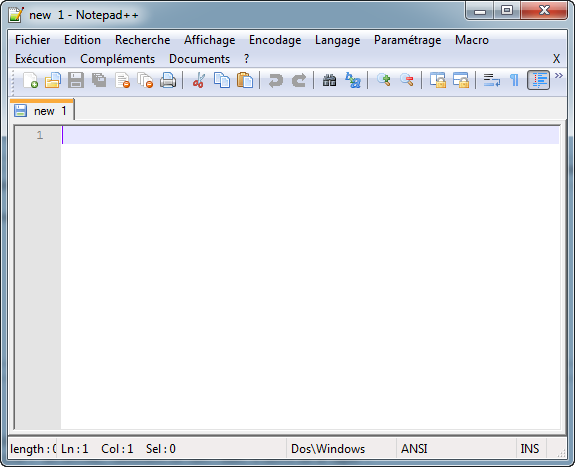
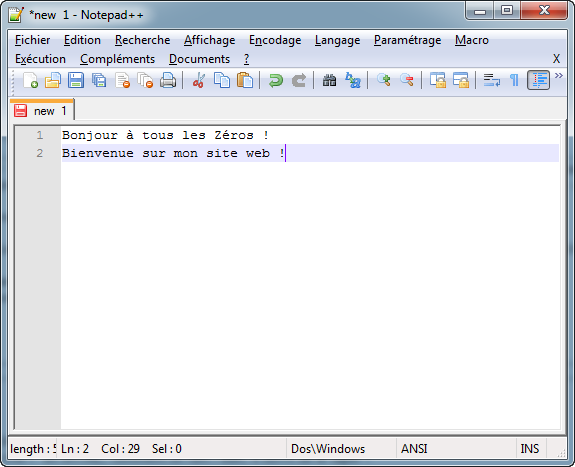
3+9Allez, mettons-nous en situation ! Comme je vous l'ai dit, nous allons créer notre site dans un éditeur de texte. Vous avez dû en installer un suite à mes conseils dans le premier chapitre : qu'il s'appelle Notepad++, PSpad, jEdit, vim, TextWrangler… peu importe. Ces logiciels ont un but très simple : vous permettre d'écrire du texte !

Dans la suite de ce cours, je travaillerai avec Notepad++. Je vais donc l'ouvrir (figure suivante).

Ouverture de Notepad++

Bon, qu'est-ce qu'on fait maintenant ? Qu'est-ce qu'on écrit sur cette feuille blanche ?

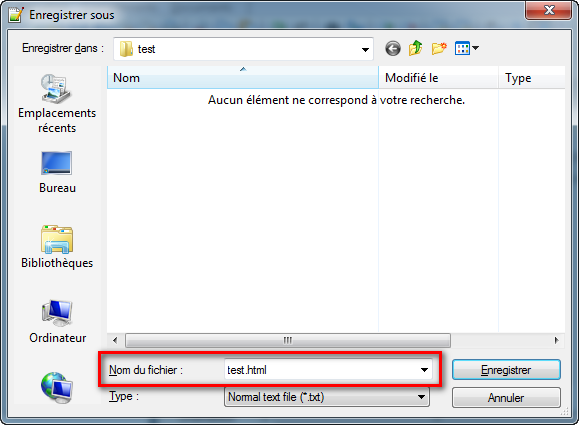
On va faire un petit essai. Je vous invite à écrire ce qui vous passe par la tête, comme moi à la figure suivante.



