UNIVERSITE PARIS-SORBONNE PARIS IV

U.F.R. d'anglais

Mémoire de maîtrise

Présenté par

Ludivine MAS

Publishing on-line :
design, computer graphics and implementation.

Sous la direction du Professeur Liliane Gallet-Blanchard

Année universitaire 2001-2002

FOREWORD

This work was written in parallel with a Web site creation, that of CATI's <http://www.cati.paris4.sorbonne.fr>. CATI (short for *Cultures Anglophones et Technologies de l'Information*) is a Research Centre at Paris-Sorbonne Paris IV that works in the field of Humanities Computing. To improve its existing Web site was a good opportunity to apply principles exposed here. That is why references to CATI's Web site could be found throughout this work.

Furthermore, we will not pretend to go into technical details as far as electronic publishing is concerned. However, if some of you are eager to learn a few practical tips, a CD-Rom is provided at the end of this study.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

Introduction.

The first thing we are struck by in the Web site creation is its accessibility to everyone although it belongs to Information Technology, that is to say to a technical domain.

The fact of the matter is emphasized by the growth of more and more didactical and easy to use development tools. Many sites that offer assistance in building a Web site[1] allow anyone to create his/her own site without any HTML[2] knowledge. That is the reason why we can argue that even if building a professional Web site still requires specific skills, these are not however an elite's privilege. Incidentally, the webmastering profession is made up of people coming from curricula that seem to be completely opposed to this activity (e.g. humanities).

We must keep in mind that computing and more particularly multimedia are domains that cause passion for them and therefore that involve commitment. An interested enthusiast will know how to acquire the webmaster's skills despite his literary background.

Moreover, coming from different horizon remains a real advantage because a webmaster must look at the contents he/she has to publish on-line with a critical eye in order to define a web interface which will suit the site's objectives and the needs of the target audience.

That is why a multidisciplinary profile is really appreciated now in the labour market and is considered as a force by employers. However, new diploma and curricula are created to form people to those new jobs and although self-taught enthusiasts are lauded for the time being, we can easily imagine that employers will soon prefer graduates to them.

---

[1] Most of the sites offering web site hosting. For example, WebBuilder by Lycos Multimania. Retrieved 08 September 2002 <http://www.multimania.lycos.fr/build/webbuilder/>

[2] Short for *HyperText Markup Language,* the authoring language used to create documents on the World Wide Web. Retrieved 08 September 2002 <http://www.webopedia.com/>

But to acquire technical skills may be considered as a secondary problem in comparison with a good apprehension of the World Wide Web potential.

Publishing on-line is not and will never be the same as publishing a book. Web sites, thanks to hypertext links in particular, offer us the possibility of collecting information throughout different pages without following a linear process. So, designing a Web site must take into account the random character of internauts' visits. This is possible by providing a visual identity and an informative content that will help users to know where they are and what kind of information they are going to find on the Web site.

Indeed, we must consider the Internet as a new communicative media with incredible possibilities. Of course, the World Wide Web represents a tremendous promise to interact with others, exchange information or communicate. But who has not once visited a Web site and read "Welcome ! You are the 20th visitor" ? Is this what is called "information superhighway" ?

There are millions of potential visitors worldwide; millions of internauts, but with specific needs, expectations or means. Platforms, browsers, types of connexions may differ, as well as interests. Naturally, we cannot pretend that our Web site will concern all of the internauts or will be seen by all of them. However, if we believe in our project of Web site building, we must give it a chance to convey its message to the largest audience. That is the reason why a preliminary analysis is necessary to identify our Web site purpose and its target audience. Only then can we reasonably design our Web site.

A Web site creation is not a punctual event, it is an everlasting thought. Interactivity but also reactivity are the key terms of the Internet. Everyday, thousands of Web sites appear or die. A Web site should provide an up-to-date content and should evolve over time and adapt its structure to the new technologies.

Some technical and organisational tips, which are impossible to circumvent, will come with this Web comprehension in order to anticipate and guarantee the right implementation of a Web site.

Publishing on-line becomes then a matter of some persons eager to share knowledge worldwide and aware of technical constraints.

I. DESIGN

A. Interface design.

1. Preliminaries.

Designing a Web site requires a preliminary study. As a matter of fact, even if we all agree to say that the Internet is a universal media that gives everyone access to information, we cannot pretend that our Web site will concern and attract all of them.

In order to design an adapted and attractive Web site, we must already define its purpose and objectives. Which are the main messages we want to convey? How will we make them meaningful to internauts? Once these objectives have been stated, we must identify the target audience by answering the following question: for whom is our Web site intended? In other words, which profile we can draw of our potential visitors and which needs and expectations they have. We will then be able to define a design strategy that matches our site contents with visitors' expectations.

a. The purpose statement and the site objectives.

The purpose statement is usually the trigger which determines a site design. It has to be written and be available at all times during web development. A statement written in general terms is enough; at this stage, we just need an outline that will be detailed in the objectives list.

Let us take the example of the CATI's Web site we want to create. Its purpose statement could be "To provide information about our Research Centre's works and seminars.". But in order to start discerning the kind of visitors our site will be intended for and the contents it will present, we must detail this purpose statement in more precise terms.

In the *Web Style Guide*, Lynch and Horton[1] advise us to begin our statement of objectives by a sentence such as: "We expect our Web site to accomplish these goals in the next twelve months:»

CATI's Web site objectives might be written in these terms:

"The Web site will provide information about the Centre's research areas, namely Humanities Computing and Hypermedia presentation applied to cultural history subjects by the publication of seminars and conferences. The Web site will carry complementary contents to seminars given at Université Paris-Sorbonne Paris IV and will give all the information that foreign students may need. It will also present all the projects in which the Research Centre CATI is involved. We expect that the Web site will soon tend to be a distance study platform.»

This list of objectives allows us to have some clues concerning the target audience of the Web site. Actually, the cited participants in this site seem to be: students, professors, professionals and institutions. But we will see in the next section that we must learn more about them to design a suitable Web site.

b. The target audience.

The target audience is a key issue for it affects on the way a Web site is built in terms of graphics and contents. Many factors may be taken into account to define people we want to attract: their age or sex, the religion aspects, their ethnic backgrounds, computers and modem equipment, geographic location, interests, needs and expectations. The latter also largely depend on users' navigation habits. Are they occasional web surfers who will prefer a graphic-based site, aesthetically impeccable and who will not have other specific demands?

---

[1] Patrick J. Lynch and Sarah Horton. *Web Style Guide*. 2nd ed. New Haven: Yale University Press, 2002.

Or, are they expert and frequent users who will prefer detailed text menus for example to make search and fast loading easier?

Recent studies underline the fact that the more users are expert, the less demanding they are as far as graphics are concerned. But on the other hand, they pay attention to features that enable their visit to be more efficient and fast (e.g. search tools).

Novice and occasional users, unlike to expert ones, rely on design graphics to understand what the site will offer to them. They are fond of symbols and icons such as printing or help icons, FAQ (Frequently Asked Questions), etc.

As far as CATI's Web site is concerned, we can expect to deal with expert and frequent users for they all belong to academic institutions where computer science is largely relayed. Attention then should be brought to the quality of contents and to navigation facilities. However, it will be interesting to take their equipment into account. As a matter of fact, universities usually work on a particular browser or platform and do not have the possibility to renew their equipment. It should be wise not to use the latest technology to build our Web site for old computers may not support it.

c.   Design strategies.

Now that we have decided what kind of goals we want to reach and that we have clearly identified our target audience, we must define a strategy to combine all these elements. To define the design strategy we have to adopt, means in fact defining the communicational channel we will use. Lynch and Horton[2] distinguish a few themes for intranet information delivery which are dependent upon two variables: the linearity of our presentation and the duration of user contact time (*see figure 1*).

---

[2] Patrick J. Lynch and Sarah Horton. *Web Style Guide*. 2nd ed. New Haven: Yale University Press, 2002.
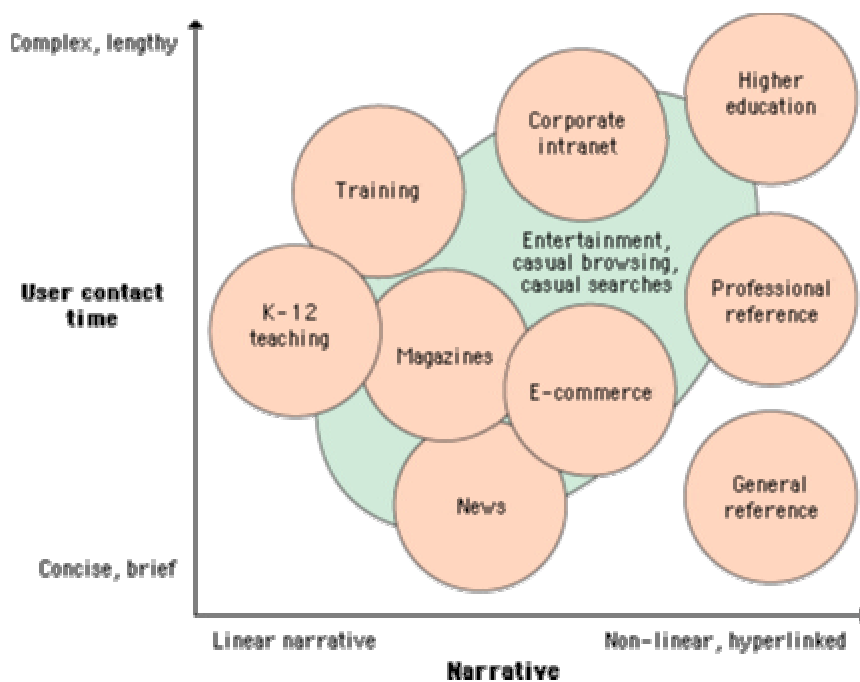
Figure 1 – **Diagram of design strategies for Web sites**

This figure sketches out the main sort of applications a Web site can have.

A training application (used for distance studies for example) is very linear in design and does not offer possibilities to navigate throughout the site. The user enters a precise and normally timed process (nearly one hour) that he has to follow from start to finish. This kind of application often presents contents under the form of quiz.

A teaching application is also focused on contents but offers more digressions via links to other resources. Such an application fits CATI's site objectives. As a matter of fact, one of the announced objectives is to give students complementary information to their courses or to CATI's research areas. That is why, we must think of providing for a specific page where students could find links to Web-based resources. Moreover, as we can see on figure 1, the user contact time is usually longer for this application for its contents is often particularly dense. Many people prefer reading long texts from a paper format than an electronic source and we must thus provide a printing version.

A continuing education application cannot stand too linear a design. On the contrary, it lies upon a profusion of topics, that are most of the time simply listed and that provide in-depth information. Users of education applications are generally expert ones who prefer easy search and fast downloading to attractive graphics.

Reference Web sites also provide fast search and easy downloading of files. That is why the user contact time is shorter.

The four themes we have just detailed concern Web sites for which visitors arrive with a specific goal in mind. But as we can see on Figure 1, other applications exist (such as entertainment sites, news sites, e-commerce…) whose goal is to make loyal casual browsers to their site. For our work deals more particularly with teaching application, we will not detail those last themes but the reader may consult *The Manual Style Guide* for further information.

2. Navigation.

Navigation is considered as the key issue in the process of site design. A good navigation system must be thought about so as to minimise the users' mental efforts by providing a comprehensible path he/she has to follow in order to reach the information he/she comes and fetches. Thus, navigation allows the user to have an idea of the site structure and the contents he/she might find thanks to a site map or a table of contents page for example. Hypertext links change their colour whether they have been visited or not. This allows the user to get his/her bearings within the site and to remember the pages he/she has already consulted. As a matter of fact, it is important to provide context, that is to say to let the user know where he comes from and where he/she can go. A navigation panel will come up to this problem.

Before listing some of the navigation tools that have to be present on a Web site, we must consider the navigation system of a site as a completely autonomous means of surfing.

In fact, the navigation has to be designed without relying on the "Back" and "Forward" buttons of browsers because these do not refer to the same action according to the type of browser we use.

a. Navigation bars and "breadcrumbs".

Navigation bars are the main element of a site navigation system. Their major advantage is their presence on every page of the site, which allows a visitor coming from another site to browse the whole site.

According to the size of a Web site (in terms of pages quantity) and its complex structure, one or two navigation bars can be found. We will then speak of the main navigation bar and the secondary navigation bar.

The main navigation bar is usually placed on the top area of the page in order to be legible by any user (even those who use a standard office 14-inch or 15-inch display monitor). The user must not scroll to find a way of surfing within the site. The main categories of the Web site must appear in this bar. As this navigation bar will be present on every page, it must also provide a link to the home page if the latter is not accessible by the presence of a logo.

The secondary navigation bar details the sub-categories of the site and must provide a link to the home page of each category. It is often placed on the left hand side of the page.

Navigation bars can be either text links or graphics, but it is strongly recommended to choose text links at least for one of them (usually the secondary one) to guarantee legibility and accessibility to all. Most of the time, these navigation bars come with what we call "breadcrumbs».

Breadcrumbs are in fact location indicator devices; they are also called "where-am-I indicators" for they help the visitor locate where he/she is in the site architecture, or which

path he/she has followed to get on the current page. Hence, breadcrumbs bring a real added value to the navigation panel for they convey browsing information and they are another means for users to go "up" in the site architecture by clicking on one of the listed links. As a matter of fact, breadcrumbs are composed of a list of items usually separated by a special character among the few accepted (namely arrows ">", vertical bars "|" or colons ":").

The first item represents the site home page or a category home page; the following ones are usually categories and sub-categories. The last one corresponds with the name of the current page (*see figure 2*).

