

# TD3 : modélisation et manipulation XML

UCBL - Base de données pour la bioinformatique - 2023 / 2024

Objectif du TD : étudier une DTD et écrire des requêtes XPath/XQuery

Ce jeu de données décrit des activités dans des stations de sports d'hiver.

```
1 <!DOCTYPE inscriptions [  
2 <!ELEMENT inscriptions (activites, touristes)>  
3  
4 <!ELEMENT activites (activite?)>  
5 <!ELEMENT activite (nom)>  
6 <!ELEMENT nom (#PCDATA)>  
7 <!ATTLIST activite code ID>  
8  
9 <!ELEMENT touristes (touriste*)>  
10 <!ELEMENT touriste (nom, adresse, inscription*)>  
11 <!ELEMENT nom (#PCDATA)>  
12 <!ELEMENT adresse (#PCDATA)>  
13 <!ELEMENT inscription (#PCDATA)>  
14 <!ATTLIST touriste num ID #REQUIRED>  
15 <!ATTLIST inscription codeA IDREF #REQUIRED>  
16 <!ATTRIBS inscription annee CDATA #IMPLIED>  
17  
18 ]>  
2 <!DOCTYPE inscriptions SYSTEM "sports-hiver.dtd">  
3 <inscriptions>  
4   <activites>  
5     <activite code="SA">  
6       <nom>Ski alpin</nom>  
7     </activite>  
8     <activite code="SN">  
9       <nom>Snowboard</nom>  
10    </activite>  
11    <activite code="SC">  
12      <nom>Scrabble</nom>  
13    </activite>  
14  </activites>  
15  <touristes>  
16    <touriste num="t1">  
17      <nom>JC Dusse</nom>  
18      <adresse>Paris</adresse>  
19      <inscription codeA="SA" annee="2004"/>  
20      <inscription codeA="SC" annee="2004"/>  
21    </touriste>  
22    <touriste num="t2">  
23      <nom>Gigi</nom>  
24      <inscription codeA="SN" annee="2005"/>  
25    </touriste>  
26  </touristes>  
27 </inscriptions>
```

## Exercice 1 Compréhension du jeu de données

1. La DTD contient des erreurs et ne peut valider le document XML. Si l'on considère que le document XML est correct, quelles modifications faut-il apporter à cette DTD afin que le document lui soit conforme ?

## Exercice 2 Requêtes XPath

Écrire en XPath les requêtes suivantes :

1. Les activités proposées (éléments *activite* complets)
2. Le nom des touristes
3. Les inscriptions du touriste identifié par *t1*
4. La seconde activité
5. Le nombre de touristes avec une inscription en 2004
6. Le nombre d'activités dont le code contient un *N* mais dont le texte ne commence pas par *ski*

## Exercice 3 Requêtes XQuery

Écrire en XQuery les requêtes suivantes :

1. Une liste non ordonnée (HTML) des activités dont le code contient un 'A'.
2. Pour chaque touriste, son nom suivi du nombre d'activités différentes auxquelles participe ce touriste sur le nombre total d'activités disponibles (e.g., "JC Dusse : 2/3"). Le résultat sera trié par nombre d'activités décroissant. Les fonctions `fn:count(item1, ..., itemn)` et `fn:distinct-values(val1, ..., valn)` peuvent être utiles.