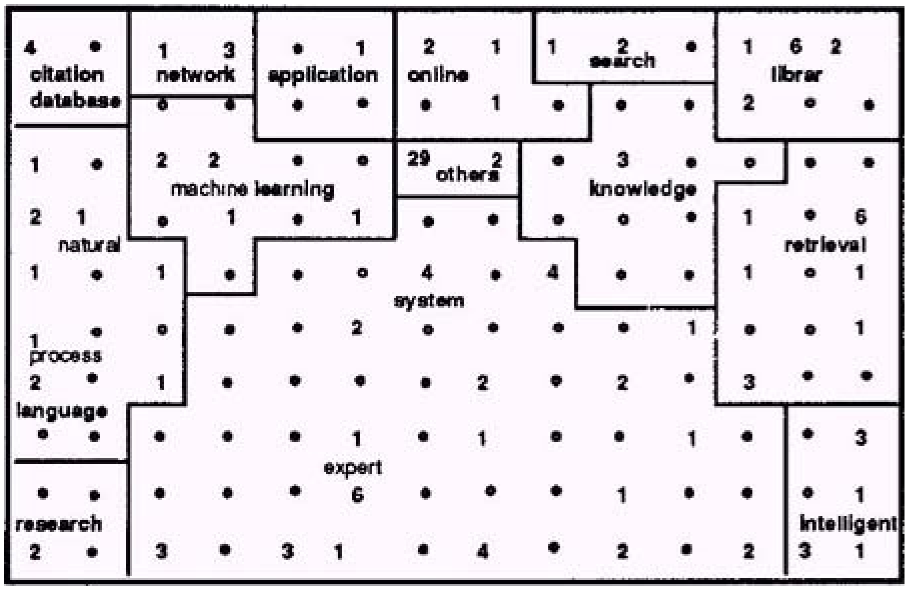
The difficulty that this action presents for the user is that it leads us to think of solutions of the type where the system presents a general vision of content from a database and the user selects where to continue searching deeper. Accordingly the cognitive charge that this falls on is a lot less if the user selects items (openings, themes...) than if he formulates the search.

The presentation technique of complete databases by antonomasia is the map (understood as cartographic representation, not as textual hierarchy): in one single screen the user obtains a global image of the collection and from this he interacts in order to drop down to more specific levels following a hierarchic and/or associated route.

2.1. Self organizing maps (SOM)

Lin (1992, 1997) proposes visualization *a posteriori.* Applying the incremental learning algorithm of Teuvo Kohonen — researcher of neural networks— , Lin creates self organizing maps (SOM) based on signs. The areas are proportionate in size to the frequency of appearance of the terms and can be shown in different colours. The terms that co-occur with most frequency are placed in the same area or one nearby. In order to see what each point refers to, it is clicked on and the information of the document is obtained.

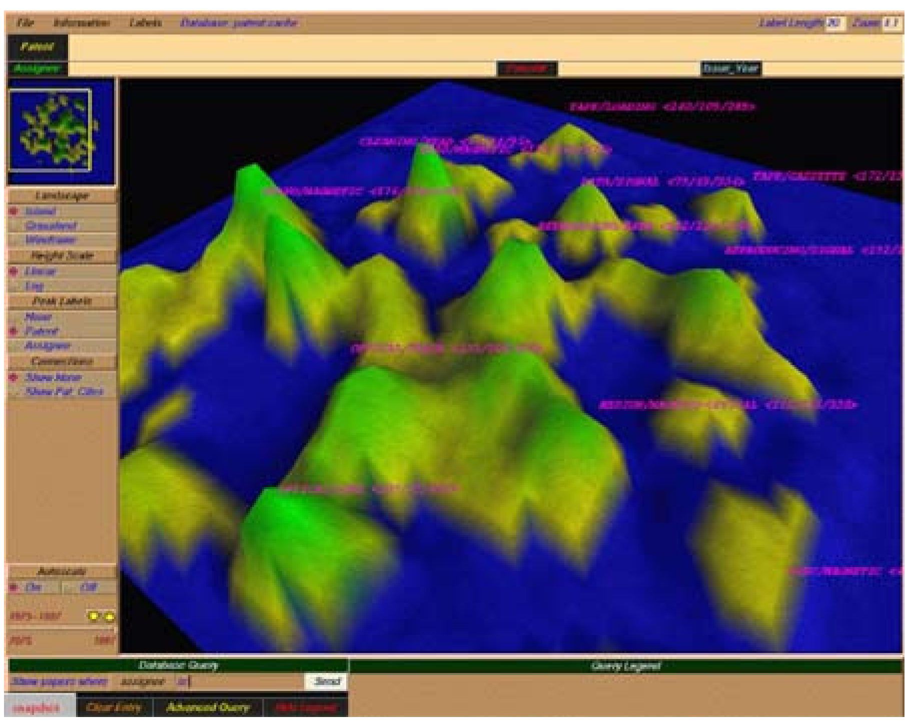
Figure 1. SOM type map (Lin, 1991), made from the results of a search in the LISAdatabase, http://faculty.cis.drexel.edu/~xlin/fulltext/ACM91.pdf



2.2. VxInsight

The Sandia National Laboratories use the metaphor of the islands of Polynesia in their system VxInsight to present scientific bibliography based on citation. In this system, the instructions understand mountain ranges, and it aims to be a way of detecting new areas of study and those connections between them (flying overhead in a virtual reality environment) and of seeing the subcomponents of the disciplines (flying at a lower level), even the titles of the articles located in the zone (flying at an even lower level). The use of these metaphors is not new. Figure 2 shows an example.

Figure 2. VxInsight. From the Sandia National Laboratories, Visualizing using the mountain terrain metaphor. http://www.cs.sandia.gov/projects/VxInsight/snapshot.html



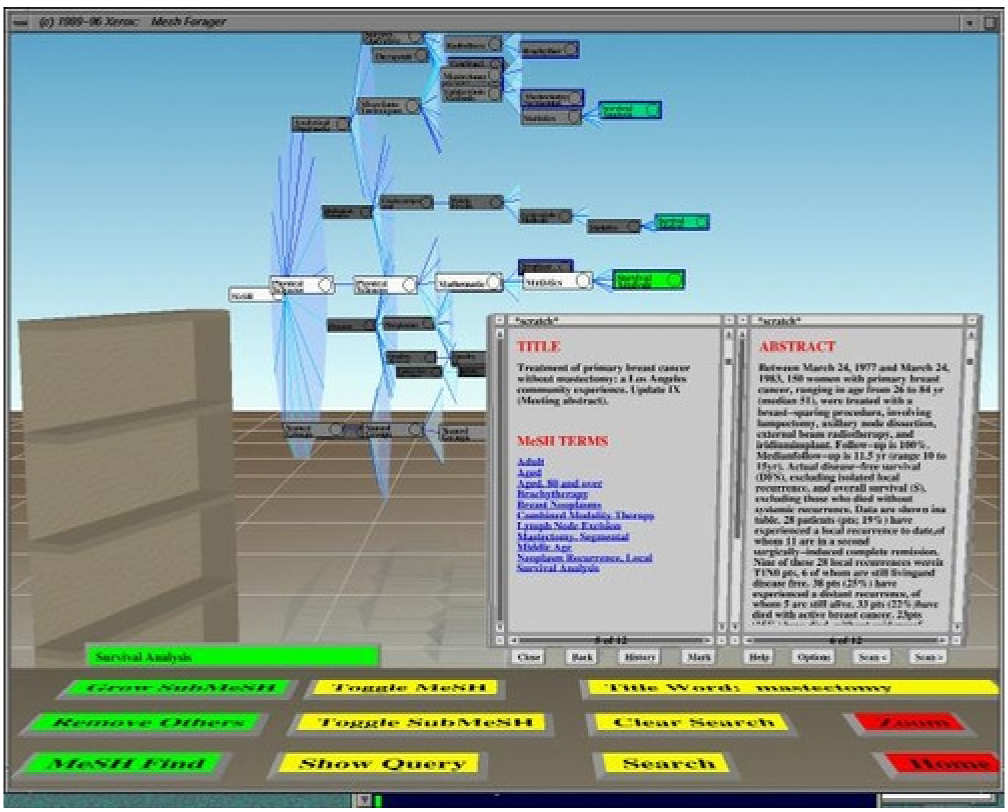
2.3. Cat-a-Cone

http://www.sims.berkeley.edu/~Hearst/cac-overview.html

Cat-a-Cone also uses the presentation of hierarchical categories, but they do this simultaneously with documents resulting from a search and use 3-D animation where hierarchies are presented in the form of cones that rotate in order to enable us to see what is behind them. As can be observed in figure 3, its interface shows a hierarchical theme in which the user can browse and select material, and at the same time the records of the retrieved documents are shown on the screen in the form of an open book.

