

HyWebMap, a system to create interactive networks

Imad SALEH*, **Nasserddine BOUHAÏ***, **Fabrice PAPY****

Université Paris 8, Equipe Hypertexte dynamique, Département Hypermédia,
Documentation***

2 Rue de la liberté, 93526 Saint-Denis Cedex 2,

Tél. : 01 49 40 67 58, Fax : 01 49 40 67 83

Email : isaleh@wanadoo.fr, nasser@labart.univ-paris8.fr, papy@labart.univ-paris8.fr

ABSTRACT: The Web is an open, evolutionary and heterogeneous information space. To exploit this heterogeneous space and better orientate the reader, we developed the HyWebMap system which allows: orientation in the space visited by using a benchmark tool, a browser, a history tool and a global image display based on an interactive map. The reader can create a personal space starting from the spaces visited and enrich it by adding and renaming links. HyWebMap allows the reader to add nodes and annotations to personalise his space. HyWebMap notifies and assists the reader when a change occurs in his personal space.

KEY-WORDS: virtual network, agents, annotation, navigation, reorganization, nodes, links, Web tools

1. Introduction

In this article, we introduce the software Web tool HyWebMap. It has been developed at the "Paragraphe" Laboratory at University of Paris VIII using the Java Language. The version 1.0 of HyWebMap can be downloaded from the site: <http://www.labart.univ-paris8.fr/hwm.html>.

To develop and create this type of system for the web, we asked ourselves the following questions:

What are the readers' cognitive abilities? What are their goals and their strategies? How can the software tools help reach these goals and help them decide on their choices? What is the pertinent data that is important to a site reader? How can software tools help re-organise the site visited depending on their point of view and their main interest? How can software tools help their personal knowledge grow by exploiting other existing knowledge?

2. Problem of disorientation and cognitive overhead in a hypertext

Disorientation in hypertext space can be seen as a result of navigating within a hypertext. However, some authors relativize this disorientation because it only occurs in large hypertexts and it depends on the author's experience in creating hypertext networks.

The proliferation of search engines and other meta-engines that count millions of Internet sites and that can retrieve from one search thousands of so-called "pertinent" answers, shows perfectly well this disorientation phenomenon that affects the internaut. The long and confusing voyage through the list of website addresses inevitably leads to a distortion in the search itself and produces the type of disorientation often experienced in the first hypertext networks.

It is true that Web portals can be an attractive qualitative alternative compared to the too long URLs, which provided by search engines. These portals are indeed more pertinent but they are also strong vectors of disorientation for the visitor. The information offered is rapidly aggregated around a few generic terms and gets rapidly out of control by its size. The internaut is simultaneously close and at a distance from his search target.

A common strategy in hypertext systems consists in providing the reader with a map of the network so that they can have a global vision of the structure of the hypertext. This method allows readers to easily locate their current position within the network at any given moment, therefore reducing the disorientation factor. Hypertexts systems such as NoteCards, Intermedia, StorySpace, Extended Books and Neptune; supply graphic navigators maps, which aid in localisation within a text. A navigator can be global, displays all the nodes and links of the hypertext network, or local, displays only the links attached to a specific node. Previously we have explained that although cartographic representation of the hypertext network structure can make navigation easier for the reader and therefore reduce disorientation this is not the only solution available [BRO 88a, BRO 88b, BRO 89].

In addition to studying the techniques to solve the problems of disorientation, arising from the reader returning to a previous node and the problems of navigation history, researchers have studied readers' behaviour to explain this disorientation [BRO 88a]. They have highlighted 3 levels [BAL 96]:

- Navigating in a software space

The reader may not be aware of all of the options available when navigating in a hypertext system. Professionals likely to use a hypertext system can come from many different sectors for which the information in hypertext format is available, however they do not have the in-depth knowledge necessary to exploit a hypertext system effectively.

- Navigating in a conceptual space

The reader does not know how the author has organised the information in order to give it a hypertext form. For instance, how the author defines the hierarchy and the reasoning behind his choice of links. The perception the reader can have of how the information is organised can be in conflict with the author's and lead to inappropriate link activation.

- Navigating in the text

The reader lost in a hypertext network can be compared to a driver lost on the road making a series of left and right turns hoping to find his original itinerary. The hypertext reader is constantly requested to activate links, which create digressions causing him to lose the direction of his original navigation. The numerous diversions the reader makes can take him somewhere in the network that might no longer correspond to his original search.