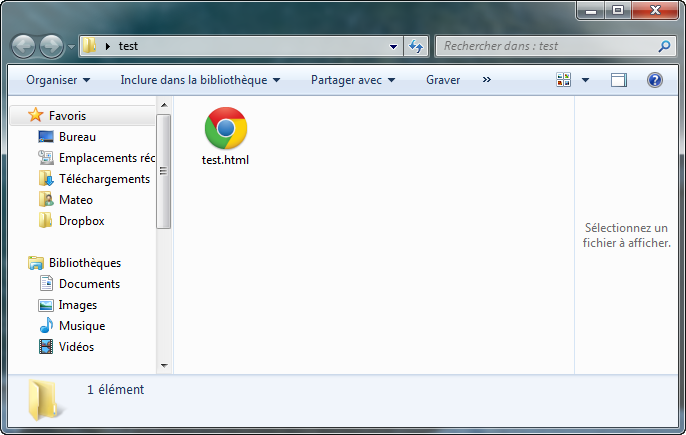
Du texte dans Notepad++

Vous pouvez écrire les mêmes phrases que moi ou ce que vous voulez ; le but est d'écrire quelque chose.

Maintenant, enregistrons ce fichier. Pour cela, c'est très simple : comme dans tous les programmes, vous avez un menu Fichier > Enregistrer. Une boîte de dialogue vous demande où enregistrer le fichier et sous quel nom. Enregistrez-le où vous voulez. Donnez au fichier le nom que vous voulez, en terminant par .html, par exemple test.html, comme indiqué à la figure suivante.

Enregistrement d'un fichier sous Notepad++

Je vous recommande de créer un nouveau dossier dans vos documents qui contiendra les fichiers de votre site. Pour ma part j'ai créé un dossier test dans lequel j'ai mis mon fichier test.html.

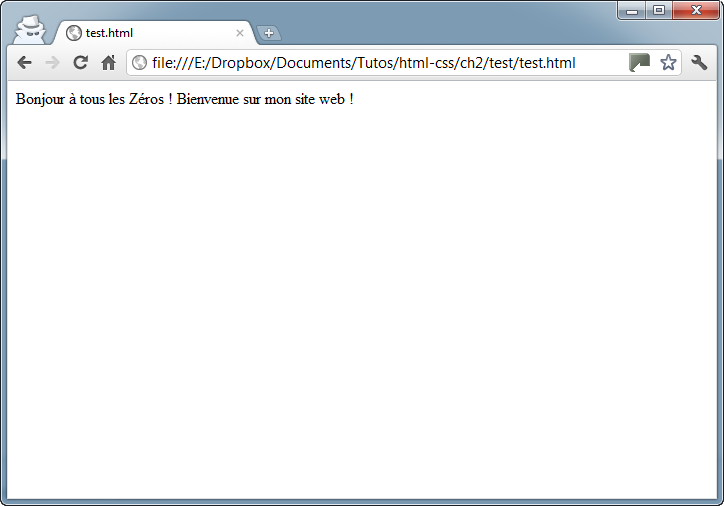
Ouvrez maintenant l'explorateur de fichiers dans le dossier où vous avez enregistré votre page. Vous y verrez le fichier que vous venez de créer (figure suivante).

Le fichier dans l'explorateur

L'icône qui représente le fichier dépend de votre navigateur web par défaut. Ici, l'icône est celle de Google Chrome, mon navigateur par défaut, mais le fichier a peut-être une autre icône chez vous. Voici par exemple les icônes qui apparaissent lorsque votre navigateur principal est Firefox ou Internet Explorer (figure suivante).

Icône fichier FirefoxIcône fichier Internet Explorer

Faites simplement un double-clic sur ce fichier et… votre navigateur s'ouvre et, comme à la figure suivante, affiche le texte que vous avez écrit.

La page web affichée

Cela ne marche pas bien, on dirait ! Tout le texte s'affiche sur la même ligne alors qu'on avait écrit deux lignes de texte différentes !?

En effet, bien vu !   
Le texte s'affiche sur la même ligne alors qu'on avait demandé à l'écrire sur deux lignes différentes. Que se passe-t-il ?

En fait, pour créer une page web il ne suffit pas de taper simplement du texte comme on vient de le faire. En plus de ce texte, il faut aussi écrire ce qu'on appelle des **balises**, qui vont donner des instructions à l'ordinateur comme « aller à la ligne », « afficher une image », etc.