The objective of this system is to facilitate document retrieval using subject headers and to not only offer the thematic structure but also the results in the same interface (Hearst, Karadi, 1997).

Figure 3.Cat-a-Cone. From Xerox Parc, presents thematic hierarchy with cones showingthe document that the user marks at the same time.
http://www.sims.berkeley.edu/~Hearst/papers/cac-sigir97/survival.jpg



Visualisation of the set of results (preview)

In the case of visualization to present search results, similar solutions can be opted for to those presented when a global vision is offered of the collection, but with two peculiarities: on the one hand, the number of documents will be less (therefore, the algorithm calculation to bring about visualizations is simplified); on the other hand, the relationships between documents will vary according to those which have been selected, which means that the presentation will vary in each set of results (it is a dynamic presentation, not static as could occur when visualizing the entire collection) and this complicates the algorithm calculations.

2.4. BEAD (Chalmers; Chitson, 1992)

This is a project for visualizing complex multivariable documents in information retrieval. In this case, the comparisons among documents are made by drawing upon a great number of variables. These conditions are not suitable for the methods of visualization and analysis of known graphics. for this the prototype presented of browsing information based on graphics used in a three dimensional space in which articles are represented as particles (this refers to the metaphor of particle magnetism) :

The relationship between articles is represented by relative position in space; the force between particles tends to place the most similar articles nearer than those least similar. The result is a three dimensional scene that can be used to visualize models in a multidimensional space (see figure 4).

Figure 4. BEAD. In this example of BEAD, the main cluster of retrieved documents is presented with the descriptor "Information Retrieval",upon a zoom can be carried out, http://www.dcs.gla.ac.uk/~matthew/papers/sigir92.pdf

In this way, BEAD uses the techniques of physical simulation in order to model documents as particles in three dimensional spaces, specifically it is the technique used to simulate the burning of metals in order to improve the algorithm of multi-dimensional scale (MDS) for grouping of documents. Once the documents are grouped in clusters, the user introduces search terms and the system responds with colour codes in each document according to relevance with respect to the query. The user can zom in on the articles that appear interesting. Unfortunately, the result of the experiment undertaken in 1992 was too slow: 18 minutes for 100 particles.

2.5. SPIRE (Spatial Paradigm for Information Retrieval)

In the SPIRE (Spatial Paradigm for Information Retrieval) project of the Pacific Northwest National Laboratories (http://www.pnl.gov/infoviz), they have developed diverse metaphors of visualization. We shall comment on two of these. The first, Galaxies (see figure 5), makes documents appear as stars and they are grouped together as if they were constellations, and then these are grouped together, and all uses the co-occuring statistic as a base. In this way, one look is enough to easily see the themes and to select the one that interests you, and once inside this zone to delve deeper to see which documents each cluster includes.

The other metaphor is that of landscape and it is used in ThemeView (previously called Themescape). In this case, the body of documents is taken and represented by valleys and mountains based on the statistical frequency of the key words: the more relevant a document, the higher the mountain that represents it; valleys represent the absence of key words.

The innovation that the authors have incorporated is that the body of documents is about 20,000 and can be presented by maps in line (see figure 6).

Figure 5. Galaxies. Presents documents grouped together by similarlity, http://www.pnl.gov/infoviz/galaxy800.gif

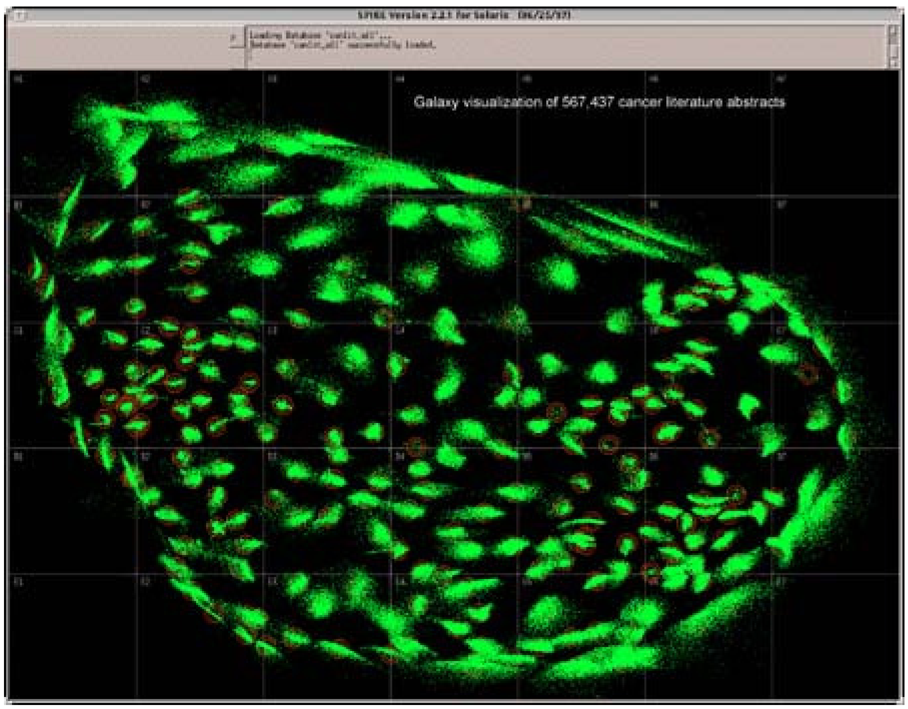
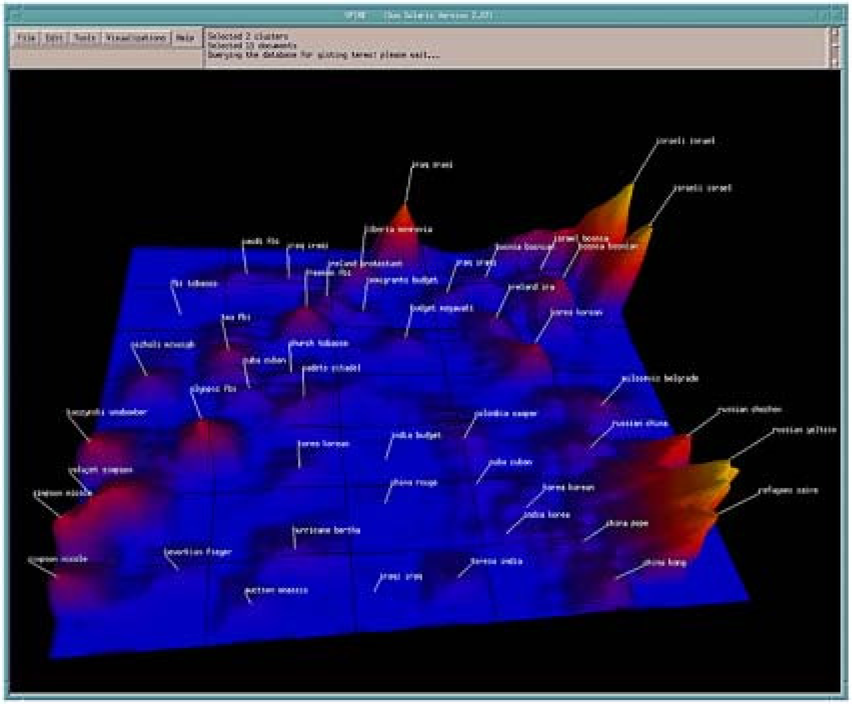


