

## Table des matières

<u>HtmlDoc</u> .....	1
<u>Installation</u> .....	1
<u>Utilisation de HTMLDOC</u> .....	1
<u>Configuration des entrées "Input"</u> .....	1
<u>Les paramètres de sortie "Output"</u> .....	1
<u>L'onglet Page</u> .....	2
<u>L'onglet TOC</u> .....	2
<u>Onglet Color</u> .....	2
<u>Onglet Font :</u> .....	2
<u>Onglet PS :</u> .....	2
<u>Onglet PDF :</u> .....	2
<u>Onglet Encrytion</u> .....	3
<u>Onglet Option</u> .....	3
<u>Premier essai</u> .....	3
<u>Première étape : paramétrage des entrées</u> .....	3
<u>Résolution des problèmes souvent rencontrés</u> .....	5
<u>Mise en page</u> .....	5
<u>Retraits</u> .....	5
<u>Texte autour d'une image</u> .....	6
<u>Conclusion</u> .....	8

# Htmldoc

**Htmldoc** est, au départ, un programme développé par une société commerciale pour ses besoins internes. Elle la finalement diffusé sous Licence GNU/GPL afin que chacun puisse profiter librement de cet outil.

Le format html, dérivé du SGML, permet de structurer un ensemble d'informations mais ne peut prendre en compte des paramètres de mise en page. Les seuls formats (libres) qui puissent gérer un document structuré et avec mise en page sont le Postscript et le PDF. La fonction de Htmldoc est donc de pouvoir générer des documentations parfaitement imprimables sans perdre la structure du document. D'un côté, Htmldoc utilise un fichier source en HTML et il va en sortir un document Postscript ou PDF. Ce programme reconnaît la majorité des balises HTML 3.2 et peut générer des fichiers PDF au format 1.2 et 1.3, encryptés ou non.

## Installation

A moins de vouloir compiler HTMLdoc, dans ce cas il vous faudra la bibliothèque de widgets FLTK, l'installation est relativement simple. Htmldoc est aussi disponible pour Windows. Après avoir téléchargé et décompressé l'archive sur le site d'ESP (<http://www.easysw.com/htmldoc>), vous pouvez lancer l'installation en lançant le script d'installation **htmldoc.install** en tant que root. Htmldoc est opérationnel. Pour lancer le programme il suffit de saisir dans une console : **htmldoc**.

## Utilisation de HTMLDOC

### Configuration des entrées "Input"

Le premier Onglet concerne la *configuration des entrées* (Input). Le *Type de document* vous permet de préciser que vous souhaitez simplement transformer une page web en fichier PDF (type **web page**) et l'autre option "**book**" vous permet de suivre la structure du document HTML afin de pouvoir générer automatiquement un index de la documentation. Dans ce mode vous devez faire attention à bien respecter la structure du document sinon Htmldoc peut avoir une erreur fatale.

Le bouton "**add files**" vous permet d'insérer les fichiers sources en HTML, vous pouvez en spécifier plusieurs, ils seront lus dans l'ordre indiqué (c'est pour cela que vous avez des boutons "**Move up et Move down**" au cas où vous auriez besoin de modifier l'ordre des fichiers.

Le champ "**logo image**" vous permet de spécifier une image d'en-tête pour votre documentation, elle sera répétée sur chaque page, nous pourrions paramétrer par la suite l'emplacement de ce logo sur chaque page.

Le champ "**Title File/image**" vous permet de préciser un logo pour la page d'accueil mais cela peut être aussi un document précis au format HTML qui contiendra votre page de garde.

Vous pouvez enregistrer votre travail grâce au bouton "**save as**", htmldoc créera un fichier **.book** dans le répertoire indiqué.

Le bouton "**Generate**" permet de lancer la création du document, mais il faudra d'abord spécifier les *paramètres de sortie* avant de lancer cette commande.

### Les paramètres de sortie "Output"

Les différents formats de sortie :

Dans cet onglet vous pouvez choisir entre obtenir une sortie au format **Postscript** (PS) , **Portable Document Format** (PDF) ou **HTML**. Dans le champ "**output path**", vous pouvez préciser ici un nom de fichier.

