

Le XHTML, et la validation

Le XHTML

Le **XHTML**, eXtensible HyperText Markup Language, est un langage identique au langage HTML mais respectant la syntaxe définie par **XML**. (*Extensible Markup Language*). Ce qui donne un document plus correct, plus propre et répondant au standard de W3C. Le principe est le suivant:

1. bonne imbrication des balises

Les balises doivent être correctement imbriquées selon leur hiérarchie.

Incorrect :

```
<h1><b>Imbrication incorrecte</h1></b>
```

Correct :

```
<h1><b>Imbrication correcte </b></h1>
```

incorrect

```
<ul>
  <li>café</li>
  <li>Thé
    <ul>
      <li>thé vert</li>
      <li>fhé noir</li>
    </ul>
  <li>lait</li>
</ul>
```

Correct

```
<ul>
  <li>Café</li>
  <li>Thé
    <ul>
      <li>Thé vert</li>
      <li>Thé noir</li>
    </ul>
  </li>
  <li>lait</li>
</ul>
```

2. toutes les balises doivent être correctement fermées

Incorrect

<p> ceci n'est pas correct <p> j'ai oublié de fermer la balise p</p>

Correct

<p> ceci n'est pas correct </p> <p> j'ai oublié de fermer la balise p</p>

3. le nom des attributs et balises en minuscule

Incorrect

<FRAMESET COLS="20%","70%">

Correct

<frameset cols="20%","70%">

4. balises vides avec terminaison

Incorrect

<input type="text" name="nomzone" size=20>
voici un exemple >

Correct

<input type="text" name="nomzone" size=20 />
voici un exemple />

Incorrect

A break:

A horizontal rule: <hr>
An image:

Correct

Un saut de ligne:

une règle horizontale: <hr />
une image:

5. les attributs doivent être entre guillemets

Incorrect

<table width =80%>

Correct

<table width ="80%">

<http://www.w3.org/QA/2002/04/valid-dtd-list.html#DTD>

<http://developpez.com/tutoriels/xhtml/>

6. tous les attributs doivent avoir une valeur explicite

Incorrect

```
<input type="checkbox" checked />
```

Correct

```
<input type="checkbox" checked="checked" />
```

Incorrect

```
<frame noresize>
```

Correct

```
<frame noresize="noresize" />
```

Autres exemples :

HTML	XHTML
checked	checked="checked"
readonly	readonly="readonly"
disabled	disabled="disabled"
selected	selected="selected"
noresize	noresize="noresize"

7. l'attribut id remplace l'attribut name pour les balises : a, frame,

8. tous documents xhtml doit avoir un DOCTYPE.

<http://www.w3.org/QA/2002/04/valid-dtd-list.html#DTD>

<http://developpez.com/tutoriels/xhtml/>

Validation par Le W3C

Le W3C, Worl Wide Web Consortium, est un organisme travaillant sur la mise en place de standards Web. Ce qui nous intéresse plus particulièrement, ici, est leur système de validation de code source.

Par valider le code source, le système du W3C, va vérifier la syntaxe de votre code conformément aux règles établies. Valider sa page permet : ! En voici quelques raisons :

- d'avoir un code portable, compatible avec la plupart des navigateurs.
- d'avoir un code standardisé et connu de tout le monde.
- une meilleure accessibilité.

Pour avoir un site valide W3C, il faut :

- Renseigner le navigateur sur la norme utilisée via le DOCTYPE du document (*ex*: XHTML 1.0 STRICT).
- Renseigner le navigateur sur le jeu de caractères utilisé (*ex*: UTF-8).
- Se conformer à la syntaxe définie par le DOCTYPE utilisé
- Passer un test de validation W3C du document XHTML à l'adresse <http://validator.w3.org/>

Modalité de validation

1- DOCTYPE

Le **DOCTYPE**, Document Type, déclare au navigateur la norme utilisée ou le type de document. Il doit être placé en tout début du code source, avant la balise `<html>`.

HTML 4.01 Strict

Cette DTD contient toutes les balises HTML, à l'exception de certaines balises comme font, center .. Les frames ne sont pas supportés par cette DTD. Les éléments de mises en forme du texte ne sont pas utilisés; on utilise à la place des feuilles de style CSS.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
"http://www.w3.org/TR/html4/strict.dtd">
```

HTML 4.01 Transitional

Cette DTD contient toutes les balises HTML, y compris les balises comme font et les balises de mise en forme. Les frames ne sont pas supportés par cette DTD.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
```

<http://www.w3.org/QA/2002/04/valid-dtd-list.html#DTD>
<http://developpez.com/tutoriels/xhtml/>

4.01 Frameset

Est équivalente à HTML 4.0.1 Transitional mais elle utilise la balise frameset

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Frameset//EN"
"http://www.w3.org/TR/html4/frameset.dtd">
```

XHTML 1.0 Strict

Cette DTD contient toutes les balises HTML, à l'exception de certaines balises comme font. Les frames ne sont pas supportés par cette DTD. Les éléments de mises en forme du texte ne sont pas utilisés; on utilise à la place des feuilles de style CSS. Les balises doivent respecter la syntaxe XML

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

XHTML 1.0 Transitional

Cette DTD contient toutes les balises HTML, y compris les balises comme font et les balises de mise en forme. Les frames ne sont pas supportés par cette DTD. Les balises doivent respecter la syntaxe XML

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
```

XHTML 1.0 Frameset

Est équivalente à XHTML 1.0 Transitional mais elle utilise la balise frameset

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">
```

XHTML 1.1

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
```

2- Jeux de caractères

Les normes HTML demandent de respecter le codage des caractères **ASCII 7 bits**, c'est-à-dire que les caractères accentués ne sont pas autorisés.

Pourtant, les navigateurs actuels reconnaissent les caractères accentués, ainsi vous pouvez entrer des caractères accentués directement sous votre éditeur de texte, mais votre page sera vraisemblablement illisible dans la plupart des pays du monde ...

Pour coder un caractère accentué, on entre une combinaison précédée du caractère & et terminée par un point-virgule (;). L'annexe vous donne la liste des jeux de caractère HTML définie par la norme ISO-8859-1

<http://www.w3.org/QA/2002/04/valid-dtd-list.html#DTD>
<http://developpez.com/tutoriels/xhtml/>