

Règles de spécification XML

Le XML impose des règles de syntaxe très spécifiques par rapport au Html. En outre, on retrouvera ces mêmes règles de syntaxe dans tous les langages dérivés du XML comme le XHTML ou le WML par exemple.

Le XML est un langage de balises [Markup Language].

Mais au contraire du Html où les balises sont définies, vous devez inventer vos balises. Rappelez-vous, le XML est « Xtensible ». Il faut donc écrire soi-même le nom des balises utilisées.

Il y a quelques règles pour la composition des noms (mais elles ne déroutent pas les habitués du Javascript) :

- Les noms peuvent contenir des lettres, des chiffres ou d'autres caractères.
- Les noms ne peuvent débuter par un nombre ou un signe de ponctuation.
- Les noms ne peuvent commencer par les lettres xml (ou XML ou Xml...).
- Les noms ne peuvent contenir des espaces.
- La longueur des noms est libre, mais on conseille de rester raisonnable.
- On évitera certains signes qui pourraient selon les logiciels, prêter à confusion comme "-", ";", ".", "<", ">", etc.
- Les caractères spéciaux pour nous francophones comme é, à, ê, ï, ù sont à priori permis, mais pourraient être mal interprétés par certains programmes.

On profitera de cette liberté dans les noms pour les rendre le plus descriptifs possible comme par exemple `<gras_et_italique>`.

Les balises sont sensibles aux majuscules et minuscules [case sensitive].

Ainsi, la balise `<Message>` est différente de la balise `<message>`. La balise d'ouverture et la balise de fermeture doivent donc être identiques. Par exemple;

`<Message> ... </message>` est incorrect et `<message> ... </message>` est correct.

Une tendance se dégage pour n'écrire les balises qu'en minuscules, limitant ainsi les erreurs possibles.

Toute balise ouverte doit impérativement être fermée.

Fini les écritures bâclées du Html où l'on pouvait dans certains cas omettre la balise de fin comme pour le paragraphe `<p>` ou l'élément de liste ``.

Ainsi en Html, ce qui suit est affiché correctement :

```
<p>
<ul>
<li>Point 1
<li>Point 2
```

Le XML est beaucoup plus strict. On devrait avoir :

```
<p>
  <ul>
    <li>Point 1</li>
    <li>Point 2</li>
  </ul>
</p>
```

Les éventuelles balises uniques ou appelées aussi balises vides, comme
, <meta> ou en Html, doivent également comporter un signe de fermeture soit balise/. Ainsi une balise <meta/> est correcte en XML.

Les balises doivent être correctement imbriquées.

Le XML étant très préoccupé par la structure des données, des balises mal imbriquées sont des fautes graves de sens.

Ainsi l'écriture suivante est incorrecte, car les balises ne sont pas bien imbriquées :

```
<parent><enfant>Loïc</parent></enfant>
```

L'écriture correcte avec une bonne imbrication des éléments est :

```
<parent><enfant>Marine</enfant></parent>
```

Tout document XML doit comporter une racine.

En fait, la première paire de balises d'un document XML sera considérée comme la balise de racine [root].

Par exemple :

```
<racine>
    ... suite du document XML ...
</racine>
```

Si on ose faire un lien avec le Html, votre élément racine était <body> ... </body>.

Tous les autres éléments seront imbriqués entre ces balises de racine.

Par exemple :

```
<parents>
    <enfants>
        <petits_enfants> ... </petits_enfants>
    </enfants>
</parents>
```

Les valeurs des attributs doivent toujours être mises entre des guillemets.

Le XML peut avoir (comme le Html) des attributs avec des valeurs. En XML, les valeurs des attributs doivent obligatoirement être entre des guillemets, au contraire du Html où leur absence n'a plus beaucoup d'importance.

Ainsi, l'écriture suivante est incorrecte, car il manque les guillemets.

```
<date_anniversaire=071185>
```

La bonne écriture est :

```
<date_anniversaire="071185">
```

IMPORTANT : Le XML est un langage strict. Votre document doit impérativement respecter la syntaxe du XML. On dira alors que le document est "bien formé" [Well-formed]. Seuls les documents "bien formés" seront affichés correctement. À la moindre erreur de syntaxe, le document ne sera pas ou ne sera que partiellement affiché.