More specifically, an empirical study [FOS 89] has shown how the reader becomes disorientated. This can be due to, the size of the network, learning time factors and unclear goals strategy, etc. Observable symptoms of disorientation can be; the number of windows open simultaneously, returning constantly into the same nodes, quick window closing, etc. This observation should not cause one to see disorientation as a major drawback uniquely occurring during a hypertext consultation. Disorientation can also happen with a book. If the reader looks at the bibliography of that book and then looks at the bibliographies of the books entered under that first bibliography, the same phenomenon of disorientation would happen as well.

Finally, the cognitive overhead phenomenon [AST 93, BEL 99] can be added to the disorientation problem. This phenomenon comes from the incapability of the reader to remember the essential part of the information linked to a node, or the difficulty in rapidly identifying the nature of the links activated. Six key cognitive variables have been identified when navigating in hyperspace: *intellectual ability, analysis ability, flexibility in analysis, synthesis, abstract reasoning, and independence of the cognitive field* [ROU 97, CAS 96]. In addition to the problems of disorientation and cognitive overhead in Web space,

we can add the occurrence of navigation obstacles which impede the reader from obtaining his goals. Navigation obstacles can occur because of:

- the considerable increase in traffic of heterogeneous information.
- the existence of numerous obsolete links either internal or external.
- the organisation of the pages not adapted to the readers' needs.

We can see that the reader needs to locate himself, orientate himself and navigate more efficiently and comfortably in a space of information in perpetual evolution, an evolution that does not seem to slow down.

2. Description of the HyWebMap system

2.1. The Web Agent of the HyWebMap

In the HyWebMap system, we have created a software tool (called an "agent") that is able to analyse any HTML page or Uniform Resource Locator [URL] chosen by a reader. The agent lists the principle characteristics of the page or URL, deleting any obsolete internal or external links from the site and allows the reader to save this information in a personal database.

The agent takes into consideration the full addresses. It analyses all the HTML pages and saves the relative internal and external links allowing access to a future paragraph on the same HTML page.

Propagation of the links :

Phase 1 : From an original page, the agent finds all the available links on that page, for example, the first $x = 1$ links indicated by the arrows in figure 1. The agent stocks all the addresses of those links in the readers' personal database.

Phase 2 : The agent explores all the listed links when analysing the next series of pages, (links 1.1 to 1.3, figure 1). The new addresses of links found are also saved into the readers' database.

Phase 3: The agent continues to search for and acquire links, constantly updating the readers' database, until it reaches the bottom of the current HTML structure (i.e. no further links are found by the agent).

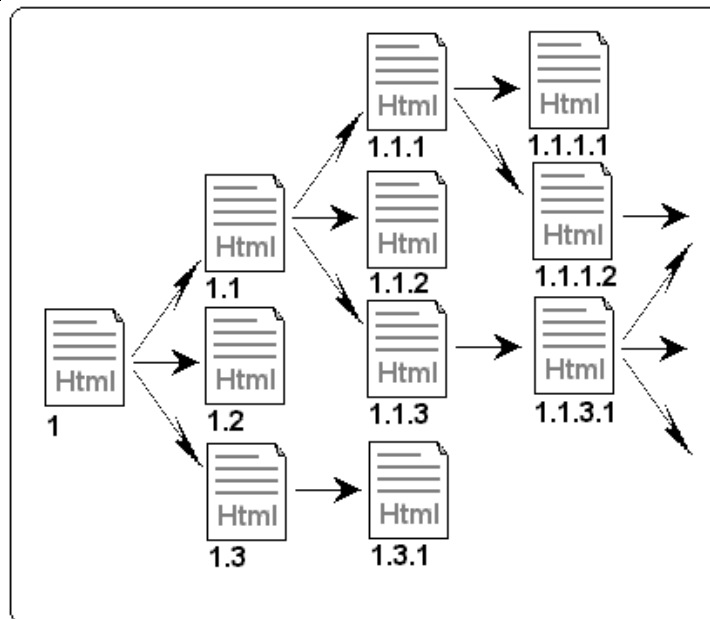


Figure 1. Propagation of the links

The agent can detect the initial links from the source page to other pages in the same site (internal links) or to different sites (external links). We can see the exponential nature of such an exploration. Let's imagine that every page has x links, examining the links with a depth of $n = 3$, results in an arborescence of x^3 different pages the agent must examine. However, if the depth of the hypertext is very large ($n = 10$), the agent will have x^n pages to examine which will take much more time and cause the system to become inefficient. To solve this problem, the agent can be parameterised, enabling the reader to define the itinerary of the agent in depth (vertical) and /or width (horizontal).