Figure 2 – **An example of breadcrumb**

We can notice that the last item is not a link unlike the others.

Breadcrumbs are usually placed on the top left corner of the page, but under the global navigation. Keith Instone, on the occasion of the third Annual Information Architecture Summit (March 16-17, 2002), defined three types of breadcrumbs:

- Location breadcrumbs.
- Path breadcrumbs.
- Attribute breadcrumbs.

Location breadcrumbs allow users to understand where they are in the site hierarchy. Designers make a widespread use of this type of breadcrumbs, especially for Web sites that

are built with static pages. Actually, when browsing within a dynamic site, it is more pertinent to tell users how they get to the current page by using **path breadcrumbs**. As there are different ways to navigate on a dynamic Web site (according to the choices we make), there is more than a single path possible. The last type of breadcrumbs is called **attribute breadcrumb**. Attribute breadcrumbs are in fact a way of presenting meta-information under the form of breadcrumbs.

As far as our site is concerned (CATI's), a location breadcrumb will be enough as it is made of static pages and it is built according to a strict hierarchical model. For example:

<u>Home</u> > <u>Seminars</u> > <u>Maîtrise (4th year)</u> > MAN 413

Breadcrumbs and navigation bars are part of **the central navigation** where the main body of the information is represented. However **a global navigation** has to be also present on every page in order to give users access to key pages such as Home, Contact us, Site map…

This other navigation bar is usually text links and is placed at the top and bottom of the pages. Finally, **a feature navigation** can also be provided to announce events or punctual information, but it must not be continually present.

As we have already seen, navigation can appear under the form of text links, but a few rules have to be respected to provide a comprehensible navigation to our visitors.

b. Hypertext links.

On the World Wide Web one of the well-understood rules is that blue underlined text is a hyperlink that allows to navigate within a site. This implies that users do not have to mouse over on a text to understand that it is in fact a link to another page.

Hypertext links must clearly indicate where the users will get, and must change aspect when visited in order to show them the pages users have already consulted.

The HTML convention for hypertext links states that blue underlined text indicates unvisited links and purple or maroon text indicates visited links. However, these default colours are not standards anymore since the use of Cascading Style Sheets (*see page 61*) has developed. In any case, hypertext links should have the same colour and the same effect on every page. For example, a "Back" button cannot link to the previous page on a page and to the home page elsewhere.

Moreover hypertext links must indicate the nature of the document the user will get. Does the text link to an external Web site or within the site? Does it open a .pdf document or a video file? When linking to large files, it would be useful to provide a kilobyte size in brackets along the hypertext link so users know the downloading time it will take. Talking about quick download time, it represents the biggest advantage of text links. And they remain legible even if users turn off graphics in their browser. Of course, hypertext links can be boring to look at if they constitute the only navigation scheme. That is the reason why we recommend using them along with graphics.

<center>c. Search facilities and sitemap.</center>

Users' behaviour tends to be more autonomous than it was at the beginning of the Internet. They get used to Web browsing and thus want to experience their own way to navigate within a site. **Search** allows them to find what they are really looking for and gives them a feeling of independence. According to Jakob Nielsen's studies, "Search lets users control their own destiny"[3].

---

[3] "Search: Visible and Simple". <u>Jakob Nielsen's Alertbox, May 13, 2001</u>. Retrieved 08 September 2002 <http://www.useit.com/alertbox./>.

For the less accustomed users to the Web, search may appear as the last resort to go on surfing. That is why search should be available from every page on the site and should be placed in a noticeable position (usually on right hand side and top area of the page). Jakob Nielsen's studies show that search must be kept as simple as possible. Users hardly manage to use advanced search or to express pertinent queries. Search should be a box whose input field should be wide enough to contain a simple query. Users want to get to the information they are looking for as quickly as possible. Hence they prefer a search box that allows them to enter directly their query rather than a "Search"' link that adds another step. The best proof of this, is the example given by Nielsen : "When I changed the useit.com home page to include a search box instead of a link, search engine use increased by 91%"[4]. No comment.

However providing a search box is not enough, we must keep it efficient by the relevance of search results.

Another means of searching within a Web site that a designer may put at users' disposal is an overview diagram of the site; in other words **a site map** or site index.

A site map is in fact a separate page which is accessible by clicking on the corresponding link "Site map" present on every page. It allows users to visualise the site's information architecture in a single glance. In spite of their key role, sitemaps still do not belong to users' navigation habits. According to Jakob Nielsen's survey, less than 20% of users are inclined to turn to the site map of a site. Once on the site map page, their success in finding the destination they want to get among the numerous links depends a lot on the kind of site map it is. As a matter of fact, Bernard and Chaparro make the distinction between three types of sitemap menu structures:

---

[4] "Search: Visible and Simple". <u>Jakob Nielsen's Alertbox, May 13, 2001</u>. Retrieved 08 September 2002 <http://www.useit.com/alertbox./>.

- An **index site map** where hyperlinks are catalogued alphabetically (*see figure 3*)[5].

- A **full site map** where hyperlinks are organised by categories (*see figure 4*).

- A **restricted site map** which is also organised by categories but the user has to click on a particular category to see details. (*see figure 5*).

Figure 3 - Alphabetized Index sitemap

Figure 4 - Full categorical sitemap

Figure 5 - Restricted categorical sitemap

After comparing search performance, Bernard and Chaparro's conclusions are akin to Nielsen's. Providing a full site map is preferable because users can see the information they are looking for at first glance. As soon as this information is two clicks from the site map (in the case of a restricted site map for example), erroneous destinations become more frequent.

---

[5] These figures are illustrations used by Michael Bernard in his article "Sitemap Design: Alphabetical or Categorical?»

That is why we can say that giving users the possibility to visualise the site's information architecture is a prime necessity. Both search boxes and site maps make their navigation easier and prevent them from getting to a dead-end page.

As a conclusion, we will emphasise how important a good navigation system is. It must be a combination of navigation schemes in order to answer users' needs.

However, beside the navigation tools we have dealt with in this section, there are small details that participate in the success of a good navigation panel. For example, to give a HTML relevant title to the current page allows users to read in the top title bar of their browser where they are. Even a personalised 404 error page that keeps the site's layout allows users to understand that they still are on the Web site and redirect them to other internal pages.

But in fact, navigation will be coherent insofar as it lies on a coherent site structure.

B. <u>Site design.</u>

1. Site architecture.

a. Organising information.

In order to provide a functional and attractive Web site, we must organise our contents under the form of chunks of information. As a matter of fact, this division of information is necessary for users to keep our site structure in memory and by the way to know how to get more detailed information through navigation. Actually, our site navigation must answer a hierarchy of importance and rank information.

The commonly used model of site structure is based on what is called the concept of "layers"; that is to say, visitors advance through our Web site from a general overview (the home page) down through major submenus (or categories) and content pages. Information is then organised from the general to the specific and corresponds to the mental path a user follows to get the information he/she is looking for.

In their *Web Style Guide*, Lynch and Horton schematise hierarchical organizations as shown in figure 6[6]:

Figure 6 – **A Web site in a hierarchical structure**

To make the parallel with the concept of layers, the **home page** represents the first layer that introduces our Web site and catches the attention of visitors. **Major submenus** compose in fact the second layer. This level can often be reached thanks to the main navigation bar, and introduces the themes that are going to be exposed in our Web site. It still remains general and only broaches subjects. The third-layer pages are the equivalent of the **content pages**; they provide detailed contents. We may consider that if visitors arrive at this level, they are seriously studying our Web site.

Let us take a case study (CATI's) in order to understand this notion of layers and hierarchy. CATI's site structure may be diagrammed as follows:

---

[6] <u>Web Style Guide</u>. Retrieved 02 Oct. 2002 < http://info.med.yale.edu/caim/manual/contents.html>.
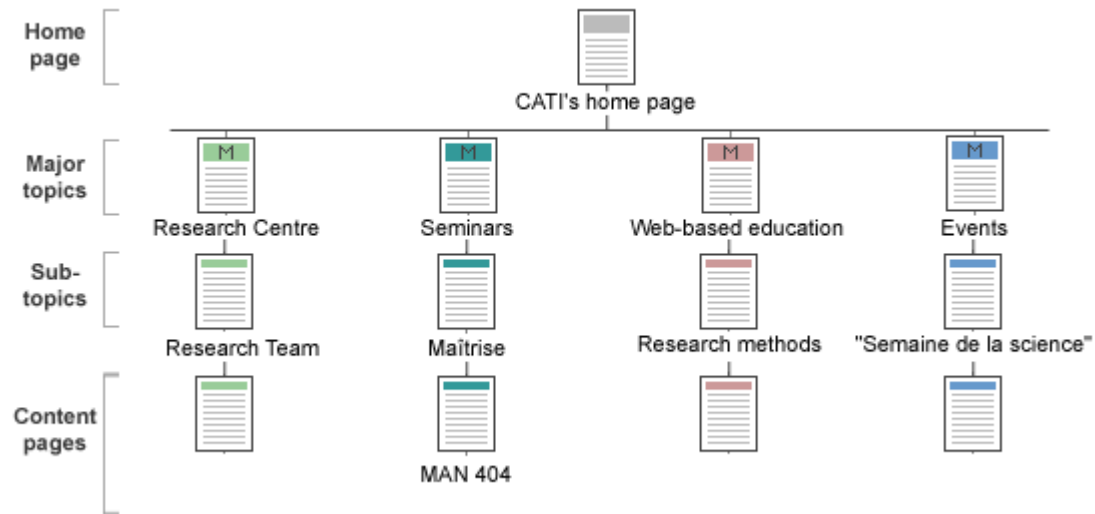
Figure 7 – **CATI's site structure**

As this academic Web site has got rich contents, we can notice that four layers are necessary to reach detailed information. But this means of dividing information into topics and subtopics presents also the advantage of quicker download. Hence, several small pages will be loaded more easily than large pages. Our visitors will appreciate this saving of time.

b. The folder structure.

Organising information into logical units means also organising our Web files into sub directory folders that share the same logic.

The basic structure of a Web site includes few elements that compose a **directory tree**.

The main site home page is usually named *index.html* for it is the default name used by web browsers. This home page must be located in the root directory, that is to say at the base of the site. Sub directory folders (corresponding to topics for example) must be created. They will contain a main page (*e.g.* seminars.html) and subsequent pages.

Our web site graphics must be put in a graphics folder in the root directory if our site is a small one or in each sub directory folder if it is a larger one. Finally, a styles folder must be created to contain style sheets[7].

The diagram below shows what CATI's folder structure may look like:
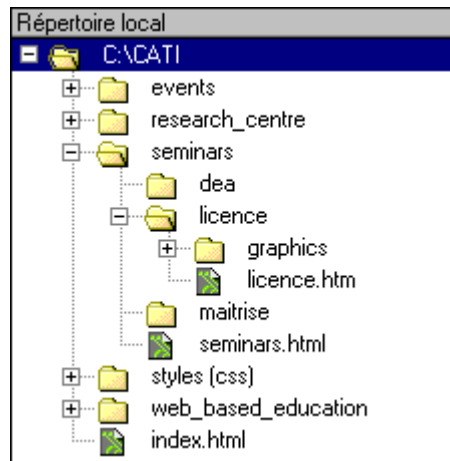
Figure 8 **- CATI's folder structure**

2. **Site elements.**

a. The home page.

As we have just seen, the home page sits at the top of the hierarchical organisation of our Web site. Careful attention has to be given to it as it represents the main entry into our site. The home page must rapidly indicate to visitors what they are going to find on our site. Moreover, as it may be the first interaction with them, we must not try their patience with long loading time. We consider that a loading time that exceeds ten seconds is beyond tolerance and may discourage users from visiting our Web site.

This technical constraint leads us to wonder about the relevance of using graphic or text menus for our home page. As a matter of fact, using strong graphics presents the advantage of grabbing attention on the one hand, but slows down loading time on the other. Whereas a home page based on text menus often seems to lack visual appeal despite the fact

---

[7] We will deal with Cascading Style Sheets on page 61.

that it downloads faster. Of course, the choice we will make between graphic or text menus will depend a lot on our target audience and our site purpose. Casual users will prefer a visually attractive home page, whereas frequent users will prefer rapid access to contents.

However, finding a balance between graphics and texts is possible. A good approach is to put a graphic banner up at the top of the home page in order to provide a visual identity. And the rest of the page may be then composed of text-based links. This compromise will satisfy educational Web sites whose major needs is to share efficiently a maximum of information.

That is the reason why we have chosen this approach for CATI's home page (*see figure 9*).

Figure 9 - **CATI's home page.**

As far as the home page is concerned, there are a few counsels to follow. First of all, a home page must not have vertical scrolling. The main information must be visible at first glance.

A logo placed at top left will allow users to know where they are and an introductory text will explain them what they will find on the site. Navigation is of course a key element of the home page but must not provide a "home" button, contrary to interior pages.

b. Useful elements.

Beside the home page and interior pages that must be well designed from an organisational and visual point of view, some functional elements must not be skipped.

If we keep in mind the hierarchical organisation of our Web site, we notice that we have submenus pages that present a category and provide more detailed information to access content pages. The major submenus pages act as mini home pages of the topic. Moreover, their strategic location in the hierarchy requires an adequate set of links in order not to have dead-end pages or documents that cannot be reached. Each submenu needs a link to the home page and links to content pages.

Some people may need help to navigate on our Web site. That is why, key pages such as a sitemap or a Frequently Asked Questions (FAQ) pages are sometimes necessary depending on the complexity of our Web site. Both will help users understand the services our site offers to them.

