

Feuille de travaux pratiques n° 1

XSLT, SAX & DOM

Exercice 1 Mondial

La base de donnée MONDIAL est une base de donnée proposée par l'Université de Göttingen¹. Cette base a été construite à partir de données du Web telles que :

- CIA World Factbook² ;
- l'«International Atlas» de Kümmerly Frey, Rand McNally et Westermann ;
- des données de la base de données géographique TERRA de Karlsruhe (Allemagne).

Cette base existe en XML. Elle est disponible (DTD, XML et d'autres documents complémentaires) dans Madoc, ainsi que des compléments sur les outils qui n'ont pas été abordés en cours.

On cherche à écrire un programme permettant d'obtenir un document XML d'exemple pour le contexte sur les espaces maritimes vu en cours. Le document à générer doit respecter la DTD «em.dtd» (un peu adaptée) suivante :

```
1 <!-- ajout des id partout... des noms peuvent contenir des espaces -->
2 <!ELEMENT em (liste-pays, liste-espace-maritime)>
3
4 <!ELEMENT liste-pays (pays)*>
5 <!ELEMENT liste-espace-maritime (espace-maritime)*>
6
7 <!ELEMENT pays (fleuve)*>
8 <!-- ATTLIST pays
9 id-p ID #REQUIRED nom-p CDATA #REQUIRED
10 superficie CDATA #REQUIRED nbhab CDATA #REQUIRED
11
12 <!-- ELEMENT fleuve (parcourt)+>
13 <!-- ATTLIST fleuve
14 id-f ID #REQUIRED nom-f CDATA #REQUIRED
15 longueur CDATA #REQUIRED se-jette IDREF #REQUIRED
16
17 <!-- ELEMENT parcourt EMPTY
18 <!-- ATTLIST parcourt
19 distance CDATA #REQUIRED id-pays IDREF #REQUIRED
20
21 <!-- ELEMENT espace-maritime (cotoie)+>
22 <!-- ATTLIST espace-maritime
23 id-e ID #REQUIRED nom-e CDATA #REQUIRED
24 type (océan|mer|inconnu) 'inconnu' <!-- type #REQUIRED -> 'inconnu' + ajout de 'inconnu' --
25 >
26 <!-- ELEMENT cotoie EMPTY
27 <!-- ATTLIST cotoie id-p IDREF #REQUIRED
```

Question 1

Donner le programme permettant d'obtenir le résultat précédent en utilisant XSLT, dans un premier temps sans utiliser «xsl:for-each» ou des règles nommées. (éventuellement, une seconde version plus procédurale peut-être proposée).

Question 2

Donner les programmes PHP permettant d'obtenir le résultat précédent en utilisant les outils suivants : SAX³, DOM sans utiliser XPath (avec au moins une version sans utiliser "getElementsByTagName()"), DOM en utilisant XPath (sans utiliser «//»). Pour les versions DOM et la version SAX, le document produit devra être généré d'abord en DOM avant d'être retourné. Pour chaque version, faites en sorte que le traitement soit optimal en temps de traitement.

1. <http://dbis.informatik.uni-goettingen.de/Mondial/>

2. <http://www.odci.gov/cia/publications/factbook/>

3. Vous pouvez utiliser la bibliothèque Sax4PHP : <http://sourceforge.net/projects/sax4php>

Question 3

Discuter des différentes versions proposées (XSLT, DOM, DOM/XPath, SAX) en termes de performance, d'adéquation au traitement, de clarté du code, de maintenance, etc.

Question 4 Facultative

Reprendre cette recherche en utilisant :

1. Java et JDOM. Pour cela, vous pourrez vous appuyer sur le site Web de JDOM. Deux versions sont attendues : avec et sans filtres. Là encore, le résultat sera construit en JDOM.
2. les objets PHP : XMLReader pour le traitement et XMLWriter pour la génération. Comparez avec les solutions précédentes.
3. l'API SimpleXML de PHP.

Travail à rendre :

Vous devrez rendre votre travail à l'enseignant sous forme électronique (archive de nom "noms-binome.tar.gz"), vous donnerez :

- votre code (bien présenté et commenté) ;
- un fichier "readme" et un fichier shell permettant à l'enseignant d'exécuter votre code sur sa machine ;
- un petit document PDF de 5 à 7 pages présentant votre travail (analyses, algorithmes, conclusions), les subtilités de vos traitements et vos difficultés.

Une zone de dépôt sur Madoc sera disponible à cet effet. Une séance de tests pourra être organisée.

Quelques références :

- Standard XPath : <http://xmlfr.org/w3c/TR/xpath>
- Standard DOM : <http://www.w3.org/DOM/>
- Documentation XMLLINT : <http://xmlsoft.org/xmllint.html> ;
- Documentation XSLTProc : <http://xmlsoft.org/XSLT/xsltproc.html>
- Documentation PHP DOM : <http://www.php.net/manual/fr/ref.dom.php>
- Documentation PHP «SAX» : <http://www.php.net/manual/fr/ref.xml.php>
- Documentation PHP SimpleXML : <http://php.net/manual/fr/book.simplexml.php>
- Documentation XMLReader/XMLWriter : <http://fr2.php.net/manual/fr/book.xmlreader.php>
& <http://fr2.php.net/manual/fr/book.xmlwriter.php>
- Site JDOM : <http://www.jdom.org/downloads/docs.html>

RAPPEL :

- Il est aussi possible d'exécuter du PHP dans une fenêtre «terminal» en faisant : `«php mon_exo.php»`.
- Il est possible d'exécuter le code XSLT en ligne de commande en utilisant Saxon⁴, par exemple : `«java -jar saxon9he.jar -xsl :mon_exo.xsl -s :mondial.xml»`. L'ajout du paramètre `"-TP :profile.html"` permettra d'obtenir un profilage de l'exécution en HTML. XSLT peut aussi être exécuté par PHP.
- Éventuellement, pour évaluer le temps de traitement, il faut faire précéder la ligne par «time», par exemple : `«time php mon_exo.php»` ou `«time java MaClasse»`. Le temps de traitement effectif de votre programme est la ligne «user».

4. Documentation : <http://www.saxonica.com/documentation/> . Téléchargement : <https://sourceforge.net/projects/saxon/files/>.