

TP 1 – XML et DTD : éléments de base

Objectif : introduire le langage XML, créer un document XML bien formé, créer un document XML valide, définir une description DTD d'une structure d'un document XML .

1. Créer un document XML bien formé

- 1.1. Dans un fichier de texte nommé biblio.xml saisir le document XML ci-dessous. Ce document décrit plusieurs types de livres.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<biblio>
  <livre>
    <!-- Élément enfant titre -->
    <titre>Les Misérables</titre>
    <auteur>Victor Hugo</auteur>
    <nb_tomes>3</nb_tomes>
  </livre>
  <livre>
    <titre>L'Assomoir</titre>
    <auteur>Émile Zola</auteur>
  </livre>
  <livre lang="en">
    <titre>David Copperfield</titre>
    <auteur>Charles Dickens</auteur>
    <nb_tomes>3</nb_tomes>
  </livre>
</biblio>
```

- 1.2. Visualisez votre document XML dans le browser

2. Créer un document XML valide

- 2.1. Dans un nouveau fichier de texte nommé biblio.dtd, définissez une DTD pour le document biblio.xml.

- un élément livre doit être composé des trois éléments dans l'ordre : titre, auteur et nb_tomes de manière optionnelle.
- l'attribut « lang » de l'élément livre ne prend que les valeurs « en » ou « fr » par défaut.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!ELEMENT biblio (livre)+>

<!ELEMENT livre (titre, auteur, nb_tomes ?)>

<!ELEMENT titre (#PCDATA)>
<!ELEMENT auteur (#PCDATA)>
<!ELEMENT nb_tomes (#PCDATA)>
<!ATTLIST livre lang (en|fr) "fr">
```

- 2.2. Intégrez l'appel à la DTD dans le prologue de biblio.xml. Tester.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE biblio SYSTEM "biblio.dtd">
```

- 2.3. Vérifiez la validité de votre document XML. Comment le faire ?
- 2.4. Pour s'assurer que le parseur effectue la vérification, testez les cas d'erreurs suivants :
 - valeur d'attribut lang erroné, différent de fr ou en
 - un élément titre ou auteur manquant, ou pas dans le bon ordre, et nb_tomes plus d'une fois.

3. Afficher un document XML

Proposez la façon d'affichage d'un document XML. CSS:

```
livre {
  display: block; margin-top: 18pt; margin-left: 20%; margin-right: 20%;
  padding: 8px; background-color: #ffffcc; border: 1px solid red; text-align: center
}
titre {
  display: block; padding: 4px;
  font-family: "sans-serif"; font-size: 14pt; font-weight: bold; color: maroon
}
auteur { display: block; padding: 4px; font-size: 8pt }
nb_tomes { display: block; padding: 4px; font-size: 8pt; font-style: italic }
```

4. Questions.

- 4.1. Si je dispose d'une grande quantité de code HTML. Faut-il le convertir en XML ? Comment ?
- 4.2. Comment les navigateurs Web interprètent la signification des balises du code XML ?
- 4.3. Est-ce qu'il est possible de placer des données binaires dans une section CDATA ?
- 4.4. Comment savoir si un document XML est bien formé ? valide ?
- 4.5. Comment savoir combien d'éléments doit contenir une DTD ?

5. Construire une DTD et XML correspondant en réponse à un problème

Vous désirez de représenter les documents réglementaires de construction. Les documents réglementaires se différencient par l'actualité (nouveaux et modifiés). Ils sont regroupés dans la base des documents par la thématique (e.g. accessibilité des handicapés, acoustique, aération, amiante, ascenseurs, assurance de construction, sécurité incendie). Chaque document possède d'un titre complexe qui se compose d'un numéro, d'une date, d'un nom, parfois d'un sigle et est relatif à un ou plusieurs domaines spécifiques. La date est requise ; elle comprend la date de la création d'un document et la date de sa publication dans un Journal Officiel.

Comment vous pouvez décrire la structure de ces documents (dtd + document de test en xml correspondant) ?

(Version + développée)

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!ELEMENT base (document)+>

<!ELEMENT document (thematique*,actualite, titre)>
<!ELEMENT thematique (#PCDATA)>
<!ATTLIST thematique typeThematique ("accessibilité des handicapés" | "acoustique" | "aération" |
"amiante" | "ascenseurs" | "assurance construction" | "sécurité incendie") #REQUIRED>
<!ELEMENT actualite (#PCDATA)>
<!ATTLIST actualite typeActualite (modifie|nouveau) "nouveau" >
<!ELEMENT titre (sigle?, numero, date, title,relatif)>
<!ELEMENT sigle (#PCDATA)>
<!ELEMENT numero (#PCDATA)>
<!ELEMENT date (#PCDATA)>
<!ATTLIST date typeDate (dateCreation|dateCreation) #REQUIRED>
<!ELEMENT nom (#PCDATA)>
<!ELEMENT relatifa (#PCDATA)>
```