

XML

UE : STRUCTURES DE DONNÉES: AVANCÉ

BLOC 2

CHRISTOPHE DAMAS



DÉTAILS PRATIQUES

- **Evaluation**

- Projet intégré (10% de l'examen de juin)
- Examen intégré (90 % en juin, 100% en septembre)
 - Sur machine
 - Moodle accessible (syllabus, ...)

- **Support**

- Syllabus
- Slides
- Exercices

1. Bases XML / DTD
2. DTD / XML Schema
3. XML Schema
4. XML et JAVA
5. Projet SD
6. Projet SD
7. Projet SD
8. XSL
9. XPath
10. XSL – XPath : together
11. XQuery
12. XQuery

XML

1 : BASES XML

BLOC 2

CHRISTOPHE DAMAS

E XML t a n s i b l e a r k u p a n g u a g e





XML = Langage de balisage des données.

Balise = Définition du contenu au niveau sémantique

EXEMPLE - DONNÉES

Rue de l'arbre, 43
1300 Wavre

**Ceci correspond à une adresse postale ;
la rue est « rue de l'arbre »,
le numéro est 43,
la ville est Wavre
et le code postal 1300.**



EXEMPLE - DONNÉES

String



Rue de l'arbre, 43
1300 Wavre

EXEMPLE – MÉTA DONNÉES

<adresse>

<rue>Rue de l'arbre </rue>

<numero> 43 </numero>

<codePostal> 1300 </codePostal>

<ville> Wavre </ville>

</adresse>

XML – EN BREF

- ✓ **A été conçu pour manipuler des données (transporter et stocker)**
- ✓ **Permet un nombre illimité de balises; chacun peut en créer.**
- ✓ **Rend les données auto-descriptives**
- ✓ **Est international (unicode)**
- ✓ **Est indépendant au point de vue :**
 - Plate-forme
 - Vendeur / constructeur
 - Langage de programmation
- ✓ **Est facilement lisible par un humain**
- ✓ **Est une recommandation du W3C**

XML ET HTML

PGCD (XML, HTML) = ML
... càd Markup Language

SGML = Standard Generalized Markup Language

**Standard de 1986 pour la normalisation de
document électronique**

QUELLES DIFFÉRENCES ENTRE XML ET HTML?

HTML	XML
Langage de présentation des données	Langage de structuration des données (importance de la sémantique)
Interprétation dépendante des browsers	Information toujours identique
Nombre de tags définis par le langage	Nombre illimité de tags définis par l'utilisateur
Conçu uniquement pour le Web	Conçu pour toutes manipulations (transport ou stockage de données)

PROLOGUE D'UN DOCUMENT XML

1) Instruction de traitement <?

- Interprétée par l'application servant à traiter le document XML.
- Elle ne fait pas totalement partie du document.

Les instructions de traitement qui servent le plus souvent sont

Obligatoire! 1^{ère} ligne du document xml

la déclaration XML:

```
<?xml version="1.0" encoding="ISO-8859-1"
standalone="yes"?>
```

PROLOGUE D'UN DOCUMENT XML

2) Déclaration de type de document

Permet de définir la structure d'un document

```
<!DOCTYPE elementRacine SYSTEM "monFichier.dtd">
```

*Facultatif mais après la déclaration XML
et avant l'élément racine!*

VOCABULAIRE - 1

Une balise = tag **< >**

→ délimiteur des données (méta donnée)

Un élément **< > </ >**

→ balise d'ouverture + (donnée ou sous-éléments) + balise de fermeture

```
<titre>Les Misérables</titre>
```

```
<livre>
```

```
    <titre>Les Misérables</titre>
```

```
    <auteur>Hugo</auteur>
```

```
</livre>
```


VOCABULAIRE - 2

Un élément racine = root

→ élément base du document xml; il englobe tous les autres.

```
<maitre>
```

...

```
</maitre>
```

Un élément vide

→ élément sans donnée

```
<vide></vide>
```

```
<vide/>
```

Un attribut

→ descriptif d'une balise ouvrante (nom= " valeur "). Une balise d'ouverture peut contenir un ou des attributs mais une seule fois chaque attribut.

```
<personne sexe="F">Lisa</personne>
```

```
<animal race="chien"/>
```

```
<?xml version="1.0" encoding="ISO-8859-1" standalone="yes"?>
```

