

Chapitre 1

La vie d'un site Web

Les fichiers pages

Nous voyons dans ce chapitre toutes les étapes que suit un site Web au cours de son existence. D'abord, il faut concevoir et planifier le site et répartir l'information entre les différentes pages qui le constituent. Ensuite, chaque fichier page doit être créé ; nous verrons quels outils logiciels utiliser pour cela. Il faut alors tester le fonctionnement correct de ces pages.

On doit ensuite rendre le site accessible sur le Net, le « mettre en ligne », comme on dit. Pour cela, il faut un hébergeur. Après, il faut attirer des visiteurs donc faire connaître le site et le rendre visible des principaux moteurs de recherche, c'est le « référencement ».

Maintenant, tout au long de l'existence du site, il faut modifier et mettre à jour les pages qui le nécessitent ; c'est la tâche essentielle de maintenance.

Du côté de l'internaute, les choses sont assez simples. Le visiteur n'a besoin que d'un ordinateur, d'un contrat avec un fournisseur d'accès Internet et d'un logiciel navigateur.

Du côté de celui qui procure le site, il y a plus d'intervenants, une même personne ou organisation pouvant dans certains cas tenir plusieurs rôles.

Il y a d'abord le propriétaire du site, le plus souvent une entreprise ou une association. Il joue le rôle juridique important de responsable du contenu du site. Si l'on voulait comparer avec un journal, on dirait que le propriétaire est le directeur de la publication.

Continuons notre comparaison. Il y a ensuite les auteurs qui créent les pages : auteurs des textes, metteurs en page, graphistes qui définissent le style du site, programmeurs qui gèrent les coulisses du site, etc. Ils sont sous la direction du Webmaster qui joue donc le rôle de rédacteur en chef. Ce livre se concentrera sur les aspects techniques du travail du Webmaster.

Il y a en plus un intermédiaire indispensable, l'hébergeur. C'est une entreprise qui possède de très puissants ordinateurs fonctionnant comme serveurs reliés au réseau mondial. Elle met à disposition de ses clients propriétaires de sites un quota de place sur disque renfermant les informations constituant chaque site.

L'hébergement est normalement payant, en fonction de la place occupée et des services annexes demandés. Il y a aussi, pour les amateurs qui veulent avoir un site personnel, des hébergeurs gratuits à condition d'accepter quelques publicités intercalées dans le site. En France, les hébergeurs supportent une part de la responsabilité juridique du site. Quelquefois, certains propriétaires sont leur propre hébergeur : c'est le cas de certaines grosses entreprises ou certaines institutions officielles (universités, etc.).

Le point fondamental est qu'**une page Web est un fichier texte**. Lorsque, au cours de la visite d'un site, on a besoin d'une certaine page, les choses se déroulent en deux temps :

- Le navigateur émet sur le réseau une requête de la page voulue. Le serveur lui envoie le fichier correspondant et le navigateur le stocke dans une zone du disque qu'on appelle le cache Internet.
- Ensuite, le navigateur interprète ce fichier pour visualiser la page sur l'écran de l'internaute.

Cette division a plusieurs avantages :

- le navigateur interprète le fichier page une fois qu'il est dans le cache sur l'ordinateur du visiteur, donc il n'est pas gêné par les éventuels ratés de transmission ; s'il s'en produit, ils sont réglés pendant la 1^{re} étape ;
- si on redemande une page qu'on a visualisée récemment, il y a des chances qu'elle soit encore dans le cache, donc on gagne le temps de sa transmission ;
- pour le navigateur, il n'y a pas de différence entre interpréter un fichier lu dans le cache après son téléchargement depuis le net ou un fichier pris n'importe où sur le disque ; cela offre la meilleure manière possible de vérifier les pages avant leur mise en ligne ; voici les étapes suivies par un fichier page :

1. écriture du fichier page ;
2. essai du fichier en local dans un ou plusieurs navigateurs ; pour cela, sous Windows (le procédé est analogue sur les autres plate-formes), il suffit de faire glisser à la souris l'icône du fichier vers la fenêtre du navigateur ;
3. une fois que le fichier donne satisfaction, transmission du fichier vers le serveur (chez l'hébergeur).