Des cases à cocher vous permettent de spécifier si vous souhaitez avoir une page de garde (**Title page**); et si vous souhaitez utiliser la compression jpeg pour le format **PDF**, les tirettes vous autorisent la modification du taux de compression.

## L'onglet Page

Dans cet onglet vous allez pouvoir indiquer le format de page pour la sortie **PDF** ou **PS** (universal, A4, letter). La case "**2-sided**" permet de générer des documents recto verso et la case "**Landscape**" permet d'orienter le document en mode paysage.

Les quatre cases correspondent aux marges du document.

La section "**Header et Footer**" (En-tête et pied de page) se subdivisent en 3 parties, qui correspondent au côté gauche, milieu et côté droit du document. Vous avez plusieurs possibilités pour le remplissage de ces champs :

- ◆ **Blank** : vide, rien à insérer
- ◆ **Title** : insertion de l'image ou du document **Title page** de l'onglet **Input**.
- ◆ **Chapter Title** : le titre de la section sera insérer (disponible en mode Book)
- ◆ **Heading** : Titre du document
- ◆ **Logo** : insertion du logo définit dans l'onglet Input
- ◆ **1,2,3, ii, III, abc, ABC** : Numérotation classique
- ◆ **1N 1C** : numérotation relative (ex: page 1/4)
- ◆ **Date, time** : insertion date et/ou heure

## L'onglet TOC

Cet onglet n'est disponible que si vous êtes dans le mode **Book** (Onglet Input). Vous pouvez spécifier ici l'apparence de votre page d'index, qui affiche la structure complète du document.

Vous pouvez préciser le nombre de niveaux, et comme pour l'onglet précédent, les en-têtes.

## Onglet Color

Permet de paramétrer certains attributs de page :

- ◆ **Body color** : pour modifier la couleur des pages du document (par défaut blanc)
- ◆ **Body image** : permet d'insérer un filigrane dans le document
- ◆ **Text Color** : attention cela affecte la couleur de la police sauf si une autre est spécifiée dans le document source. (par défaut, respecte la couleur du source)
- ◆ **Link color** : le format PDF gère les liens, vous en préciser la couleur (par défaut bleu)
- ◆ **Link style** : vous pouvez avoir des liens soulignés ou non.

## Onglet Font :

Dans cet onglet vous pouvez préciser la taille des polices, l'espacement entre les lignes, le type de Police (Times, Helvetica ou courier) pour le corps, les en-têtes et pied de page et l'encodage (pour nous Français, c'est ISO 8859-1)

## Onglet PS :

Si vous avez opté pour une sortie en **Postscript**, vous pouvez spécifier si vous souhaitez un postscript de niveau 1, 2 ou 3.

## Onglet PDF :

Cet onglet permet de paramétrer certaines options disponibles pour le format **PDF**.

- ◆ **Version du PDF** : 1.1, 1.2 et 1.3 sont disponibles
- ◆ **Page mode** : permet de préciser le type de document à générer : le mode Outline donnera un document avec la table des matières visible, le mode document, affichera seulement les pages du document, le mode fullscreen est utilisé pour faire des présentations.
- ◆ **Page Layout** : vous permet de spécifier comment les pages seront visualisées dans Acrobat. Une page ou plusieurs colonnes.
- ◆ **First page** : précise la page qui doit d'abord être affichée, table des matières ou page1 ou chapitre

1.

- ◆ **Page effect** : acrobat peut générer des animations de transitions à la lecture, c'est ici que vous pouvez indiquer le type d'effet à générer. Si vous spécifiez un effet, vous pouvez alors en préciser la durée grâce aux deux tirettes.
- ◆ **Links** : les liens seront respectés si vous cochez cette option.

## Onglet Encrytion

Vous pouvez protéger vos oeuvres de la copie, des modifications, de l'impression grâce à cet onglet. Il suffit de préciser que vous souhaitez encrypter votre document et de lui indiquer les modes. N'oubliez pas de donner un mot de passe.