<ListeMembres>

<Personne Nationalite="Belge">

<Nom>Duchmol</Nom>

<Prenom>Hubert</Prenom>

<Coordonnees>

<Tel>02/764.46.54</Tel>

<Adresse>

<Rue>Rue de l'arbre, 43</Rue>

<CodePostal>1300</CodePostal>

<Ville>Wavre</Ville>

</Adresse>

</Coordonnees>

</Personne>

<Personne>

<Nom>Durant</Nom>

<Prenom>Claude</Prenom>

<Coordonnees>

<Tel>02/774.48.49</Tel>

<Tel>02/774.48.50</Tel>

<Adresse>

<Rue>Rue de l'Eglise, 34</Rue>

<CodePostal>1200</CodePostal>

<Ville>Bruxelles</Ville>

<Pays>Belgique</Pays>

</Adresse>

</Coordonnees>

</Personne>

</ListeMembres>

Instruction de traitement

Elément « Root » (apparaît une seule fois)

Balise (tag) d'ouverture

Balise (tag) de fermeture

<Pays>Belgique</Pays>

Elément

RÈGLES DE COMPOSITION

Tout identificateur XML doit :

1. Commencer par une des 26 lettres de l'alphabet ou l'underscore (_).
2. Se poursuivre avec des lettres, des chiffres, des underscores (_), des tirets (-) ou des points (.).
3. Absolument ne pas contenir d'espace.
4. Ne pas contenir de deux points (:) car ceux-ci sont réservés.

**Attention XML est sensible à la casse;
A n'est pas a! b n'est pas B, ...**

RÈGLES DE FORMULATION

Un document XML est bien formé/formulé ssi il

- I. Respecte des règles de composition
- II. Renferme en première ligne la déclaration XML.
- III. Contient un seul élément racine.
- IV. Se compose d'éléments qui se correspondent de manière hiérarchique.

COMMENTAIRE

Idem qu'en HTML, car ils reprennent la syntaxe du SGML.

Ils commencent donc par `<!--` et se terminent par `-->`. Ils peuvent être placés à n'importe quel endroit tant qu'ils se trouvent à l'*extérieur* d'une balise.

Exemples de commentaires valides :

```
<!-- ceci est correct -->  
<elt> <!-- ceci est correct aussi -->  
Un peu de texte </elt>
```

En raison de la compatibilité XML/SGML, la chaîne de caractères `-` est *interdite* dans un commentaire.

EXERCICES – VOCABULAIRE 1

<ceci>cela</ceci>	cela est
<ceci>bla</ceci>	est une balise
<ici>ça</ici>	est une donnée
<tout>ça</tout>	toutçatout est
<cela/>	cela est
</cela>	cela est
<ceci>	ceci est
<personne cela="la"/>	cela est

EXERCICES – VOCABULAIRE 2

Corrigez les éléments suivants:

`<element><element/>`

`<element></element attribut="valeur">`

`<element/><element>`

`<element attribut=valeur/>`

`<element attribut="valeur">`

EXERCICES - COMPOSITION

Les identifiants (correspondant à des éléments ou des attributs) suivants sont-ils écrits correctement?

	OUI – NON	Pourquoi?
2do		
things2Do		
Type6		
prénom		
élève		
c.p.		
ville-pays		
-fruit		
_fruit		
nom complet		
prenom:nom		

EXERCICE – BIBLIOTHEQUE

Ecrivez un document XML bien formé contenant les informations suivantes:

Hygiène de l'assassin (isbn : 978-2-253-11118-4) est le premier roman d'Amélie Nothomb. Le personnage principal du livre est le célèbre romancier Prétextat Tach, né le 24 février 1908, qui est atteint du syndrome d'Elzenveiverplatz. Amélie Nothomb est née le 13 août 1967.

Un aller simple (isbn : 2-07-040368-8) est un roman de Didier Van Cauwelaert.

Le portrait de Dorian Gray (isbn : 9783196729047) est un roman d'Oscar Wilde né le 16 octobre 1854 et mort le 30 novembre 1900. Le personnage principal de ce roman est Dorian Gray reconnu comme étant un jeune homme d'une très grande beauté.

ELÉMENT OU ATTRIBUT

Elément	Attribut
Importance de l'ordre	Ordre importe pas
Plusieurs fois présents dans un élément	Une seule fois présent dans un élément
Avoir une sous structure	Aucune sous-structure
Imposer un contenu vide	Imposer une valeur par défaut
	Enumération possible
	Imposer l'unicité du contenu
Représenter un objet	Représenter les propriétés d'un objet