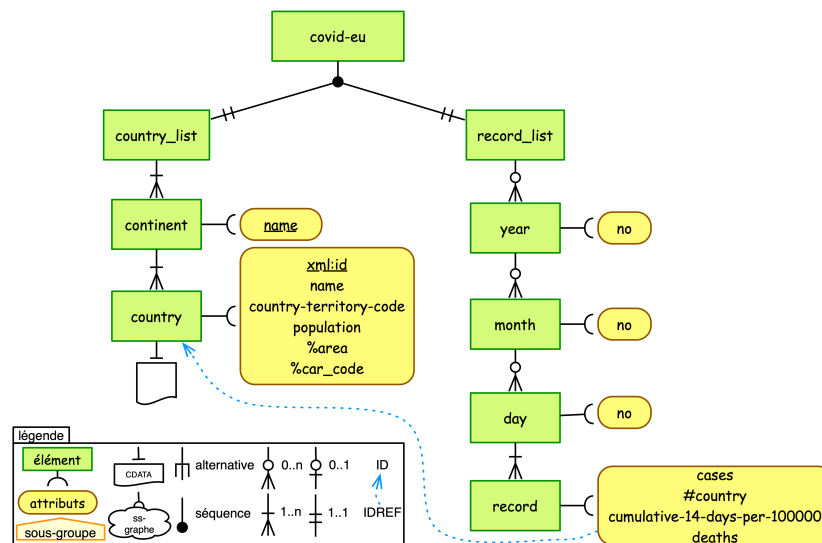


Feuille de travaux pratiques n° 1

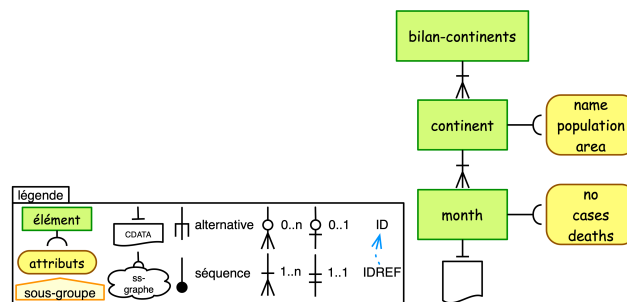
XSLT, SAX & DOM

Exercice 1 COVID-19 dans le monde

Le thème du TP est issu de l'Open Data de l'union européenne¹. On s'intéresse plus particulièrement au jeu de données relatives au coronavirus COVID-19². Ce jeu comprend les données relatives aux cas de covid et aux décès dans 214 pays du monde, chaque jour depuis décembre 2019. Ces données sont disponibles en format XML. Le document d'origine est assez mal fait, aussi je me suis permis de le transformer pour en faire une ressource un peu mieux structurée³. Le document "covid-tp.xml" correspond aux données en date du 11 novembre 2020. Vous pouvez régénérer le document sur des données plus récente. Le résultat respecte la DTD "covid.dtd".



On cherche à écrire un programme permettant d'obtenir la situation mensuelle pour chaque continent (triée par ordre chronologique). Le document à générer doit respecter la DTD «info.dtd» suivante :



Sur le document "covid-tp.xml", cela donnera le document "info.xml" présent sur Madoc.

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <!DOCTYPE bilan-continent
3   SYSTEM "info.dtd">
4 <bilan-continent>
5   <continent name="Asia" population="4.542059903E9" area="3.08063076E7">
6     <month no="2019-12" cases="27" deaths="0"/>
7     <month no="2020-01" cases="9766" deaths="213"/>
8     <month no="2020-02" cases="73470" deaths="2679"/>
9     <month no="2020-03" cases="85871" deaths="3990"/>
10    <month no="2020-04" cases="333821" deaths="11321"/>
11    <month no="2020-05" cases="592831" deaths="11862"/>
  </continent>
</bilan-continent>
```

1. <https://data.europa.eu/euodp/fr/home?>

2. <https://data.europa.eu/euodp/fr/data/dataset/covid-19-coronavirus-data>

3. <https://github.com/edesmontils/Covid-eu-enrichment>

