



Module 1

Rappels fondamentaux sur XML

Contenu du module

- Concepts de document
 - Le balisage
 - Document bien formé vs. document valide
- Les applications d'XML
 - Analyse, transformation (XSLT), formatage (XSL-FO)
- Les modèles de données : XMLSchema
- Validation de données : DTD, XMLSchema

Historique

- XML
 - eXtensible Markup Language
 - Langage extensible de balisage de documents
 - Une recommandation du W3C
- XML 1.0
 - 10 février 1998
- XML 1.1
 - 4 février 2004
- Langage de description de la structure logique des documents, adapté aux applications Web (mais pas que !)

Concepts de document

- Le document électronique aujourd'hui:
 - Tables de base de données relationnelle
 - Bon de commande
 - Formulaire électronique
 - Manuel de procédures administratives
 - Fichier de configuration d'applications
 - Données échangées entre applications
 - etc...
- XML permet de décrire la structure et le contenu de tout type de document.

Le document XML

- Structure Logique et Physique du Document.
 - Structure hiérarchique et imbrication des éléments structurels.
 - Relation et dépendance entre les éléments structurels.
 - Exemple: un livre contient une introduction suivi d'un ou plusieurs chapitres pouvant contenir à leur tour des sous-chapitres.
- XML décrit:
 - La structure logique du document et des contraintes sur la structure du documents
 - La structure physique du document (les entités du documents)

Le balisage

- XML utilise le balisage descriptif.
 - Décrit la sémantique du contenu.
 - Basé sur la structure hiérarchique du document.
 - Sépare le contenu des instructions de traitement (y compris le formatage).
 - Permet validation et navigation de la structure du document.
- Le Balisage en XML.
 - `<GI>contenu</GI>`
 - Un **élément** XML !
 - Balise d'ouverture : `<GI>`
 - Balise de fermeture : `</GI>`
 - La balise d'ouverture peut contenir des **attributs** qui qualifient l'élément.
 - `<titre isbn="1-8956-7826-3">XML : Tout savoir !</titre>`

Document XML bien formé vs. document valide

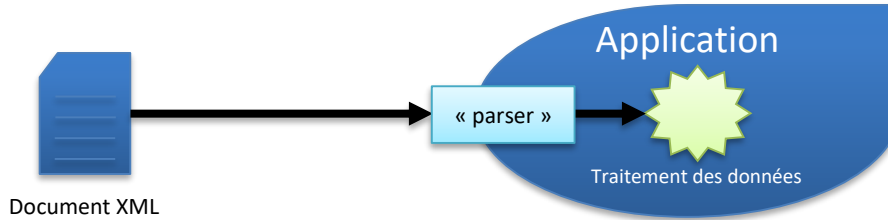
- Documents XML bien formés :
 - Contient au moins un élément.
 - Possède un élément racine.
 - Qui englobe tous les autres.
 - Imbrication correcte des balises.
- Documents Valides :
 - Sont associés à une grammaire !
 - Elle permet de valider la sémantique.
 - Deux technologies de grammaires XML :
 - DTD : Document Type Definition
 - W3C : En même temps qu'XML !
 - XMLSchema
 - W3C : mai 2001

Les applications d'XML

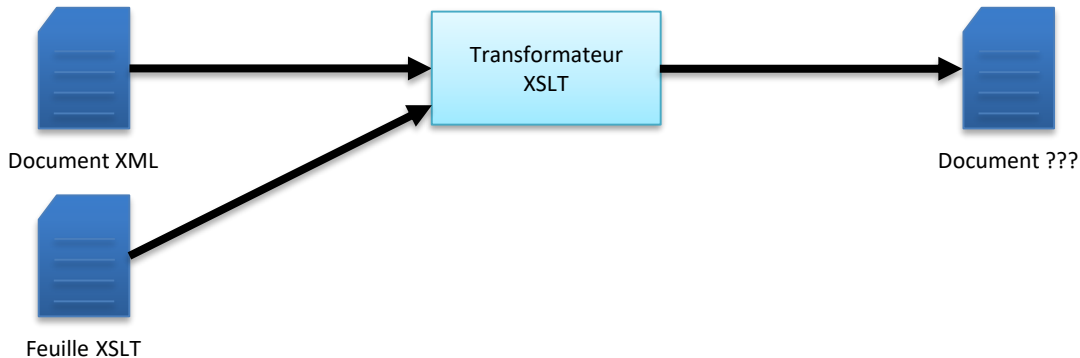
- XML sans outils de traitement ne sert strictement à rien !
- Besoin de technologies associées pour manipuler les structures de document !
 - Analyse
 - Pour exploitation dans des applications informatiques.
 - A l'aide d'un analyseur : « parser »
 - Transformation.
 - XML vers ...
 - Formatage de données
 - Une sortie structurée.
- Transformation.
 - XSLT : XML Stylesheet Language Transformation
- Formatage.
 - XSL-FO : XML Stylesheet Language – Formatting Objects