Bon, tout cela était trop facile. Évidemment, il a fallu que ces satanés informaticiens s'en mêlent et compliquent les choses. Il ne suffit pas d'écrire « simplement » du texte dans l'éditeur, il faut aussi donner des instructions à l'ordinateur. En HTML, on utilise pour cela des **balises**.

**Les balises**

Les pages HTML sont remplies de ce qu'on appelle des balises. Celles-ci sont invisibles à l'écran pour vos visiteurs, mais elles permettent à l'ordinateur de comprendre ce qu'il doit afficher.  
Les balises se repèrent facilement. Elles sont entourées de « chevrons », c'est-à-dire des symboles < et >, comme ceci : <balise>

À quoi est-ce qu'elles servent ? Elles indiquent la nature du texte qu'elles encadrent. Elles veulent dire par exemple : « Ceci est le titre de la page », « Ceci est une image », « Ceci est un paragraphe de texte », etc.

On distingue deux types de balises : les balises en paires et les balises orphelines.

**Les balises en paires**

Elles s'ouvrent, contiennent du texte, et se ferment plus loin. Voici à quoi elles ressemblent :

<titre>Ceci est un titre</titre>

On distingue une balise ouvrante (<titre>) et une balise fermante (</titre>) qui indique que le titre se termine. Cela signifie pour l'ordinateur que tout ce qui n'est *pas* entre ces deux balises… n'est pas un titre.

Ceci n'est pas un titre <titre>Ceci est un titre</titre> Ceci n'est pas un titre

**Les balises orphelines**

Ce sont des balises qui servent le plus souvent à insérer un élément à un endroit précis (par exemple une image). Il n'est pas nécessaire de délimiter le début et la fin de l'image, on veut juste dire à l'ordinateur « Insère une image ici ».

Une balise orpheline s'écrit comme ceci :

<image />

Notez que le / de fin n'est pas obligatoire. On pourrait écrire seulement <image>. Néanmoins, pour ne pas les confondre avec le premier type de balise, les webmasters recommandent de rajouter ce / (*slash*) à la fin des balises orphelines. Vous me verrez donc mettre un / aux balises orphelines et je vous faire de même, recommande, c'est une bonne pratique.

**Les attributs**

Les attributs sont un peu les options des balises. Ils viennent les compléter pour donner des informations supplémentaires. L'attribut se place après le nom de la balise ouvrante et a le plus souvent une valeur, comme ceci :

<balise attribut="valeur">

À quoi cela sert-il ? Prenons la balise <image /> que nous venons de voir. Seule, elle ne sert pas à grand chose. On pourrait rajouter un attribut qui indique le nom de l'image à afficher :

<image nom="photo.jpg" />

L'ordinateur comprend alors qu'il doit afficher l'image contenue dans le fichier photo.jpg.

Dans le cas d'une balise fonctionnant « par paire », on ne met les attributs que dans la balise ouvrante et pas dans la balise fermante. Par exemple, ce code indique que la citation est de Neil Armstrong et qu'elle date du 21 Juillet 1969 :

<citation auteur="Neil Armstrong" date="21/07/1969">

C'est un petit pas pour l'homme, mais un bond de géant pour l'humanité.

</citation>

Toutes les balises que nous venons de voir sont fictives. Les vraies balises ont des noms en anglais (eh oui !), nous allons les découvrir dans la suite de ce cours.