Another site element which unfortunately is not enough taken into account on Web sites is the notion of time and date. It would be useful to foresee indication of newness whenever a page is updated or a new content added. This may be indicated in two ways: either a "NEW" graphic is put in evidence next to the new information or date is included on each page to announce a possible updating.

3. Intellectual properties issues.

a. The copyright notice.

According to Brad Templeton, copyright is "the legal **exclusive** right of the author of a creative work to control the copying of that work"[8].

A Web site does not necessarily need a copyright notice in order to be copyrighted. As a matter of fact, the Berne convention, which is followed by all major nations, stipulated that a work is protected whether it has a copyright notice or not. However, this text of law is most of the time unknown to the general public. And the absence of copyright notice on a site may be interpreted as a permission to copy whatever they want.

In order to avoid misunderstandings, we recommend to provide a copyright notice and to write it properly according to its acknowledged form:

Copyright [dates] by [author/owner]

We can use © instead of the copyright term but this symbol seems to disappear since the United States have decided to use the word "**Copyright**" for technical reasons (some browsers do not correctly display the © symbol). It is also possible to use them both.

The date that has to be included in the notice, actually refers to the year of its publication.

And finally, we must include the **name of the owner** of the copyright. We must be careful for the author of a work is not necessarily the legal owner. As a matter of fact, an employee who creates something will not be the owner of the copyright insofar as he/she works on behalf of his/her company.

Moreover, there is another element that we can notice on some Web sites and that does not appear in our copyright model. It is the reservation of rights phrase. "All rights reserved" is in fact not legally needed except in Bolivia and Honduras, which are countries following the Buenos Aires convention. Here is an extended but legal example of copyright notice:

Copyright © 2002 by CATI. All Rights Reserved.

---

[8] « Ten myths about copyright explained ». Retrieved 25 Sept. 2002
<http://www.templetons.com/brad/copymyths.html>

The fact that a Web site is copyrighted does not always mean that we cannot copy some elements from it provided that we ask for the permission to the owner of the copyright. As a general rule, asking for permission seems to be the best way to avoid troubles.

b.   The "fair use" concept.

Within limits, some cases of copying are allowed by law. But the copying must respect what is called the "**fair use**" concept. Its principle lies in the fact that ideas cannot be copyrighted whereas their structure can. In other words, the law cannot go against the freedom of expression. But it may become a real drudgery for educational institutions to ask for the author's permission each time they want to criticise or cite an excerpt of his/her work. The fair use concept is difficult to define; courts are used to analyse individually each case of copying submitted to them. However, fair use must respect at least three legal conditions. First, the example copied must be short and attributed to its author. And we cannot use more than is necessary for our commentary. Then, the copying should not harm the commercial value. As a matter of fact, if several copies exist on the market, users need not buy the original anymore. Finally, the use of the copying is examined. If the copying answers educational needs, fair use can be invoked.

Once again, when in doubt, we have better ask for the permission of the copyrighted work.

c.   Copyright issues in practice.

When creating a Web site, we may encounter some problems of references. We may want to link to other sites or illustrate our contents with images coming from them or even scanned images. But to which extent do we have the right to use external resources?

As far as linking to other Web sites is concerned, controversies give way to acceptance. Except in one case we will explain later, we may freely set up HREF links to other Web sites for it corresponds to the essence of the World Wide Web itself. The Internet

aims to link people with each other, to exchange ideas and to jump from one subject to another. Hence it is possible to make such links insofar as they are used properly and are not accompanied by irrelevant comments that would libel the author or his/her work.

Nevertheless, there is a particular case in linking to external sites that recommends to ask for permission first. In fact framed HREF links may cause some troubles.

Some Web sites use what we call "frames"; in other words the designer has divided its page into two parts at least. These parts present specific contents. On the left, we commonly find the site menu, and in the main frame, the body of text. But in the case of commercial sites, there might be an advertisement in one frame. Someone who has been linked to this site may not answer for this advertisement and may not approve of this association. That is the reason why we recommend to obtain permission to set up "frame" links to other Web sites.

As far as images are concerned, permission has to be always obtained to publish them on our Web site. As part of the fair use concept, some uses may be tolerated if they mention the source. Even our own photographs may be discussed if they include people. We have to ask for their permission first before publishing the photographs.

The World Wide Web gives us the impression of freedom. It seems that it has no boundaries, no constraints and that we can feel free to copy resources. But if we compare the Internet with a book, we easily understand that copying the whole work or its illustrations may be prejudicial to its author's interests. So we must be careful and ask for permission whenever it is possible.

C. <u>Page design.</u>

Now that we have a more precise idea of how our Web site will be organised and which categories it will bring up, we are ready to focus on page design.

What must our Web pages contain? Are there design principles to follow? Before answering these questions, we must first deal with the ticklish problem of page size.

1. Page size versus screen resolution.

Who has not once experienced a Web site whose pages were bigger than his/her screen resolution? Who was not forced to use horizontal scrollbar to see the entire content of some sites? On the contrary, some of us who have a high-resolution screen, may have been frustrated by a Web site claiming "This site best viewed at 800x600".

As a matter of fact, from now on, 800x600 pixels is the common standard page size. This standard has been obtained by analysing screen resolutions of Web surfers and by determining an average. Until recently (February 2001), statistics announced that the commonly used screen resolution was 800x600 with 53% of the Web surfers. Thus designers and authors started to develop fixed-size Web pages limited to 800x600 pixels in order to satisfy the majority of internauts. However, a recent study led by OneStat shows that people tend to have higher screen resolution than in 2001. Here are the screen resolutions of Web surfers in July 2002[9]:

Screen resolutions of Web surfers
OneStat - July 2002

---
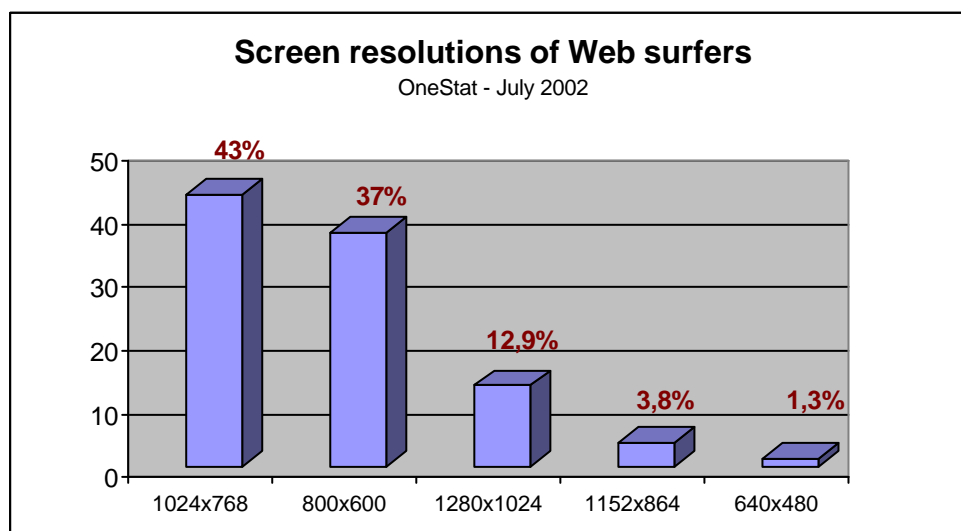
[9] This statistics is extracted from a study by OneStat. OneStat is considered as number one provider of real-time Website analysis. Retrieved 27 Sept. 2002 <www.onestat.com>

Figure 10 – **Diagram of screen resolutions.**

We can notice that more internauts use a 1024x768 screen resolution now. Does it mean that designers must increase their page size to suit this new trend? Of course not. More than 50% would have hard times surfing on the Net if pages were developed in a fixed-size (in this case 1024x768 pixels). On the contrary, this study must restate the problem of page size by emphasizing the necessity of accommodating a wider range of Web surfers. We must then consider the possible flexibility of page design.

Fixed-size Web pages use what is called **absolute positioning**. In other words, each element is assigned to a particular and fixed location on the page.

But Web page authors can also use **relative positioning** that allows pages to contract and expand according to users' screen resolution. The size of elements are no longer defined in pixels but in per cent. Pages that use relative positioning are called fully flexible pages or "liquid" pages.

Nevertheless, as satisfying a majority of people always seems to be difficult, many designers keep on developing fixed-size pages. But we are welcome to try if we are convinced of the advantages of flexibility. Whatever our opinion is, we must choose the page size of our Web site before starting the building process.

2. The layout grid.

a. Definition.

The first step in designing our Web pages is to define a layout grid that will be used as a model for all pages of our site. Establishing a layout grid consists in dividing a page into rectangular zones that stand for the elements of our page (texts, headings, graphics) we want to use (*see figure 11*).
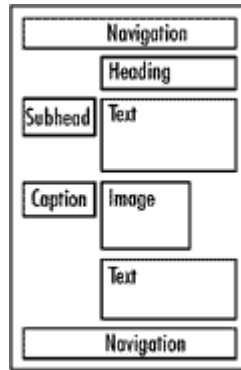
Figure 11 - **A page grid.**

A good layout grid is a grid that manages to find a balance between graphics and text while following the few design principles we will detail on page 35, namely consistency and readability. As a matter of fact, the spatial organisation of graphics and texts may carry different meanings or may have different impact on users. The way we arrange elements must prioritise information and make the reading of the page easier.

A consistent visual hierarchy can be obtained by using visual contrasts and by organising elements in a logical and predictable way. This balance directs the reader's eye into the content.

By the way, it is interesting to note that reading pattern begins at top left of the page (for occidental languages) and zigzags down the page to the bottom right. Few visitors read the whole content of a page; they just scan the page and draw their attention to enhanced elements. Moreover, only the top four inches of the page are visible at first glance on a 15 inch standard office computer monitor. That is why important information must be located at the top of the page. Visitors must not scroll down to reach our major themes. In the same way, navigation, which is an important feature of our pages, must be placed in these four visible inches but must not take too much place.

b. Examples of page grids.

According to Lynch and Horton "consistency and predictability are essential attributes of any well-designed information system"[10]. That is to say that repetition of elements is seen as an advantage for it provides landmarks for users. Here are two examples extracted from the *Web Style Guide*:

Figure 12 – **Diagram of good and bad Web page layouts.**

On the left, we cannot find any feeling of an overall design. Sometimes images and texts are centred, sometimes left-justified. None of the elements seem to answer a visual and organisational coherence. On the contrary, the second page layout corresponds to an organised design grid. Elements are all left-justified and their placement seems to follow a logical and regular order: Navigation > Heading > Image > Subheading > Text, and so on. This proves here that repetition is not boring but functional.

This regularity in design is however dictated by design principles that will help us organise contents within the zones we have just defined when creating a layout grid.

---

[10] Web Style Guide. Retrieved 02 Oct. 2002 < http://info.med.yale.edu/caim/manual/contents.html>.

Actually, a layout grid can be more detailed. The annex I presents a precise grid of CATI's Web site, where not only zones have been defined but also fictitious texts have been included. This kind of page layout (also called "zoning") represents a good starting point of graphic layout creation.

3. Design principles.

a. Consistency.

If consistency is often decried when it is used off-line (for example in books), it is highly recommended on-line. As far as a Web site is concerned, consistency can be applied to navigation, tone, style as well as terminology. The goal is to maintain consistency throughout the site, not only to help users to find their way but to make the site maintenance easier. As a matter of fact, it will be easier for other designers to make changes on a page without starting from scratch. A layout grid and consistent content will be then crucial.

Consistency can be found at three levels. First, we can find page-to-page consistency. For example, logos and navigation must appear in the same location on all pages otherwise users wonder if they are still on the same site, they waste time, lose patience and may quit. Then, consistency may refer to homogeneity within a page. This homogeneity can be guaranteed by the definition of a style for our Web site. Fonts, inter-paragraph spacing, size of subheads will then be applied consistently. We will see in the Web graphics part (*see page 61*) that Cascading style Sheets will help us to reach this visual harmony.

Finally, we can also speak of consistency between our site and the Web standards that must be respected if possible. For instance, a commonly used Web standard is to keep unvisited hypertext links blue and visited links purple. Logos are commonly located in the upper left corner of pages. Terminology often refers to industry standards; hence, we will not call a search button "Find" if the standard is "Search".

Providing a consistent page design is providing landmarks for users.

b. Readability.

Readability must not be taken without due consideration. To dispose text properly is not so easy as we can imagine. We must direct users' eye attention to important features while providing them structured contents. A Web page that contains large texts with no enhancement will appear as a wall of gray and may discourage users from reading.

As Lynch and Horton underline it "Good typography depends on the visual contrast between one font and another, and the contrast between text blocks and the surrounding empty space"[11]. As a matter of fact, boldface fonts can attract users' attention when they are placed among normal fonts. However, inopportune or frequent use of bold text may detract contents from its important character.

Moreover, a patchy typography makes it difficult for users to establish an organisation of our information and predict the location of major elements such as headings and subtitles. In order to make the decoding of text easier, we recommend to justify text on the left and to shorten line length of text blocks. Hence, studies revealed that the eye's span of movement is about 3 inches wide. A wider text, then, may force users to make a particular effort. Even if this effort is not consciously made, it slows down reading and makes it less pleasant. About 40 to 60 characters per line represents a reasonable average.

Finally, we must place the most important information at the top of the page because of page length problems. The method of **inverted pyramid**[12], which is commonly used in magazines and newspapers, is highly recommended too. It consists in starting with the conclusion. This journalistic method is now applied to Web pages for studies have shown that few people scroll down the page.

---

[11] Web Style Guide. Retrieved 02 Oct. 2002 < http://info.med.yale.edu/caim/manual/contents.html>.

