

Intro au XML



Introduction

Qu'est-ce qu'XML (Extensible Markup Language/langage de balisage extensible)?

Créé par le W3C (World Wide Web Consortium)

- C'est un langage
 - C'est un langage de balisage
 - C'est un langage extensible
-
- **HTML** : langage de balisage spécialement créé pour la conception des pages web
 - **XML** : décrire la structure et l'organisation d'un site web pour faciliter son exploration par robots d'indexation des moteurs de recherche

```
<?xml version="1.0"?>
<questionnaire>
  <question>
    Qui était le premier
    empereur romain ?
  </question>
  <réponse>
    Auguste
  </réponse>
  <!-- Note : tu auras besoin
    de plus de questions.-->
</questionnaire>
```

XML

Structure du XML

Le prologue

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

Les attributs

```
<savants>
```

```
  <!-- Quelques physiciens importants -->
```

```
  <savant id="phys0">
```

```
    <identité prénom="Galileo" nom="Galilei"/>
```

```
    <date: naissance="1564-02-15" décès="1642-01-08"/>
```

```
    <contributions>Relativité du mouvement & système héliocentrique</contributions>
```

```
  </savant>
```

```
  <savant id="phys1">
```

```
    <identité prénom="Isaac" nom="Newton"/>
```

```
    <dates naissance="1643-01-04" décès="1727-03-31"/>
```

```
    <contributions>Gravitation universelle, principe d'inertie, décomposition de la lumière  
& calcul infinitésimal</contributions>
```

```
  </savant>
```

```
  <savant id="phys2">
```

```
    <identité nom="Einstein" prénom="Albert"/>
```

```
    <dates naissance="1879-03-14" décès="1955-04-18"/>
```

```
    <contributions><![CDATA[Relativités restreinte & générale, nature corpusculaire de la  
lumière & effet photoélectrique]]></contributions>
```

```
  </savant>
```

```
</savants>
```

Normes d'écriture :

```
<savants>
```

```
  <savant>
```

```
    blabla
```

```
  </savant>
```

```
</savants>
```

Pourquoi le XML?

Le XML permet:

- de stocker et de structurer l'information (avec l'arborescence)
- de pouvoir d'éditer facilement (fichier texte)
- d'être compréhensible (commentaires possible)
- d'être extensible (rien n'est figé)
- d'être utilisable par les logiciels (pour python, le package standard xml)

Ces limites sont:

- pas de typage de données (short, long...)
- assez lourd dû aux balises (par rapport à binaire, tabulaire) et pour la transmission de données

Les schemas

Le schéma est la grammaire du document XML. Il décrit sa structure. Il est utilisé pour la valider.

Document Schema Definition Languages (DSDL) est un cadre qui décrit les tâches de validation de différents types qui peuvent s'appliquer à un document XML.

Un DSDL correspond à un langage informatique déclaratif de description de document XML. Comme SGML, XML est un langage de balisage générique, dont le vocabulaire et la grammaire ne sont pas définis a priori. Ces règles s'expriment en un schéma.

Contrairement à SGML, XML accepte plus d'une syntaxe de schéma. DTD, XML Schema, Relax NG, Schematron sont des DSDL.



DSDL est spécifié par la norme ISO/CEI 19757.

Cas du DTD

Le DTD (Document Type Definition), défini par W3C, est le seul moyen officiel de validation de fichier XML.

Un fichier DTD indique:

- les noms des attributs assortis de leurs caractéristiques
- les cardinalités (? pour 0 ou 1, * pour 0 à plusieurs et + pour un à plusieurs)



Exemple de DTD

```
<!ELEMENT doc (livre* | article+)>
<!ELEMENT livre (titre, auteur+)>
<!ELEMENT article (titre, auteur*)>
<!ELEMENT titre(#PCDATA)>
<!ELEMENT auteur(nom, adresse)>
<!ATTLIST auteur id ID #REQUIRED>
<!ELEMENT nom(prenom?, nomfamille)>
<!ELEMENT prenom (#PCDATA)>
<!ELEMENT nomfamille (#PCDATA)>
<!ELEMENT adresse ANY>
```


DTD


11

Cas du XSD

Un autre schéma est le XSD, XML Schema Definition (aussi soutenu par W3C).

Il s'agit d'un fichier XML qui est un schéma décrivant d'autres fichiers XML.

UNIVERSITÉ DE FRANCHE-COMTÉ



XML Schema : exemple (1)

```
<xsd:schema xmlns:xsd="http://www.w3.org/1999/XMLSchema">

  <xsd:element name="purchaseOrder" type="PurchaseOrderType"/>
  <xsd:element name="comment" type="xsd:string"/>
  <xsd:complexType name="PurchaseOrderType">
    <xsd:sequence>
      <xsd:element name="shipTo" type="USAddress"/>
      <xsd:element name="billTo" type="USAddress"/>
      <xsd:element ref="comment" minOccurs="0"/>
      <xsd:element name="items" type="Items"/>
    </xsd:sequence>
    <xsd:attribute name="orderDate" type="xsd:date"/>
  </xsd:complexType>

</xsd:schema>
```

Schema

29

Conclusion

Le XML est un langage. Il est utilisé dans de nombreux domaines.

Avec sa structure, il s'adapte pour décrire au mieux les informations.

Il possède un système de validation.

Il est notamment utilisé dans:

- www
- xdocument
- les logiciels
- fichiers de configuration
- ...