Reprenons notre éditeur de texte (dans mon cas Notepad++). Je vous invite à écrire ou à copier-coller le code source ci-dessous dans Notepad++. Ce code correspond à la base d'une page web en HTML5 :

<!DOCTYPE html>

<html>

    <head>

        <meta charset="utf-8" />

        <title>Titre</title>

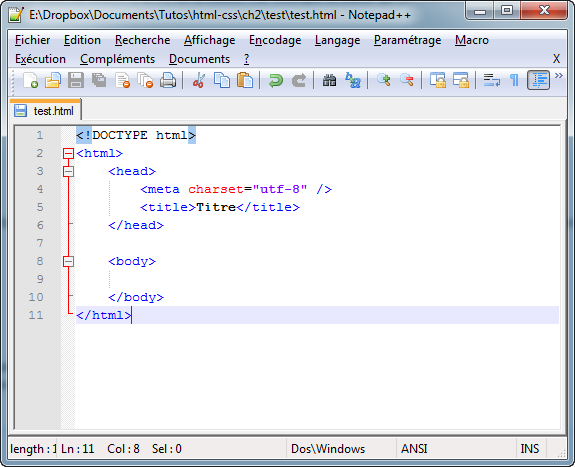
    </head>

    </body>

</html>

J'ai mis des espace<body>s au début de certaines lignes pour « décaler » les balises. Ce n'est pas obligatoire et cela n'a aucun impact sur l'affichage de la page, mais cela rend le code source plus lisible. On appelle cela l'**indentation**. Dans votre éditeur, il suffit d'appuyer sur la touche Tab pour avoir le même résultat.

Copié dans Notepad++, cela donne la figure suivante.

Code HTML5 minimal dans Notepad++

Vous noterez que les balises s'ouvrent et se ferment dans un ordre précis. Par exemple, la balise <html> est la première que l'on ouvre et c'est aussi la dernière que l'on ferme (tout à la fin du code, avec </html>). *Les balises doivent être fermées dans le sens inverse de leur ouverture.* Un exemple :

* <html><body></body></html> : **correct**. Une balise qui est ouverte à l'intérieur d'une autre doit aussi être fermée à l'intérieur.
* <html><body></html></body> : **incorrect**, les balises s'entremêlent.

Euh, on pourrait avoir des explications sur toutes les balises que l'on vient de copier dans l'éditeur, m'sieur ?

Bien sûr, c'est demandé si gentiment.   
Ne prenez pas peur en voyant toutes ces balises d'un coup, je vais vous expliquer leur rôle !

**Le doctype**

<!DOCTYPE html>

La toute première ligne s'appelle le **doctype**. Elle est indispensable car c'est elle qui indique qu'il s'agit bien d'une page web HTML.   
Ce n'est pas vraiment une balise comme les autres (elle commence par un point d'exclamation), vous pouvez considérer que c'est un peu l'exception qui confirme la règle.

Cette ligne du doctype était autrefois incroyablement complexe. Il était impossible de la retenir de tête. Pour XHTML 1.0, il fallait écrire :  
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">.   
Dans le cadre de HTML5, il a été décidé de la simplifier, pour le plus grand bonheur des webmasters. Quand vous voyez une balise doctype courte (<!DOCTYPE html>), cela signifie que la page est écrite en HTML5.

**La balise </html>**

<html>

</html>

C'est la balise principale du code. Elle englobe tout le contenu de votre page. Comme vous pouvez le voir, la balise fermante </html> se trouve tout à la fin du code !

**L'en-tête <head> et le corps <body>**

Une page web est constituée de deux parties :

* L'en-tête <head> : cette section donne quelques informations générales sur la page comme son titre, l'encodage (pour la gestion des caractères spéciaux), etc. Cette section est généralement assez courte. Les informations que contient l'en-tête ne sont pas affichées sur la page, ce sont simplement des informations générales à destination de l'ordinateur. Elles sont cependant très importantes !
* Le corps <body> : c'est là que se trouve la partie principale de la page. Tout ce que nous écrirons ici sera affiché à l'écran. C'est à l'intérieur du corps que nous écrirons la majeure partie de notre code.

Pour le moment, le corps est vide (nous y reviendrons plus loin). Intéressons-nous par contre aux deux balises contenues dans l'en-tête…

**L'encodage (charset)**

<meta charset="utf-8" />

Cette balise indique l'encodage utilisé dans votre fichier .html.

