

TP 2 – Le web structuré

1. XML

Qu'est-ce que XML ?

- XML signifie langage de balisage extensible (eXtensible Markup Language).
- XML est un langage de balisage similaire à HTML.
- XML a été conçu pour stocker et transporter des données.
- XML a été conçu pour être auto-descriptif.
- XML est une recommandation du W3C (World Wide Web Consortium).

La différence entre XML et HTML

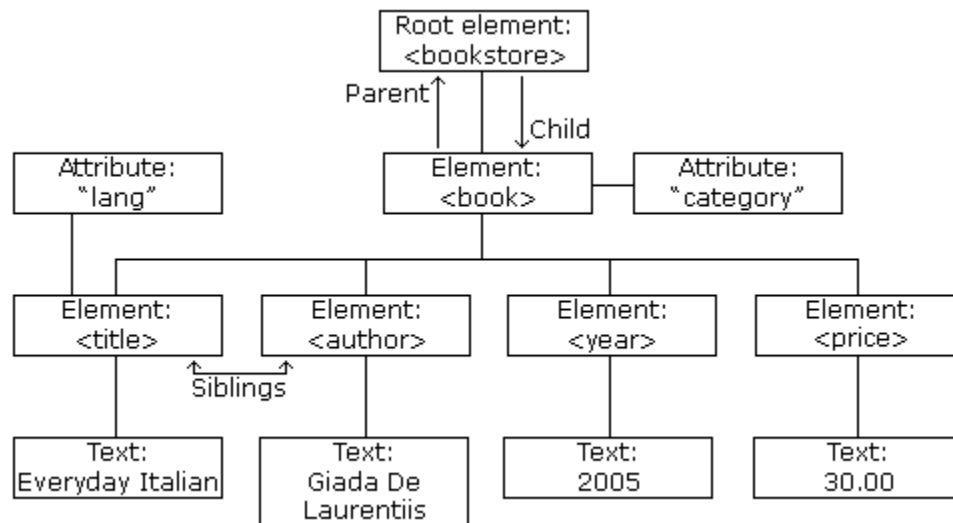
XML et HTML ont été conçus avec des objectifs différents :

- XML a été conçu pour transporter des données, en mettant l'accent sur ce que sont les données.
- HTML a été conçu pour afficher des données, en mettant l'accent sur l'apparence des données.
- Les balises XML ne sont pas prédéfinies comme le sont les balises HTML.
- XML est souvent un complément à HTML.

Exemple en XML

```
<author gender="male" nationality="Algérien">  
  <firstname>Yasmina</firstname>  
  <lastname>Khadra</lastname>  
  <birthdate>1955-01-10</birthdate>  
</author>
```

La structure arborescente XML



Les documents XML sont formés sous forme d'arbres d'éléments.

Un arbre XML commence par un élément racine et se ramifie à partir de la racine vers les éléments enfants.

Tous les éléments peuvent avoir des sous-éléments (éléments enfants) :

```

<root>
  <child>
    <subchild>.....</subchild>
  </child>
</root>
  
```

Les termes parent, enfant et frère/sœur (sibling) sont utilisés pour décrire les relations entre les éléments.

Les parents ont des enfants. Les enfants ont des parents. Les *siblings* sont des enfants du même niveau.

Tous les éléments peuvent avoir un contenu textuel (comme "Yasmina") et des attributs (comme nationality="Algérien").

Syntaxe auto-descriptive

XML utilise une syntaxe très auto-descriptive.

Un prologue définit la version XML et le codage des caractères :

```
<?xml version="1.0" encoding="UTF-8"?>
```

Exercice 1 :

Écrire un code XML pour représenter un extrait d'un magasin de livres (bookstore) qui affiche les informations suivantes pour chaque livre dans le magasin de livres :

- Le titre du livre
- L'auteur ou les auteurs
- L'année de publication
- Le prix du livre

Chaque livre a une catégorie, et chaque titre contient la langue du livre.

Exemple de livre

- "Harry Potter" est un livre écrit par J.K. Rowling et publié en 2005. Harry Potter" est un livre captivant écrit par l'auteure britannique J.K. Rowling. Il a été publié en 2005 et a rapidement conquis le cœur des lecteurs du monde entier. L'histoire suit les aventures d'un jeune sorcier nommé Harry Potter, qui découvre sa véritable identité de sorcier et embarque dans un voyage extraordinaire à Poudlard, l'école de sorcellerie. Le livre est devenu le premier volet d'une série de sept tomes, chacun représentant une année scolaire de Harry à Poudlard. Le livre est disponible au prix de 29,99 \$.
- "Python Crash Course" est un livre informatique écrit par Eric Matthes et publié en 2019 dans sa deuxième édition. Ce livre est une ressource idéale pour les débutants qui souhaitent apprendre la programmation en utilisant le langage de programmation Python, qui est largement utilisé et apprécié pour sa simplicité et sa polyvalence. Il offre une introduction complète à Python et couvre un large éventail de sujets, y compris les bases de la programmation, les structures de données, le développement web et la création de jeux. La version imprimée de ce livre coûte environ 30\$.
- "Guns, Germs, and Steel: The Fates of Human Societies" est un livre d'histoire écrit par Jared Diamond et publié en 1997. Ce livre explore l'histoire de l'humanité en analysant les facteurs qui ont influencé la montée et la chute des civilisations à travers le temps. L'auteur examine les rôles des armes, des germes (maladies), et de l'acier (technologie) dans la manière dont les sociétés ont évolué et se sont développées. La version imprimée de ce livre est de 15\$.