2.2. Creating a database

After the web agent has completed the analysis of the site chosen by the reader, it creates a relational database containing tables of **nodes**, **links**, **anchors**, **profiles**(i.e. the characteristics of a Web page: the number of characters, images, external and internal links, etc) and **properties** (describes the informational characteristics of a node (page) : title, date of creation, key words, authors etc.). In addition, the web agent copies the pre-existent relation *linking*(Page-A, Page-B, Link L) which includes all the links of Page-A that are attached by a link L to a Page-B.

2.3. Navigation with the Web map (virtual network)

HyWebMap was inspired by the map approach because we believe that conceptual maps are the more pertinent when they are made by the readers themselves. When searching and saving pertinent information on the web, the reader always needs tools to help him externalise and structure his own mental process which is in interaction with the information found. These tools must enable the reader to elaborate a personal map from an existing map.

Following the definition of the database the reader can see a map (a “**virtual network**”) of the visited space (nodes and links). The navigation can be global, HyWebMap displays all the nodes and links of the visited space, or local, displays only the links attached to a specific node. The readers can modify the map by adding nodes and personal links. Personal links can be added by using the copy/paste functions or by graphically linking the nodes on the map (simply drawing an arrow between two nodes). This results in the creation of a hypertext link. Personal links are represented by red arrows in order to be clearly distinguished from the links of the map visited and are not integrated in the source pages. They are only stored on the reader’s hard drive and are active only when the map is selected. Figure 2 shows the architecture of the HyWebMap system.

Nota bene: A web page has a sequence of MIME (Multipurpose Internet Mail Extension) fragments. It can be a sequence of texts, sounds, and video. HyWebMap considers each fragment as a node.

2.4. Construction of a personal space

With the HyWebMap system, it is possible to create a personal virtual meta network from other virtual networks (Figure 3). The readers can then widen their meta networks with personal links and nodes. The Meta network is saved on the reader’s hard drive under a reader-defined named and it takes the “**.hwm**” extension by default. At this stage, the meta network is like a virtual network.

The reader can :

- 1) navigate locally in the virtual network without being connected to the Internet which reduces access time. When the reader wants to visualise a page, HyWebMap calls the navigator installed on the reader’s computer. The reader is then free to switch at any time between his navigator and his virtual network.

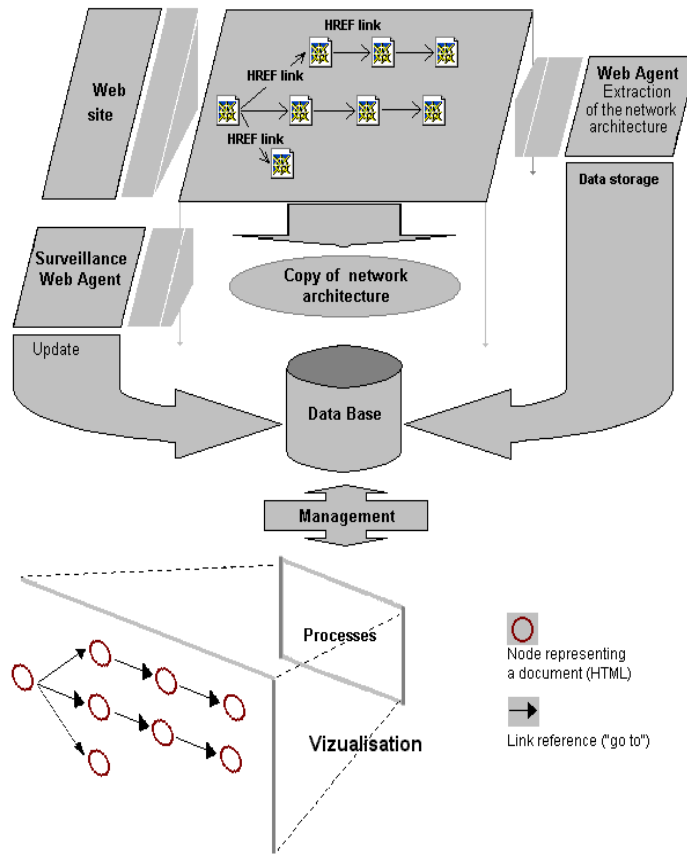


Figure 2. HyWebMap Architecture

2) re-organise the structure of a site visited to match his personal knowledge of the domain without changing the structure of the original site.