[12] Jakob Nielsen. "Inverted Pyramids in Cyberspace". Jakob Nielsen's Alertbox, June 1996. Retrieved 02 October 2002 <http://www.useit.com/alertbox./>.

Beginning with the conclusion allows them to get the most important point of the text even if they do not scroll. Interested users will scroll to get more details.

As we have just seen, page length may trouble readability. Here opinions differ on certain points. If long pages are easier for users to download and for designers to maintain, they do not favour reading on-line.

As a matter of fact, long pages force users to use the vertical scrollbar to navigate. And the big inconvenient of the latter is that it disorientates users by making them lose local context. Most of the time, navigation, which is commonly placed at the top of the page, is scrolled off the screen. We should favour shorter pages as soon as they are intended to be read on-line. On the other hand, longer documents may represent a good material for print. However, we cannot make it a rule because it can vary depending on whether the page contains large graphics or not and on users' behaviours. Are they going to read on-line or print documents? Maybe a good compromise would be to divide long documents into chunks of information. The different pages thus obtained would be linked via hypertext links. But in order not to forget people who prefer reading on paper, our solution would propose a link to a separate page where the whole content would be presented under the form of a printable version.

In the next part of this work, we will learn some technical constraints that we will have to take into account. Creating a Web site is in fact a mix of the idea we have of it and the technical possibilities that are at our disposal.

II. WEB GRAPHICS

A. __Introduction__

According to Marc Bergère, "In the webdesign, graphics and ergonomics should serve the contents. A good graphic web interface attracts internauts […]. The contents will make them come back"[13]. That is the reason why, we must not leave anything to chance.

Data processing is in a constant evolution and therefore numerous platforms, screen resolutions and browser versions compose the computer population. In other words, a Web site will not look the same to all visitors. All the difficulty of the designer's work will rest in finding a technical response to these differences in order to guarantee compatibility and legibility to all (or at least to as many people as possible). However, the designer should not sacrifice the site's identity and content for all that.

In this part we will often analyse display differences at first and then, look at all the possibilities a designer has to get round them and to optimise his/her Web site.

The first thing internauts will look at is the colour harmony. We must not forget that colours do have a particular meaning for each of us. Some may like the graphic interface to such a point that they will make it theirs and go on surfing within our Web site and others may feel uncomfortable in front of it.

We will learn how to define our graphic charter while respecting technical constraints and visitors' expectations and how to optimise our graphics by choosing the best graphic file format. Then, we will see some HTML rules to make loading easier, graphics more interactive and we will try to reach a visual harmony for all the internauts thanks to the use of fonts and Cascading Style Sheets.

Some paragraphs may refer to specific techniques that cannot be completely detailed in a popular work like this one. But, the reader may consult the CD-Rom attached to study a few

---

[13] Bergère, Marc. Chromoweb. Retrieved 28 August 2002 <http://www.smartpixel.net/chromoweb/uks/indexgb.html>.

points thoroughly (*e.g.* How to make an animated or transparent GIF? How to make a favorites icon? …)

B. <u>Colour</u>

Colour on the Web is actually an important component of site design. We must be very careful to choose the palette that will compose our site's graphic charter. Do not we use to say that the experience of colours is very subjective?

As a matter of fact it depends a lot on cultural factors and sensitivity. But choosing colours for an on-line use depends also on objective constraints and particularly display monitors.

1. Display monitors and browser-safe colours.

If internauts use higher and higher resolution monitors, there are still many people using monitors set to 256 colours.

In order to guarantee a Web site that looks the same whether users have got higher resolution monitors or not, we must use the 216 browser-safe colours created by programmers for Netscape from a mathematical colour tube. Using this palette ensures that colours display without dithering on any platform, at any resolution. Which are these 216 colours?

In HTML, we use only hex values to encrypt colours (e.g. #FFFFFF for white). The 216 browser-safe colours hex values can only contain 00, 33, 66, 99, CC, FF.

The browser-safe colours chart (*cf. annex II*) will help us to select colours for our graphic charter. However, we must respect colour symbolism and harmony when associating some of these 216 colours.

2. Colour symbolism.

Colour symbolism may vary according to culture, trends, age, and gender. That is why our site's target audience must have been clearly identified.

If our audience belongs to eastern cultures, it will associate colours with meanings completely different from western cultures' ones. For example, mourning and death is associated to white in eastern cultures and to black in ours. We must be aware of those differences in order to choose colours that will not offend our audience.

The chart below sums up colour symbolism by culture:

Colour	Symbolism
Red	**China** - symbol of celebration and luck, used in many cultural ceremonies that range from funerals to weddings. **India** - colour of purity (used in wedding outfits). **Western cultures** - Christmas colour when combined with green; Valentines Day when combined with pink; indicates stop (danger) at traffic lights. **Eastern cultures** – signifies joy when combined with white.
Yellow	**Asia** – sacred, imperial. **Western cultures** – joy, happiness.
Blue	**China** – associated with immortality. **Colombia** – associated with soap. **Hindus** – the colour of Krishna. **Jews** – holiness. **Middle East** – protective colour. Blue is considered to be the safest global colour.
Orange	**Ireland** – religious significance (Protestant). **Western cultures** – inexpensive goods, Halloween (with Black).
Green	**China – France** – studies indicate this is not a good colour for packaging. **India** – the colour of Islam. **Ireland** – religious significance (Catholic). **Some tropical countries** – associated with danger. **Western cultures** – indicates go (safe) at traffic lights, environmental awareness.
Purple	**Western cultures** – royalty.
Gray	**Western cultures** – loneliness, lifeless and loveless times or settings.
Brown	**Colombia** – discourages sales. **India** – The colour of mourning.
White	**Eastern cultures** – mourning, death. **Japan** – white carnations signify death. **Western cultures** – purify (used in weddings).
Black	**Western cultures** – mourning, death.
Saffron	**Hindu** – sacred colour.
Pastels	**Korea** – trust. **Western cultures** – spring, Easter, pale blue stands for an infant boy; pale pink stands for an infant girl.

Colour trends carried on by clothing fashion may also be helpful to understand our audience preferences. Moreover, gaudy and happy colours would be appropriate for a site with children as the target audience, but would seem inappropriate for adults.

Even gender, according to Natalia Khouw's theory about colour and gender, may influence the perception we have of colours. Men and women do not seem to have preferences for light versus dark colours but women do have for soft colours to bright ones. On the contrary, men would prefer bright colours to soft ones.

As we can see colours are pregnant with meaning and their symbolism is a key issue in our life. The Feng Shui [14], for example, confers virtues on colours such as stability and confidence on blue.

In site design, we must also keep in mind this colour symbolism to send the right message to our visitors.

Colour symbolism chart

Colour	Symbolism
Red	Excitement, energy, passion, desire, speed, strength, power, heat, love, aggression, danger, fire, blood, war, violence, all things intense and passionate.
Yellow	Joy, happiness, optimism, idealism, imagination, hope, sunshine, summer, gold, philosophy, dishonesty, cowardice, betrayal, jealousy, covetousness, deceit, illness, hazard.
Blue	Peace, tranquillity, calm, stability, harmony, unity, trust, truth, confidence, conservatism, security, cleanliness, order, loyalty, sky, water, cold, technology, depression, appetite suppressant.
Orange	Energy, balance, warmth, enthusiasm, vibrant, expansive, flamboyant, demanding of attention.
Green	Nature, environment, healthy, good luck, renewal, youth, vigour, spring, generosity, fertility, jealousy, inexperience, envy, misfortune.
Purple	Royalty, spirituality, nobility, ceremony, mysterious, transformation, wisdom, enlightenment, cruelty, arrogance, mourning.
Gray	Security, reliability, intelligence, staid, modesty, dignity, maturity, solid, conservative, practical, old age, sadness, boring.
Brown	Earth, home, outdoors, reliability, comfort, endurance, stability, simplicity.
White	Reverence, purity, simplicity, cleanliness, peace, humility, precision, innocence, youth, birth, winter, snow, good, sterility, marriage (Western cultures), death (Eastern cultures), cold, clinical, sterile.
Black	Power, sexuality, sophistication, formality, elegance, wealth, mystery, fear, evil, anonymity, unhappiness, depth, sadness, remorse, anger, underground, good technical colour, mourning, death (Western cultures).

---

[14] The ancient art of placement. It addresses the relationship between human beings and their environments and offers keen insights and practical methods that sustain harmonious balance in places where we live and work.

Once these subjective factors understood, we must be aware of technical constraints that make the choice of colours more difficult.

3. Colour harmony.

As a matter of fact, using the browser-safe colours does not guarantee successful graphics. A rule to harmonize colours consists on choosing **a dominant colour** (for big surfaces) and **a tonic colour** (for small surfaces).

The tonic colour, also called "interactive colour" is more particularly used to catch the internaut's attention on interactive and important areas (buttons, hypertext links…).

Let's take the example of a 3-colour harmony (often used for web graphics), here are the choices we have. We can base our graphic charter on:

- a complex separated 3-colour harmony

- a complex consecutive adjacent 3-colour harmony

- a primary/secondary or secondary/primary colour harmony

- an equilateral colour harmony

- or an isosceles colour harmony.

Black, grey or white can be used in any combination without restriction.

4. Colour into practice: CATI's Web site example.

As the reader will see on the CATI's Web site layout (see *Annex I*), the graphic charter that has been chosen lies on a primary/secondary colour harmony.

We can distinguish two dominant colours (blue and orange) and a tonic one (red). All of them belong to the browser-safe colours chart (see *figure 1*).

Figure 1 – **CATI's Web site colours**

Blue is one of the two dominant colours for it is considered as the safest global colour that conveys an idea of confidence, harmony, cleanliness and communication. Moreover, it is usually accepted unanimously and does not bear negative connotations whatever the internauts' cultural background may be.

Orange, the second dominant colour, has been chosen for it conveys the idea of warmth and balance without presenting any disadvantage too.

Red, the tonic colour, provides enhancement to the layout and indicates interactive areas such as the hyperlinks of the left navigation bar.

We must not forget that CATI's Web site is an institutional site, intended to offer huge contents (articles, courses…). That is the reason why we have chosen the white background whose neutrality improves legibility.

Once we have decided which colours will be used for our graphics, we must try to find the best graphic file format to edit them on the World Wide Web.

C. <u>Graphic file formats</u>.

Graphics in their natural state are not adapted to the Web: they are too big to be quickly downloaded. Ideally, an image should weigh 5 Kb maximum and the sum of all images on a page should not exceed the maximum weight of 25 Kb.

That is why we have to compress images, using the standard compression formats such as GIF (Graphic Interchange Format) and JPEG (Joint Photographic Expert Group).

The GIF and JPEG compressions are considered as universal standard formats but many formats are created each year like PNG (Portable Network Graphics) and WIF (Wavelet Image File). However these new formats often do not come up to common browsers' standards. Thus, we will only deal with GIF and JPEG in this chapter.

1. GIF Files.

The Graphic File Interchange (GIF), a proprietary format developed by Compuserve in the 1980's, is actually the most used on the World Wide Web. The main reason of its popularity is that all Web browsers support its format for it is limited to the 256 colours palette but can work with less colours too (2, 4, 8, 16, 32, 128 or 256).

a. How does it work?

The GIF file format uses a lossless compression called LZW (Lempel Zev Welch) which is based on a substitution method of compression. As a matter of fact, the algorithm used seeks out parts of the image that are the same to substitute their sequence of digits for a single number. That is how a file size can be reduced by 30-50% or more according to the amount of repetitions there are in an image. A flat colour will compress at a high rate thanks to the GIF compression.

Graphics software allow **dithering** on GIF files. Dithering is an algorithm used to mix pixels in order to give the impression of a gradation. Sometimes it is necessary to make the image look better but it increases the file size for repetitive colours are then less numerous.

b.   Interlaced GIF.

A GIF file can also be interlaced. In other words, an **interlaced GIF** appear first in a poor resolution, allowing the internaut to have an idea of what the fully rendered picture will look like. The image improves in resolution till it has completely loaded.

Interlacing thus represents a real advantage for GIF compression for JPEG cannot be interlaced as we will see in our next section.

a basic GIF download   *an interlaced GIF download*

Figure 2 – **Interlacing**

Common browsers (such as Netscape Navigator and Microsoft Explorer) support interlacing. But even if some internauts navigate on other browsers, they will see your GIF image anyway but in its basic format.

c.   Transparent GIF.

Another advantage that can lead users to choose the GIF compression is transparency. The **transparent GIF** is also a variant of the basic GIF. It allows the internauts to see the background colour of the browser window they are in. And once again, only the GIF compression allows this propriety.

| *background image* | *transparent GIF* | *result* |

Figure 3 – **Transparency**

You will learn in the "Web graphics" section of our CD-Rom how to make a good quality transparent GIF.

d. Animated GIF.

Finally, GIF can also be **animated** (*cf. the tutorial on our CD-Rom*). Animated graphics may look attractive but they often distract the internaut's attention from the contents of your Web site.

The fact that animation arouses interest (sometimes too much as we have seen) remains an advantage to present strong messages. E-business, for example, uses animated GIF for its advertising banners in order to convey commercial messages.

However, relevant uses of animated GIF are rare. Web style guides often recommend to minimise animation on a site for it is longer to download than normal graphics and does not usually bring anything to contents.

2. JPEG graphics.

The other commonly used graphics file format on the Web is the JPEG (Joint Photographic Experts Group) compression.

a. How does it work?

The JPEG compression uses a mathematical technique called Discrete Cosine Transformation, which, contrary to GIF, is a lossy method; that is to say that the compressed picture loses data and will not appear exactly the same after the compression. However, we will see later that this loss of data depends on the result we want to get.