## Onglet Option

**Html editor** : Vous pouvez préciser un éditeur html externe , celui-ci sera lancé quand vous ferez un double clic sur le nom du fichier HTML dans l'onglet **Input**.

**Browser width** : paramètre important car cela va permettre à htmldoc de simuler une largeur de page.

**Search path** : dans certains cas, les images sont insérées avec des liens relatifs dans le code HTML, vous pouvez préciser ici un lien absolu à suivre.

## Premier essai

Pour faire un premier test, je vais utiliser une documentation que j'ai faite sur la réalisation de pages HTML. Cette documentation possède la structure suivante :

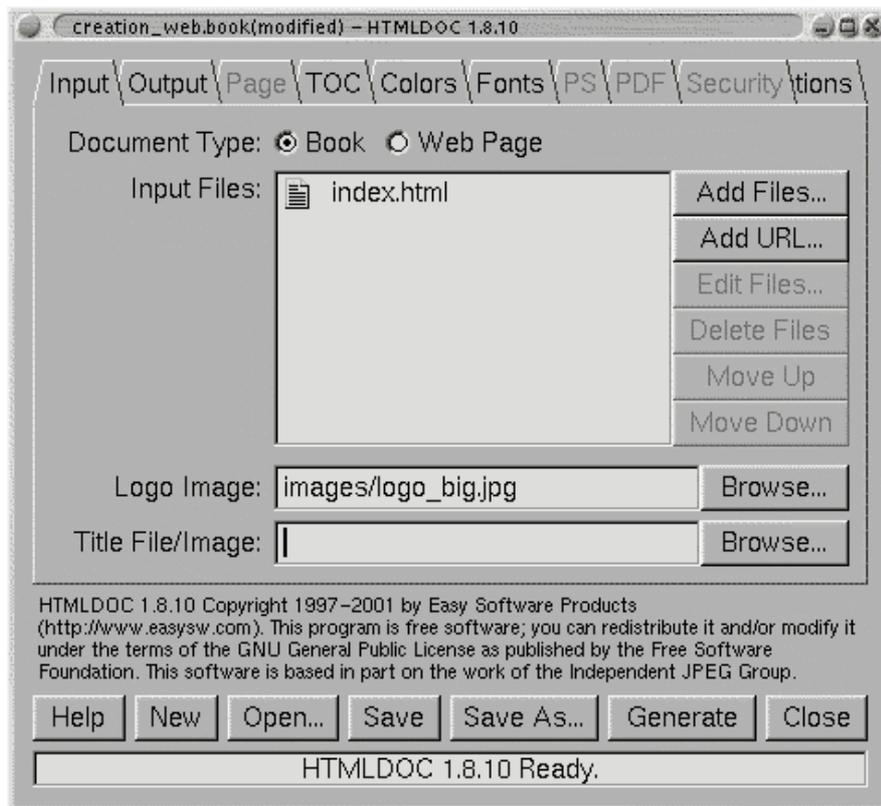
### Plan général

- ◇ Introduction
- ◇ De la création à la diffusion vers plusieurs formats.
- ◇ Création d'une arborescence
- ◇ Définir les paramètres de base d'une page
- ◇ Saisie du texte du document
- ◇ La structure du document
  - . Les en-têtes
  - . Les paragraphes et alignement du texte
  - . Les marges et retraits
  - . Les listes
  - . Les tableaux
  - . Les polices
  - . Les liens
  - . Les ancres
  - . Les lignes de séparation
- ◇ Les formats d'images
  - . Le Format JPEG
  - . Le format Gif
  - . Le format PNG
- ◇ Récupération, découpage et intégration d'une image
  - . Problèmes de copyright et sources d'images
  - . Découpage et redimensionnement
  - . Effet d'ombre portée
  - . Intégration dans la page

Elle a été rédigée avec le composer de Netscape, qui génère un code HTML3.2 suffisant pour ne pas perturber Htmldoc. Cette documentation est disponible au format html [ici](#). Pour bien comprendre la notion de structure de document je vous conseille de lire le didacticiel à ce sujet : Création de documents HTML.

