Le langage XML Schema

http://www.w3.org/TR/xmlschema-0/http://www.w3.org/TR/xmlschema-1/http://www.w3.org/TR/xmlschema-2/

103

Introduction (1)



- Un schéma écrit en XML Schema définit une classe de documents
- Les XML Schémas améliore les DTD :
 - Un XML schéma est un fichier XML :
 - pas un nouveau langage, pas de parseur dédié.
 - Si on sait lire un fichier XML, on peut lire un schéma
 - Propose des types de données utiles (en plus des chaîne de caractères)
 - Booléen, entier, décimal, chaîne, date,...

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM – mthiam@univ-thies.sn

104

Introduction (2)



- Un XML schema est composé :
 - D'un prologue :
 - indique quels espaces de noms vont être utilisés
 - Espace de noms : pointeur vers un document (pré-) définissant un vocabulaire qui pourra être utilisé dans le document courant
 - D'un corps :
 - Liste d'assertions d'éléments, de types simples ou complexes,...

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

105 105

Schémas : Schema, prologue Syntaxe



- <?xml version="1.0" encoding="ISO-8859-1"?>
 - Première assertion d'un document XML
- <xs:schema

```
xmlns:xs="http://www.w3.org/2001/XMLSchema"
targetNamespace="http://www.lri.fr/~thiam/FCXML"
xmlns="http://www.lri.fr/~thiam/FCXML"
elementFormDefault="qualified" version="1.0">
```

- Définition de l'espace de nom xs qui permet d'utiliser le vocabulaire prédéfini pour XML schema
 - Le vocabulaire XML schema devra être préfixé par xs!

UFR SET - UT / Dep. Info

106

 $Pr\ Mouhamadou\ THIAM-mthiam@univ-thies.sn$

106 106

Schémas : Schema, prologue Détails



- targetNamespace indique à quel espace de noms appartiennent les éléments définis dans le schéma.
- xmlns="..." indique l'espace de noms par défaut
- elementFormDefault indique si les éléments doivent
 - être préfixés : qualified
 non préfixés : unqualified
 par leur espace de noms

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

107 107

Schémas : Schema, corps



- Le corps permet de définir des éléments :
 - <xs:element name="library" />
 - Un élément peut avoir plusieurs attributs optionnels :
 - Type="..." : pour définir des types simples ou complexes.
 - Des restrictions de cardinalités généralisant celles des DTD :
 - minOccurs="x" ou maxOccurs="x" avec $x \in N^* \cup \{unbounded\}$

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM – mthiam@univ-thies.sn

Type simple et type complexe



- Un **type complexe** est un type qui peut contenir des **déclarations** d'attributs et d'éléments.
- Un **type simple** ne peut pas contenir de déclarations d'attributs ou d'éléments.

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

109 109

110

TYPES SIMPLES

110

Définition

- Un type simple est défini par :
 - un ensemble de valeurs,
 - une représentation lexicale,
 - un ensemble de facettes qui caractérise l'ensemble de valeurs.
- Un type simple peut être :
 - un type atomique, un type liste ou un type union,
 - primitif ou dérivé,
 - prédéfini ou défini par l'utilisateur.

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

111 111

Type atomique, type liste ou type union



- Un type atomique est un ensemble de valeurs indécomposables dites valeurs atomiques.
- Un type liste est un ensemble de séquences de longueur finie de valeurs atomiques.
- Un type union est un ensemble de valeurs atomiques appartenant à plusieurs types atomiques.

