

Introduction

HTML : actuellement le format standard des documents sur le Web

Plusieurs inconvénients:

- nombre d'éléments et d'attributs limités
- extension => révision de la spec HTML
- conversion vers HTML => perte d'information

XML

"Extensible Markup Language"

- Sous ensemble de SGML (Standard Generalized Markup Language)
- Permet la distribution de documents SGML sur le Web

Recommandation W3C depuis le 10 Février 1998

Spécification à: <http://www.w3.org/TR/REC-xml>

XML comparé à SGML

Différence principale :

- Le parsing de SGML nécessite une DTD (Document Type Definition)
- Pas le parsing de XML

Par exemple en XML:

- Tags de fin obligatoires
- Tags des éléments vides finissent par /

```
<ul>
  <li>un item</li>
  <li>un autre item</li>
</ul>
<p>Un paragraphe avec une image </p>
```

=> DTD sert uniquement à la validation

XML comparé à HTML

- XML est plus bas niveau (syntaxe abstraite)
- HTML est une application (syntaxe concrète) de SGML
- XHTML est une application de XML

XML comparé à HTML (suite)

HTML a une sémantique :

```
<BODY>  
  <H1>Du texte</H1>  
  <P>Un autre morceau de texte</P>  
  <P>Plus de texte</P>  
</BODY>
```

XML comparé à HTML (suite)

XML n'en a pas :

```
<MYDOC>  
  <ELEM1>Du texte</ELEM1>  
  <ELEM2>Un autre morceau de texte</ELEM2>  
  <ELEM3>Plus de texte</ELEM3>  
</MYDOC>
```

Utilisations de XML

1. Affichage de documents => XML + style
avantage par rapport à HTML + style : conserve toute l'information
2. Traitement spécialisé => XML + outil spécialisé
possibilités illimitées

DOM

Outils spécialisés = outils génériques + modules spécialisés

Echange de modules nécessite une interface standard : "Document Object Model" (DOM)

Aussi utilisé pour "DHTML" : HTML + CSS + scripting + DOM

DOM (suite)

API (Application Programming Interface) standard d'accès à un document et son style

- indépendante du langage de programmation
- indépendante de la plateforme

DOM Level 1 :

- Accès et manipulation de la structure d'un document XML ou HTML

Recommandation W3C depuis le 1 Octobre 1998
Spécification à: <http://www.w3.org/TR/REC-DOM-Level-1>

DOM (suite)

DOM Level 2 :

- CSS
- Range
- Filtres et Iterateurs
- Evénements utilisateur et de transformation
- Namespaces

Working Draft W3C du 28 Décembre 1998

Spécification à: <http://www.w3.org/TR/WD-DOM-Level-2>

XML Namespaces

Composition de documents par modules:
Forms, Tables, Math, etc...

Chaque élément et attribut est identifié par un nom: prefix + URI

Exemple:

```
<?xml version="1.0"?>
  <bk:book xmlns:bk='urn:loc.gov:books'
           xmlns:isbn='urn:ISBN:0-395-36341-6'>
    <bk:title>Cheaper by the Dozen</bk:title>
    <isbn:number>1568491379</isbn:number>
  </bk:book>
```

Recommandation W3C depuis le 14 Janvier 1999

Spécification à: <http://www.w3.org/TR/REC-xml-names>

XML Linking

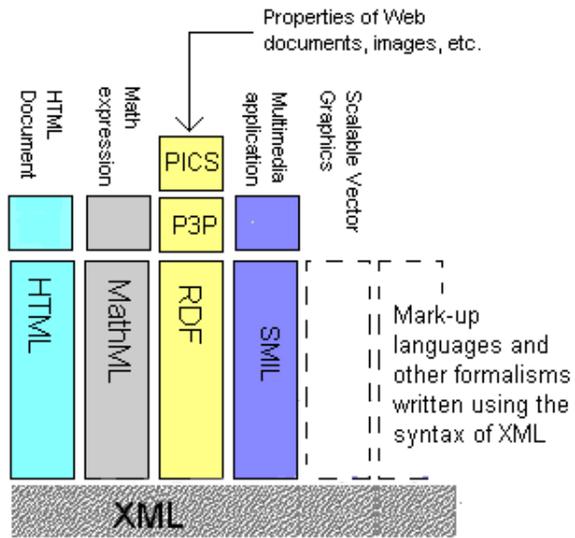
Mécanisme généralisé de liens :

- liens hypertextes, similaires à ceux d'HTML mais sans ancors
- inclusions
- collections
- ...

Autres technologies associées à XML

- XSL, XTL, CSS : transformation et style
- Schemas : nouvelle forme de DTD
- RDF
- ...

Architecture Générale



Pour plus d'info voir [href="http://www.w3.org/XML"](http://www.w3.org/XML).