



## M2 : Intégration et travail coopératif

Bertrand DAVID

Ecole Doctorale ED IIS  
DEA ISCE

Xml  
La transformation avec Xslt  
Les parseurs  
les outils pour Xml

Transformation de documents  
dans les applications multimédia

XML

eXtensible Markup Language

Un langage universel et standardisé pour la représentation  
textuelle de données structurées

Recommandation du W3C ( World Wide Web Consortium) qui gère  
la définition de standards du Web

Mais aussi,

des constructeurs : IBM, SUN, Microsoft,...

des laboratoires : INRIA ( France ) , MIT  
(US),...

## Motivations

- ◆ Utiliser un standard (indépendance vis-à-vis d'un constructeur)
- ◆ Réutiliser le contenu, aborder l'édition multi-supports (séparation du contenu et de la structure de présentation)
- ◆ Faciliter les traitements automatiques
- ◆ Disposer d'un outil simple et universel pour les échanges d'informations

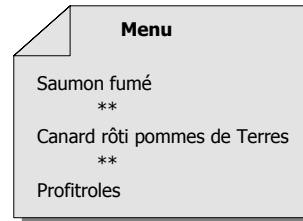
### Avantages

- Code lisible sur n'importe quel éditeur de textes
- Création de balises spécifiques (<> *HTML*)

## Composition d'un document XML

- ◆ Prologue :
  - ◆ balise de début
    - <?xml version="1.0" ?>
  - ◆ déclaration d'un document type
    - <!DOCTYPE doc1 SYSTEM « doc1.dtd »>
- ◆ code XML :
  - ◆ Élément (avec ou sans attribut)
    - <balise1> </balise1>
    - <Code number="TYC001" />
  - ◆ Commentaire
    - <!-- ceci est un commentaire -->
  - ◆ instruction à exécuter
    - <? Nom\_du\_programme arguments ?>
  - ◆ sections CDATA
    - <![CDATA[ du texte .... ]]>

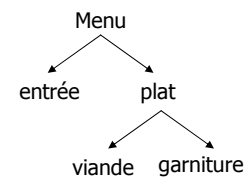
## Exemple Xml



```
<?xml version="1.0" ?>
<menu date="12-02-00" ?>
  <entrée> Saumon fumé</entrée>
  <plat>
    <viande> Canard rôti</viande>
    <garniture> pommes de Terres</garniture>
  </plat>
  <dessert> Profiteroles</dessert>
</menu>
```

Seul le contenu est décrit  
(pas de présentation)

La structure d'un document  
Xml est **un arbre**



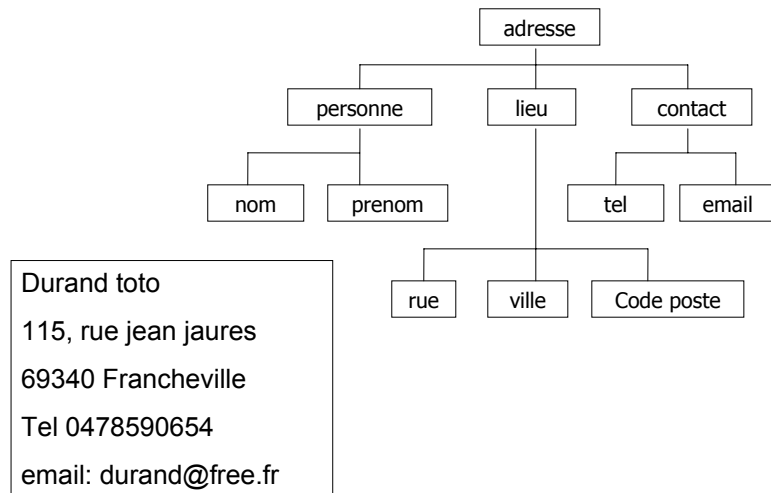
## Règles pour Xml

- ◆ Une seule racine pour un document Xml
- ◆ Les balises doivent être complètement imbriquées  
(pas de croisement)

```
<balise_1> data1 <balise_2> data2 </balise_1> </balise_2>
```

- ◆ Une balise de début et une balise de fin <espèce>..... </espèce>
- ◆ Une balise « vide » doit être délimitée par un slash avant le crochet fermant  
<vide/>
- ◆ ...