UFR SET - UT / Dep. Info

 $Pr\ Mouhamadou\ THIAM-mthiam@univ-thies.sn$

Types prédéfinis ou définis par l'utilisateur

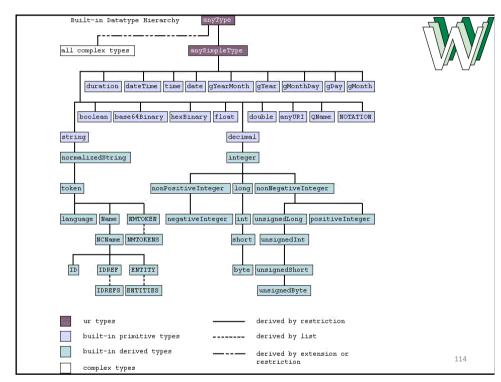


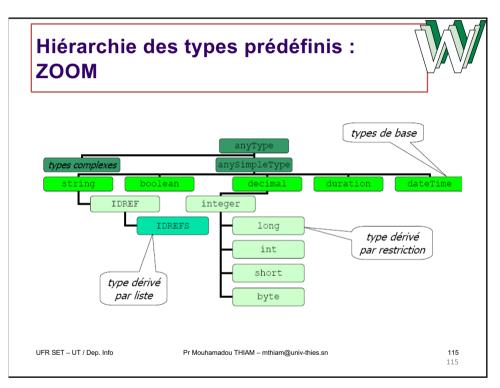
- Un type prédéfini peut être primitif ou dérivé.
- Un type peut être **défini** par dérivation de types existants.
- Un type défini peut être anonyme ou nommé.

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

113 113





Facettes des types de base :



_								
_		\neg		m		n		\sim
_					_			\sim
	VΙ	ıu	ч		$\mathbf{\circ}$		L	les

<u>Datatype</u>	ordered	bounded	cardinality	numeric
string			X	
boolean			Х	
float	Х	Х	Х	Х
double	Х	Х	Х	Х
decimal	х			х
duration	х		Х	
dateTime	Х		Х	

UFR SET - UT / Dep. Info

116

 $\label{eq:problem} \mbox{Pr Mouhamadou THIAM - mthiam@univ-thies.sn}$

116 116

Facettes des types de base : contraignantes ou non-fondamentales



	string	boolean	decimal	float	duration	dateTime
length	x				X	X
minLength	x					
maxLength	x					
pattern	x	x	x	x	X	X
enumeration	x		x	x	X	X
whiteSpace	X	х	x	x	X	X
minInclusive			x	x	X	x
minExclusive			x	x	X	x
maxInclusive			x	x	X	X
maxExclusives			X	X	X	X
totalDigits			X			
fractionDigits			X			

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

117 117

Types prédéfinis temporels



duration	1 an, 2 mois, 3 jours, 10 heures, 30 minutes: P1Y2M3DT10H30M
dateTime	13h20, le 31 mai 1999 (UTC): 1999-05-31T13:20
	13h20, le 31 mai 1999 (UTC + 5h): 1999-05-31T13:20:00-05:00
time	13h20 du jour courant (UTC) : 13:20
date	31 mai 1999 : 1999-05-31
gYear	1999: 1999
gYearMonth	mai 1999: 1999-05
gMonth	mois de mai de l'année courante :05
gMonthDay	31 mai de l'année courante :05-31
gDay	31 du mois courant :31

UFR SET - UT / Dep. Info

 $\label{lem:problem} \mbox{Pr Mouhamadou THIAM - mthiam@univ-thies.sn}$

118 118

Types de base : string, boolean, float et decimal



string	Ensemble des séquences de caractères de longueur finie	
boolean	Ensemble de valeurs = {true, false, 1, 0}	
float	Ensemble de valeurs = $m \times 2^e$ où $ m \le 2^{24}$ et -149 $\le e \le 104$ + zéro positif et négatif infinité positive et négative	-1E4 1267.43233E12 12.78e-2 12 0 -0 INF -INF
	pas un nombre	NaN
decimal	Nombre décimaux de précision arbitraire. Ensemble de valeurs = $i \times 10^n$ où i et n sont des entiers et $n \ge 0$	-1.23 12678967.543233 +100000.00 210

UFR SET – UT / Dep. Info Pr N

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

119

119

Types prédéfinis dérivés de decimal: un extrait



120

120

integer	Dérivé du type decimal par fractionDigits = 0	-1 0 12678967543233 +100000
long	Dérivé du type integer par minInclusive = -9223372036854775808 et maxInclusive = 9223372036854775807	-1 0 12678967543233 +100000
int	Dérivé du type integer par minInclusive = -2147483648 et maxInclusive = 2147483647	-1 0 126789675 +100000
short	Dérivé du type int par minInclusive = -32768 et maxInclusive = 32767	-1 0 12678 +10000
byte	Dérivé du type short par minInclusive = -128 et maxInclusive = 127	-1 0 126 +100

UFR SET – UT / Dep. Info Pr Mouhamadou THIAM – mthiam@univ-thies.sn

Types dérivés par restriction

- Intervalle
- Énumération
- motif

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

121

Intervalle : Exemple



• Ensemble des altitudes des points de la Terre (le plus haut sommet a une altitude de 8850m).