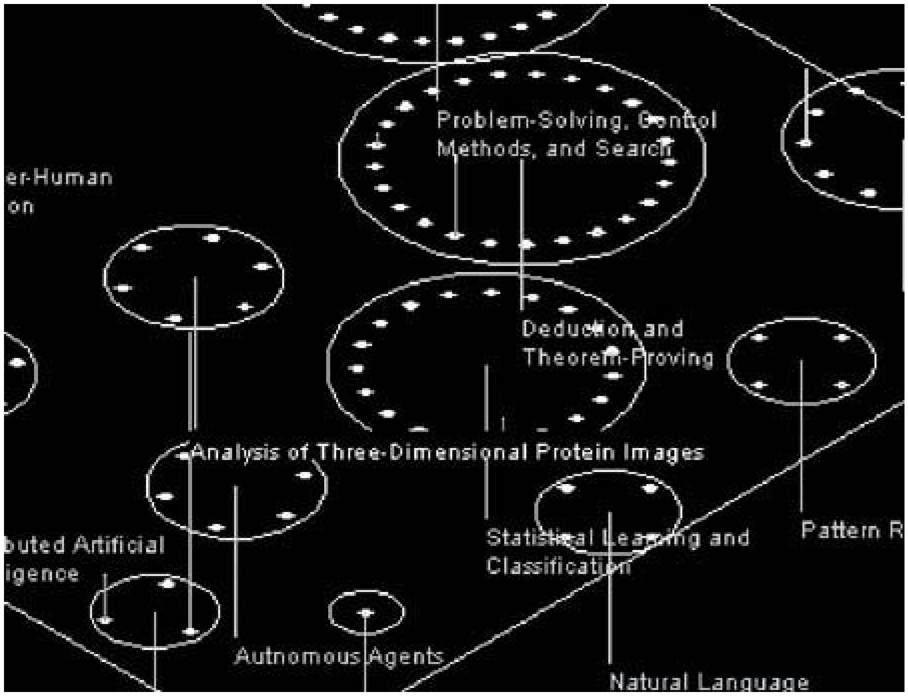
Figure 6. Themeview. From the Pacific Nortwest National Laboratories, representation of valleys and mountains based on the statistical frequency of keywords, http://www.pnl.gov/infoviz/themeview800.gif



2.6. JAIR

JAIR (Foltz, 1997) was created to offer access to the *Journal of Artificial Intelligence Research* (JAIR) to specialists interested in the subject. Accordingly, we start from the idea that this is not a referential, but a complete text system, and directed at people who know the specific subject vocabulary about which information is given. In this case, a bidimensional respresentation of the documents was opted for even though a third dimension was contemplated in order to offer additional information. By means of the 2D representation the user can centre his interests on a section of the results (focus) and have the impression that he is exploring an object instead of seeing only an image.

Figure 7. JAIR. Classifies the articles in categories taken from the ACM system of classification, http://www.infoarch.ai.mit.edu/JAIR/JAIR-help.html



In JAIR, the design of the presentation of icons has not taken into account the

queries stored, nor a similarity of terms to create vectors, but has opted for classifying the articles in categories taken from the classification system of the 1991 ACM. With this reference a hierarchical structure of two levels has been created: the first positioning among categories and a second level of documents within categories.

The same article can be included in two categories. This structure is similar to that of the "cone tree" (Robertson, Mackinlay y Card, 1991) but without being multilevel. The articles are fitted into the presentation in an equidistant way so that the centre of each category and the radius of this is proportional to the square root of the size of the category, which means the area is proportional to the size (see figure 7). Kruskal's algorithm of mulitdimensional scale (1964a, 1964b) is used to position the centres of the categories and find a configuration in which the order of dissimilitude is shown in the distance between categories, but it has to be modified to stop overlapping between categories. The result is that the more articles share categories the more they are found there.

JAIR has also resorted to the use of colours to emphasise elements of visualization. As far as user options are concerned, we can point to complete text search, the listing or search history, the option of modifying searches so that they are found in the search history in order to create new queries, and the ability to visualize the reference, the summary or the complete text according to what is desired in each document. Besides this, the automatic access of a document is shown in a list of related articles.

2.7. NIRVE

A very significant example of document representation obtained by a query of a data retrieval system by means of clustering is the case of NIRVE (NIST Information Retrieval Visualization Engine) (Cugini, Laskowski y Sebrechts, 2000) (http://zing.ncsl.nist.gov/~Cugini/uicd/NIRVE-home.html). In this system, each recovered document includes the title of the document, a unique identifier, its relevance with respect to the query, the size of the document and the number of occurrences of each descriptor that is assigned to it.

The screen is arranged in two windows: a document window with titles, clusters, descriptors and concepts; another one of operating controls that cannot be shown in a normal way in the document space (that is to say, in one the graphic and in the other the options). The user can vreate a set of descriptors of the query using a concept and a name and a colour can be assigned to that concept. The association of decriptors and concepts is shown with the interactive legend in the bottom part of the document window (below the graphic).

A profile is created for each document according to the frequency with which the descriptors appear for this concept; in this way a cluster would be a set of documents that share a subgroup of concepts. Each cluster is represented as a box with coloured bars that indicate the profile of the concept by means of its documents; the more documents the cluster contains, the bigger the box is represented (see figure 8).

Figure 8. NIRVE. Representation in 3D of the clusters in form of coloured boxes that indicate the concept by means of the documents that shape it, and whose size is an indication of the number of documents that it contains. http://www.itl.nist.gov/iaui/vvrg/Cugini/uicd/NIRVE-paper.html

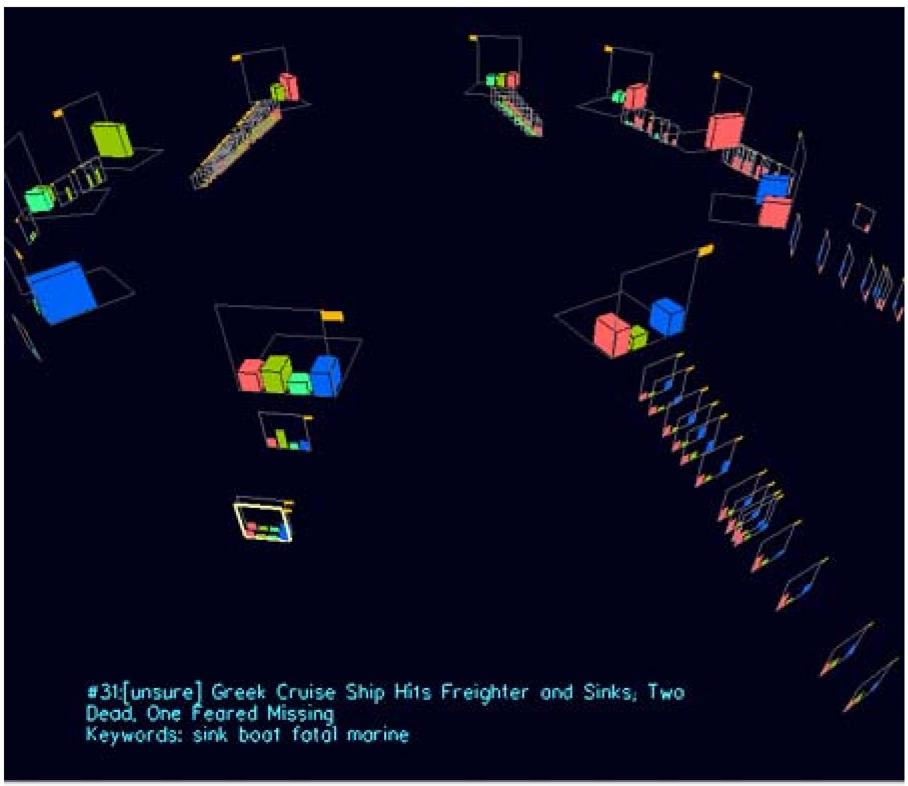
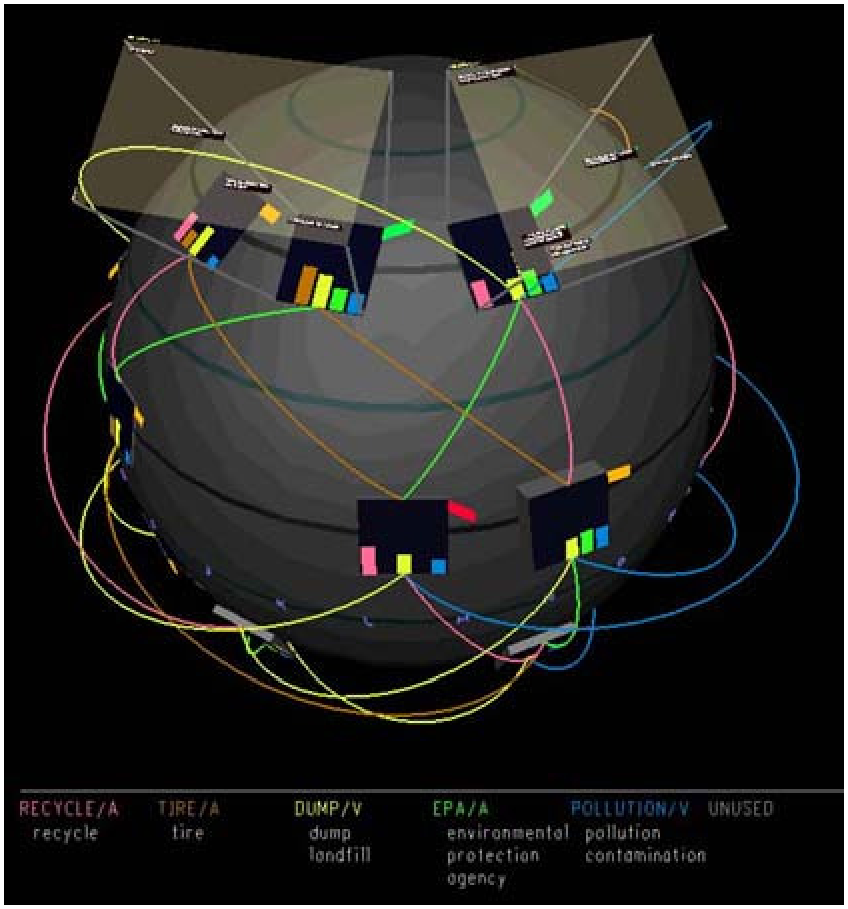