Pour définir la liste des éléments, des attributs,... ainsi que les règles qui les régissent, Xml utilise une DTD (Document Type Definition)

**Exemple de DTD (1)**

ED IIS - DEA ISCE

Intégration &amp; Travail Coopératif

7

**Exemple de DTD (2)**

```

<?xml version="1.0" >
<adresse>
  <personne>
    <nom>durand</nom>
    <prenom> toto </prenom>
  </personne>
  <lieu>
    <rue> 115, rue jean jaures </rue>
    <ville> Francheville </ville>
    <code postal> 69340 </code postal>
  </lieu>
  <contact>
    <tel> 0478590654 </tel>
    <email>durand@free.fr</email>
  </contact>
</adresse>
  
```

```

<?xml version="1.0" >
element adresse (personne,lieu,contact)

element personne (nom,prenom)
element nom PCDATA
element prenom PCDATA

element lieu (rue,ville,code postal)
element rue PCDATA
element ville PCDATA
element code postal PCDATA

element contact (tel,email)
element tel PCDATA
element email PCDATA
  
```

ED IIS - DEA ISCE

Intégration &amp; Travail Coopératif

8

## Feuille de style CSS

Cascading Style Sheet

- ◆ Liste de règles qui décrit les éléments
- ◆ Pour chaque élément

- Couleurs et image
- Polices de caractère
- Textes
- Boîtes
- ...

## Feuille de style XSL

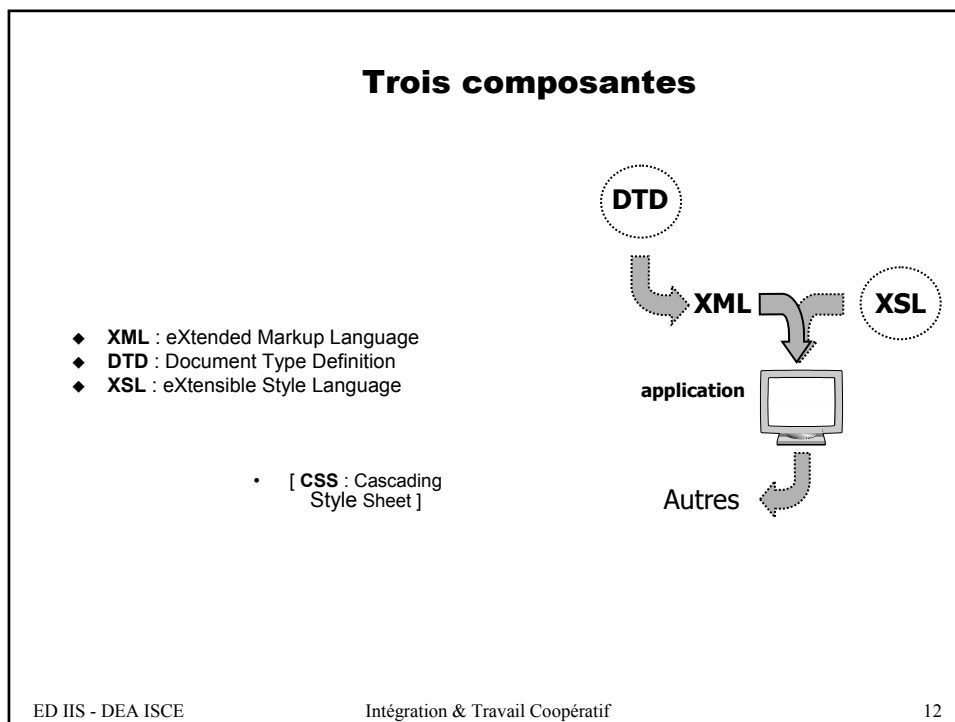
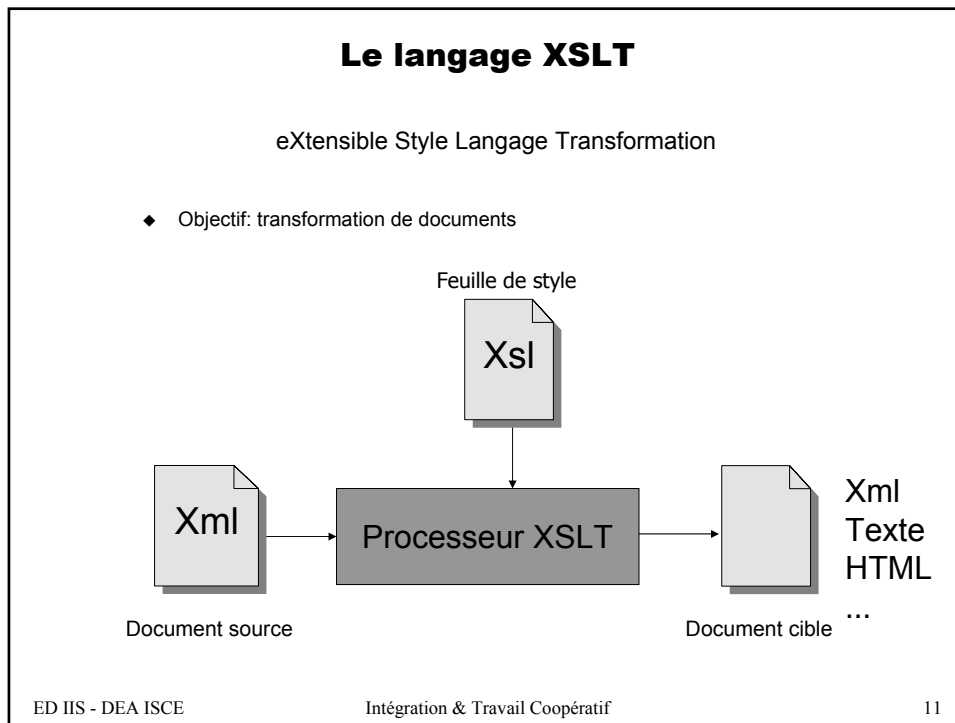
eXtensible Style Language