Désignation des fichiers

Si votre site a une certaine extension, vous allez organiser les différents fichiers qui le constituent en plusieurs répertoires et sous-répertoires. Ainsi un de ces fichiers aura chez vous la désignation : disque:répertoire_principal_de_votre_site\sous-répertoire \nom_fichier.extension (le séparateur « \ » correspond à DOS et Windows ; vous auriez « : » sur Mac et « / » sous Unix/Linux). Chez votre hébergeur, ce sera :

disque:arborescence/.../répertoire_principal_attribué_à_votre_site/sousrépertoire /nom_fichier.extension (le séparateur est / car les hébergeurs sont souvent sous Unix, et les serveurs s'adaptent dans tous les cas).

Le navigateur utilise une autre désignation de votre fichier Web qu'on appelle une URL (Universal Ressource Locator), de la forme : protocole_site/sous-

répertoire/nom.extension où protocole est http:// pour un site Web normal (http=HyperText Transfer Protocol). Le Web utilise d'autres protocoles ou pseudo-protocoles :

https://	http sécurisé
ftp://	transferts de fichiers (File Transfer Protocol)
news:	connexion à des forums
telnet:	votre ordinateur simule un terminal sur une machine éloignée
mailto:	envoi d'un courrier électronique
javascript:	exécution d'une instruction JavaScript
file://	fichier sur le disque local. En fait, il suffit de donner la désignation du fichier, le navigateur comprendra qu'il est local.

La désignation du site implique de déterminer le serveur (l'ordinateur de l'hébergeur), et le site voulu parmi les sites qui sont chez cet hébergeur. Le serveur est théoriquement défini par ce qu'on appelle son adresse IP, groupe de chiffres du genre 81.66.76.178. Il est évidemment plus facile pour les visiteurs d'utiliser un nom symbolique qui rappelle le contenu du site, par exemple www.laposte.com. Alors, comment trouve-t-on le site ? Il y a sur le réseau des serveurs spéciaux, connus des navigateurs, qui fournissent la correspondance nom du site – localisation. Ces serveurs s'appellent des DNS (Domain Name Server).

Donc, le navigateur ayant besoin d'un fichier, il émet d'abord une requête vers un DNS pour localiser l'hébergeur. Il lui envoie alors sa requête de fichier. Le serveur de l'hébergeur reconnaît le site voulu parmi ses clients et il se positionne sur le répertoire principal attribué à ce site. Ensuite, s'il y a des sous-répertoires, il n'y a plus qu'à descendre l'arborescence et fournir le fichier. Au passage, les \ ou autres séparateurs sont transformés en /.

Types de fichiers et extensions

Le type de fichier le plus souvent demandé est une page Web normale (que nous allons apprendre à construire). Un tel fichier a, au choix, l'extension .htm ou .html. Le Webmaster est entièrement maître de décider de son choix, mais le visiteur doit s'y conformer : s'il demande .html alors que le fichier implanté par le Webmaster a .htm, il aura le message « Impossible d'afficher la page ». Le Webmaster doit lui aussi se conformer à ce choix lorsqu'il implante un lien, y compris pour les liens aux autres pages de son propre site.

Conseil : choisissez une fois pour toutes .htm : cela sera cohérent et, ni vous ni vos visiteurs n'auront d'hésitation au moment d'appeler une page.

Des pages Web peuvent se référer à d'autres types de fichiers, d'extensions :

.gif, .jpg, .png : fichiers images ; normalement, ces types sont les seuls fichiers images visualisables sans logiciels de compléments par les navigateurs courants.

.css : fichiers de styles pour la mise en forme.

.js : fichiers de programmes JavaScript externes.

.pdf : fichiers documents pdf . Il faut qu'un logiciel de lecture de pdf soit installé sur l'ordinateur du visiteur. Même problème pour d'autres documents comme .doc(x)

(Word) ; si un logiciel convenable n'est pas installé, la plupart des navigateurs proposent de télécharger le document sans le visualiser.

.asp, .php et autres : un lien peut demander l'exécution d'un programme soit en local, soit sur le serveur ; le fonctionnement est tributaire de l'existence du logiciel voulu chez le visiteur ou l'hébergeur.

Majuscules et minuscules

Beaucoup de serveurs fonctionnent sous Unix / Linux. Dans ces systèmes, les majuscules et minuscules ne sont pas équivalentes dans une désignation de fichier : si vous demandez page.htm alors que le fichier est Page.htm, la recherche ne réussira pas et vous aurez « Impossible d'afficher la page ».