3) build virtual thematic networks in relation to his centres of interest.

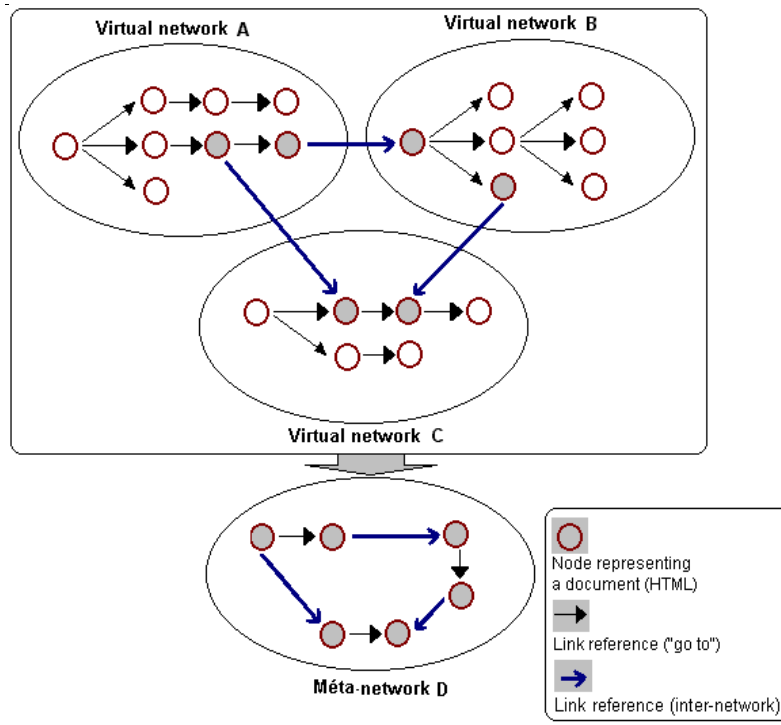


Figure 3. Virtual meta network

2.5. typologies of nodes and links in HyWebMap

The HyWebMap virtual networks are organised in a hierarchical (inverted "tree") structure.

Five types of nodes and seven types of links are defined in the system to assure the cohesion of the virtual networks:

Types of nodes:

ROOT NODE: a main node, located at the top of the hierarchical structure, defined by the reader, from this point the reader can access his virtual network.

PARENT NODE: a node directly below the root node. It can be linked to one or several root nodes and can have one or more children.

CHILD NODE: a node directly below a parent node. It can be linked to one or more parent nodes.

ORPHAN NODE: a node not linked to any other nodes.

EXTERNAL NODE: *a node that does not belong to the original site but can be accessed by a referenced link.*

This typology can identify the structure of documents and interpret the semantic meaning between the nodes. The functionality of the nodes is the result of HyWebMap posing the following questions :

Where am I? What are the subsequent types of nodes? Why is there a link between two types of nodes?

Types of links:

ANCESTOR LINK: *link between a root node and a parent node.*

PARENT LINK: *link between parent node and a child node.*

CHILD LINK: *it is an arrival link from the parent link but has no other links towards other nodes.*

REFERENCE LINK: *it is a free link between nodes (represents the link "go to").*

INTERNAL LINK: *integrated internal link located on the same page (region).*

EXTERNAL LINK: *link between one node of a site and a node of another site. This type of link creates a site-to-site link.*

BROKEN INTERNAL LINK: *a link that no longer exists on the same site (orphan). It is a point of departure but not a point of arrival.*

BROKEN EXTERNAL LINK: *a link that no longer exists between a node of the readers' virtual network and a node of an external site.*

This typology of explicit links helps the reader to clarify and interpret a list of destinations usually defined by HTML addresses or URL's that would be too long and without sense. Displaying the source links helps answer the following question :

What else could I have obtained here?

Especially in the case of the interpretation of pages (nodes) obtained as a result of a search, direct access to the entrance link can provide a convenient access to pertinent information sought after by the reader.

2.6. Update and surveillance agents of the virtual network

We have developed surveillance agents that constantly verify the availability of nodes in the virtual network from their original sites. When a node is no longer available, the agents detects its type:

- if it is a **root node**, the agents alert the reader by asking him to choose another root node.