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM – mthiam@univ-thies.sn

122 122

121

Enumération : Exemple 1



• Liste des noms des pays africains :

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

123 123

Enumération : Exemple 2



• Liste des codes des pays :

UFR SET - UT / Dep. Info

 $Pr\ Mouhamadou\ THIAM-mthiam@univ-thies.sn$

124 124

Motif: Exemple 1



• Différents motifs de ISBN à 10 chiffres

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

125 125

Motif: Exemple 2



Différentes formes d'un nom XML :

UFR SET - UT / Dep. Info

126

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

126 126



- L'opérateur xsd:list définit un nouveau type simple
- valeurs sont les listes de valeurs du type simple donné par l'attribut itemType.
- pas de listes générales (# langages de programmation)
- Uniquement listes de valeurs séparées par des espaces
- Souvent utilisées comme valeurs d'attributs.
- Pas de caractères d'espacement dans type simple donné par itemType

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

127

Exemple 1



• Liste de pays :

```
<xsd:simpleType name="liste-pays">
     <xsd:list itemType="nom-pays-af"/>
</xsd:simpleType>
```

• Contenu conforme à ce type :

```
<pays-africains>
    Sénégal Gambie Mauritanie
</pays-africains>
```

UFR SET - UT / Dep. Info

128

 $Pr\ Mouhamadou\ THIAM-mthiam@univ-thies.sn$

Exemple 2



• Type pour liste d'entiers

```
<xsd:simpleType name="IntList">
     <xsd:list itemType="xsd:integer"/>
</xsd:simpleType>
```

• Type pour liste de 5 entiers

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

129

Types union



 L'opérateur xsd:union définit un nouveau type simple dont les valeurs sont celles des types listés dans l'attribut memberTypes.

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

130

Exemple 1



• Ensemble des noms ou des codes des pays :

• Eléments dont le contenu est conforme à ce type :

```
<pays>AL</pays>
<pays>Algerie</pays>
```

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

131 131

Exemple 2



• Définition du type Unbounded

• Définition du type IntegerOrUnbounded

UFR SET - UT / Dep. Info

132

 $Pr\ Mouhamadou\ THIAM-mthiam@univ-thies.sn$



TYPES COMPLEXES

133

133

Type complexe



 Un type complexe définit des compositions d'éléments et/ou d'attributs :

```
<xsd:complexType ...>
```

Suite de déclarations d'attribut, d'élément ou de composition d'éléments

```
</xsd:complexType>
```

- Une composition d'éléments est réalisée à l'aide de 3 constructeurs:
 - sequence : séquence d'éléments typés,
 - choice : un élément parmi une liste d'éléments possibles,
 - all : tas d'éléments typés.

UFR SET - UT / Dep. Info

134

Pr Mouhamadou THIAM – mthiam@univ-thies.sn

Déclaration d'un élément



- Un élément dans la définition d'un type complexe est spécifié sous la la forme d'un élément xsd:element dont les attributs peuvent être
 - name : nom de l'élément,
 - type : type de l'élément,
 - ref: référence à un nom d'élément déclaré au niveau global (réutilisation),
 - minOccurs : nombre minimum d'occurrences de l'élément (1 par défaut),
 - maxOccurs: nombre maximum d'occurrences de l'élément (1 par défaut),
 - ..
 - le contenu est la définition d'un type.