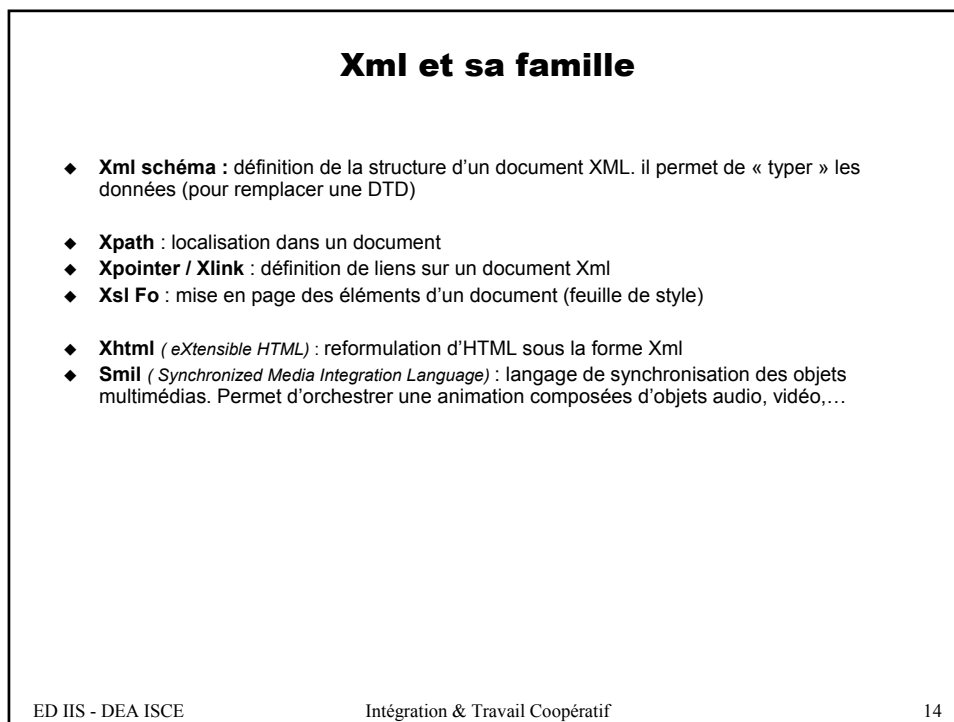
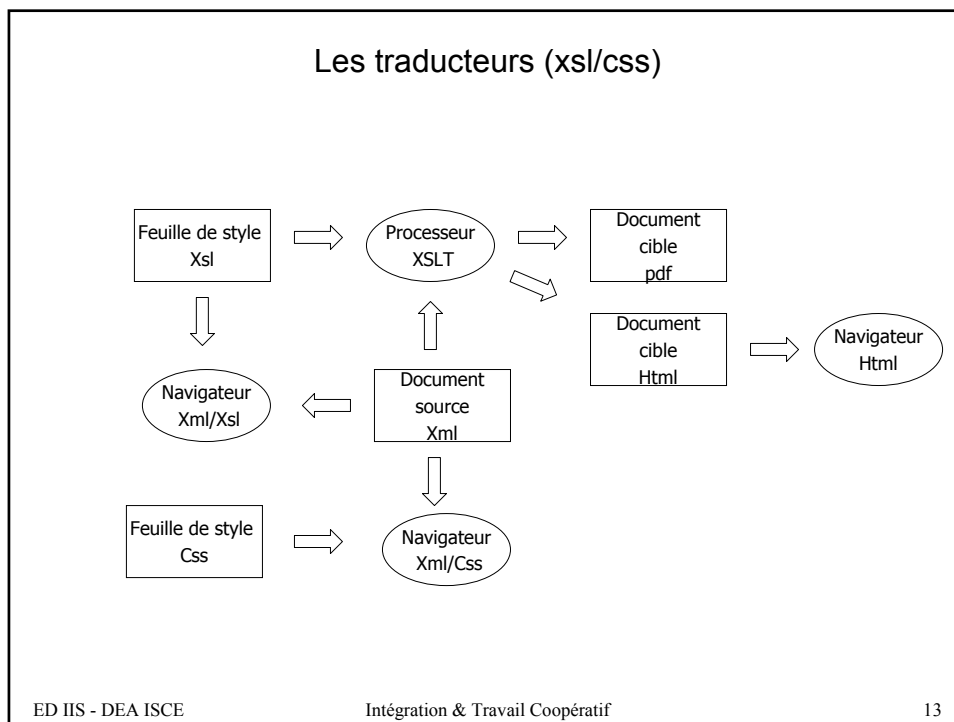
Feuille de style = ensemble de règles (plus puissant que CSS)

