



CENTRE DE RESSOURCES INFORMATIQUES

Chapitre 1: Présentation de XML

Dr Youssou KASSE

Introduction

- Pour donner un aperçu rapide de XML, on peut le définir tel un langage (ou une méthode) permettant de structurer des données en employant un format texte comme support.
- XML (*eXtended Markup Language*) est une spécification du W3C depuis février 1998,

Pourquoi XML?

- SGML, qui permettrait de définir de nouveaux langages de balisage spécifiques est complexe.
- Par exemple `<A>` est ambigu car on ne sait pas s'il faut l'interpréter comme :
 - `<A>`
 - ou `<A>`.

Pourquoi XML?

- HTML possède des inconvénients:
 - Jeu de balises restreint;
 - Mélange des descripteurs physiques et logiques;
- Html a été conçu pour afficher de l'information, XML a été créé pour structurer de l'information.

Quelques applications

- **MathML** *Mathematical Markup Language* notation mathématique sur le web
- **SMIL** *Synchronized Multimedia Integration Language* pour la création multimédia, il spécifie comment et quand des éléments multimédia peuvent apparaître dans une page web

Quelques applications

- **SVG** (Scalable Vector Graphic) est une application XML dont l'objet est la description d'objets graphiques en 2 dimensions.
- **WML** *Wireless Markup Language* le langage de balisage pour l'internet mobile

Présentation de XML

```
1 <?xml version='1.0' encoding='iso-8859-1'?>
2 <mescontacts>
3     <contact>
4         <nom>FAYE</nom>
5         <prenom>Moustapha</prenom>
6         <email>moustapha.faye@uadb.edu.sn</email>
7     </contact>
8     <contact>
9         <nom>DIONE</nom>
10        <prenom>Amy</prenom>
11        <email>amy.dione@gamil.com</email>
12    </contact>
13 </ mescontacts>
```

Présentation de XML: Principe

- Les documents XML sont destinés essentiellement à être traités par des programmes informatiques appelés **processeurs XML**.
- Les documents XML doivent être facilement lisibles par l'homme
- Un document XML doit obligatoirement être **bien formé**

Présentation de XML: Document bien formé

- Tout document XML doit avoir un et un seul **élément racine** : un élément qui contient tous les autres.
- Chaque élément doit avoir une **balise de début** et la **balise de fin correspondante** (sauf pour la balise vide: `<A/>`).

Présentation de XML: Document bien formé

```
1 <?xml version='1.0' encoding='iso-8859-1'?>
2 <mescontacts>
3   <contact>
4     <nom>FAYE</nom>
5     <prenom>Moustapha</prenom>
6     <email>moustapha.faye@uadb.edu.sn</email>
7   </contact>
8   <contact>
9     <nom>DIONE</nom>
10    <prenom>Amy</prenom>
11    <email>amy.dione@gamil.com</email>
12  </contact>
13 </mescontacts>
```

Présentation de XML: Document bien formé

- Les balises doivent être correctement imbriquées entre elles .
- `<A><A/>`
- Les valeurs associées aux attributs doivent être encadrées par les délimiteurs « " »

Présentation de XML: Document bien formé

```
1 <?xml version= "1.0" encoding= "iso-8859-1"?>
2 <mescontacts>
3     <contact telephone= "777777777">
4         <nom>FAYE</nom>
5         <prenom>Moustapha</prenom>
6         <email>moustapha.faye@uadb.edu.sn</email>
7     </contact>
8     <contact telephone= "666666666">
9         <nom>DIONE</nom>
10        <prenom>Amy</prenom>
11        <email>amy.dione@gamil.com</email>
12    </contact>
13 </mescontacts>
```

Présentation de XML: Document bien formé

- Les noms des balises doivent commencer par une **lettre** ou "_", les autres caractères peuvent être des chiffres, des lettres, "_", "." ou "-".
- Les noms des balises ne doivent pas commencer par xml.
- **ATTENTION:** XML est key sensitive

Présentation de XML: Document bien formé

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

```
<personne>
```

```
  <nom>
```

```
  <prenom>
```

```
</nom>
```

```
</prenom>
```

```
<personne>
```

Ko : Mauvaise imbrication

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

```
<personne>
```

```
</Personne>
```

Ko : Personne différent de personne

Présentation de XML: Document bien formé

```
<?xml version="1.0" encoding="ISO-8859-1"?>  
<xmlmine>  
</xmlmine>
```

Ko : Pas de nom de balise commençant par xml

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

Ko : Balise non fermée

```
<?xml version="1.0" encoding="ISO-8859-1"?>  
<personne/>
```

Ok : Élément vide