Conseil : n'utilisez que des minuscules dans les noms de répertoires, les noms de fichiers, les extensions. C'est pour la même raison de cohérence que précédemment et vous vous éviterez ainsi bien des hésitations.

Page d'accueil du site

En fait, on peut appeler un site sans préciser de nom de fichier ; cela revient à demander la page d'accueil du site grâce au mécanisme suivant : ne pas préciser de nom de fichier revient à appeler le répertoire principal du site et si, sur ce répertoire, il y a un fichier de nom index.htm, index.html, default.htm, etc. (la liste des fichiers possibles forme une option du logiciel de gestion du serveur et elle est spécifiée par l'administrateur de l'hébergeur), alors c'est ce fichier qui est fourni. Moralité, fournissez votre page d'accueil sous le nom index.htm.

Attention :

Si aucun fichier du genre n'est fourni, alors le navigateur affiche la liste des fichiers présents, ce qui est très gênant, car cela facilite la vie des pirates. Une option peut l'empêcher, mais il n'est pas sûr que l'hébergeur l'active. Par conséquent, il est, à notre avis, impératif, que vous fournissiez un fichier index.htm (éventuellement page vide) dans tous les répertoires et sous-répertoires de votre site.

Les langages du Web

Les pages Web sont des fichiers texte, mais il faut préciser en quel langage ces textes sont rédigés. À part les fichiers annexes images, sons, vidéos, documents accessoires, les pages Web font intervenir quatre langages : HTML et ses variantes, CSS, JavaScript et PHP.

JavaScript et **PHP** sont des langages de programmation qui ne seront pas du tout traités dans ce livre. JavaScript est interprété par le navigateur et ses programmes s'exécutent sur l'ordinateur de l'internaute. Il permet d'implanter des animations et de dynamiser les pages. Au contraire PHP est exécuté sur le serveur ; il permet d'implanter des traitements d'échanges de données avec le serveur, notamment de traiter les réponses à des questionnaires, des inscriptions et des achats en ligne.

HTML est le langage principal des pages. Il a eu plusieurs versions successives, suivant les recommandations du W3C (World Wide Web Consortium). La version que nous étudions dans ce livre est la dernière, **HTML5** ; sa définition est presque totalement achevée. Pourquoi y a-t-il eu tant de versions successives ? Cela vient des problèmes d'incompatibilités ou au moins des différences d'interprétation des différents navigateurs qui suivaient ou non les recommandations du W3C.

Ces problèmes sont maintenant dépassés. Les éditeurs des logiciels navigateurs ont maintenant compris qu'il faut se plier aux normes et les dernières versions des navigateurs (notamment Internet Explorer >=9) sont conformes à HTML5. Nous signalerons les quelques incompatibilités résiduelles, mais les Webmasters sont maintenant débarrassés de ce gros souci : avec le renouvellement des matériels, il ne reste plus de navigateurs anciens que marginalement.

XHTML est une variante qui a été proposée pour introduire des règles d'écriture plus rigoureuses que le laxisme qui règne avec HTML. Mais, de fait, les navigateurs tolèrent tout de sorte que l'on ne parle pas de XHTML5 ; nous indiquerons les règles rigoureuses et nous recommandons de s'y conformer pratiquement, en somme nous recommanderons de faire du XHTML5 tout en disant que c'est du HTML5.

Les instructions de HTML s'appellent des balises. Pour implanter un élément, on écrit une ou plusieurs balises.

CSS est le langage de mise en forme des pages. Le fait que nous ayons deux langages a pour but de totalement séparer ce qui est l'implantation d'éléments dans la page (textes, images, tableaux, etc.) de ce qui est la présentation ou la mise en forme de ces éléments. Les premières versions de HTML mélangeaient les instructions d'installation d'éléments et les instructions de mise en forme. La norme actuelle impose qu'il n'y ait dans le fichier HTML5 que du contenu sémantique, c'est-à-dire qu'il n'implante que des éléments qui apportent une information significative, tandis que les spécifications CSS mettront en forme ces éléments. De la sorte, même en l'absence des éléments de mise en forme, le visiteur aura toute l'information que le site est censé procurer (avec une présentation peu esthétique peut-être).

Il y a trois versions successives de CSS ; nous traitons CSS3, la plus perfectionnée, et nous signalerons les dispositions qui risquent de ne pas être obéies par les navigateurs actuels.

