

# XML – DTD

Michel Beigbeder

25 septembre 2018

# XML, eXtensible Markup Language

Références rapides: <http://www.mulberrytech.com/quickref/>

Un cours plus complet, bien écrit, en français:

<http://www.teluq.ca/inf6450/index-fr.htm>

- ▶ XML, DTD
  - ▶ Semaine 1: Introduction, création de documents XML, principes
  - ▶ Semaine 2: Définition de type de document (DTD), Validation, Relax NG, XML Schema
  - ▶ Semaine 3: Espaces de noms
- ▶ XML pour les documents
  - ▶ Semaine 4: Introduction au XHTML, HTML 5 et SVG
  - ▶ Semaine 5: XML comme format de documents (SGML, DocBook, ODF)
  - ▶ Semaine 6: EPUB
- ▶ Technologies XML
  - ▶ Semaine 7: XPath
  - ▶ Semaine 8–9: XSLT
  - ▶ Semaine 10: XQuery, CSS
- ▶ Semaine 11–12: Traitement du XML en Java: SAX, DOM

- ▶ Version 1.0 de février 1998
- ▶ Qualification et description de données (textuelles)
  - ▶ Cf. exemples 0 à 6
- ▶ Balise (ouvrante, fermante, autofermée), élément
- ▶ `[-\.:[:ALPHA:]] [-\.:[:ALPHA:][:DIGIT:]]*`
- ▶ Cf. exemple 7
- ▶ Attribut
- ▶ Cf. exemple 8

# Structure d'un fichier XML

- ▶ Prologue, arbre (arborescence) non vide d'éléments et aussi commentaires (`<!-- -->`) et instructions de traitement (`<!`  
`>`)

Cf. exemples 9 à 12

- ▶ Jeu de caractères et encodage

Cf. exemples 13 à 15

Les caractères `<` `>` `"` `'` `&` lorsqu'ils sont dans les données doivent être introduits par leur entité `&lt;` `&gt;` `&quot;` `&apos;` `&amp;`;

# Imposer une structure plus forte: DTD

- ▶ DTD: *Document Type Definition*
- ▶ Tous les éléments doivent être décrits en terme de contenu et d'attributs
- ▶ Instruction de description des éléments
  - ▶ `<!ELEMENT nom_d_element contenu>`
  - ▶ *contenu* est soit EMPTY, soit ANY, soit une disjonction (opérateur |) ou une séquence (opérateur ,) d'éléments quantifiés (quantificateurs: ?+\*)

Cf. exemples 23, 25

- ▶ Instruction de description des éléments
  - ▶ `<!ATTLIST nom_d_element nom_d_attribut type obligation [valeur_par_défaut] >`
  - ▶ *type* est soit CDATA, ID, IDREF, IDREFS, une énumération (e.g. (rouge|noir) avec une valeur par défaut)
  - ▶ *obligation* est soit #REQUIRED, #IMPLIED, #FIXED

Cf. exemples 26a, 26b