- if it is a **parent node**, the deleting is more complex because it plays a semantic role in the organisation and understanding of the network. HyWebMap offers two solutions:

manual solution - The agents alert the reader by asking him to choose another parent node or the reader can choose the default, which is to keep the obsolete node in the network with its profiles and annotations intact and specifying that it is an empty node.

semi-automatic - The agents check the properties relation and the key words field . If the parent node to be deleted has key words, the agents launch the search engine Altavista (www.altavista.com) with these key words as the search argument. The first five Web pages obtained are proposed to the reader. The reader is free to choose among these five as a replacement for the deleted parent node.

If the parent node to be deleted has no key words, the manual solution, described above, is used.

- if it is a **child node**, it is deleted locally without any re-organisation of the virtual network.

2.7. Renaming links

In the HyWebMap system, we have introduced the concept of renaming links between nodes. This function allows the reader to personalise his network. He can name the links between nodes according to his knowledge of the content of the pages. This function enables the reader to have a more reflexive and planned navigation.

2.8. The HyWebMap Navigation Tools

Navigation in HyWebMap either is global or local. As global, it displays on screen all the nodes and links of the visited space. As local, it only displays on the screen the selected node with all the links and nodes attached to it as well as its characteristics and its level in the virtual network (figure 4). Each visited node is shown by a red frame. Visited links are represented by red arrows. This representation thus simplifies orientation for the reader by differentiating visited and non-visited nodes.

HyWebMap has several tools to help the reader navigate in his network:

- **Navigation by history**: A classic tool, it enables the reader to go back to a node already visited.

- **Navigation by level of arborescence:** The reader can choose the level he wants to visit in the arborescence of his virtual network.

- **Navigation by date:** The HyWebMap system saves the navigation history of the virtual network for two weeks. The reader can ask to see the visited nodes during a period of time defined between two dates. HyWebMap presents a virtual navigable network of the nodes already visited, which gets rid of duplicates. That is not the case for all standard navigators which only present a list of nodes visited in the time zone.

Netscape Communicator and Microsoft Explorer navigators offer functions like history and bookmarks which enable the readers to locate and organise the addresses considered to be interesting. Those addresses are represented under two types of lists. The two navigators enable the creation of categories and sub-categories of addresses. However, those techniques have proven to be insufficient in solving the problem of personalised navigation on the Web and the creation of personal hypertext network from favourite web pages.

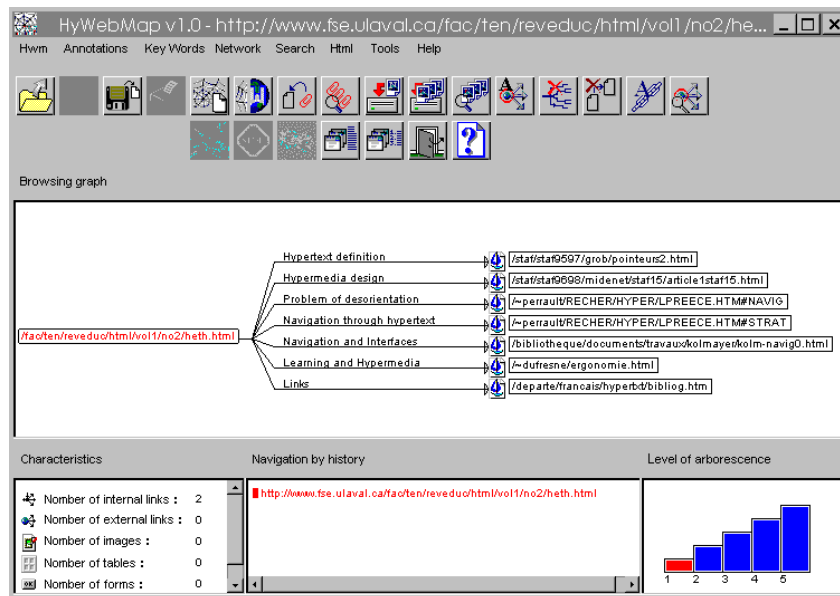


Figure 4. Present a node selected with its characteristics.

2.9. Annotations, key words attached to the nodes

The personal virtual network presents the reader's web space. HyWebMap enables all readers to widen their virtual networks by creating their personal annotations on one or several nodes. This annotation can be a text, links or a text

between links. This results in a virtual network annotated with a network of notes.

An annotation has 3 distinct fields :

A **comment field** to make comments concerning texts, pictures or a particular link.