## Première étape : paramétrage des entrées

Nous allons d'abord lui indiquer le fichier source en HTML, ici c'est index.html, je vais aussi lui indiquer une image que je vais insérer dans chaque page de ma documentation comme en-tête.



Dans l'onglet **Output**, nous lui précisons que nous souhaitons une sortie au format PDF, que nous voulons utiliser la compression JPG (pour une diffusion sur le web) et que nous ne voulons pas de page de garde. N'oublions pas de lui préciser le nom du fichier de sortie : votre\_nom.pdf

Dans l'onglet **Page**, nous précisons que nous souhaitons avoir un format A4 et nous allons diminuer la marge gauche à 0.5. En en-tête nous indiquons que nous souhaitons avoir notre image "Logo" au centre et que nous souhaitons obtenir une numérotation relative en pied de page.

Dans l'onglet **Index**, nous précisons les mêmes paramètres pour l'en-tête et le pied de page. En titre, nous mettrons table des matières.

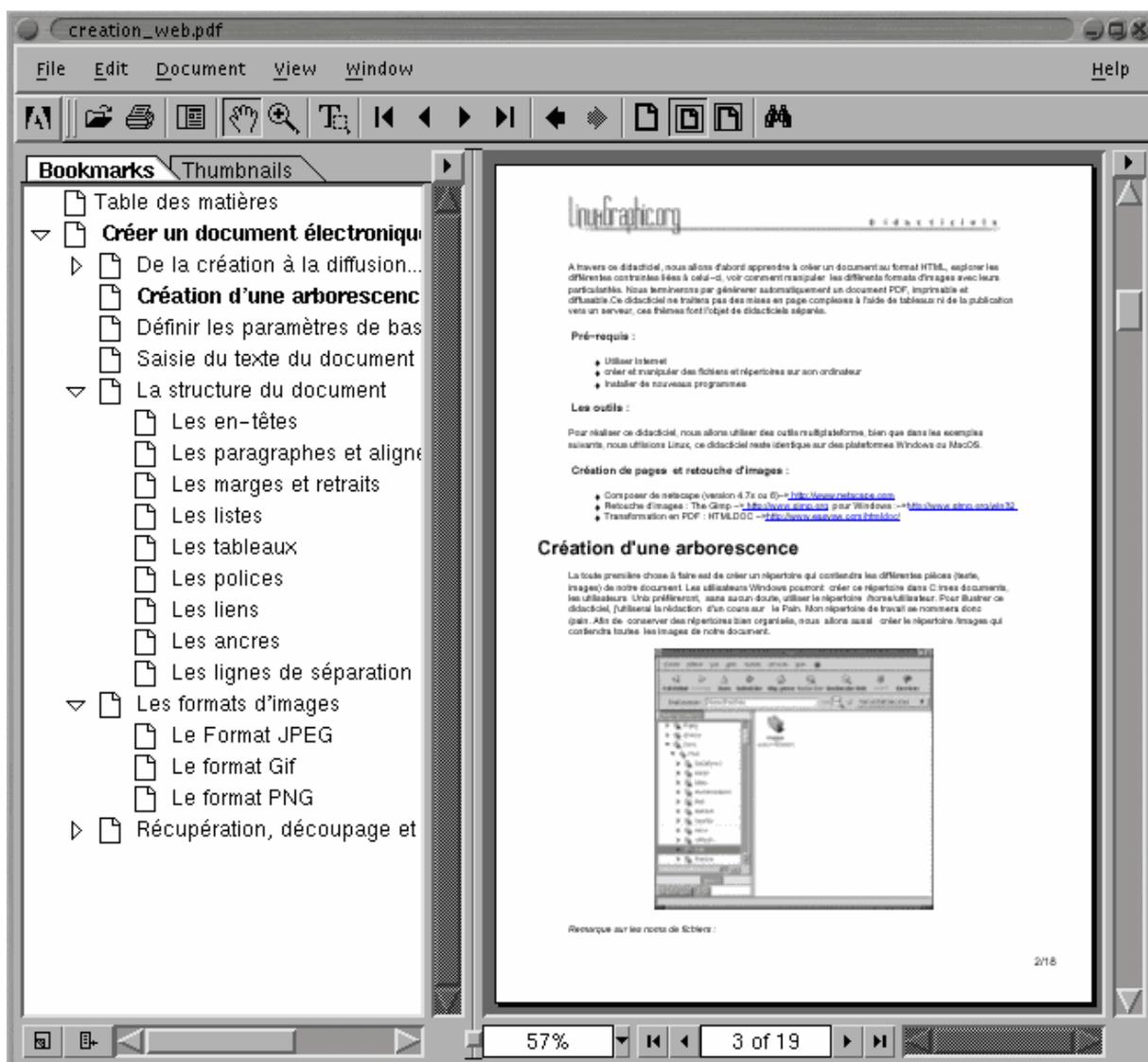
Dans l'onglet **Colors**, nous laisserons les couleurs comme elles sont.