XML Syntax Rules

- Les documents XML doivent avoir un élément racine.
- Le prologue XML.
- Tous les éléments XML doivent avoir une balise de fermeture <element></element>.
- Les balises XML sont sensibles à la casse.
- Les éléments XML doivent être correctement imbriqués.
- Les valeurs des attributs XML doivent toujours être mises entre guillemets.
- Certains caractères ont une signification spéciale en XML. Si vous placez un caractère comme "<" à l'intérieur d'un élément XML, cela générera une erreur car le parseur l'interprète comme le début d'un nouvel élément.

Il y a 5 références d'entités prédéfinies en XML :

<	<	less than
>	>	greater than
&	&	ampersand
'	'	apostrophe
"	"	quotation mark

- La syntaxe pour écrire des commentaires en XML est similaire à celle de HTML :

```
<!-- This is a comment -->
```

- Les espaces blancs sont préservés en XML.
- Les documents XML qui respectent les règles de syntaxe énoncées ci-dessus sont dits être des documents XML 'Bien Formés'.

XML élément

- Un élément peut contenir :
 - Du texte
 - Des attributs
 - D'autres éléments
 - Ou une combinaison des éléments ci-dessus
- Les éléments XML sont extensibles.
- Un élément sans contenu est dit être vide

Exercice 2 :

Créer un document XML en suivant les instructions suivantes :

1. Définissez un élément racine appelé movies.
2. Créez les éléments suivants :
 - title : Cet élément doit contenir le titre d'un film.
 - director : Cet élément doit contenir le nom du réalisateur du film.
 - year : Cet élément doit contenir l'année de sortie du film.
 - genre : Cet élément doit contenir le genre du film (par exemple, "action", "comédie", "science-fiction", etc.).
3. Ajoutez des attributs à l'élément movie. Par exemple, vous pouvez ajouter un attribut "rating" pour spécifier la note du film (par exemple, "1", "2", "3", etc.).
4. À l'intérieur de l'élément movie, créez un élément description qui doit contenir un résumé du film.
5. Ajoutez du texte aux éléments créés.
6. Assurez-vous que votre document XML est bien formé en respectant les règles de syntaxe XML.

Exercice 3 :

En exploitant la phrase suivante « L'émir Abd el-Kader est né le 6 septembre 1808 », proposez trois syntaxes XML différentes pour représenter la phrase.

XML Namespaces- L'attribut xmlns

En XML, les noms des éléments sont définis par le développeur. Cela entraîne souvent des conflits lors de la tentative de fusion de documents XML provenant de différentes applications XML. Cet XML contient des informations sur un tableau HTML :

<pre><table> <tr> <td>Pomme</td> <td>Ananas</td> </tr> </table></pre>	<pre><table> <name>Table</name> <width>80</width> <length>120</length> </table></pre>
---	---

Si ces fragments XML étaient ajoutés ensemble, il y aurait un conflit de noms. Les deux contiennent un élément <table>, mais les éléments ont un contenu et un sens différents. C'est pourquoi il est essentiel d'utiliser **des préfixes de noms d'espace** pour différencier ces éléments

si vous devez les fusionner dans un même document XML, en particulier lorsque les éléments proviennent de contextes ou d'applications XML distincts. Cela évite les conflits et assure la clarté de la structure XML globale.

- Lors de l'utilisation de préfixes dans XML, un **espace de noms** pour le **préfixe** doit être défini.
- L'espace de noms peut être défini à l'aide d'un attribut xmlns dans la balise de démarrage d'un élément.
- La déclaration de l'espace de noms a la syntaxe suivante : xmlns:préfixe="URI"

```
<root>

<h:table xmlns:h="http://www.w3.org/TR/html4/">
  <h:tr>
    <h:td>Apples</h:td>
    <h:td>Bananas</h:td>
  </h:tr>
</h:table>

<f:table xmlns:f="https://www.w3schools.com/furniture">
  <f:name>African Coffee Table</f:name>
  <f:width>80</f:width>
  <f:length>120</f:length>
</f:table>

</root>
```

Exercice 4 :

Observez la syntaxe suivante. Quelle est la différence avec la précédente ?

```
<table xmlns="http://www.w3.org/TR/html4/">
  <tr>
    <td>Apples</td>
    <td>Bananas</td>
  </tr>
</table>
```