

INF4375 – Paradigmes des échanges Internet

Introduction au XML

Jacques Berger

Objectifs

L'utilité du XML

Le vocabulaire relatif au XML

La syntaxe du XML

Les outils

Prérequis

Aucun

XML

eXtensible Markup Language

Dernière version : 1.1

Couramment utilisée : 1.0

XML

Norme du W3C (World Wide Web Consortium)

Création de langages

Balises

Utilité

Modéliser des données

- Une exécution
- Un algorithme
- Une structure de données
- Une configuration
- Une image
- Un graphique
- Une formule mathématique
- Une molécule chimique
- Un document
- Un livre
- Un message
- etc.

Origine

GML

Generalized Markup Language

Date des années 60

SGML

Standard Generalized Markup Language

ISO 8879:1986

Origine

Les dérivés du SGML

HTML (HyperText Markup Language)

XML

MathML

MusicML

etc.

Abstraction

Un arbre syntaxique

Éléments auto-descriptifs

Contient de l'information

Types de documents

Documents narratifs
Sans schéma

Documents structuraux
Avec schéma

XML

Extension habituelle : .xml

MIME Type : application/xml

Vocabulaire

Document

Un fichier ou un ensemble de fichiers XML
Les données, la structure, les instructions

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

...

```
</document>
```

Vocabulaire

Élément

Un noeud de l'arbre (abstraction)

Peut contenir des données ou d'autres éléments

```
<element>  
  <contenu>texte</contenu>  
</element>
```

Vocabulaire

Élément vide

Un élément qui ne contient rien

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

```
<vide />
```

Vocabulaire

Élément racine

Le premier élément du document

L'élément qui contient tous les autres éléments

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

...

```
</document>
```

Vocabulaire

Élément texte

Élément dont le contenu n'est que du texte

```
<element>Texte</element>
```


Vocabulaire

Élément enfant

Un élément à l'intérieur d'un autre élément

```
<parent>  
  <enfant>  
    ...  
  </enfant>  
</parent>
```

Vocabulaire

Attribut

Une donnée dans la déclaration d'un élément

```
<element attribut1="valeur1" attribut2="valeur2" />
```

Vocabulaire

Namespace

Une façon d'identifier qu'un élément appartient à un ensemble d'éléments spécifique

```
<xsl:template>
```

```
...
```

```
</xsl:template>
```

Vocabulaire

Racine du document

Un élément abstrait, non visible, situé avant l'élément racine

Ne pas confondre avec l'élément racine

Vocabulaire

Noeud

Tout est un noeud

Élément

Commentaire

Instructions

Texte

Règles de forme

Document bien formé

Respecte la norme du W3C

Document valide

Respecte son schéma

Note : un document bien formé et sans schéma est valide

Règles de forme

Un seul élément racine par document.

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

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

Règles de forme

Balise fermante obligatoire

```
<element>Texte</element>  
<element />
```

```
<element>  
  <enfant>  
</element>
```


Règles de forme

Les identifiants sont sensibles à la casse

Éléments

Attributs

`<element> ≠ <Element>`

Règles de forme

L'indentation est optionnelle

```
<document><element1><element2>Data</element2><element2>More  
Data</element2></element1></document>
```

Règles de forme

Certains caractères spéciaux doivent être échappés

Exemple : '>' devient >

Consultez la norme pour la liste complète

Processing-instruction

Les documents XML commencent généralement par une ligne de ce genre :

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

Il s'agit d'une processing instruction

Processing-instruction

La processing-instruction sert aux logiciels qui utiliseront ce document XML

Il s'agit de métadonnées pour le document

Processing-instruction

On y spécifie :

- La version XML utilisée

- L'encodage de caractères des données

- Si le document doit être lié à un schéma

- Autres...

Outils

Microsoft XML Notepad 2007

XML Marker

Liquid XML Studio 2009

XML Mind Editor

Liens

XML 1.0

<http://www.w3.org/TR/2008/REC-xml-20081126/>

XML 1.1

<http://www.w3.org/TR/2006/REC-xml11-20060816/>