A **key words field** (not really part of the node) key words can be created by the reader and can be used during the interrogations on the network.

A **variation field** used for synonym or syntax correction. Also used for format modification.

For a given node, the usage of each of these fields is optional. They are also completely independent. This distinction between annotations is very useful when performing a search. The annotations are in HTML format. Each time a node is revisited, the attached annotation appears in the navigator.

2.10. The generation of a website from the HyWebMap virtual network

We have created a function to enable the generation of a virtual network into HTML format. It can save and export the virtual network to HTML format. The mechanics of this function are completely transparent to the reader. This is very useful because the consultation of the network can be done by a standard navigator.

The site takes the extension **“.html”** when it is downloaded. HyWebMap generates 2 frames. On the left, there is the navigation frame (name given to the link between two nodes) and on the right, the frame presents the profiles of the source page: name of the network, title of the source page, URL of the source page, key word reader(user), name of the author, key words author, date of creation (the last three are displayed only if the page has meta-tags to detect them). HyWebMap displays the personal annotations at the bottom of the page. Figure 5 shows the generation of the virtual network onto a HTML site.

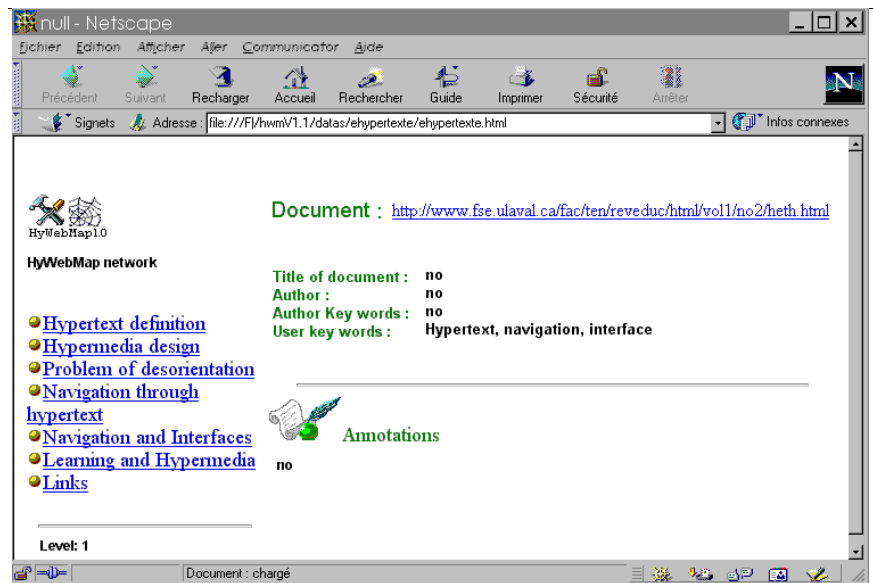


Figure 5. Generation in HTML

3. Domains in which HyWebMap can be used

HyWebMap, a functional system is used in several domains:

Education : Teachers can re-organise their courses in accordance with the themes of each course section. The students can then create their own virtual networks according to their main interests. Presently, several teachers from the Hypermedia, Documentation, Education Sciences Departments at the University of Paris VIII use HyWebMap as a tool to prepare and organise their classes.

Research : Several students use HyWebMap as a system to search for information on the web. They create virtual thematic networks around their research.

- **Web portals** : Portals created around a specific theme.
- **Long Distance Teaching.**

4. Interactive networks experiences

We have tested HyWebMap in real learning situation with teachers and students in “ 3rd year of education sciences department (Paris 8 University)”, on

september 2000. All 13 students had prior experience with information searching on Internet.

Protocol:

Two students groups were created. Each one had to visit an exhaustive list of web sites.

- The first group of 6 students used 5 computers. Their task was to produce interactive networks on one of the following themes : *Hypertext, Environment, French Universities, Health, long distance teaching.*

- The second group used 5 computers with *Netscape Communicator*. Their goal was the same of the first group.

The work of the two groups was carried out on 2 days. First group worked the first day with HyWebMap and the next day with Netscape Communicator. The second group started with Netscape on the first day and finished with HyWebMap on the second day.

Results:

We will briefly present results obtained during utilization of HyWebMap. We have analyzed 1611 operations. We only comment on navigation tasks recorded on each computer.