Avec l'onglet **Font**, nous précisons que nous souhaitons Helvetica pour le corps du texte.

Dans l'onglet **PDF**, nous allons laisser les paramètres par défaut

Dans l'onglet **Options**, nous allons mettre une valeur de 900 comme largeur du navigateur car cette documentation contient de grandes images qui occuperaient trop de place. Généralement une valeur de 800 suffit.

Nous lançons l'opération de génération du PDF en cliquant sur le bouton **Generate** et voici ce que nous obtenons avec **Adobe Acrobat** en ouvrant le fichier PDF créé :



Vous pourrez constater que mon document fait 18 pages, que la structure définie par le format HTML est parfaitement conservée et reprise dans la table des matières sur le côté gauche d'Acrobat Reader. Tout pourrait être parfait et s'arrêter ici, mais HTMLDOC a quelques inconvénients sur la mise en page que nous allons résoudre.

## Résolution des problèmes souvent rencontrés

### Mise en page

Le problème de HTMLdoc est qu'il ne respecte pas toujours une mise en page parfaitement optimisée. Par exemple, si il manque un peu d'espace pour mettre une image, il la mettra sur la page suivante en vous laissant un grand blanc. Cela vous oblige à reprendre le fichier source en HTML est à essayer d'enlever des retours de chariots pour faire tenir l'image dans la page. La solution la plus radicale est de réduire la taille de l'image dans le source HTML.

### Retraits

Un autre problème rencontré est une mauvaise gestion des retraits.

ou qui contiennent de nombreuses couleurs, car l'indexation sur 256 couleurs pensera toutement l'image d'origine, de plus la taille de l'image en GIF sera nettement plus élevée que son équivalent au format jpeg. En revanche, pour la création de traits fins, de logos avec peu de couleurs, d'images avec de grands aplats de couleurs, l'utilisation de ce format sera avantageuse. En effet, la méthode de compression n'étant pas destructive, la croix de 2 pixels que le format jpeg a détériorée sera parfaite avec l'utilisation de ce format, et ce avec un poids numérique réduit car cette image peut être indexée sur 2 couleurs (noir et blanc).  
Exemple de carte utilisant le format GIF, les bordures sont nettes et le poids reste raisonnable car elle est indexée sur 150 couleurs



L'image ci-dessus pèse 14 ko, son équivalent en jpeg pèserait environ 19 ko pour une qualité comparable, mais il y aurait quand même des pixels parasites autour du texte.  
Le seul inconvénient du GIF est que ce format utilise une méthode de compression LZW qui est brevetée par la société Unisys et qui est soumise à un paiement de royalties de 5000\$ pour les logiciels qui l'utilisent.

**Le format PNG**

Les limitations et les problèmes de brevets sur le format GIF ont poussé le W3C au développement d'un nouveau standard pour Internet : le PNG (Portable Network Graphic). Les principales caractéristiques de ce format sont :

- ◆ Support de milliers de couleurs
- ◆ Méthode de compression non destructive pour l'image (le niveau de compression est aussi réglable)
- ◆ Support de la transparence Alpha sur 256 niveaux
- ◆ Support des animations dans un format dérivé : le MNG
- ◆ Standard libre, pas de royalties à reverser

Ce format est nettement supérieur au GIF mais il n'est pas encore largement adopté sur Internet car il est relativement récent. Les navigateurs récents comme Netscape, Mozilla, Konqueror le supportent parfaitement (sauf la couche alpha qui n'est bien supportée dans Nets Mozilla) à l'exception du logiciel Internet Explorer (5) de Microsoft qui ne respecte pas la valeur gamma du format PNG.