```
12     <month no="2020-06" cases="1.119998E6" deaths="25129"/>
13     <month no="2020-07" cases="1.929728E6" deaths="39033"/>
14     <month no="2020-08" cases="2.799266E6" deaths="45643"/>
15     <month no="2020-09" cases="3.512145E6" deaths="51806"/>
16     <month no="2020-10" cases="3.004369E6" deaths="47999"/>
17     <month no="2020-11" cases="954814" deaths="16183"/>
18 </continent>
19 <continent name="Europe" population="7.66212338E8" area="2.316130284E7">
20     <month no="2019-12" cases="0" deaths="0"/>
21     <month no="2020-01" cases="15" deaths="0"/>
22     <month no="2020-02" cases="1139" deaths="23"/>
23     <month no="2020-03" cases="437423" deaths="27943"/>
24     <month no="2020-04" cases="857419" deaths="105434"/>
25     <month no="2020-05" cases="623098" deaths="38863"/>
26     <month no="2020-06" cases="480790" deaths="15973"/>
27     <month no="2020-07" cases="486132" deaths="10732"/>
28     <month no="2020-08" cases="772080" deaths="10085"/>
29     <month no="2020-09" cases="1.429338E6" deaths="14654"/>
30     <month no="2020-10" cases="4.812004E6" deaths="42505"/>
31     <month no="2020-11" cases="2.975928E6" deaths="39082"/>
32 </continent>
33 <continent name="Africa" population="1.30690303E9" area="3.0334283E7">
34     <month no="2019-12" cases="0" deaths="0"/>
35     <month no="2020-01" cases="0" deaths="0"/>
36     <month no="2020-02" cases="3" deaths="0"/>
37     <month no="2020-03" cases="5134" deaths="166"/>
38     <month no="2020-04" cases="31598" deaths="1425"/>
39     <month no="2020-05" cases="104887" deaths="2480"/>
40     <month no="2020-06" cases="251822" deaths="5807"/>
41     <month no="2020-07" cases="516291" deaths="9433"/>
42     <month no="2020-08" cases="336450" deaths="10274"/>
43     <month no="2020-09" cases="228136" deaths="6088"/>
44     <month no="2020-10" cases="302274" deaths="7015"/>
45     <month no="2020-11" cases="128655" deaths="3259"/>
46 </continent>
47 <continent name="America" population="1.013601796E9" area="4.19801063E7">
48     <month no="2019-12" cases="0" deaths="0"/>
49     <month no="2020-01" cases="11" deaths="0"/>
50     <month no="2020-02" cases="86" deaths="0"/>
51     <month no="2020-03" cases="190000" deaths="3687"/>
52     <month no="2020-04" cases="1.113933E6" deaths="70904"/>
53     <month no="2020-05" cases="1.570678E6" deaths="86923"/>
54     <month no="2020-06" cases="2.396208E6" deaths="87158"/>
55     <month no="2020-07" cases="4.126614E6" deaths="106910"/>
56     <month no="2020-08" cases="4.044723E6" deaths="112185"/>
57     <month no="2020-09" cases="3.282705E6" deaths="89303"/>
58     <month no="2020-10" cases="3.821622E6" deaths="83443"/>
59     <month no="2020-11" cases="1.762451E6" deaths="25609"/>
60 </continent>
61 <continent name="Oceania" population="4.146738E7" area="8.5085783E6">
62     <month no="2019-12" cases="0" deaths="0"/>
63     <month no="2020-01" cases="7" deaths="0"/>
64     <month no="2020-02" cases="19" deaths="0"/>
65     <month no="2020-03" cases="5298" deaths="21"/>
66     <month no="2020-04" cases="2812" deaths="95"/>
67     <month no="2020-05" cases="503" deaths="15"/>
68     <month no="2020-06" cases="705" deaths="2"/>
69     <month no="2020-07" cases="8750" deaths="87"/>
70     <month no="2020-08" cases="11296" deaths="432"/>
71     <month no="2020-09" cases="3925" deaths="320"/>
72     <month no="2020-10" cases="8564" deaths="79"/>
73     <month no="2020-11" cases="4547" deaths="32"/>
74 </continent>
75 <continent name="Other" population="0" area="0">
76     <month no="2019-12" cases="0" deaths="0"/>
77     <month no="2020-01" cases="0" deaths="0"/>
78     <month no="2020-02" cases="705" deaths="6"/>
79     <month no="2020-03" cases="-9" deaths="1"/>
80 </continent>
81 </bilan-continents>
```

Question 1

Donner le programme permettant d'obtenir le résultat précédent en utilisant XSLT, sans utiliser «xsl :for-each» ou des règles nommées.

Question 2

Donner les programmes PHP permettant d'obtenir le résultat précédent en utilisant les outils suivants : SAX⁴, DOM sans utiliser XPath (avec au moins une version sans utiliser "getElementsByTagName()"), DOM en utilisant XPath (sans utiliser « // »). Pour chaque version, faites en sorte que le traitement soit optimal en temps de traitement.

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 d'une part le rapport (un petit document PDF de 5 à 7 pages présentant votre travail, analyses, algorithmes, conclusions, etc., les subtilités de vos traitements et vos difficultés) et d'autre part une archive de nom "noms-binome.tar.gz" où 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.

Deux zones de dépôt sur Madoc seront disponibles 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 Saxon⁵, par exemple : «java -jar saxon9he.jar -xsl :mon_exo.xsl -s :ma_source.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».
- Pour faire des tests de performance (temps d'exécution), il est possible d'utiliser l'outil "hyperfine"⁶.
- Pour comparer deux documents XML, vous pouvez utiliser XMLDiff⁷.

4. Vous pouvez utiliser la bibliothèque Sax4PHP : <https://github.com/edesmontils/Sax4PHP>

5. Disponible sur Madoc. Documentation : <http://www.saxonica.com/documentation/>

6. <https://github.com/sharkdp/hyperfine>

7. <https://github.com/Shoobx/xmldiff>