We may wonder what kind of data the JPEG compression discards…

The JPEG algorithm divides the image into 8x8 pixels squares of data and turns them into a set of curves that match together to make up the image. A high compression will eliminate parts of the data which are less significant (the smaller curves). Actually, these less significant parts are those that the human eye cannot perceive. As a matter of fact, JPEG images are true colour images. If a GIF file is limited to 256 colours, a JPEG one can use 16 million colours – much more than the human eye can distinguish.

b. Why use JPEG?

JPEG stores full colour information and by the way is more appropriate to compress gray-scale images, photographs and "realistic" scenes.

The biggest advantage to JPEG is that it is the user who chooses the level of compression he/she wants. Of course, the more compression he/she applies to an image, the lower quality and the file size. Let us look at the example below:

We can note that the original image (*see figure a*), weighing 358 Kb in its natural state, can be minimised to 4 Kb (*see figure c*) while keeping a good visual quality. But this example underlines also the fact that GIF cannot compete with JPEG as far as realistic scenes are concerned.

Imagine we have a file size to respect: our picture must weigh 13 Kb at the maximum. Just take a look at the example above and you will see that the GIF image (*see figure d*) loses quality (the gradation of the sunset is altered) whereas the high quality JPEG (*see figure b*) remains faithful to the original photograph.

The compression ratio is not always so spectacular. We must keep in mind that the JPEG compression grounds its technique on the human eye's limits. And we better perceive

small changes of colour than changes in brightness. That is the reason why a gray-scale JPEG file will be only 10-15% smaller than the original file at the same visual quality.

Moreover, this lossy compression method is not suitable for graphics with text or sharp edges for it tends to blur them.

3. When use GIF or JPEG?

a. Conclusion.

Choosing the most adapted graphic file format to minimise images file size remains difficult. Issues at stake are especially faster download speeds and non-degraded image quality.

First of all, we must determine how images are intended to be used. As a matter of fact, a medical use cannot accept approximate visual quality and would prefer JPEG compression which supports full-colour images. Someone who wants to store a great quantity of images in an indexed form will also choose JPEG for its huge compression ratio no matter what quality he/she gets. But on the other hand, GIF will be superior to JPEG for a Web use and especially for titles, buttons, logos, transparent images…

Then we must analyse the graphics we want to compress. Is it a true-colour image or a flat colour one? Is the background composed by flat colours or by gradations?

Finally, we must weigh-up the pros and cons of each file to choose the most suitable one. The summary table below will help us to see the advantages of both formats.

b. Summary table.

JPEG	GIF
• stores full colour information (24 bits/pixel, *i.e.* 16 million colours).	• stores 8 bits/pixel (*i.e.* 256 or fewer colours).
• makes image file smaller for archiving (indexes, previews…)	
• good at storing full-colour or gray-scale images of natural, real world (e.g. photographs, fantasy drawings, computer drawn images, medical images…)	• good at storing lettering, images with sharp edges, simple cartoons, line art, flat colours…
	• supports dithering, interlacing, transparency and animation.
• High compression ratios possible depending on the expected image quality (faster download speed).	
• adapted to large images (at least 100 pixels)	• adapted to small images.

D. Graphics and the HTML code.

    1. HTML Tags.

        a. Height and width tags.

Optimising the page and graphics downloading can be made easier by using height and width attributes into the HTML image source tags. Here is an example of what the HTML code then looks like:

<IMG SRC="graphics.gif" HEIGHT="70" WIDTH="220">

If the image source tag contains these attributes, the browser can reserve the accurate space that graphics need even before they have been downloaded. Therefore, it allows the

internauts to have a precise idea of what your layout will look like and to start reading its contents.

On the Web, efficiency and time saving are the two keywords that we must keep in mind to design a Web site that will be adapted to users' needs. But contrary to received ideas, specifying width and height attributes do not speed up the downloading of graphics; it just speeds up the legibility of your page. However, legibility can also be improved by the use of ALT tags.

b. ALT Tags.

The ALT tag (ALT for "alternative text") is also an attribute of the image source tag; it provides an alternate message to viewers who, for some reason that we will detail later, cannot see our graphics. An example of the HTML image source tag is presented below:

<IMG SRC="events.gif" ALT="Campus events">

If you browse graphically, the ALT text appears under the form of a tooltip (*see figure 4*) but if you browse non-graphically, it displays the text instead of the image and thus allows a flowing navigation (*see figure 5*).

Figure 4 – **The tooltip effect**

Figure 5 – **Browsing non-graphically**

Most of the time, people believe that ALT tags play the role of "caption" for an image because of their tooltip effect. But, this effect is nothing less than the tip of the iceberg for there are at least three major reasons to use ALT texts.

The first reason concerns the internauts' navigation habits. Some of them use to browse non-graphically to speed up downloading or to directly access to the contents.

The second reason essentially refers to the notion of accessibility. The blind (or even some sighted people) access a Web site using audio-based technologies that recognise text only.

The last reason is a technical one and deals with the problem of site referencing. Actually, automated indexing programs interpret ALT text as inherent in the page contents.

We then understand why the World Wide Web Consortium's guidelines highly advise to: "Provide equivalent alternatives to auditory and visual content"[15].

To provide the same information (normally conveyed by an image) via ALT tags is called "providing an equivalent experience". That is to say, an ALT tag cannot be a complement to the image (or even a description); it must be a textual replacement of it and moreover, a short and meaningful one. As a matter of fact, some browsers do not interpret word-wrapping or line breaks within ALT texts, so if the latter are inserted into a table, they might break the page layout when browsed non-graphically.

As far as meaning and coherence are concerned, they are easy to check. An infallible test consists in disabling image loading in our browser and read our page aloud, left-to-right, top-to-bottom. Each time we encounter an image without an ALT tag, we say the word "image". This test allows us to figure out how audio-based browser technologies will utter our message. Thus, it quickly comes home to us that creating a Web site must also mean creating a Web site accessible by everyone, especially for a universal media like the Internet.

---

[15] World Wide Web Consortium (W3C). "Techniques for Web Content Accessibility Guidelines 1.0". Retrieved 28 August 2002 <http://www.w3.org/TR/WCAG10-TECHS/>

Now that we are aware of the importance of correctly filling the ALT tag, let us see some particular cases that may be problematic:

- **ALT tag for purely decorative image** (that is to say for non-essential elements).

In this case we had better insert a blank ALT tag (ALT="") into the image source tag.

- **ALT tag for spacer or "punctuation" image**.

Some images, even if they do not convey meaning, are used to organise our content. For example, a list bullet may be replaced by "*" in an ALT tag not to lose its hierarchical role. A divider may be replaced by a sequence of hyphens (-) or underline characters (_).
(ALT="----------------------").

- **ALT tag for navigational symbol** (for example an upwards pointing arrow).

To replace such key elements (for they are used to navigate within the site), we have to find the appropriate corresponding word, such as "Up" or "Down". But we must avoid nonsense that can appear because of adjacent images. As a matter of fact, an audio-based browser will read "Downup" (in the case of two adjacent images: one with the upwards pointing arrow and the other with the downwards pointing arrow), if we forget to leave space within our second ALT text. The best result would be " Down | Up ", obtained by such an ALT tag: ALT=" | Up".

- **ALT tag for Image map hotspots.**

We will learn more about image maps in the next section, but we must already know that each hotspot must contain a meaningful ALT tag as a hotspot plays the same role of a separate image.

In conclusion, as far as ALT tags are concerned, we can remember us the importance of adding **short and meaningful ALT texts for every image** within our Web site.

2. Image maps.

The last subject we will deal with in this more technical part is Image maps (or Imagemaps). An image map is in fact a single graphic into which we can incorporate multiple links according to the hotspots we first define.

Image maps utility seems obvious to everyone in the case of graphic illustration like the example presented below (see *figure 6*) for it allows designers to define regions thanks to shape descriptors (rectangle, circle, polygon) in spite of the linear character of an image.

In this example, the gas refrigeration method is schematised. We can easily imagine that some parts of this illustration need to be detailed in other pages of the site. The three pink areas symbolise three hotspots that are URLs targets.
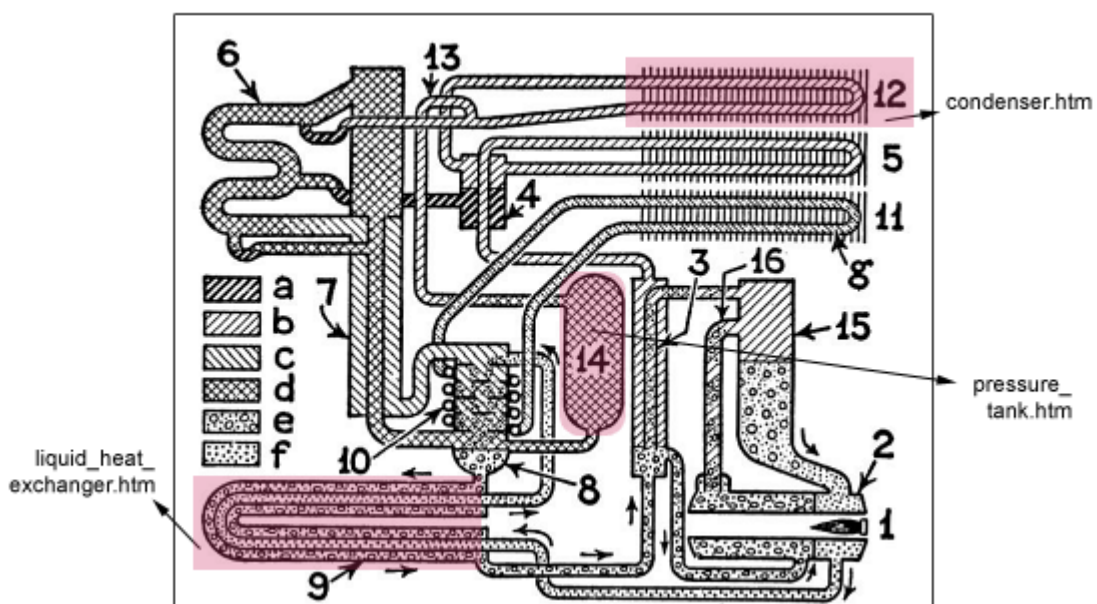

Figure 6 – **Imagemaps (Gas refrigeration)**

Graphics, thus, appear to be more didactical and interactive. However, the internaut has to examine the whole image to detect all the possible URLs targets and, as a consequence, all the possibilities he has to go on surfing on our site because links do not appear clearly.

That is the reason why imagemaps are more commonly used for navigation bars where "live" areas (hotspots) usually correspond with the navigational buttons that are designed in the layout.

For example, the navigation bar of CATI's Web site could be an imagemap with six hotspots leading to six different pages:

Patrick Lynch recognizes the "space-efficient graphic impact" and "visual appeal"[16] of image maps. And as a matter of fact, professionals tend to use image maps more frequently for they are easy to implement and supported by all the browsers.

But once again, we must not forget to provide a textual equivalent of the imagemap for people who cannot handle graphics.

E. **Towards a visual harmony.**

1. **Typefaces.**

Typefaces on a Web site must be carefully chosen as text may be the only thing visible on the screen if users switch off graphics in their browsers for faster download.

The use of different typefaces on a page organises the content and structures the hierarchy between elements of the text (headline, subtitle and so on). But as Marc Bergère underlines it, we must think "useful". Too many fonts on a Web site may be confusing but we may choose

---

[16] Patrick J. Lynch and Sarah Horton. *Web Style Guide*. 2nd ed. New Haven: Yale University Press, 2002.

one "useful" font (among the Web-safe faces as we will see) and a "decorative" one; in other words, a font for the content and another one for titles.

Before giving in to the pleasure of creating fun pages, let us have a look at the problems we may encounter. First, we must take platform differences into account and then, we must guarantee legibility for all. But we will also see that the way we use fonts may be a subject that arouses controversy.

a. Platform display differences.

In his article entitled "Web Type 101, a primer"[17], Daniel Will-Harris warns us right from the beginning: "… what you see isn't necessarily what they get", thus crossing the famous WYSIG (What You See Is what you Get) willingly used in Web design.

As a matter of fact, platform display differences and even browser differences may change the way fonts will look on the Web. According to the platform we have, different fonts may be installed on it.

We will call "Web-safe" faces fonts that the viewer should have. For example, a Mac user will be able to read Times (usually the browser default) and Helvetica whereas a PC user will be able to read Times New Roman (usually the browser default) and Arial. Actually, the constant evolution of platforms and the use of browsers that come with special fonts (*e.g.* Verdana and Comic Sans for IE4) allow us to draw up a more exhaustive list of fonts that are commonly found on personal computers:

- **Common Windows fonts** are: Arial, Comic sans, Courier New, Georgia, Helvetica, Times New Roman and Verdana.

- **Common Mac fonts** are: Arial, Chicago, Courier, Geneva, Helvetica and Times.

---

[17] "Web Type 101, a primer". <u>Efuse</u>. Retrieved 12 September 2002
<http://www.efuse.com/Design/web_fonts_basics.html>.

As we can imagine, the problem will lie in the compatibility between the two platforms and may limit our choice to fonts common to both. But is it really a choice?

