

RÉSUMÉ DU COURS DE XML

— Présentation —

Motivation : Un seul langage pour décrire et échanger des données.

XML est un *méta-langage* universel pour les **données** sur le web.

Avantage : *standard largement partagé*

Inconvénient : *balisage répétitif donc fichiers volumineux*

```
<librairie>
  <livre categorie="enfants">
    <titre langue="en">Harry Potter</titre>
    <auteur>J K. Rowling</auteur>
    <annee>2005</annee>
    <prix>29.99</prix>
  </livre>
  <livre categorie="programmation">
    <titre langue="fr">Apprentissage OCaml</titre>
    <auteur>Catherine Dubois</auteur>
    <annee>2004</annee>
    <prix>39.95</prix>
  </livre>
</librairie>
```

(On peut visualiser un fichier XML sous forme d'arbre)

Le prologue `<?xml version="1.0" encoding="UTF-8"?>` est **facultatif**, mais doit toujours être en première ligne si existant.

Les éléments sont constitués d'un nom (`<livre>`), éventuellement d'attributs (`langue="en"`) et d'un contenu (vide, d'autres éléments et/ou du texte).

Attention : si un texte contient des délimiteurs XML (`<` , `>` , `&`), l'inclure dans une section CDATA : `<![CDATA[texte]]>`.

— Espaces de nom —

On peut définir des **espaces de nom** pour éviter les ambiguïtés :

```
<root>
  <h:table xmlns:h="http://www.w3.org/TR/html4/">...</h:table>
  <f:table xmlns:f="http://www.w3schools.com/furniture/">...</f:table>
</root>
```

Cette adresse ne sert qu'à donner un nom unique à l'espace de nom.

Les définitions `xmlns:h` et `xmlns:f` peuvent être placés sur l'élément `root`.

Pour valider les fichiers :

- le document est-il **bien formé** ? (structure d'arbre)
- le document est-il **bien valide** ? (dialecte XML spécifique)

Un document est bien formé si :

- il possède une racine unique
- il possède au moins un élément
- chaque balise ouverte est fermée (les balises sont sensibles à la casse)
- l'imbrication est correcte
- la valeur des attributs est comprise entre guillemets

— Méta-description de documents —

DTD : Document Type Definition

XSD : XML Schema Definition

cf. le polycopié (difficile à résumer ...)