- ◆ Transformer un arbre
- ◆ Produire un résultat formaté (pour visualisation sur écran, impression,...)

### Objectif :

- Production de texte
- Personnalisation d'un document
- Réorganisation





## Les parseurs

- Parseur (parser)  
parseur= analyseur syntaxique

Deux types de parseur Xml

- SAX (*Simple Api for Xml*) génère des événements (beginDocument, startElement, endElement,..) en fonction des balises rencontrées (parseur événementiel).
- DOM (*Document Object Model*) produit une représentation du contenu sous la forme d'une arborescence.

## SAX

### Fonctionnement

- lit séquentiellement le document Xml en entrée
- reconnaît et interprète les balises
- transmet chaque construction identifiée à l'application à travers l'API SAX

### Utilisation

- permet la lecture et la transmission rapide des documents
- fournit à l'application les données Xml

## DOM

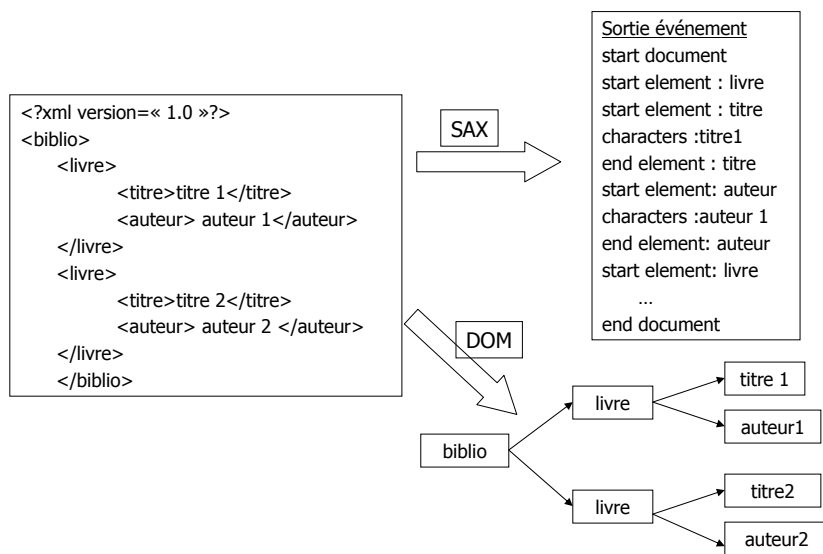
### Fonctionnement

Le document Xml est chargé en totalité sous la forme d'un arbre en mémoire

### Utilisation

DOM fournit des fonctions pour naviguer dans l'arbre constitué, modifier sa structure, atteindre, créer et supprimer des objets (protocole d'accès aléatoire).