Outils logiciels

Nous l'avons déjà dit, les principaux fichiers pages Web (les .htm) sont de simples fichiers textes. Il en est de même des fichiers annexes .css (les feuilles de styles qui gèrent la mise en forme), .js (programmes JavaScript exécutés chez l'internaute) et .php (programmes PHP exécutés sur le serveur).

Donc le seul outil logiciel nécessaire pour créer des fichiers de ces quatre sortes est un simple éditeur de texte.

Un logiciel de traitement de texte comme Word ne convient pas ou alors il faut sauvegarder les fichiers avec l'option « texte simple » (le terme exact dépend de la version). En outre, il ne faut absolument pas utiliser l'option « sauvegarder en tant que page Web » avec l'extension .htm : elle produira un fichier dix fois trop grand et qui ne fonctionnera pas comme prévu.

Le bloc-notes fourni avec Windows (ou les équivalents Linux ou Mac) peut convenir, mais il est un peu rudimentaire. **Wordpad** exige lui aussi l'option texte simple. Une bonne solution est d'utiliser un éditeur de texte téléchargeable gratuitement comme **Notepad2** ou **Notepad++**. Pour le télécharger, cherchez le nom dans un moteur de recherche comme Google. Il est plus perfectionné que le bloc-notes standard : il met les mots-clés en évidence en s'adaptant à différents langages de programmation et il permet d'agir sur la taille d'affichage et d'autres mises en forme (du fichier en cours d'écriture, pas de la présentation de la page Web telle qu'elle sera visualisée).

Sur Mac, on peut conseiller **BEdit**. Dans tous les cas, pour télécharger n'importe lequel de ces logiciels, cherchez sur votre moteur de recherche préféré avec comme mot-clé le nom ou le genre du logiciel et le mot téléchargement.

Éditeurs Web

Une autre sorte de logiciels disponibles est ce qu'on appelle les éditeurs Web ; ils sont censés permettre de construire des pages Web sans connaître les langages HTML, CSS, etc. Nous ne les étudions pas dans ce livre dont le but est de vous apprendre HTML5 et CSS3. Toutefois, nous citerons FrontPage et Dreamweaver (payant ; c'est un système très complet). Plus simples et téléchargeables gratuitement, il y a FrontPage Express, NVU, Netlor Studio et WYSIWYG Web Builder.

Autres logiciels

Un logiciel qui nous rendra service dans ce livre est **Web-Design-Toy** ; c'est un outil d'apprentissage qui permet de tester des balises HTML isolées pour voir leur effet. L'écran est en deux parties : à gauche, vous tapez le texte HTML à essayer et vous voyez le résultat à droite.

Parmi les autres logiciels téléchargeables gratuitement, nous serons amenés à utiliser **MS Gif Animator** (ou un logiciel analogue) qui permet de créer des images

animées. En outre, il y a des sites de validation auxquels vous pourrez soumettre vos pages Web ; ils vous indiqueront alors si vos pages sont correctes ou vous montreront vos erreurs.