Navigation by virtual network	73%
Back	1%
Next	1%
Return to Home	4%
History	5%
Navigation by level	15%
Navigation by date	1%

As predicted, virtual networks of HyWebMap were often used intensively because such networks are *browsable*.

During interviews about manipulations with Netscape Communicator, students expressed their difficulties on the one hand to organize and present collected pages and on the other hand to browse in these pages (students note often the absence of "map"). Thus to simulate a network, students have frequently used "Bookmarks" with "links sub-categories". But bookmarks present informations like simple lists. They don't allow to rename links between HTML pages. To insert their own annotations, some students (45%) were constrained to use Netscape Composer to create their "comment page".

Finally, nearly 85% of students expressed that they would opt for HyWebMap if they had to choose between Communicator and HyWebMap. Their choice was guided by available tools in HyWebMap : assistant agents, annotations of pages, personal networks, generation of web sites,...

Others opt for Netscape Communicator because they are accustomed to of this browser (we can note here that the remarks of students concern, infact, problems of acquisition of new tools functionalities.

Tests are currently led about interaction and navigation in personal virtual networks created with HyWebMap. These tests are led by students in High Schools and universities in the context of educational tasks.

5. Discussion

Many tools can create Web site maps: WebAnalyzer (<http://www.WebAnalyzer.com>), CHEOPS (<http://www.crim.ca/ipsi/cheops>), Mapuccino(www.ibm.com/java/mapuccino), MAPA(<http://www.DynamicDiagrams.com>), etc. The most *innovating* one is Paul Kahn's MAPA [Kah 95, 97]. It can build a 3 D map from a URL in a Java applet. It uses colors to separate the hierarchical levels of the html pages. MAPA is not only a navigation tool, it can also be a tool to manage Web sites. HyWebMap like MAPA has skills to navigate and manage Web sites. The differences between the two systems are that HyWebMap's reader can re-organise the visited space by creating his own network depending on his point of view without affecting initial nodes and links. He can append nodes and annotations into this virtual network and rename links. HyWebMap can generate new sites automatically, browsable by any navigator.

6. Conclusion

HyWebMap beyond its technical characteristics allow the user to be placed in its main role of reader/author in a situation that facilitates the ability to build and enrich his own knowledge with any knowledge space.

Different professional and educational situations which have tested HyWebMap encourage us to continue tests in the field of information.

By, now, we develop extensions on HyWebMap in order to ameliorate interface in using 3D technologies which appears more relevant for browsing and building virtual spaces.

7. Bibliography

- [AST 93] ASTOLFI J-P., " Comment les élèves apprennent-ils ?", *Sciences Humaines*, n° 32, octobre 93, pp. 26-29.
- [BAL 96] BALPE J.P., LELU A., SALEH I., PAPY F., "Techniques avancées pour l'hypertexte", Editions Hermès Paris 1996.
- [BEL 99] BELISLE C., Zeiliger R., Cerratto T., "S'orienter sur le Web en construisant des cartes interactives avec Nestor", Actes H2PTM 99, Editions Hermès 1999.
- [BRO 88a] BROWN P.J., "Hypertext : the way forward", éditions J.C. Van Vliet, *Document manipulation and typography, Proc. of the International Conference on Electronic Publishing*, Nice, avril 1988, pp.183-191.
- [BRO 88b] BROWN P.J., "Linking and searching within hypertext", *Electronic Publishing*, vol. 1 (1), avril 1988, pp. 45-53.
- [BRO 89] BROWN P.J., "Do we need maps to navigate round hypertext documents?" *Electronic Publishing*, vol. 2 (2), juillet 1989, pp. 91-100.
- [CAS 96] Castelli C., et al. "Getting lost in Hyperspace : Lessons learned and future direction" ", in ED-MEDIA 96/ED-TELECOM 96, 1996.
- [FOS 89] FOSS C.L., "Detecting lost Users: Empirical studies on browsing hypertext", *89H385299*.
- [ROU 97] ROUET J.F., "Le lecteur face à l'hypertexte", in apprendre avec le multimédia Où en est-on ? Sous la direction de GRIMONT A., Paris Editions Retz, 1997.
- [Kah 95] Kahn P., "3Global an local hypermedia designe in the encyclopedia african " in *Hypermedia Design*, Fraïsse, Garzotto, Isakowitz, Nanard an Nanard ed. 1995 Spring, London.
- [Kah 97] Kahn P. et al. "Real space and cyberspace : a comparison of museum maps and electronic publication maps " in ICHIM 97, 1997, Paris.