Figure 5: Spoke and Wheel Model

The representation of NIRVE has been tested with different metaphors, for example that of the globe, in which the icons of the clusters are fitted on the surface of a globe (a sphere, see figure 9) where the latitude is determined by the number of cluster concepts (the more concepts there are, the nearer the north pole). The icons are arranged in such a way that the clusters with most concepts in common are closer to each other. If two clusters are different by only one term, then an arc is drawn that joins them and that will have the colour assigned to a different concept.

Figure 9. NIRVE, Globe metaphor, http://www.itl.nist.gov/iaui/vvrg/Cugini/uicd/NIRVE-paper.html



When the user opens a cluster, a 2D rectangle is projected that contains the titles of the documents that form it; these titles are ordered in a way that the most similar documents are displayed close horizontally and the most relevant ones are found higher up in the vertical. If the user selects a title, he can open the complete text in the web browser and see how the descriptors appear prominently in the colour that corresponds to the concept to which reference is made. Also, documents can be marked and even entire clusters indicated if they are of more or less interest.

2.8. PRISE

PRISE (Cugini, Piatko, Laskowski, 1996) applies clustersing and vectoring techniques , and achieves with them visual presentations different from the usual listing. The idea is to show the listing of documents in different ways, with a line in form of a spiral following the order of relevance, or with a three axis graphic , each of which corresponds to a descriptor, thus showing the space where they meet; or lastly by means of a representation of "vicinities" between terms. The figures 10, 11 and 12 show an example of each type.

Figure 10. PRISE. Presentation of documents ordered by relevance according to a line or spiral, http://zing.ncsl.nist.gov/~Cugini/uicd/viz.html

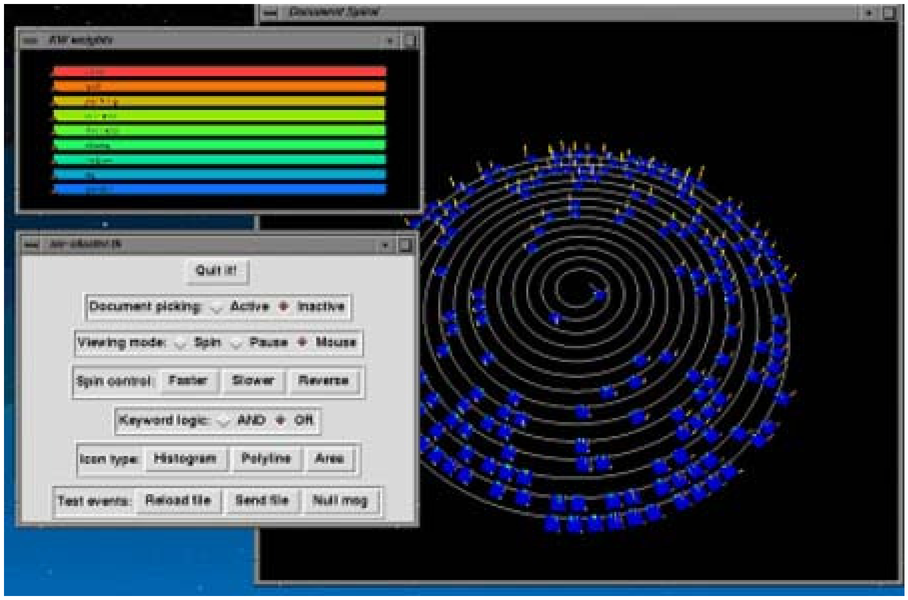


Figure 11. PRISE. Shows each descriptor in an axis and creates a three-dimensional space in which the documents are represented, http://zing.ncsl.nist.gov/~Cugini/uicd/viz.html

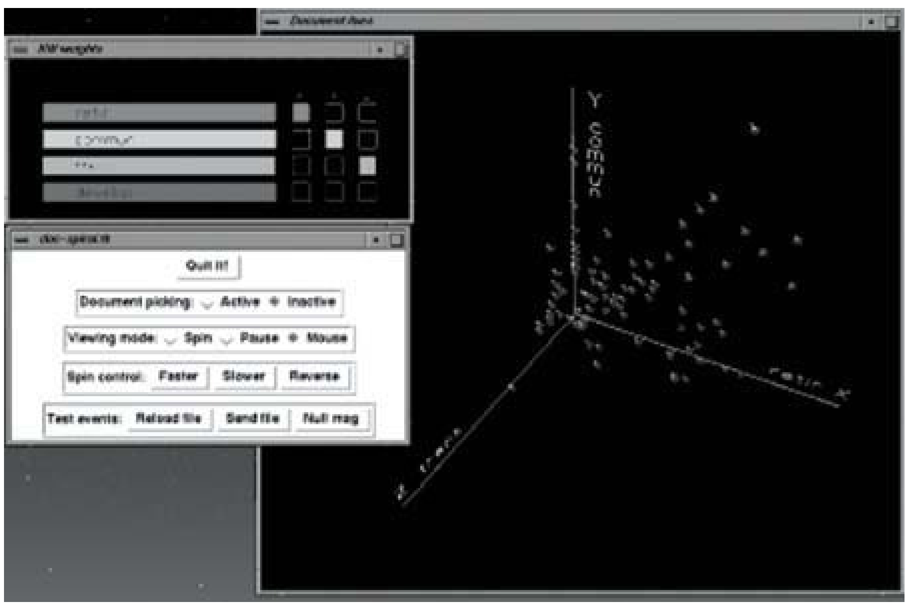
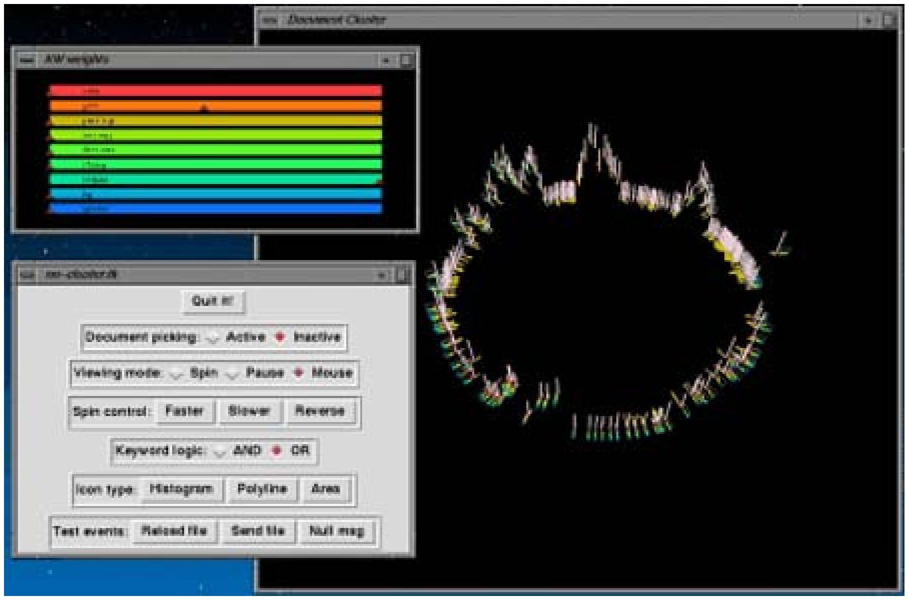


Figure 12. PRISE. Groups the documents in this representation according to the degree of vicinity, http://zing.ncsl.nist.gov/~Cugini/uicd/viz.html



2.9. Perspective Wall

Perspective Wall (Mackinlay, Robertson, Card, 1991) (http://people.cs.vt.edu/~North/infoviz/perspectivewall.pdf) is one of the best-known inside the world of visualization. It is developed by Inxight. This system simulates a three-part wall that allows the user to bring information to the front of the screen, while the other two parts remain in the background. The data is shown in two dimensions and the user can order the documents on the abcissa axis by another different criteria (see figure 13) .