Moreover, the text display is about 30% larger on a PC than on a Mac as we can see in the example below:
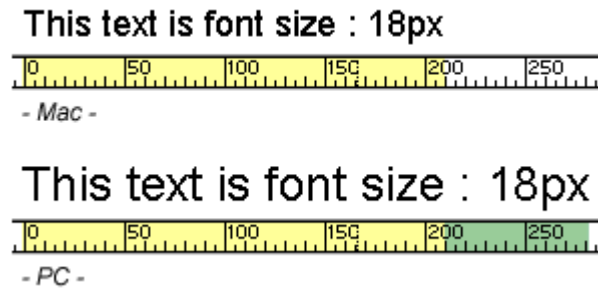
Figure 7 – **Platforms display differences**

This difference is due to the fact that Mac OS uses 18 pixels to render 18 point text whereas a Windows system uses 24 pixels to render the same text. That is why some texts may be illegible to Mac users if they have been designed in a small size on a PC (*e.g.* the use of Arial Bold size 1 is almost illegible). We will see later how to avoid the problem of legibility.

b. How to choose the right fonts for a Web site?

First of all, we must keep in mind Marc Bergère's piece of advice: we have better choose a "decorative" font and a "useful" one.

A decorative font often gives a Web site its identity. As a matter of fact, we can hardly imagine a Web site intended for children whose content would be entirely written in Times New Roman. It would be boring.

In order to get round the Web-safe faces, we can use graphic text to a certain extent. Headlines, buttons, titles will be more attractive and will not weigh up a page if we apply what we have learnt about graphic file compression.

First, we must choose the GIF file format as it is the most adapted to lettering: it will not alter the sharp edges of fonts. Then, we must anti-alias types to the background colour. **Antialiasing**, according to Webopedia[18] is "a software technique for diminishing jaggies – stairstep-like lines that should be smooth. Antialiasing reduces the prominence of jaggies by surrounding the stairsteps with intermediate shades of gray (for gray-scaling devices) or color (for color devices). Although this reduces the jagged appearance of the lines, it also makes them fuzzier.» In other words, antialiasing gives the text a smoother appearance. See the example below:

Figure 8 – **Antialiasing** (the character on the right is antialiased)

Finally, our graphic text must include ALT tags so that internauts can read the text before the graphic loads.

Our decorative font chosen, it is time for us to decide which **useful font** we will use throughout our Web site. We already know that we must choose it among the short list of Web-safe faces. But which are the characteristics of each of them?

- <u>Times New Roman</u> is a serif[19] type, similar to those used in newspapers and easy to read on screen.

- <u>Times</u> is almost similar to Times New Roman but it has no italic version.

---

[18] *Webopedia: the #1 online encyclopedia dedicated to computer technology*. Retrieved 18 August 2002 <http://www.webopedia.com>

[19] a serif is « a small line added as embellishment to the basic form of a character. Typefaces are often described as being *serif* or *sans serif* (without serifs)." *Webopedia*.

- <u>Arial</u>, a sans serif type, looks modern but is not very legible because of its narrow side.

- <u>Verdana</u>, a sans serif face too, is easy to read and larger than Arial.

- <u>Comic Sans</u> looks funnier and may suit for non-professional purpose.

- <u>Georgia</u> is a serif face designed for the screen which guarantees a better legibility.

- <u>Helvetica</u> is a sans serif face that can be compared to Arial.

We must select two or three fonts among these Web-safe faces in a preference order and a serif or sans serif propriety. This order will be very useful if the first selection we made is not installed on the viewer's computer. In this case, the browser will display our second selection, and so on. If none of our three selections are installed, the browser will display a serif (or sans serif according to our criteria) font among those present on the viewer's computer.

Choosing a unique font increases the risk that the browser displays its default font (Times New Roman or Times).

We can specify our preferences in a Web page thanks to the <font face> tag. Here is an example of the HTML code: <font face="Verdana, Arial, Helvetica, sans-serif">.

Let us notice that if the first font is not common to both platforms, the second one should be common.

c. A matter of controversy.

The most problematic issue of typefaces is actually the font size displayed by browsers. Not only a PC platform displays text nearly 30% larger than a Mac platform does, but also some browsers allow resizing and some internauts, unable to read text at a small size, use to specify larger text.

The platform display differences can be controlled by using relative sizing. Relative sizing (*e.g.* <font face="Arial" size="**-2** ">) adjusts the size font relative to the normal screen

display. Whereas the ability that users have to specify larger text cannot be avoided without the creation of Cascading Style Sheets (CSS) which give the designer all the control needed on fonts. Typefaces defined in CSS will be displayed in the same manner on any platform. (*cf*. next chapter). But here comes the matter of controversy: do designers have the right to overlook the users' liberty of specifying text at a comfortable size for them, in order to guarantee the same layout for all?

Maybe a good compromise would be to use Cascading Style Sheets that present many other advantages as we will see while specifying font sizes large enough to be legible on Mac platforms and comfortable to be read.

2. Cascading Style Sheets (CSS).

To reach a visual harmony is not only a question of aesthetics but a question of identity. For instance, if our visitors are used to see blue pages on our Web site and once get to a white one, they may wonder if they are still within our Web site. But this homogeneity, which is necessary for users not to loose their bearings, can turn to nightmare if it has not been thought of before the Web site building. Let us imagine a Web site that contains hundreds of pages. Its designer has not anticipated possible changes in the presentation of texts or even colours. How long will it take him/her to change these elements within hundreds of pages? Can he/she be sure not to forget any pages? This nightmarish example leads us to focus on Cascading Style Sheets (later abbreviated to CSS) that seem to answer this kind of problem. Style sheets are in fact files that define the layout of a document and describe how it is presented on screen or in print.

a. A CSS overview.

As far as a Web page presentation is concerned, HTML offers no laudable effects. As a matter of fact, we only have a few possibilities at our disposal to change the font size for

example. Moreover, as we have seen on page 58, we cannot swear to it that the text will be identically displayed according to the platform we use. Another example that shows the HTML limits is text ident. To ident text, HTML only proposes wangling solutions to us like putting a transparent image that will artificially shift the text.

With the appearance of style sheets roles have been redistributed. The HTML returns the primary functions it was designed for, that is content and structure, whereas CSS deal with presentation without altering the structure. But what are exactly Cascading Style Sheets? What can we use them for?

CSS control the rendering of HTML elements such as fonts, colours, leading or margins thanks to a collection of rules they are containing. CSS are easy to use for they need not change the HTML code; they are just added to it but in an independent way. Furthermore, style sheets allow us to avoid the problem of display differences on platforms; they are more and more supported by browsers. Actually, most browsers that are distributed today support CSS. This is the case for Internet Explorer 4 (and later versions), as well as Netscape Navigator 4 (and later versions).

b. Advantages of Cascading Style Sheets.

The separation of content and presentation between HTML and style sheets thus presents some major advantages. The biggest one (at least from an author's point of view) is that presentational preferences are defined in the style sheet once and may be applied to the entire site. No need to specify in each page the font colour or the kind of horizontal rule we want to use. The style sheet will affect the whole site without changing the HTML.

Moreover, the characteristic of specifying preferences in a unique location presents another advantage but this time, from the user's point of view. As a matter of fact, style sheets appreciably reduce the download time. Obviously, it will take more time to load a page which

contains ten horizontal rules that are directly encoded in HTML, than to load a single file that contains all the style information. The HTML tag <HR WIDTH="50%" SIZE="2" ALIGN="CENTER"> repeated the number of times the author wants this horizontal rule to appear, is replaced by a single line for the whole site:

HR { width : 50% ; size : 2 ; align : center }

Cascading Style Sheets also offer more flexibility in terms of desktop publishing quality layout. They allow us to get the exact rendering we want. If some pages of our Web site are destined to be printed, CSS will improve the rendering on paper by taking into account the display differences that exist between computer screen and paper.

But let us have a look at the syntax of CSS and the practical applications we can make of them.

c. An introduction to Style Sheets.

As we have just learnt it, a style sheet is composed of rules that determine how style must be applied to the document. A rule may be schematised like this:

Selector { property : value ; property : value ; …   }

The **selector** is the link between the HTML document and the style. All HTML tags can become selectors (e.g. H1, BODY, BLOCKQUOTE…) and then are called "type selectors" for they concern all elements of their type. For example, the following rule will affect all level 1 headings and will turn their font colour to blue throughout a Web page:

H1 { color : blue }

But there are also "attribute selectors" that allow us to apply style to particular parts of our page and not to the whole page. Attribute selectors will be discussed later in this section.

The **property** is the quality that an element possesses such as color, font-size, background, letter-spacing and so on. 120 properties exist and values can be assigned to them.

Values are in fact a specification of the property. For example, it could be "bold" if assigned to the font-weight property, or "15 pt" if assigned to the font-size one.

If we just want some paragraphs to be blue and bold, we can define attribute selectors. CLASS will be the attribute that will affect some BODY elements. A period ("."), used in the selector, announces a class name. In our example the rule will become:

.bluebold {color:blue , font-weight:bold }

In order to assign this style to the paragraph we want to change, we just have to add **class=bluebold** in the font tag.

The reader may wonder where he/she has to put all these rules in his/her Web site? When a style sheet affects only a Web page, it can be "glued" to the document. That is to say, we have to put the style sheet inside a style element at the top of our document. Rules must be enclosed by the <STYLE></STYLE> tags. See the example below:

```
<html>
<head>
<title>CSS</title>
<meta http-equiv="Content-Type" content="text/html;
charset=iso-8859-1">

<style>
H1 {color:green}
.bluebold {color : blue; font-weight : bold}
</style>

</head>
<body bgcolor="#FFFFFF" text="#000000">
<h1>Level 1 heading </h1><br>
<font class=bluebold>I want this text to be blue and
bold.</font>
</body>
</html>
```

However, if we want to apply the style sheet to the whole site, it must be in a separate file with the extension .css. This file will be applied to pages whenever it was specified:

<link type="text/css" rel="stylesheet" href="cati.css">

In this example, all pages including this style element will use our style sheet named cati.css.

There are commonly used tasks with CSS. As a matter of fact, CSS are often used to define fonts, margins or links. We cannot detail all the properties and values that can be assigned to a selector. We will just have a look at the example below that shows the formatting of a small text by using CSS:

Here is the HTML code and its style definition:

```
<html>
<head>
<title>CSS</title>
<style>
font {color:black; font-family:verdana ; font:10pt}
H2 {font:18pt ; color:purple ; text-decoration:underline}
UL {font:12pt ; color:#990000 ; font-style:italic ; margin-left:75}
A:hover {background:yellow}
</style>
</head>
<body bgcolor="#FFFFFF" text="#000000">
<h2>When CSS are applied to links, margins and fonts...</h2>
 <br>
   <font><a href="home.htm">The Research Centre CATI</a> works in the field of Humanities Computing
studying the relevance of hypermedia presentation to cultural  history subjects: the uses of literary or historical
databases, image analysis, spatial perception in computer assisted design and virtual reality technology. CATI
implements its research program in three directions:</font>
<ul><li>Development of information technologies in the Humanities and Social Sciences
<li>Authoring of e-documents
<li>Theoretical interpretation of IT methodology</ul>
</body>
</html>
```

In conclusion, we can say that Cascading Style Sheets give us a real opportunity to reach visual harmony and to make the site updating easier.

III. IMPLEMENTATION

Of course, this work cannot relate in detail the whole process of site building. In the two preceding parts, we have learnt enough to grasp quite well the key points that lay out a site creation such as accessibility, usability, legibility, easy download and so on. But we cannot deal here with all the technical tips that compose a webmaster's everyday life.

HTML editors become easier to use and a large amount of information or tutorials can be found on the Web to help us learn how to build a Web page. Thus, it allows us to focus on particular points. As a matter of fact, Web design books emphasize all the steps of a Web site creation without dealing with the site publishing on-line and its consequences.

We will see in this part that creating a Web site implies first that we find a Web server to host it. Then, once our Web site is on the Web, we must submit it to search engines in order to get a high ranking. That is how our site will be accessible to people. Finally, we must maintain this Web site up-to-date. In fact, a Web site requires regular attention. External links have to be checked, contents have to be upgraded. Moreover, we have to guarantee the Web site durability by providing a technical handbook that will guide our successors.

A. <u>Publishing on-line.</u>

Now that we have created our Web site on our hard drive, that the links seem to work, the graphics seem to be correctly displayed, we are ready to make this site accessible to all.

1. Web site hosting.

a. A brief overview.

In order to publish our Web site on-line, we have to place our files on a Web server. The latter is provided by a company that agrees to host our site on its Web server. Most of the time Internet Service Providers (ISP) include web site hosting in the services offered to their clients under some conditions (*e.g.* storage limits).

According to the kind of Web site we built, we have the choice between the companies that provide free hosting or those that provide paid hosting. If our Web site is a personal site with a non-commercial purpose, free Web sites can offer us a good deal. But they have the right to place advertisements on our Web site in exchange for Web hosting. The amount of disk storage may be limited and varies according to the company.

Nevertheless, if our site has a commercial goal or if we cannot stand advertisements, the best solution is the paid Web sites for they are free of advertisements. We will then have our Web site hosted at a charge of nearly $15 a month. But this case may not be suitable for professional Web sites for the URL is a sub-category of the company that provides Web hosting. Hence, if we want to publish a commercial or an institutional Web site, we had better buy our own domain name and register it. We will then have our Web site hosted with what is called a "virtual server". But we will not deal here with that case because the present work is intended for what we can call an "academical" use.

b. Some useful addresses.

Here are two useful addresses for people who want to host their Web site in France:

Lycos Multimania
<http://www.multimania.lycos.fr/myaccount/freehosting>

Tiscali
<http://www.chez.tiscali/savoirplus>

These two free Web sites provide not only free Web hosting but also a multitude of services such as: tutorials, FTP, site referencing assistance, personal search engines, HTML validator and even statistics about our Web site traffic.