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

135 135

Déclaration d'un attribut



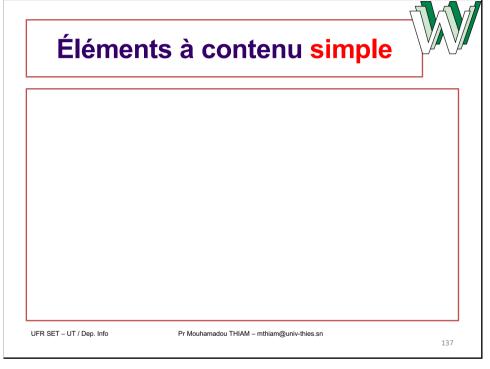
- Un attribut dans la définition d'un type complexe est spécifié sous la forme d'un élément **xsd:attribute** dont les attributs peuvent être :
 - name : nom de l'attribut,
 - ref: référence à un nom d'attribut déclaré au niveau global (réutilisation).
 - type: type des valeurs de l'attribut,
 - use : required, optional (par défaut), ...
 - default : valeur par défaut,
 - fixed : valeur fixée,
 - ٠...
 - le contenu peut être : un type simple.

UFR SET - UT / Dep. Info

136

 $Pr\ Mouhamadou\ THIAM-mthiam@univ-thies.sn$

136 136



Elément à contenu simple : exemple L'élément : <pays>Gambie</pays> peut être déclaré : <xsd:element name="pays" type="mns:nom-pays-af"/> UFR SET - UT / Dep. Info Pr Mouhamadou THIAM - mthiam@univ-thies.sn 138 138

Séquence d'éléments : type anonyme

UFR SET - UT / Dep. Info



139

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

Eléments et attributs : type

nommé

UFR SET - UT / Dep. Info

140

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

140

Éléments à contenu mixte



- pur lorsqu'il ne contient que des éléments qui, eux-mêmes, peuvent à leur tour contenir du texte et/ou des éléments.
- mixte lorsqu'il contient du texte autre que des caractères d'espacement en dehors de ses enfants.
- attribut *mixed* de l'élément xsd:complexType contient la valeur true.

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

141

142

Exemple

Elément à contenu vide



• L'élément :

```
<prix monnaie="euros" valeur="25.5"/>
```

• peut être déclaré :

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

143 143

Type nommé



• On peut nommer le type définissant un prix en euros ou en dollars :

```
<xsd:complexType name="prix-international">
  <xsd:attribute name="monnaie" type="xsd:string"/>
  <xsd:attribute name="valeur" type="xsd:decimal"/>
</xsd:complexType>
```

• et y faire référence dans la déclaration de l'élément prix :

```
<xsd:element name="prix" type="prix-international"/>
```

UFR SET - UT / Dep. Info

144

 $Pr\ Mouhamadou\ THIAM-mthiam@univ-thies.sn$

144

Elément à contenu alternatif



```
• Les éléments :
```

```
<prix><dollars>28.75</dollars></prix>
```

• peuvent être déclarés :

```
<xsd:element name="prix">
  <xsd:complexType>
    <xsd:choice>
     <xsd:element name="euros" type="xsd:decimal"/>
     <xsd:element name="dollars" type="xsd:decimal"/>
    </xsd:choice>
  </xsd:complexType>
</xsd:element>
```

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

145

Opérateur d'ensemble : Elément contenu non ordonné



- Les éléments doivent apparaître une fois dans un ordre quelconque
- Pas d'équivalent en DTD

```
<xsd:element name="book">
  <xsd:complexType>
      <xsd:all>
                                           type="xsd:string"/>
          <xsd:element name="title"</pre>
          <xsd:element name="author"</pre>
                                           type="xsd:string"/>
          <xsd:element name="year"</pre>
                                           type="xsd:string"/>
          <xsd:element name="publisher"</pre>
                                          type="xsd:string"/>
      </xsd:all>
  </xsd:complexType>
</xsd:element>
```

UFR SET - UT / Dep. Info

146

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

146

Exemple



```
<?xml version="1.0" encoding="iso-8859-1"?>
<br/>
<bibliography>
  <book key="Michard01" lang="fr">
    <author>Alain Michard</author>
    <title>XML langage et applications</title>
    <publisher>Eyrolles</publisher>
    <year>2001
  </book>
  <book key="Zeldman03" lang="en">
    <title>Designing with web standards</title>
    <author>Jeffrey Zeldman</author>
    <year>2003
    <publisher>New Riders
  </book>
  . . .
</bibliography>
UFR SET - UT / Dep. Info
                     Pr Mouhamadou THIAM - mthiam@univ-thies.sn
                                                            147
```