Les zones entourées indiquent que le retrait n'est pas respecté, ces paragraphes devraient s'aligner sur le trait bleu. Pour corriger cela, il faut ajouter un retrait supplémentaire dans les zones concernées. Cela se produit en général après une image et à la suite d'une liste à puces.

Quand c'est une image, cela se produit si elle est centrée, il suffit de laisser l'alignement à gauche pour ne plus avoir le problème ou bien d'augmenter le retrait.

Dans le cas d'une liste à puces, seule la solution d'augmenter le retrait fonctionne.

Voici mon document après correction :

ou qui contiennent de nombreuses couleurs, car l'indexation sur 256 couleurs pensera toutement l'image d'origine, de plus la taille de l'image en GIF sera nettement plus élevée que son équivalent au format jpeg. En revanche, pour la création de traits fins, de logos avec peu de couleurs, d'images avec de grands aplats de couleurs, l'utilisation de ce format sera avantageuse. En effet, la méthode de compression n'étant pas destructive, la croix de 2 pixels que le format jpeg a détériorée sera parfaite avec l'utilisation de ce format, et ce avec un poids numérique réduit car cette image peut être indexée sur 2 couleurs (noir et blanc).  
Exemple de carte utilisant le format GIF, les bordures sont nettes et le poids reste raisonnable car elle est indexée sur 150 couleurs



L'image ci-dessus pèse 14 ko, son équivalent en jpeg pèserait environ 19 ko pour une qualité comparable, mais il y aurait quand même des pixels parasites autour du texte.  
Le seul inconvénient du GIF est que ce format utilise une méthode de compression LZW qui est brevetée par la société Unisys et qui est soumise à un paiement de royalties de 5000\$ pour les logiciels qui l'utilisent.

**Le format PNG**

Les limitations et les problèmes de brevets sur le format GIF ont poussé le W3C au développement d'un nouveau standard pour Internet : le PNG (Portable Network Graphic). Les principales caractéristiques de ce format sont :

- ◆ Support de milliers de couleurs
- ◆ Méthode de compression non destructive pour l'image (le niveau de compression est aussi réglable)
- ◆ Support de la transparence Alpha sur 256 niveaux
- ◆ Support des animations dans un format dérivé : le MNG
- ◆ Standard libre, pas de royalties à reverser

Ce format est nettement supérieur au GIF mais il n'est pas encore largement adopté sur Internet car il est relativement récent. Les navigateurs récents comme Netscape, Mozilla, Konqueror le supportent parfaitement

**Texte autour d'une image**

Le dernier problème que l'on peut rencontrer est la gestion du texte autour d'une image. Les attributs d'image peuvent parfois suffire, mais ce n'est souvent pas le cas. Un exemple de ce qui peut se produire

Les en-têtes permettent de structurer facilement un document, le code suivant est utilisé :

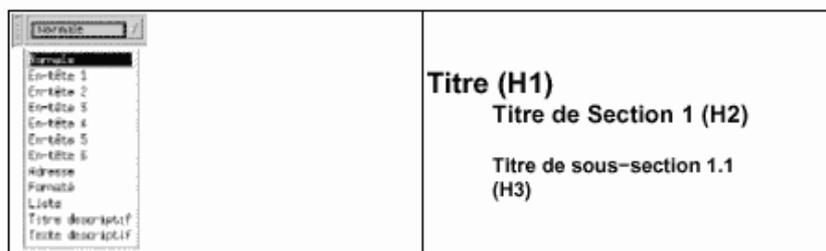


**Titre de Section 1 (H2)**

**Titre de sous-section 1.1 (H3)**

Pour pouvoir affecter cette structure au texte de votre document, il suffit de sélectionner le texte concerné, puis type d'en-tête que vous souhaitez lui appliquer. La taille des caractères s'adapte automatiquement à l'en-tête que vous lui en spécifiez une autre. Cette structure du document est très importante car elle permettra ensuite de