Figure 13.Wall perspective. The user brings the information that he is interested in seeing to the front wall and leaves the other context in the lateral walls, http://people.cs.vt.edu/~North/infoviz/7

2.10. VIBE

In VIBE (Olsen et al.; 1993; Korfhage, 1991, 1997) (http://www2.sis.pitt.edu/~WebVIBE/WebVIBE/page02.html) the presentation is made using a map. This system shows the user's search state rather than the attributes of all the collection, and does it in two dimensions. The user starts by defining at least two points of interest (POI) with keywords that represent the conceptual space he is looking for, and this creates an icon that is placed inside some coordinates. With this, the documents of the collection are compared using the theme specified by the user, they are given a vector of points according to how they correspond to this theme and are represented also by icons. To obtain bibliographic information about whichever icon simply click on it.

Accordingly, VIBE represents search results in two dimensions where the location of documents on the screen is determined by their relationship to the POIs. The similarity between the documents is reflected in the position of the rectangles (documents) to the circles (POIs). Figures 14 and 15 show an example of how the results are presented. In the retrieval all those documents that contain at least one of the descriptors indicated by the user when initiating the search are considered relevant. The documents with greater coincidence in their descriptors with those of the POI are placed closer to that POI, but always in relation to the rest of the documents.

The user can also expand the icons of the documents or documents that are useful by simply drawing a box around a document or documents that are of interest and a list is shown of the chosen selection. Clicking with the mouse on any of the documents on the list will open another window with the complete document.

One characteristic that makes the system interactive is that the user may add, change or remove the POIs from the screen. To add one, he needs to look it up on the list of POIs or indicated terms, choose one and move it with the mouse to the desired location on the results screen. On carrying out any of these changes the system automatically launches the search query and reorders the found documents to present the relationships between documents and those between POIs.

Figure 14. VIBE. This screen shows the results of a query made from five terms. Some icons of documents have been linked and the titles of these documents are shown in the dialogue box (Morse, Lewis, 1997), http://www.itl.nist.gov/iaui/vvrg/emorse/papers/smc97.html

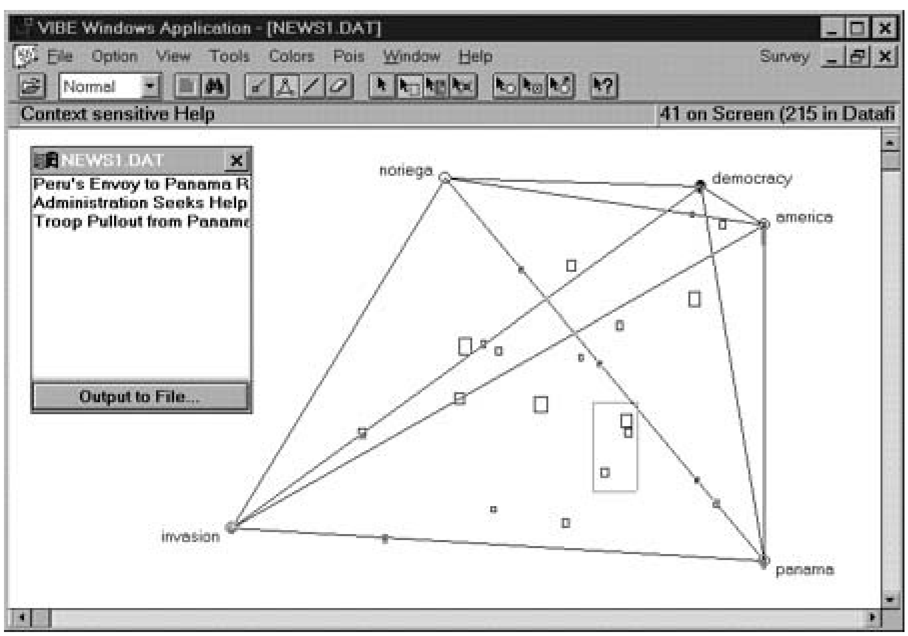
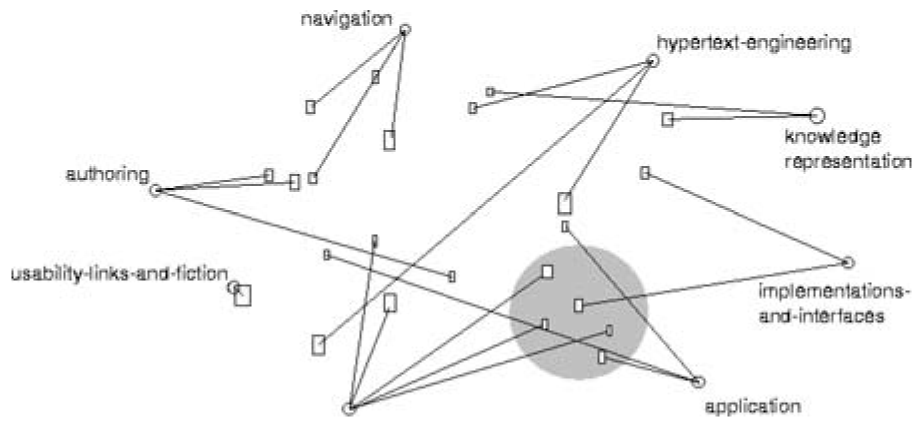


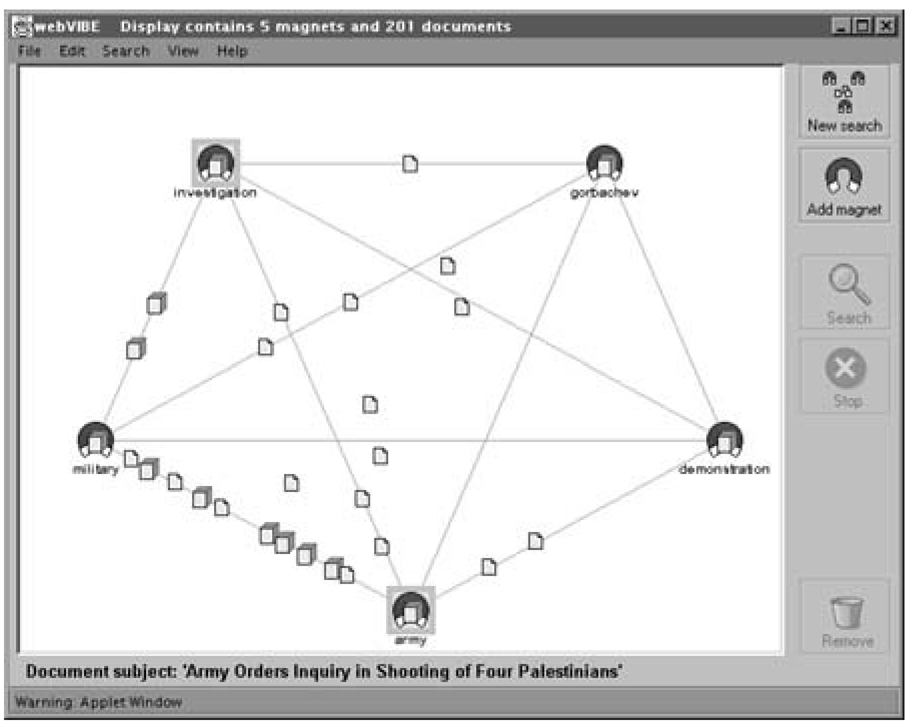
Figure 15. VIBE. Search results with VIBE (Image taken from Hearts, 1999), http://www.sims.berkeley.edu/~Hearst/irbook/10/node7.html#fig:VIBE



WebVIBE

Web-VIBE, as Morse & Lewis (1997) explain, shows the POIs through a magnetism metaphor and converts them into magnets that attract documents, that are found in a magnetic field. The position of a document is determined by the force of attraction of the terms present in a document, as shown in the figures 16, 17 and 18.