Les applications d'XML (suite)

- Analyse XML

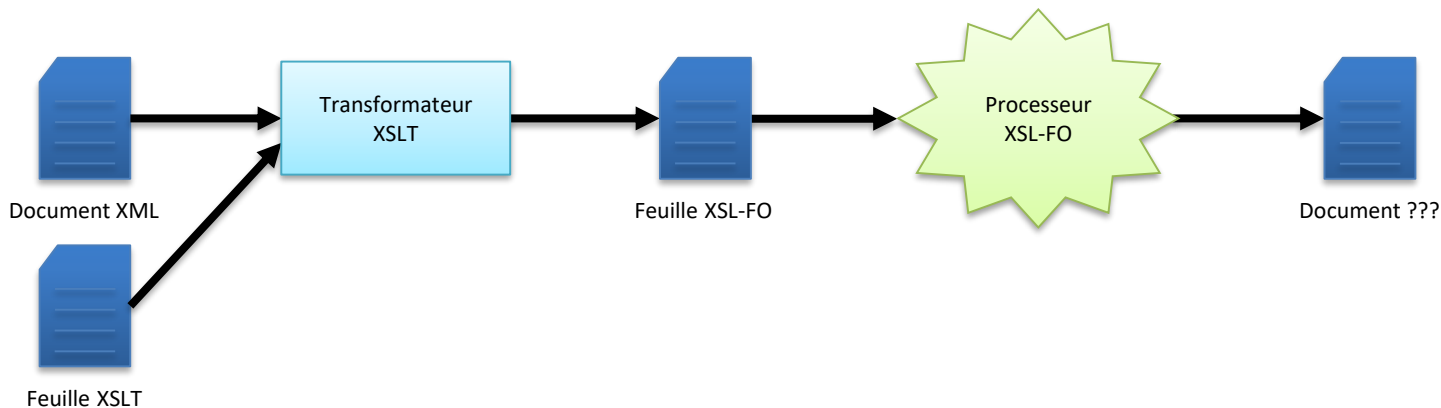


- Transformation XSLT



Les applications d'XML (suite et fin)

- Formatage (XSL-FO)



Les modèles de données : XMLSchema

- Recommandation W3C en mai 2001
- Langage de description de format de document XML
 - Permet de définir la structure et le type de contenu d'un document XML.
- Une définition se compose d'un ou plusieurs documents XML, usuellement nommée **XML Schema Definition** en anglais (ou fichier XSD).
- XMLSchema améliore la définition de format par rapport à DTD.
- Exemple :
 - Définition d'une structure « personne »

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="personne">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="nom" type="xs:string" />
        <xs:element name="prenom" type="xs:string" />
        <xs:element name="date_naissance" type="xs:date" />
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

Validation de données : DTD, XMLSchema

- DTD comme XMLSchema permettent de valider la structure d'un document XML.
 - Ce sont des définitions de grammaires XML !
- DTD permet de valider l'enchaînement des éléments, la présence d'attributs, la cardinalité des éléments.
- XMLSchema va plus loin !
 - Il permet de valider le format des données !
 - Date de naissance, numéro ISBN d'un livre, ...
 - Mieux : il permet également de définir ses propres formats de données !
- XMLSchema remplace aujourd'hui presque systématiquement DTD
 - Parce qu'il permet d'exprimer des structures de données.
 - Parce qu'il va plus loin dans le processus de validation.