Groupe d'éléments



- peut être déclarée :

• (On peut aussi grouper des attributs)

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

148 148

Utilisation des groupes d'éléments



• Pour déclarer des éléments décrivant des personnes décrites par leur nom et leurs prénoms, on peut définir le type suivant :

```
<xsd:complexType name="personne">
     <xsd:group ref="prenom-nom"/>
</xsd:complexType>
```

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

149 149

Dérivation d'un type complexe par extension

• A partir du type personne, on peut définir le type chercheur en ajoutant au prénom et au nom, le laboratoire et le pays :

UFR SET – UT / Dep. Info

150

 $Pr\ Mouhamadou\ THIAM-mthiam@univ-thies.sn$

150 150

Schéma XML



 Un schéma XML est une suite de définitions de types et de déclarations d'attributs.

 $\mathsf{UFR}\;\mathsf{SET}-\mathsf{UT}\,\mathsf{/}\;\mathsf{Dep.}\;\mathsf{Info}$

151

 $\label{eq:continuous} Pr\ Mouhamadou\ THIAM-mthiam@univ-thies.sn$

151 151

152

Exemple du document ...



```
livres> ...
  <livre edition="1">
   <titre>Le langage XML</titre>
      om>Jean</prenom><nom>Dupont</nom>
      <laboratoire>LRI</laboratoire><pays>France</pays>
   </auteur>
   <auteur>
      om>Pierre</prenom><nom>Durand</nom>
      <laboratoire>LRI</laboratoire><pays>France</pays>
   </auteur>
   <année>2004</année>
   <prix monnaie="euros" valeur="25.5"/>
   </livres>
UFR SET - UT / Dep. Info
                       Pr Mouhamadou THIAM - mthiam@univ-thies.sn
                                                                   152
                                                                  152
```

DTD



```
<!ELEMENT livres (livre*)>
<!ELEMENT livre (titre, auteur+, année, prix)>
<!ATTLIST livre edition CDATA #REQUIRED>
<!ELEMENT titre (#PCDATA)>
<!ELEMENT auteur (prenom, nom, laboratoire, pays)>
<!ELEMENT nom (#PCDATA)>
<!ELEMENT prenom (#PCDATA)>
<!ELEMENT laboratoire (#PCDATA)>
<!ELEMENT année (#PCDATA)>
<!ELEMENT prix EMPTY>
<!ATTLIST prix monnaie CDATA #REQUIRED valeur CDATA #REQUIRED>
<!ELEMENT pays (#PCDATA)>
```

UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

153 153

Schéma XML : exemple



UFR SET - UT / Dep. Info

154

 $Pr\ Mouhamadou\ THIAM-mthiam@univ-thies.sn$

154 154

Schéma XML : exemple

UFR SET - UT / Dep. Info



155

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

Schéma XML : exemple



```
<xsd:complexType name="chercheur">
   <xsd:complexContent>
     <xsd:extension base="personne">
      <xsd:sequence>
       <xsd:element name="laboratoire" type="xsd:string"/>
       <xsd:element ref="pays">
      </xsd:sequence>
    </xsd:extension>
   </complexContent>
</xsd:complexType>
<xsd:complexType name="prix-international">
   <xsd:attribute name="monnaie" type="xsd:string"/>
   <xsd:attribute name="valeur" type="xsd:decimal"/>
</r></xsd:complexType> ...
UFR SET - UT / Dep. Info
                         Pr Mouhamadou THIAM - mthiam@univ-thies.sn
                                                                      156
```

Schéma XML : exemple



UFR SET - UT / Dep. Info

Pr Mouhamadou THIAM - mthiam@univ-thies.sn

157 157

Schéma XML : exemple



UFR SET - UT / Dep. Info

 $Pr\ Mouhamadou\ THIAM-mthiam@univ-thies.sn$

158 158