Figure 16. WebVIBE. Result of a query with five terms. Two icons of points of interest (POI) are notable to indicate what must be considered as relevant criteria. (Morse, Lewis, 1997), http://www.itl.nist.gov/iaui/vvrg/emorse/papers/smc97.html



An online version exists, developed in Java, with the end user in mind (http://www2.sis.pitt.edu/~WebVIBE/WebVIBE/). At the moment it is in evaluation. You can try it on the Web and send the questionnaire to the developers to help improvements.

Figure 17. WebVIBE. Consulting screen with a Java Applet on the Web. http://www2.sis.pitt.edu/~WebVIBE/WebVIBE

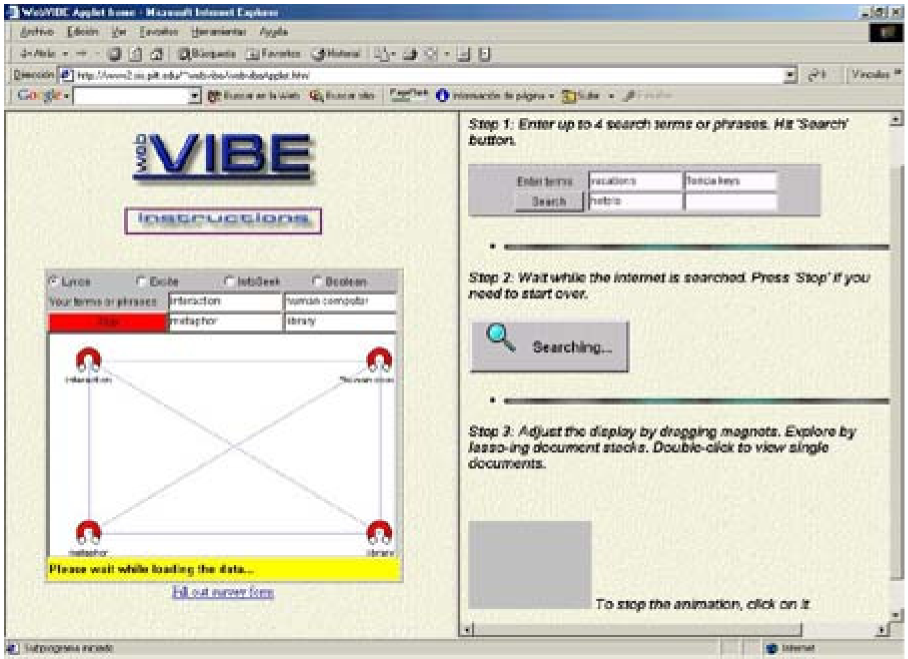
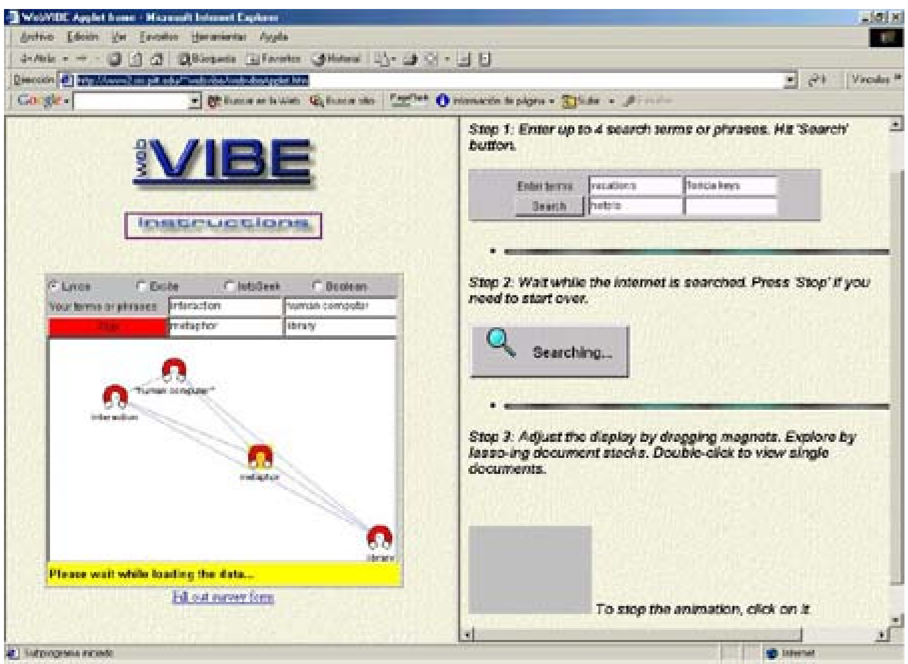


Figure 18. WebVIBE.Readjusting the magnerts (POI) by user, http://www2.sis.pitt.edu/~WebVIBE/WebVIBE



VR-VIBE

VIBE has also developed VR-VIBE (Benford et al., 1995) (http://www.crg.cs.nott.ac.uk/research/technologies/visualisation/vrvibe/), which represents the information in three dimensions and LyberWorld, designed for searching the complete text of documents (Hemmje et al., 1994).

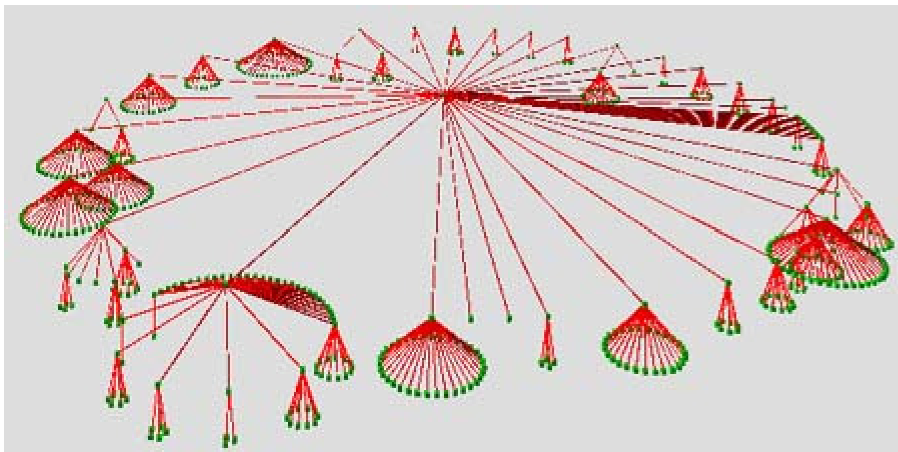
VR-Vibe has been developed by the Communications Research Group of the University of Nottingham. It has been applied to workgroup visualizations called "Populated Information Terrains". In it is defined four ways of generating visualizations:

- Co-ordinate maps in which data from orders and from abscissas is introduced.

- Hyper-structures in which links between objects are shown, for example, the scheme based on the entity-relationship model , or from the actual Web. The resulting structures can be visualized with graphical techniques in three dimensions and fish-eye techniques applied, cone-trees (figure 19) and Perspective Wall.

Figure 19. Cone Trees. Technique to visualize hyperstructures,

http://www.crg.cs.nott.ac.uk/research/applications/pits/

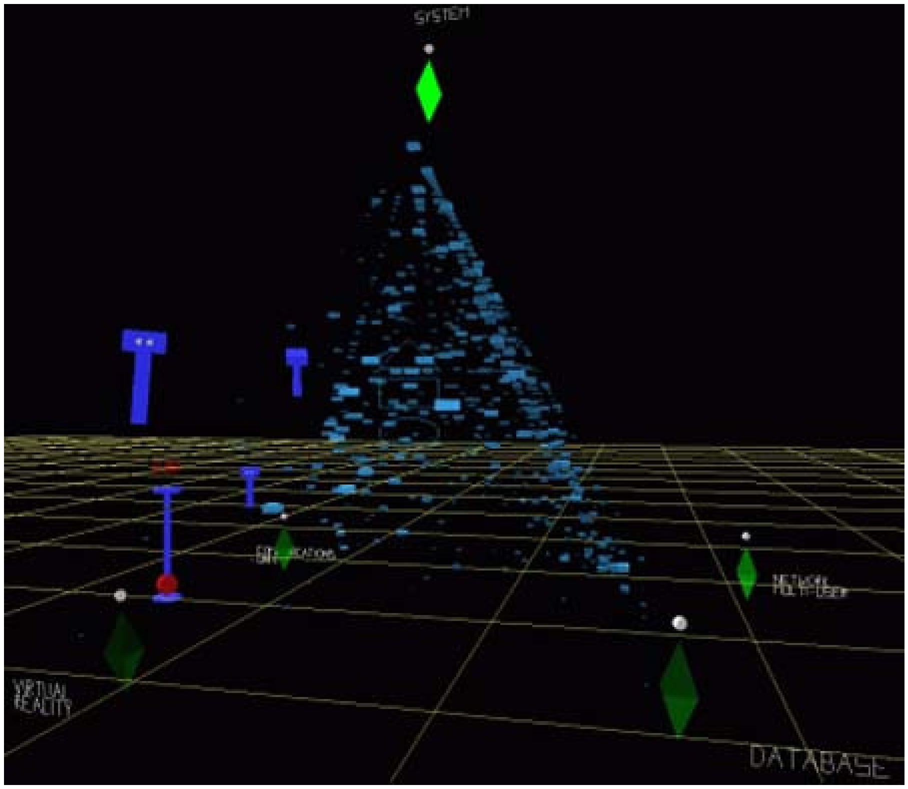


- The design of the visualization is by means of two methods: the first consists of creating real world metaphors, such as the library — in a document database —; cities, buildings and rooms — to organise information in its own way, for example its records on his computer's desktop.In the second method, people construct and organise information as they want to, like for example their folders in the "desktop" of their computer.