Once we choose a company to host our Web site, we are ready to transfer the files that are still on our hard drive to the Web server. Only then, will our Web site be accessible to all on the World Wide Web.

2.Transferring files to the Web server.

a. The File Transfer Protocol program (FTP).

The main method of publishing a Web site is to use FTP (short for File Transfer Protocol). A protocol describes in fact the way networked computers communicate to each other. As far as FTP is concerned, it is the language that computers connected to the Internet use to transfer files to and from each other. Every time we try to download a file or an audio extract from a Web site, we use FTP without even knowing how.

However, publishing a Web site on-line requires a powerful tool, that is a FTP client, to transfer our files from our hard drive to the Web server. This operation is an ongoing process for it has to be repeated each time we make changes or update pages.

In order to use a FTP program, we must make sure of possessing the login account that our Web site hosting must have provided to us. A login account is composed of the server address, a user name and a password. Without this information, we will not be able to log in to our dedicated Web server.

Dozens of FTP programs exist on the market. Some can be downloaded as shareware or registered programs. One of the most popular is WS_FTP which can be downloaded from Ipswitch[20]. In the next section, we will show how to use this FTP client and explain some of its options.

---

[20] Ipswitch. Retrieved 30 Sept. 2002 <http://www.ipswitch.com>.

b. A brief tutorial.

In order to demystify the so-called complexity of FTP use, we will focus our attention on the few steps that must be known to succeed in uploading files to our We server.

In this overview we take it for granted that we use WS_FTP as FTP Client, but the functions that we will describe can be found on other clients.

Once the program is launched, a set-up window (*see figure 1*) appears on the screen.
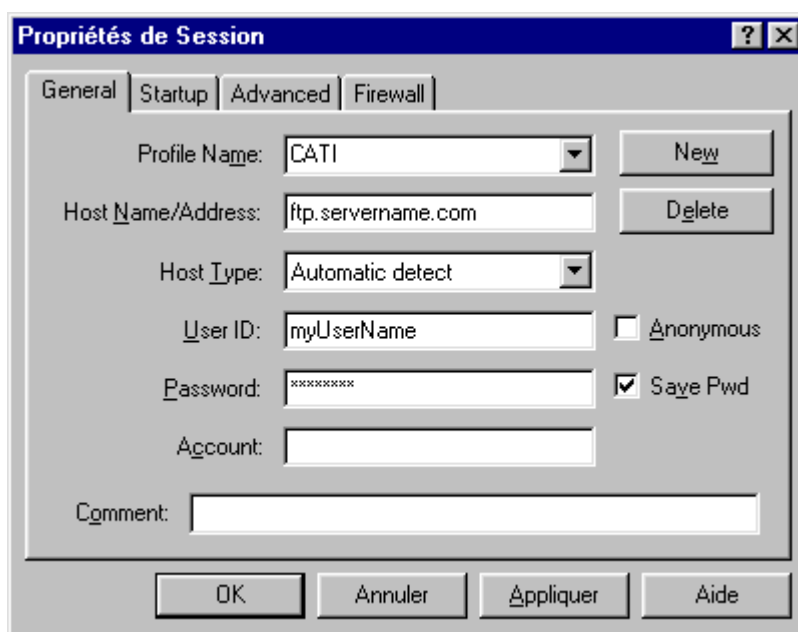
Figure 1 – **WS_FTP Setup Window**

All the fields need not be filled. Here are the few necessary fields:

The profile name is in fact the name we want to give to our connection. It is up to us to choose the name we can easily remember. For our example we chose "CATI" for we want to upload files for CATI's Web site. Then we have to fill conscientiously the **Host Name/Address** field thanks to the address provided by our Web host. Addresses are generally under the form of ftp.servername.com. We must also have received a **User ID**, that is to say a user name, and a **password** from our Web host. Both have to be entered in the proper fields.

We can notice that our password is automatically hidden by star symbols to respect confidentiality.

Actually, these four fields are enough to connect to our Web server, so we will not go into details about the other fields. As a matter of fact, once the "OK" button is hit, we are connected on the Web server. Figure 2 is a screen capture of the screen we are then on:
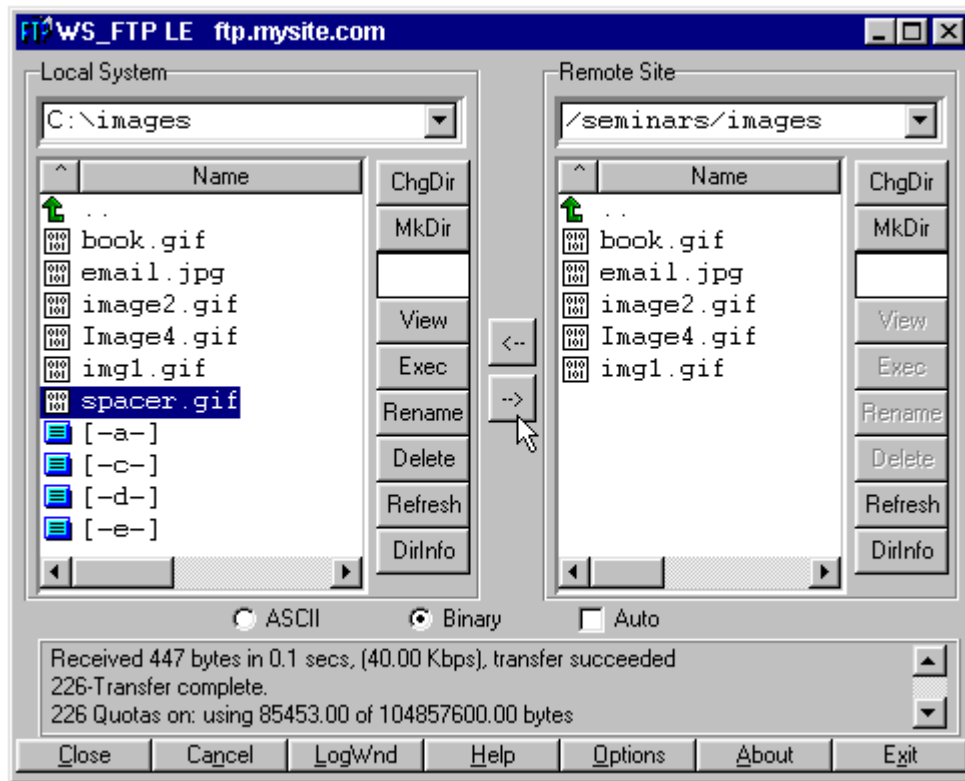
Figure 2 – **Uploading files with WS_FTP**

The window is divided into two parts. In the left frame, we can find files that are in a directory of our local system (*e.g.* our hard drive). We may have to navigate to the directory where we have placed the files we want to send. Thus, we just have to click on the green arrow on the left until we get to the right place.

In the right frame there are files which are present on our Web server. In order to transfer files to the server, we must select them in the left window and click on the "right pointing arrow" (*see figure 2*) that is between the two frames. Once the files are transferred, they will appear in the right window.

That is all. We have just published our Web site; it can be seen on the Net then.

B. <u>Site referencing.</u>

It is no use creating a Web site if no-one is aware of its presence on the World Wide Web. Nearly 85% of the 400 million people online (i.e. 400 million potential visitors for our site) use search engines to find what they are looking for on the Internet. Thanks to chosen keywords, users expect to find in the first pages of search results, the Web site that will answer their needs. If our Web site appears in this list of search results, we can say that it is ranked. However if it appears on page 100, users will not find them (despite the fact that it is ranked). The issue that is at stake then, is not to get a Web site ranking but a high ranking Web site. That is to say, our site must achieve a favourable position over sites that deal with the same subject when entering a general search term (end not our URL that is inevitably unique) in search engines.

As we will see further on, there are methods which allow us to be highly ranked but these methods have to be thought of during the whole period of design. To foresee site referencing during the Web site creation is called **Search Engine Optimisation** (SEO).
In order to optimise referencing, we will try to understand how search engines work, which keywords we must use and how we must integrate them into the HTML code before submitting our Web site.

1. Search engines.

The knowledge of search engines and especially the manner they work will help us to put in our pages the precise criteria they actually take into account and seek. There are two types of search engines:

- Spider indexed search engines

- Human indexed search engines.

Spider indexed search engines represent nowadays nearly 95% of search engines, and that is why we will essentially focus attention on them. But let us talk a little about **human indexed search engines** whose characteristic is to index web pages manually without the help of programs (spiders). In other words, a person is in charge of categorising our Web site according to the way he/she perceives it. In this case, we have no means to manipulate the search engine which is in fact just a directory. If the person referencing our site decides that it belongs to the "poetry" category instead of "science fiction", it cannot be helped.

Spider indexed search engines, as far as they are concerned, are large databases of Web sites and index automatically web pages thanks to what is called a "spider".

Webopedia defines a spider as a "program that automatically fetches Web pages. Spiders are just used to feed pages to search engines. Because most Web pages contain links to other pages, a spider can start almost anywhere. As soon as it sees a link to another page, it goes off and fetches it"[21].

When we submit our Web site to a search engine, its spider will visit it to gather information. The advantage of spider indexed search engines is that spiders may come and visit our Web site several times, whereas human indexed search engines will visit it once. It allows then an optimised referencing as it takes into account the possible changes we did on pages.

The way a spider acts depends on a few criteria as we will see. Search engine optimisation's task will be to answer these criteria and give the spider the information it looks for.

2. Keywords.

a. How a word becomes a keyword.

---

[21] *Webopedia: the #1 online encyclopedia dedicated to computer technology*. 18 Aug. 2002 <http://www.webopedia.com>

The only thing that can feed a spider is text, no matter if it is an ALT text, a title or an article. It is text. The spider will count each word on a Web page and its occurrence. The first words that appear most frequently will determine the page keywords and will be sent to the search engine. Our Web site will be indexed thanks to these keywords. That is why we must choose the right words and learn how to use them in an effective manner.

Keywords are inherent in each page. It thus allows us to submit pages individually to a search engine. Hence our Web site will be ranked according to several topics.

As we will see in the Meta tags section on page 75, keywords will also be used to enrich the HTML code. They must be pregnant with meaning and be characteristic of our site. We can thus easily understand why we have to avoid using common words like "the" or "and" in the keyword definition. However spiders are programmed not to take into account these recurrent common words.

Moreover key phrases (*i.e.* a combination of words) are welcome, what allows us not to use the coordinating conjunction "and" too frequently. For example, if our Web site deals with "humanities" and "computing", site ranking will be more effective if we combine these two words as "humanities computing". A few sites like Worldtracker[22] can provide key phrase assistance by pointing out the most searched keyword or key phrases related to our site subject.

Once our keywords chosen, we must learn how to use them properly in our pages.

b. Keyword Rich Text.

Writing a text with search engine optimisation in mind is called writing keyword rich text. The acknowledged goal is nothing but using our keywords as many times as we can, while satisfying two demanding entities: our target audience and the search engines. As a

---

[22] <u>Wordtracker: ... The Ultimate Tool to Increase Your Website Traffic</u>. Retrieved 21 Sept. 2002 <http://www.wordtracker.com>

matter of fact, if keywords are just enumerated, they may be considered as spamming by search engines and as confusing and redundant by visitors.

Keywords have to be spaced out so that the text does not sound stilted. Experts recommend writing a text of at least one hundred words insofar as spiders are able to treat one hundred Kilobytes of information per page; that is to say nearly one hundred words. The ideal average as far as keywords use is concerned, is one keyword every twelve words. But once again, trying to get high ranking must not turn to indigestible contents.

Keywords have another role to play when inserted in meta tags.

3. Meta tags.

Meta tags are HTML tags that provide information about a Web page without affecting its display. They are included in the HTML code between the <head> and </head> tags at the top of the document and may indicate the title of the page, the author, the page description and its keywords.

Contrary to popular belief, furnishing meta tags with keywords is not enough to expect a high ranking. If search engines use meta tags to get some information about our site, directories do not. Moreover, abusive use of meta tags has forced spiders to favour the role of keyword rich text rather than meta tags. As a matter of fact, spiteful people put irrelevant words such as "sex" or trademarks in their keywords tag in order to drain traffic on their site. Nevertheless, meta tags also play a role in search engine optimisation even if it is not so important as we imagine. We must consider the use of meta tags as an additional means of referencing which is very useful when combined with others.

a. The Title Tag.

If we had to choose to fill a single meta tag, it would be the title tag. Not only is it what users see at the top of their browser window, but it is the first thing a search engine displays in its page results.

The title tag must be descriptive and specific to each page and must include the keywords we have defined. It must remain short (about 150 characters) and must avoid using common words, just like keyword rich texts.

Here is for example the title tag of Webopedia home page:

<title>Webopedia: Online Dictionary for Computer and Internet Terms</title>

And here is what Google.com displays when answering the request "online dictionary Internet":

The title tag appears first, but some search engines display a small description of the site that corresponds to the Meta Description Tag's content.

b. The Meta Description Tag.

The meta description tag is also included in the HTML code between the <head> and the </head> tags. Let us take the example of Webopedia's description tag to note the exact syntax:

<meta name="description" content="An online computer dictionary and internet search engine for internet terms and techincal support.">

Some search engines display this information in their page results. See the example below:

The meta description tag must be informative and not too long. It is recommended not to exceed 250 characters for some search engines may not display the entire description beyond. In order to optimise search results we had better put our keywords at the beginning of this tag.

c. The Meta Keywords Tag.

