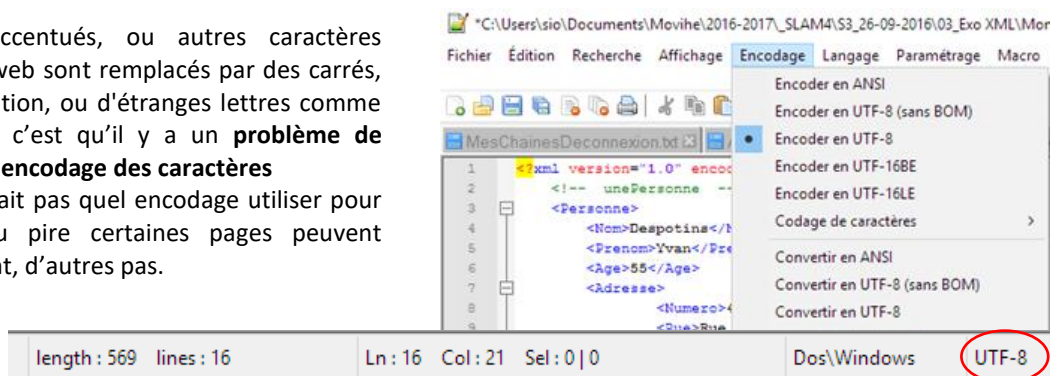


Si les caractères accentués, ou autres caractères spéciaux, de la page web sont remplacés par des carrés, des points d'interrogation, ou d'étranges lettres comme « Ã© » ou « Å§ », c'est qu'il y a un **problème de déclaration du type d'encodage des caractères**

→ Le navigateur ne sait pas quel encodage utiliser pour afficher la page. Au pire certaines pages peuvent s'afficher correctement, d'autres pas.



NotePad++ affiche l'encodage réel des caractères du fichier en cours.

1. L'encodage **réel** d'un document → à sa création

Tout document contenant du texte est enregistré avec un jeu de caractères précis.

Le **jeu de caractères, utilisé pour créer ou enregistrer le document**, correspond à **l'encodage réel du document**.

Les encodages les plus fréquents en français sont :

- l'ISO-8859-1 appelé latin1, qui permet d'enregistrer presque tous les caractères du français
- l'ISO-8859-15 appelé latin9, une variation de l'ISO-8859-1, qui rajoute le symbole « euro » et le « l'o dans l'e »
- l'UTF-8 permet d'encoder toutes les langues, du français au japonais en passant par l'arabe

Vérifier dans les paramètres de configuration de l'éditeur (Notepad++, Dreamweaver, etc.) **l'encodage utilisé par défaut pour être sûr** de l'encodage des caractères d'un fichier à sa création.

Certains éditeurs affichent également l'encodage d'un fichier ouvert et permettent de déclarer l'encodage par défaut des caractères au moment de l'enregistrement du fichier.

2. L'encodage **déclaré**

2.1 On peut ajouter une balise <meta> à chaque page HTML qui déclare ainsi le type d'affichage utilisé. Cette balise est utile si on enregistre les pages web en local, sur son ordinateur. En HTML5 cela donne : **<meta charset="UTF-8">**. Problème : ajouter cette déclarative à chaque page !

```
<!DOCTYPE html>
<html lang="fr">
<head>
  <meta charset="UTF-8">
  <title>Comment bien déclarer
```

2.2 En-têtes http : la déclaration de l'encodage d'une page HTML, peut être effectuée dans les en-têtes HTTP **envoyées par le serveur** avec la page web. Par défaut il est recommandé d'envoyer un en-tête Content-Type, associé à une directive **charset** : **Content-Type: text/html; charset=UTF-8**

Attention : L'information donnée par la balise <meta> d'une page html n'est pas prioritaire. C'est la balise Content-Type fournie par le serveur qui est prioritaire. MAIS, si jamais le serveur n'envoie aucune information sur l'encodage du document, le navigateur web cherchera l'information dans la balise<meta> de la page html. C'est pourquoi on se contente parfois de préciser l'encodage dans la balise <meta>, sans prendre la peine de configurer le serveur. Toutefois, cette pratique n'est pas recommandée ! Il suffit que l'information manque dans une seule page pour qu'alors celle-ci ne s'affiche plus correctement. Il est donc préférable, dans tous les cas de vérifier que le serveur envoie bien une directive charset correctement paramétrée.

3. Apache et la configuration charset

3.1 Comment configurer le serveur Apache pour qu'il envoie le bon charset ?

Le paramétrage et la modification des paramètres du serveur s'effectuent avec des droits d'administration. Modifier le fichier **.htaccess** du serveur Apache et vérifier l'instruction suivante : **Content-Type: text/html; charset=utf-8**

3.2 Comment visualiser les en-têtes http émises par le serveur ?

Le mode « voir la source » du navigateur ne montre pas les en-têtes HTTP, ou le code de statut HTTP. Comment faire pour les consulter ? On peut installer une extension pour Mozilla, « WebDeveloper » pour visualiser le charset paramétré sur le serveur.

<http://carcouet.paysdelaloire.e-lyco.fr/>

```
Date: Tue, 24 Jan 2017 08:50:05 GMT
Server: Apache
Expires: Tue, 24 Jan 2017 08:50:05 GMT
Cache-Control: private, no-cache, no-store, no-transform
Vary: Accept-Encoding
Content-Encoding: gzip
Content-Length: 15659
Content-Type: text/html; charset=UTF-8
```

Pour visualiser le charset paramétré sur le serveur, avec WebDeveloper : Infos / Entêtes http