- Method statistics to analyze collections of data that are grouped by similarity.

VR-Vibe goes one step further than its predecessor VIBE and creates bibliography visualizations in three dimensions. The user introduces the terms of query and the result appears represented in a 3D space in which the retrieved documents are placed between the query descriptors, more or less near according to the relevance of each one with respect to each descriptor. The colour of the documents also indicates their degree of relevance with respect to the descriptors (the clearer they are, the more relevant they are). Figure 20 shows two examples of this system.

Figure 20. VR-VIBE. Two examples of visualization of retrieved documents in a 3D environment , http://www.crg.cs.nott.ac.uk/research/technologies/visualisation/vrvibe



3. Visualization of document attributes

Finally, we present the TileBars system, which shows attributes of retrieved documents by means of visualization.

3.1. TileBars

Some systems give added information with aspects that usually are not considered. In this way, TileBars (Hearst, 1995) starting with the idea that on retrieving documents based on title and summary useful information is missed, for which it defends the idea of full text search. The author also considers that it is important that the user knows the extension of the retrieved document and the frequency of a descriptor in different parts of the structure of the text.

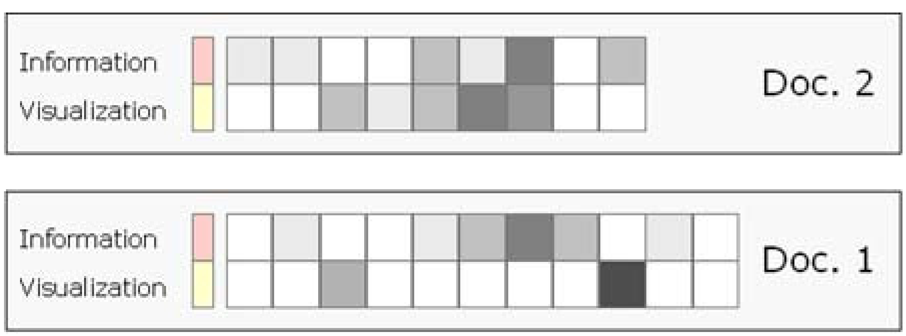
TileBars offers this information in a graphical way, showing the result of a document title search accompanied by a graphic results that consists of a rectangle in which its length shows the extent of each document and at the same time is divided into three rows, one for each search expression created by the user. Inside the rectangle appear four squares, lighter or darker according to the frequency of appearance of each expression in each part of the document (see figure 21 for the explanation of this bar and figures 22 and 23 as examples of results in a medical database and in a digital library respectively.

Previously, an algorithm called TextTiling has split the document into various parts. In this way, if the first search term appears very frequently at the beginning of a document and less frequently in the middle and again very frequently at the end, the rectangle will show, in the first row (because these correspond to the first search term), dark squares at the beginning and end of its length and fainter squares in the middle section.

This method combines the ordered list of relevant titles with a graphic that gives compementary information about each retrieved document. It is ready to work with full text documents. This system also forms part of the user's query and the obtained results.

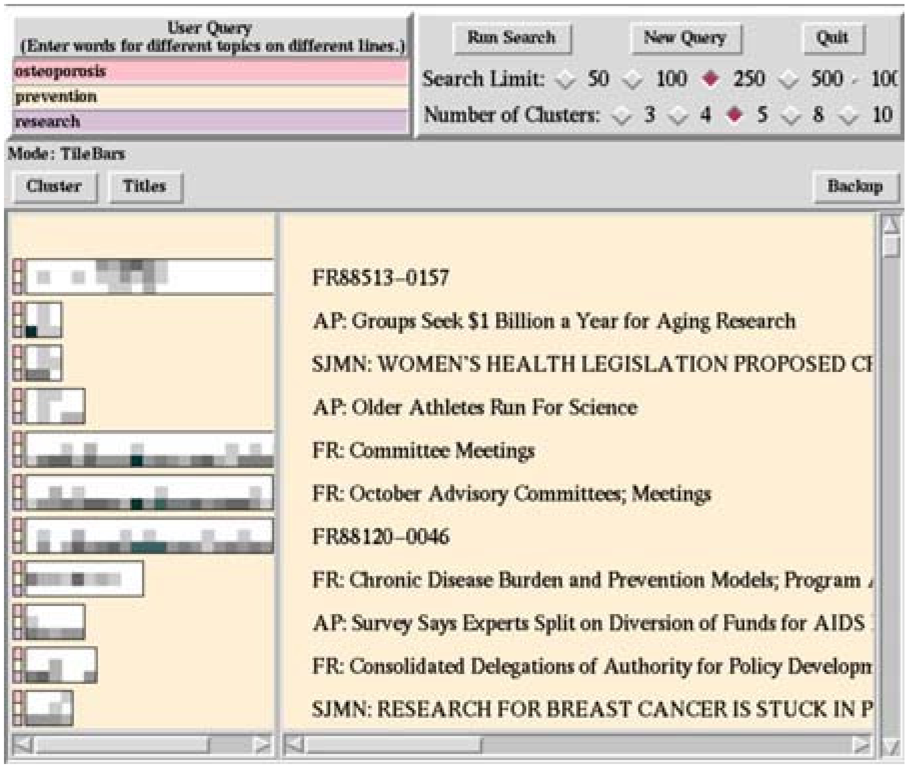
The documents are represented as horizontal bars in which the first words of the title are shown; these bars contain some segments that represent the parts of the documents; the segments are coloured with greater or lesser darkness in grey scale, which indicates the frequency with which a search term coincides with different parts of that document.

Figure 21. TileBars. The upper row indicates the frequency of the word "Information" in each section of the document, the lower one corresponds to Visualization. In document 2, less extensive than 1, there are three sections in which both words coincide, http://www.infovis.net/Revista/2002/num_104.htm



So, a white section indicates that there are no common occurrences, a grey one gives between 1 and 7 times, and black one gives 8 or more times. The concurrence that could be shared between the documents (that is to say, coincidence of terms) is shown in the segments by vertical lines.