As we have seen in the introduction, the meta keywords tag has little importance as far as search engine optimisation is concerned. Abuses have forced search engines to get information elsewhere (that is directly in the page's content). But it is still worth filling this tag as some search engines go on interpreting it. However, a few rules have to be respected. We had better use plural over singular, and avoid commas between keywords. As a matter of fact, juxtaposed words may be indexed as a key phrase and may increase search results. Finally, the meta keywords tag must not exceed 1000 characters maximum for search engines will ignore the rest.

Here is Webopedia's meta keywords tag as an example:

<meta name="keywords" content="online computer dictionary internet encyclopedia education terms technical support">

4. Web site submission.

a. Preliminaries.

Now that we have optimised our Web site in order to get high ranking, it is time to submit it to search engines. There are actually two means of submitting a Web site: manually or thanks to automated submission software. The latter are largely controversial; more and more search engines skip sites that have been automatically submitted.

The best submission method remains the manual one. It allows us to submit our home page and other key pages to different search engines. But we must take care not to spam search engines insofar as many use the same databases (s*ee section b*).

Usually, it may take from 4 to 6 weeks for a site to be indexed. However, a spider may come and visit our site earlier if another indexed Web site has got a link pointing to it.

b. Main search engines.

Here is a list of the major search engines[23]:

Altavista
(Search engines that also use this database- Looksmart)
<http://addurl.altavista.com/sites/addurl/newurl>

The Open Directory Project (DMOZ)
(Search engines that also use this database- Ask Jeeves, Lycos, Netscape)
<http://dmoz.org/add.html>

---

[23] This list is provided by Search Engine Strategies. Retrieved 21 Sept. 2002
<http://www.searchenginestrategies.biz/submittingyourwebsite.htm>

Fast

(Search engines that also use this database- Lycos)

<http://www.alltheweb.com/add_url.php>

Google

(Search engines that also use this database- Yahoo)

<http://www.google.com/addurl.html>

Inktomi

(Search engines that also use this database- AOL, Excite, MSN, Overture)

<http://submitit.bcentral.com/msnsubmit.htm>

MSN

<http://submitit.bcentral.com/msnsubmit.htm>

Yahoo

<http://docs.yahoo.com/info/suggest/>

As a conclusion, as far as site referencing is concerned, we can underline the important role of text. Sites that largely use graphics or flash animations sacrifice their chance to get a prime ranking for the benefit of aesthetics.

C. <u>Maintenance.</u>

Now that we have uploaded our Web site, another period begins, that of maintenance. As a matter of fact, a Web site requires regular updating. We must not forget that our visitors are one click from our concurrent. If we do not provide them with newness they may never come back. But this follow-up task is not only user-oriented, but it also concerns designers.

1. Maintenance and users.

a. Content changes.

Whenever we make changes on our pages, we must tell users which pages are concerned and when updating occurred. This information will make their navigation easier by helping them finding new material or skipping old contents they have already seen.

The home page is the best location to inform users of changes; a hypertext link can direct users to the new document. However, if our Web site is intended to bring up daily news or frequent communications, it may be more relevant to create a "What's New?" section accessible from the home page too. If only a paragraph is updated, a "NEW" graphics will be enough to catch users' attention.

b. Checking links.

To reach a dead-end page is always a frustrating experience. It may also discourage many of us to the point of quitting a site. That is why we must check our links regularly (and especially links to other Web sites) and correct them or remove those that no longer work. As this checking may be time consuming for large sites, we can use existing tools such as Doctor HTML[24]. It is a program that searches our Web site for mistakes. It locates, for example, dead links and reports problems encountered so as we can locate and correct them rapidly.

c. Importance of feedback.

Responding seriously to users who give feedback is part of the maintenance process. Their remarks or suggestions must be taken into account for they often underline anomalies on our Web site. Moreover, users will appreciate our interest in their opinion and may come back.

---

[24] Doctor HTML : Quality Assessment for the Web. Retrieved 02 Oct. 2002
<http://www2.imagiware.com/ RxHTML/>

Another feedback that has to be carefully analysed is that of tracking reports insofar as statistics have been implemented on our Web site. These reports provide us with the number of visits our site receives, the number of disconnects, the order in which users view our pages. Such an information is worth its weight in gold for it shows us what has to be improved on our site (*e.g.* navigation).

2. Maintenance and designers.

A Web site author must keep in mind that other designers may work on the site he/she has created. In order to make their work easier a few rules have to be respected.

a. Providing a procedures manual.

A good maintenance lies in a good anticipation of the transfer of competence. As a matter of fact, we may not be the only one appointed to Web site management. That is the reason why a procedures manual is necessary. Our successors must find, for example, information relative to the naming and file directory system we established during the development of the site. This manual must also contain our graphic charter in case others want to add graphic elements to the site while keeping its visual identity.

b. Administration tools.

A new trend in publishing on-line seems to appear. If the Web site creation remains the province of a webmaster, its updating tends to be made by non-specialists thanks to administration tools. These programs generate content. An interface provides a form that has to be completed with elements we want to publish (texts, captions…). Once submitted this form generates HTML code and automatically creates a file.

Once again, this new tendency confirms what we have stated in the introduction of this work. That is to say, publishing on-line is now the concern of each of us.

Conclusion.

Publishing on-line is in fact a skilful combination of dream and reality. First, dream because it allows thousands of people to feed the incredible knowledge database that is represented by the Web. Each of us can share his/her interests, knowledge, feelings with millions of internauts worldwide. Then, reality because publishing on-line requires few technical skills, but also presents some constraints.

The best example could be this work. At first glance, it may look like one among the numerous guides to the Web site creation that we can find everywhere. As far as the content is concerned, it may indeed. But the research process started from the idea that insofar as the Internet provides us with the biggest knowledge collection ever seen, we could make our search on the Web. That is the reason why, you could find many electronic references in our bibliography. Here is the dream. We could find almost all we needed on the Net to create this work. Nevertheless, as reading on-line is not always convenient, hundreds of pages were printed to approximate a more traditional study. Here is reality and its constraints.

Hence, we must take benefit from each medium. The Internet will not be able to replace books for example. But it can be seen as a complementary resource (and especially for educational purpose) that opens new horizons.

By the way, we can notice that in any of the processes we have been involved in (design, computer graphics and implementation), the will to satisfy needs of others is predominant. As a matter of fact, we have designed a Web site according to a target audience and its expectations. We have created graphics according to their equipment, their sensitivity (*e.g.* to colours) and finally we have published our site and tried to make it accessible to all in order to share knowledge with them.

Maybe the World Wide Web is a manner to show us how to become altruistic starting from an individualistic activity that is computing?

Bibliography

<u>Books</u>

Bernard, M. L., & Chaparro, B. S.. "Searching within websites: A comparison of three types of sitemap menu structures". *Proceedings of The Human Factors and Ergonomics Society 44th Annual Meeting in San Diego*, 2000, 441-444 <http://psychology.wichita.edu/hci/projects/sitemap.pdf >. (PDF format).

Lynch, Patrick J., and Horton, Sarah. *Web Style Guide*. 2nd ed. New Haven: Yale University Press, 2002.

Meyer, Eric A. *Eric Meyer on CSS: Mastering the Language of Web Design.* Indianapolis: New Riders Press, 2002.

Nielsen, Jakob. *Designing Web Usability : The Practice of Simplicity.* Indianapolis: New Riders Publishing, 1999.

Nogier, Jean-François. *De l'ergonomie du logiciel au design des sites web*. Paris: Dunod, 2001.

<u>Electronic resources</u>

Bergère, Marc. <u>Chromoweb</u>. Retrieved 28 August 2002 <http://www.smartpixel.net/chromoweb/uks/indexgb.html>.

Convera Canada inc. <u>Le grand dictionnaire terminologique</u>. Last updated September 2002. Retrieved 02 October 2002 <http://www.granddictionnaire.com/_fs_global_01.htm>

"Design Tips and Guidelines". <u>About</u>. Retrieved 05 September 2002 <http://webdesign.about.com/cs/designtips/>

Harris, Daniel. "Web Type 101, a primer". <u>Efuse</u>. Retrieved 12 September 2002 <http://www.efuse.com/Design/web_fonts_basics.html>.

IBM®. "Web design guidelines". <u>IBM</u>. Retrieved 02 October 2002 <http://www-3.ibm.com/ibm/easy/eou_ext.nsf/publish/572>

Imagiware, Inc. <u>Doctor HTML : Quality Assessment for the Web</u>. Retrieved 02 October 2002 < http://www2.imagiware.com/RxHTML/>

Instone, Keith. "Location, Path & Attribute Breadcrumbs". <u>Keith Instone</u>. Retrieved 07 September 2002 <http://keith.instone.org/breadcrumbs/>.

Khouw, Natalia. "The Meaning of Color for Gender". <u>Color Matters® Research</u>. Retrieved 28 August 2002 < http://www.colormatters.com/khouw.html >.

Lynch, Patrick J., and Horton, Sarah. <u>Web Style Guide, 2nd ed</u>. Last updated 04 October 2002. Retrieved 14 September 2002 <http://www.webstyleguide.com/>

Nielsen, Jakob. <u>Useit.com : Jakob Nielsen's Website</u>. Retrieved 08 September 2002 <http://www.useit.com/>

---. "Search : Visible and Simple". <u>Jakob Nielsen's Alertbox, May 13, 2001</u>. Retrieved 08 September 2002 <http://www.useit.com/alertbox./>.

---. "Site Map Usability". <u>Jakob Nielsen's Alertbox, January 6, 2002</u>. Retrieved 08 September 2002 <http://www.useit.com/alertbox./>.

---. "Inverted Pyramids in Cyberspace". <u>Jakob Nielsen's Alertbox, June 1996</u>. Retrieved 02 October 2002 <http://www.useit.com/alertbox./>.

Oppedahl and Larson. "Web Law FAQ". <u>Patents.com</u>. Last updated 27 April 1999. Retrieved 21 September 2002 < http://www.patents.com/weblaw.sht>

"Search Engine Strategies Guide". <u>Search Engine Optimization</u>. Retrieved 02 October 2002 <http://www.searchenginestrategies.biz/index.htm>

<u>Webopedia™: Online Dictionary for Computer and Internet Terms</u>. Retrieved 10 September 2002 < http://www.webopedia.com/>

World Wide Web Consortium. "Techniques for Web Content Accessibility Guidelines 1.0". <u>W3C</u>. Last updated 03 October 2002. Retrieved 04 Octoberr 2002 <http://www.w3.org/TR/WCAG10-TECHS/#gl-provide-equivalents>

ANNEXES

Annex I **: CATI's zoning.**

Annex II **: Browser-safe colours (hex values)**

#FFFFFF		#CCFFFF		#99FFFF	
#FFFFCC		#CCFFCC		#99FFCC	
#FFFF99		#CCFF99		#99FF99	
#FFFF66		#CCFF66		#99FF66	
#FFFF33		#CCFF33		#99FF33	
#FFFF00		#CCFF00		#99FF00	
#FFCCFF		#CCCCFF		#99CCFF	
#FFCCCC		#CCCCCC		#99CCCC	
#FFCC99		#CCCC99		#99CC99	
#FFCC66		#CCCC66		#99CC66	
#FFCC33		#CCCC33		#99CC33	
#FFCC00		#CCCC00		#99CC00	
#FF99FF		#CC99FF		#9999FF	
#FF99CC		#CC99CC		#9999CC	
#FF9999		#CC9999		#999999	
#FF9966		#CC9966		#999966	
#FF9933		#CC9933		#999933	
#FF9900		#CC9900		#999900	
#FF66FF		#CC66FF		#9966FF	
#FF66CC		#CC66CC		#9966CC	
#FF6699		#CC6699		#996699	
#FF6666		#CC6666		#996666	
#FF6633		#CC6633		#996633	
#FF6600		#CC6600		#996600	
#FF33FF		#CC33FF		#9933FF	
#FF33CC		#CC33CC		#9933CC	
#FF3399		#CC3399		#993399	
#FF3366		#CC3366		#993366	
#FF3333		#CC3333		#993333	
#FF3300		#CC3300		#993300	
#FF00FF		#CC00FF		#9900FF	
#FF00CC		#CC00CC		#9900CC	
#FF0099		#CC0099		#990099	
#FF0066		#CC0066		#990066	
#FF0033		#CC0033		#990033	
#FF0000		#CC0000		#990000	
#66FFFF		#33FFFF		#00FFFF	
#66FFCC		#33FFCC		#00FFCC	
#66FF99		#33FF99		#00FF99	
#66FF66		#33FF66		#00FF66	

#66FF33		#33FF33		#00FF33	
#66FF00		#33FF00		#00FF00	
#66CCFF		#33CCFF		#00CCFF	
#66CCCC		#33CCCC		#00CCCC	
#66CC99		#33CC99		#00CC99	
#66CC66		#33CC66		#00CC66	
#66CC33		#33CC33		#00CC33	
#66CC00		#33CC00		#00CC00	
#6699FF		#3399FF		#0099FF	
#6699CC		#3399CC		#0099CC	
#669999		#339999		#009999	
#669966		#339966		#009966	
#669933		#339933		#009933	
#669900		#339900		#009900	
#6666FF		#3366FF		#0066FF	
#6666CC		#3366CC		#0066CC	
#666699		#336699		#006699	
#666666		#336666		#006666	
#666633		#336633		#006633	
#666600		#336600		#006600	
#6633FF		#3333FF		#0033FF	
#6633CC		#3333CC		#0033CC	
#663399		#333399		#003399	
#663366		#333366		#003366	
#663333		#333333		#003333	
#663300		#333300		#003300	
#6600FF		#3300FF		#0000FF	
#6600CC		#3300CC		#0000CC	
#660099		#330099		#000099	
#660066		#330066		#000066	
#660033		#330033		#000033	
#660000		#330000		#000000	