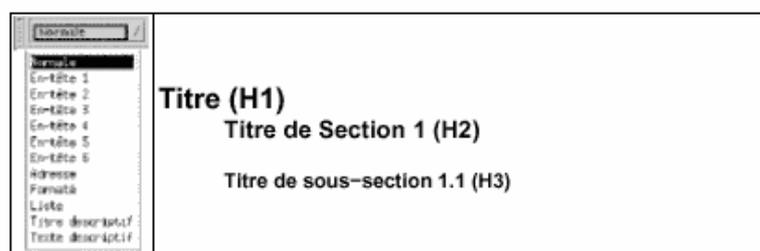
Ici, le texte titre (h1), titre de Section (h2) et titre de sous-section devraient se trouver le long de l'image et non pas au dessous. Pour corriger le problème nous allons utiliser un tableau transparent avec 2 colonnes, une contenant l'image et une autre le texte. Dans l'exemple ci-dessous, j'ai laissé le cadre du tableau pour une meilleure compréhension.



Pour pouvoir affecter cette structure au texte de votre document, il suffit de sélectionner le texte concerné, puis de choisir le type d'en-tête que vous souhaitez lui appliquer. La taille des caractères s'adapte automatiquement à l'en-tête mais vous pouvez lui en spécifier une autre. Cette structure du document est très importante car elle permettra ensuite de l'exporter plus facilement vers d'autres formats. Par défaut, les

Vous constaterez un mieux mais cela ne correspond pas tout à fait à ce que je souhaitais. En effet, le texte n'est pas collé sur l'image.

En fait, je n'ai pas spécifié une largeur spécifique (en pixels) pour ce tableau, ni la largeur des colonnes. On devouvre là un des points forts de HTMLDoc pour la mise en page : une bonne gestion des tableaux. Il suffit de préciser la taille exacte du tableau et de chaque cellule pour obtenir une mise en page parfaite. Je vais donc spécifier une largeur de 700 pixels pour le tableau, 128 pour la cellule contenant l'image et 572 pour l'autre cellule. Notre problème est résolu.



Pour pouvoir affecter cette structure au texte de votre document, il suffit de sélectionner le texte concerné, puis de choisir le type d'en-tête que vous souhaitez lui appliquer. La taille des caractères s'adapte automatiquement à l'en-tête mais vous pouvez lui en spécifier une autre. Cette structure du document est

Vous pouvez faire disparaître le cadre du tableau en lui donnant une valeur de 0 dans votre source HTML. Sachant que HTMLDoc respecte les couleurs des cellules et la taille des tableaux, on peut faire des mises en pages propres et efficaces. Ci-dessous, un exemple de mes cours pour <http://technoresto.org> où j'ai utilisé ces astuces pour obtenir des mises en page correctes.

**TechnoResto.Org**

**L'ELEVAGE DU VIN**

Lors de l'élevage du vin, certaines opérations sont indispensables...

**L'OUILLAGÉ**

Lors de sa conservation en fût, une certaine quantité de produit disparaît :

- Dans le cas de fûts neufs, le bois prend sa part.
- Dans tous les cas il y a évaporation.

Il y a donc un vide qui se crée dans le récipient et le vin est au contact de l'air où il risque la piqûre acétique et la fleur.

L'oûillage est l'opération qui consiste à maintenir les fûts toujours pleins. Cette opération s'effectue fréquemment :

- 1 fois par semaine en hiver
- 2 fois par semaine en été

Pour pratiquer l'oûillage il faut le même vin que celui contenu dans le fût.




**LE COLLAGE**

Le vin naît trouble et il peut être clarifié par filtrage, par centrifugation ou par collage. Cette opération est réservée aux vins de qualité car elle est longue et demande beaucoup de main d'œuvre. Le collage est une opération qui consiste à introduire dans le vin, pendant son séjour en fût, des produits tel que :

- la blanc d'œuf (vin rouge)
- la colle de poisson (vin blancs)
- la gélatine, le sang de bœuf, la bentonite

Ces produits ont la propriété de **floculer** au contact des tanins, ils entraînent ainsi les particules en suspension dans le vin.