Sans rentrer dans les détails, car cela pourrait vite devenir compliqué, l'encodage indique la façon dont le fichier est enregistré. C'est lui qui détermine comment les caractères spéciaux vont s'afficher (accents, idéogrammes chinois et japonais, caractères arabes, etc.).

Il y a plusieurs techniques d'encodage portant des noms bizarres et utilisées en fonction des langues : ISO-8859-1, OEM 775, Windows-1253… Une seule cependant devrait être utilisée aujourd'hui autant que possible : UTF-8. Cette méthode d'encodage permet d'afficher sans aucun problème pratiquement tous les symboles de toutes les langues de notre planète ! C'est pour cela que j'ai indiqué utf-8 dans cette balise.

Il faut aussi que votre fichier soit bien enregistré en UTF-8. C'est le cas le plus souvent sous Linux par défaut mais, sous Windows, il faut généralement le dire au logiciel.

Sous Notepad++, allez dans le menu Encodage > Encoder en UTF-8 (sans BOM) pour que votre fichier soit enregistré en UTF-8 dès le début. Cela ne s'applique qu'au fichier actuellement ouvert.  
Pour ne pas avoir à le faire pour chaque nouveau fichier, je vous conseille d'aller dans le menu Paramétrage > Préférences, onglet Nouveau document/Dossier. Sélectionnez UTF-8 sans BOM dans la liste.

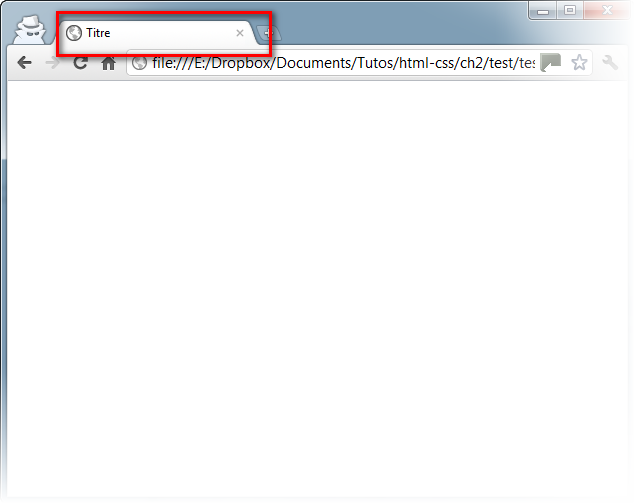
Si vous avez un problème d'affichage des accents plus tard sur votre page web, c'est qu'il y a un problème avec l'encodage. Vérifiez que la balise indique bien UTF-8 et que votre fichier est enregistré en UTF-8 (votre éditeur de texte est capable de vous le dire, Notepad++ le fait dans le menu Encodage).

**Le titre principal de la page**

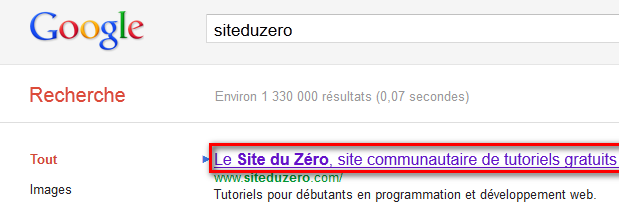
<title>

C'est le titre de votre page, probablement l'élément le plus important ! Toute page doit avoir un titre qui décrit ce qu'elle contient.  
Il est conseillé de garder le titre assez court (moins de 100 caractères en général).

Le titre ne s'affiche pas dans votre page mais en haut de celle-ci (souvent dans l'onglet du navigateur). Enregistrez votre page web et ouvrez-la dans votre navigateur. Vous verrez que le titre s'affiche dans l'onglet, comme sur la figure suivante.

Le titre de la page apparaît en haut du navigateur

Il faut savoir que le titre apparaît aussi dans les résultats de recherche, comme sur Google (figure suivante).

Le titre de la page apparaît dans les recherches Google

Autant vous dire que bien choisir son titre est important !