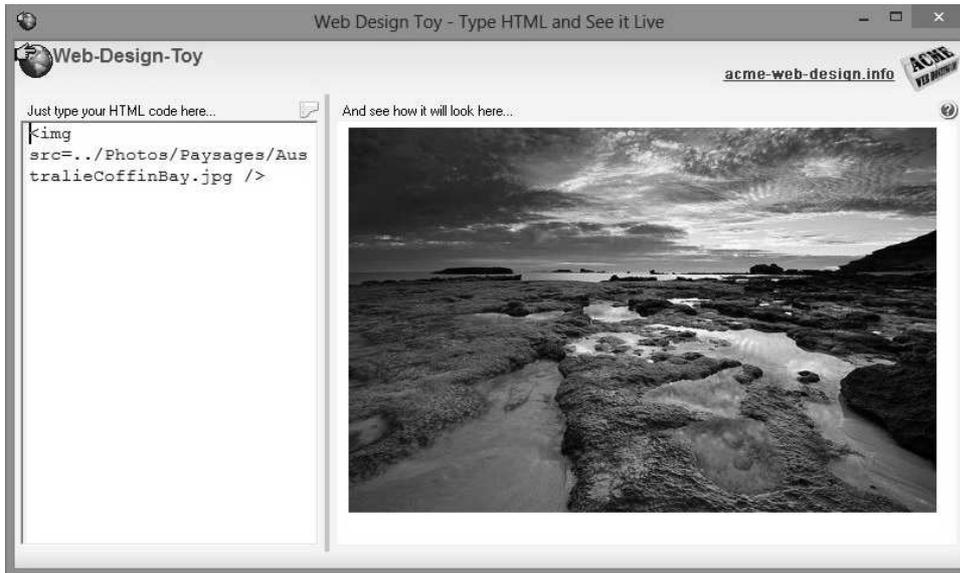


Figure 1-1 : L'écran de Web-Design-Toy

Planification du site

Avant d'écrire la moindre ligne de texte d'une page Web, vous devez planifier votre site. D'abord vous devez décider quelles informations vous voulez fournir à vos visiteurs ; c'est ce qui détermine l'originalité et la spécificité de votre site et en quoi il rend service au monde.

Ensuite, vous devez répartir ces informations entre différentes pages Web ; en effet, il est presque toujours impossible de faire tenir tout un site en une seule page Web, ou alors, il est très rudimentaire ; il est beaucoup plus réaliste de répartir les informations en plusieurs pages, chacune ayant un sujet bien délimité et cohérent. Vous devez définir ces différents sujets en veillant à cette cohérence, puis vous devez réfléchir à quels éléments répondent le mieux au sujet de la page : à part les textes, une image ne serait-elle pas utile ? ou un lien vers un site extérieur qui précise un point ...

Cela étant, vous devez déterminer comment ces différents sujets s'articulent entre eux, donc comment on passe d'une page à l'autre. Ceci met en évidence la question de la navigation dans le site et une conception pratique de cette navigation est un facteur important de l'agrément de visite d'un site. Insistons sur le fait qu'un impératif constant dans la conception d'un site Web est l'ergonomie et l'agrément. Si votre

site n'est pas facile et agréable à visiter, vous n'aurez pas de visiteurs ; en tous cas, ils ne reviendront pas.

La navigation

La navigation, c'est le moyen de passer d'une page à l'autre du site. À tout moment, le visiteur doit savoir comment il est arrivé à telle ou telle page et comment faire pour arriver à telle ou telle autre page. La navigation va dépendre principalement de la structure du site, laquelle doit refléter la structure logique des informations que vous souhaitez fournir.

Pour la structure la plus simple, la page d'accueil présente les liens vers les différentes rubriques.

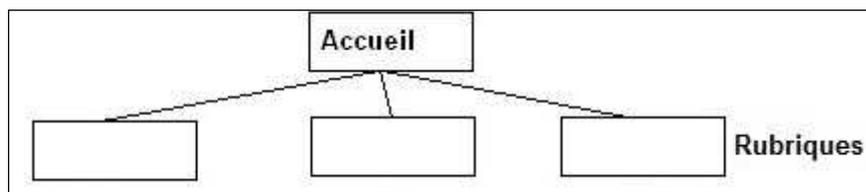


Figure 1-2 : Structure simple

Structure dite « orthogonale »

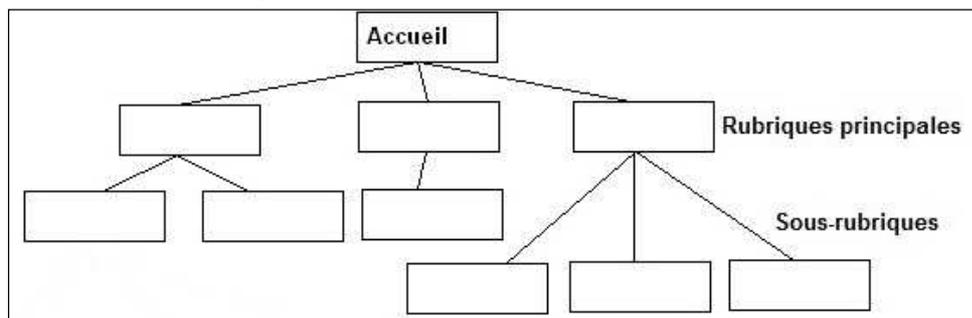


Figure 1-3 : Structure orthogonale

Dans cette structure un peu plus complexe, où les rubriques peuvent avoir des sous-rubriques, la page d'accueil présentera les liens vers les rubriques principales, tandis que chaque page de rubrique présentera les liens vers ses sous-rubriques.

Cette structure est dite orthogonale car, très souvent, les liens vers les rubriques principales sont disposés en ligne en haut de la page d'accueil alors que les liens vers les sous-rubriques sont présentés en verticale à gauche de la page.