**LE SOUTIRAGE**

C'est l'opération qui consiste à séparer le vin de son dépôt en transvasant son contenu vers un autre fût.

Le premier, réalisé rapidement, prend le nom de débourbage. Le second est effectué au printemps et permet d'éliminer une éventuelle précipitation des cristaux de tartrate. Si le vin reste en fût, un troisième soutirage a lieu en septembre.



**TechnoResto.Org**

**LES MALADIES ET ACCIDENTS DU VIN**

Le vin est une substance vivante, donc il peut être atteint de maladies ou accidents qui sont à l'origine de défauts de goût.

Contre les maladies, il existe il existe des remèdes, mais le mieux est de pratiquer des actions préventives comme :

- soigner et manipuler le vin dans de bonnes conditions d'hygiène
- utiliser correctement les outils qui sont bien entretenus.

Cependant, malgré toutes ces précautions il peut arriver que des altérations se déclarent. Il convient de distinguer :



**Les maladies d'origine microbienne :**

Nom	Caractéristiques	origine
<b>FERMENTS ANAÉROBES</b>		
<b>La Fleur</b>	Particules blanches formant un voile sur le vin	Pratique non régulière de l'oûillage qui provoque une oxydation du vin
<b>La piqûre</b>	Goût et odeur de vinaigre	Présence de certains esters
<b>FERMENTS AÉROBES</b>		
<b>L'aver</b>	Amertume du vin (maladie relativement rare)	Caves trop chaudes, collage non respecté. Maladie qui se développe en bouteille
<b>La tourne</b>	Dégagement de CO <sub>2</sub> , modification de la couleur et du goût	Présence d'un acide spécifique dans le vin

**Les maladies d'origine chimique :**

Ce sont, pour la majorité des cas, des casés.

Nom	Caractéristiques	origine
<b>La casse ferrugineuse</b>	Le vin blanc se trouble, teinte plombée, les rouges s'oxidisent et présentent un dépôt noir blanchâtre.	Écaille de fer (prise au contact du matériel de la cave)
<b>La casse cuivrée</b>	Flocus rouges formant un dépôt au fond de la bouteille.	Présence de cuivre
<b>La casse blanche</b>	Dépôt gris blanchâtre dans le vin blanc	Précipité phosphato-ferrugineux
<b>La casse brune</b>	Teinte brune du vin	Raisins atteints par la pourriture noble.

**Les goûts accidentels :**

Le vin est particulièrement sensible aux odeurs :

- **boîte moisi** : fûts mal entretenus, vendanges altérées
- **Bouchon** : mauvais liège, moisi (provient du pénicillium glaucum qui se développe en milieu humide)
- **Mercaptan** : odeur d'œuf pourri provoquée par la présence d'hydrogène sulfuré.

## Conclusion

Ce didacticiel vous a permis de découvrir qu'il était possible de réaliser de belles documentations avec de simples outils. Pas la peine d'aller courir acheter le dernier word 9 pour faire votre rapport de stage. Un simple éditeur de texte et HTMLdoc peuvent suffire, mais c'est quand même plus simple de travailler avec des outils Wysiwyg pour générer les fichiers HTML : Amaya, Netscape Composer, Staroffice offrent déjà l'essentiel pour aucune licence à payer.

Les puristes me reprocheront de ne pas avoir parlé de Lyx et de Latex qui permettent de faire des documentations complètes et structurées. Là aussi, il est possible d'obtenir des documents PDF en passant par le format Postscript. Je rappelle que le but de cette documentation était de proposer une solution multiplateformes, malheureusement Lyx n'est disponible que pour Unix.

Pour des mises en pages plus complexes, qui ressemblent d'avantage à de la PAO, il faut en passer par le module de dessin vectoriel de Staroffice, qui gère le blocs de texte, puis faire une impression en Postscript pour ensuite le transformer en PDF via la commande ps2pdf.

Copyright 2001 [Toussaint Frédéric](#)