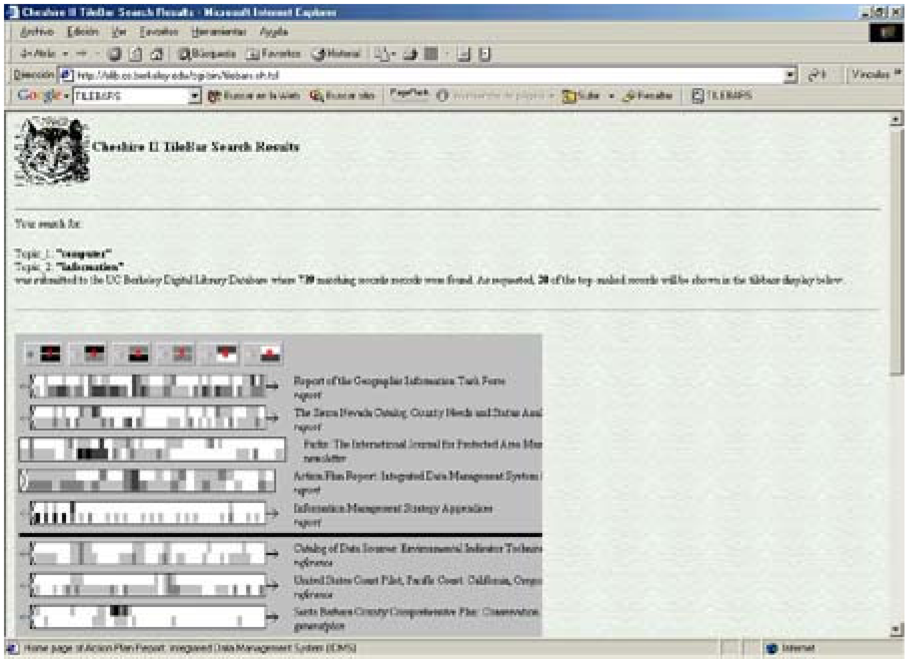
Figure 22. TileBars. Results of a medical database with a three-term query. The bars are presented on the left, and the results on the right. http://www.sims.berkeley.edu/~Hearst/tb-example.html



Recently, the Berkely Digital Library Project (Cheshire II) has implemented TileBars (http://elib.cs.berkeley.edu/TileBars/help.html) with some variations in order to adapt to very large documents. So, instead of using the cited algorithm to split the document in parts, it has opted for dividing them into pages. If the documet does not contain any of the search terms in several consecutive pages, that part is not shown complete with squares, but is substituted by a square that contains a cross.

If after this the rectangle that represents the document is still larger than that which can be represented on the screen, the width of the boxes is reduced , and as a last resort, if it is still is too big, scroll bars are included to allow stretching of the rectangle.

Figure 23. TileBars. The catalogueof the digital library Cheshire II presents its results by means of this system, http://elib.cs.berkeley.edu/TileBars/



4. Conclusions

With the techniques that this article has shown, the research aims to offer solutions to the greatest problem of representation of information, which is not so much technical but cognitive, that is to say, showing the user of the system a model that he is capable of understanding with minimum mental effort and digesting it in order to be able to manage it without problems. If the user does not manage to create an adequate mental model of the system, that is, a correct abstraction of it, the representation that is offered will not have fulfilled its objectives. Because of this, it is necessary to continue developing applications that improve communication of the system with the user, ultimately, the Human-Computer Interaction.

A great field of research in the application of visualization of data retrieval exists. In fact, it has to be said that at the moment the majority of information retrieval systems that incorporate techniques of visualization are prototypes made in research laboratories. On the other hand, the investigation into information visualization still has a lot to study and evaluate in order for satisfactory results to be obtained that can compete with more traditional products of information retrieval.

As can be observed in the given examples, the majority of systems that use visualization techniques to show information have been developed to work with full text databases, or at least with summary fields, which enables the counting of a sufficient quantity of terms to make an automatic treatment of language and establish categories by similarlty. In the case of reference databases, the information content of the document is only found in title fields (with problems of ambiguity that natural language carries) or in the descriptor norms (descriptors or materials and classification codes) that which makes establishing similar criteria more complicated.

5. Bibliography

Chalmers, M.; Chitson, P. (1992). "BEAD: explorations in information visualization". In: Belkin, N.; Ingwersen, P.; Pejtersen, A. (eds.). Proceedings of the Association for Computing Machinery Special Interest Group on Information Retrieval. SIGIR '92. Copenhague, 1992. New York: ACM, 330-337, http://www.ACM.org/pubs/articles/proceedings/ir/133160/p330-chalmers/p330-chalmers.pdf

Cugini, J.; Laskowski, S.; Sebrechts, M. (2000). "Design of 3-D visualization of search results: evolution and evaluation". *Proceedings of IST/SPIE's 12th Annual International Symposium: Electronic Imaging 2000: Visual Data Exploration and Analysis (SPIE 2000)* , http://www.itl.nist.gov/iaui/vvrg/cugini/uicd/nirve-paper.html

Cugini, J.; Piatko, C.; Laskowski, S. (1996). " Interactive 3D visualization for document retrieval". *Proceedings of the Workshop on New Paradigms in Information Visualization and Manipulation. ACM Conference on Information and Knowledge Management (CIKM '96)* , http://zing.ncsl.nist.gov/~cugini/uicd/viz.html

Dürsteler, Juan Carlos. Infovis. Consulted: 27-Jan-2005, http://www.infovis.net

Dürsteler, Juan Carlos (2000). *Visualización de información* . Barcelona: Gestión 2000. ISBN 84-8088-836-9.

Foltz, M. (1997, act. 2002). An information space design rationale. Consulted: 27-Jan-2005, http://www.infoarch.ai.mit.edu/jair/jair-rationale.html

Hearst, M. (1999). "User interfaces and visualization". In : Baeza-Yates, Ricardo; Ribeiro-Neto, Berthier. Modern Information Retrieval. New York: Adison-Wesley, pp. 257-323.

Hearst, M. (1995). "TileBars: Visualization of term distribution information in full text information access". In: Katz, I. et al. *Mosaic of Creativity: CHI '95: Proceedings of the Association for Computing Machinery Special Interest Group on Computer Human Interaction (ACM/SIGHCI). Conference on Human Factors in Computing Systems (Denver, May 1995)* . New York: ACM Press, 59-66, http://www.ACM.org/SIGCHI/chi95/Electronic/documnts/papers/mah_bdy.htm

Hearst, M.A.; Karadi, C. (1997). "Cat-a-Cone: An Interactive Interface for Specifying Searches and Viewing Retrieval Results using a Large Category Hierarchy". In: Belkin, N et al. (eds.). *Proceedings of the 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '97)* . New York: ACM, 236-245, http://www.ACM.org/pubs/contents/proceedings/ir/258525/ http://www.sims.berkekey.edu/~Hearst/papers/cac-sigir97/sigir97.html

Hemmje et al. (1994). "LyberWorld: A visualization user interface supporting fulltext retrieval". In: Croft, W.; Van Rijsbergen, C. (eds.). *SIGIR '94: Proceedings of the Association for Computing Machinery Special Interest Group on Information Retrieval (ACM/SIGIR). 17th Annual International Conference on Research and Development in Information Retrieval (Dublin, Ireland, July 1994).* London: Springer-Verlag, 249-259.

Information Visualization. Consulted: 27-Jan-2005, http://www.palgrave-journals.com/ivs/

Korfhage, R. (1997). Information Retrieval Bibliography. VIRI systems bibliography organized according to the name or type of visualization system, http://www.pitt.edu/~Korfhage/VIRI_bib.htm

Kruskal, J. (1964a). "Nonmetric multidimensional scaling: a numerical method". *Psychometrika* , 29:2, 115-129.

Kruskal, J. (1964b). "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis". *Psychometrika* , 29 :1, 1-27.

Lin, X. (1991). "A Self-Organizing Semantic Map for Information Retrieval", *SIGIR'91* , 262-269.

Lin, X. (1992). "Visualization for the Document Space". In: *Proceedings Visualization'92* , Boston: IEEE, 274-281.

Lin, X. (1997). "Map displays for Information Retrieval". *Journal of the American Society for Information Science* , 48, 40-54.

Mackinlay, J.; Robertson, G.; Card, S. (1991). "The Perspective Wall: detail and context smoothly integrated". In: P *roceedings of ACM CHI'91 Conference on Human Factors in Computing Systems* , New Orleans, Louisiana, April 28-May 2. ACM Press, 173-179.

Morse, E.; Lewis, M. (1997). "Why information visualizations sometimes fail". *Proceedings of the IEEE International Conference on Systems Man and Cybernetics (Orlando, Florida, October 11-15)* , http://www.itl.nist.gov/iaui/vvrg/emorse/papers/smc97.html

Olsen, K. A.; et al. (1993). "Visualization of a Document Collection: The VIBE System". *Information Processing & Management* , 29:1, 69-81.

Robertson, G.; Mackinlay, J.; Card, S. (1991). "Cone trees: animated 3d visualizations of hierarchical information". In: *Proceedings of ACM SIGCHI'91 Conference on Human Factors in Computing Systems, (New Orleans, Louisiana, April 28-May 2)* . New York: ACM Press, 189-194.

Shneiderman, B. (1992). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* . Reading (Ma): Addison-Wesley.

Mari Carmen Marcos is member of the DigiDoc Group of the University Institute of Applied Linguistic of the University Pompeu Fabra and the Scientific Section of Information Sciences of the Department of Political and Social Sciences of the University Pompeu Fabra . This article presents a part of the results of the HUM 2004-03162/filo project of National Plan I+D+I of the Ministry of Education and Science (Spain)