## SAX / DOM des approches différentes



## Des outils pour Xml

**Editeur XML : XML Notepad ,XML Spy**

**Editeur XSLT: XL-Style**

**Parseur : Xerces, Japx, Jdom, ...**

**Processeur XSLT : Xalan, ...**

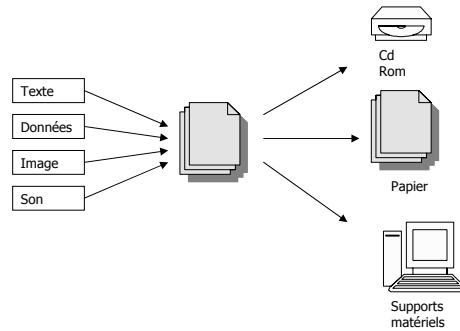
**Navigateur : IE5, Netscape, mozilla**

## Lexique

◆ <b>XML</b>	eXtensible Markup Language
◆ <b>W3C</b>	World Wide Web Consortium
◆ <b>DTD</b>	Définition de Type de Document
◆ <b>XSD</b>	XML Schema Document
◆ <b>XPATH</b>	XML Path
◆ <b>XSLT</b>	eXtensible Style Language Transformation
◆ <b>CSS</b>	Cascading Style Sheet
◆ <b>XSL-FO</b>	eXtensible Style Language Formatting Objects
◆ <b>XLINK</b>	XML Link
◆ <b>XPOINTER</b>	XML Pointer
◆ <b>API</b>	Application Programming Interface
◆ <b>DOM</b>	Document Object Model

## Transformation de documents dans les applications multimédia

Le document électronique véhicule composite et polymorphe de l'information

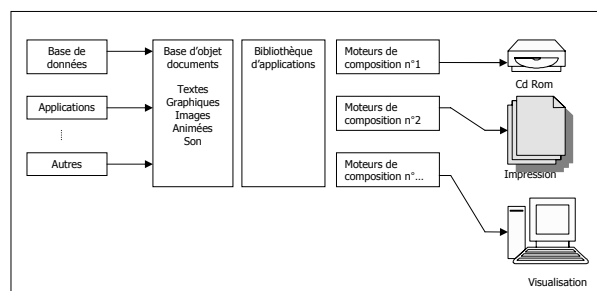


Document:

**Source :** ensemble d'informations (plus ou moins) structurées

**Destination:** présentation d'information multimédia ( organisation temporelle et spatial )

## Base documentaire unique et production à cibles multiples

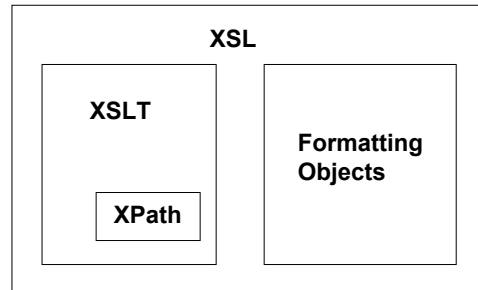


Besoins :

- Assurer l'indépendance des données vis-à-vis d'une machine, d'une application donnée
- Faciliter le passage de celles-ci d'un format à un autre
- Réutiliser le contenu, aborder l'édition multi-supports ( séparation du contenu et de la structure de présentation)
- Simplifier les traitements automatiques
- Disposer d'un outil simple et universel pour les échanges d'informations
- ...



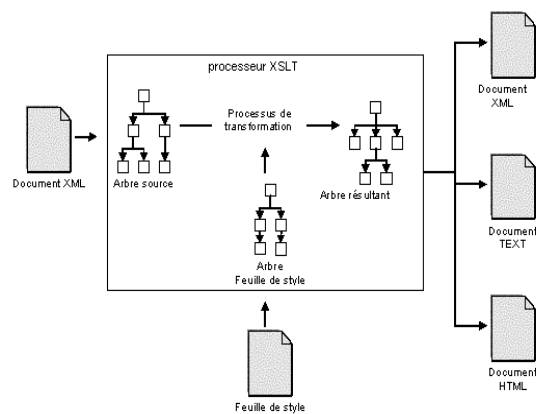
## Xsl (XML Stylesheet Language)



XSL (*XML stylesheet Language*) se décompose en deux parties :

- XSLT pour la transformation des documents XML
- XSL-FO (*XSL-Formating Object*) pour la description des informations de mise en page et générer des documents Acrobat (PDF), postscript ou RTF.

## Processus de transformation XSLT (XML Stylesheet Language Transformation)



Processus de transformation par lequel une feuille de style et un document XML sont fusionnés avec pour résultat un document formaté selon la feuille de style choisie

