

Stockage de données: les solutions simples - TP3

Hiba ALQASIR

Automne 2021

Objectif de la séance

L'objectif de cette séance est la prise en main du format de fichiers de stockage de données XML.

A la fin de la séance, vous déposerez sur le portail une archive au format zip contenant tous vos fichiers, *i.e.* un fichier texte avec vos réponses (pas de code à l'intérieur de ce fichier), et tous les fichiers xml, svg, mml ...

1 Correction de documents XML

1. Récupérer les fichiers [movies1.xml](#), [movies2.xml](#), [movies3.xml](#), et [movies4.xml](#) de Mootse.
2. Attention, ces fichiers ne sont pas bien formés, il faut les ouvrir avec un éditeur de texte et non un navigateur web.
3. Corriger ces documents pour qu'ils soient bien formés. Pour vérifier qu'un document est bien formé, il suffit de l'ouvrir avec un navigateur web, et vérifier qu'il n'y a pas de message d'erreur.

2 Créer un document XML manuellement

1. Créer un document XML structurant les données à partir du fichier [station_information.json](#) de TP1. Nommer le nouveau fichier **station_information.xml**.

Le document XML doit comprendre au moins 4 station. Pour chaque station, on souhaite connaître les informations suivantes:

- station_id
- name
- coordinates
- capacity

Le choix des balises, des attributs et de l'arborescence est libre (choisir intelligemment).

2. Faire la même chose à partir du fichier [station_status.json](#). Nommer le nouveau fichier `station_status.xml`.

Le document XML doit également comprendre au moins 4 stations. Cette fois, pour chaque station, on souhaite connaître les informations suivantes:

- `station_id`
- `num_docks_available`
- `num_bikes_available`
- `is_installed`
- `num_bikes_disabled`
- `is_renting`

3 Structuration d'informations

Les deux paragraphes suivants contiennent des informations 'en vrac'. Ils doivent être réorganisés pour afficher une structure logique et procéder dans le format XML approprié.

[*L'université Jean Monnet (UJM), est une université publique de recherche basée à Saint-Étienne, en France, officiellement créée en 1969. UJM offre aujourd'hui un éventail de formations avec plus de 276 diplômes (41 diplômes universitaires, 49 licences, 52 masters, 123 doctorats ...) dans ses 5 campus: Tréfilerie, Manufacture, la Métare, Santé Innovations, Roanne, Patrimoine à Firminy.*] ¹

```
1 <!-- proposition debut de solution -->
2 <UJM>
3   <type> université publique de recherche </type>
4   <fondation> 1969 </fondation>
5   <localisation>
6     <ville> Saint-Etienne </ville>
7     <pays> France </pays>
8   </localisation>
9   <formation>
10    <diplômes-universitaire/>
11    <licence/>
12  </formation>
13  <campus>
14    <Tréfilerie main_campus="true" />
15    <Manufacture/>
16  </campus>
17 </UJM>
```

[*Le jeu d'échecs, ou les échecs est un jeu de société opposant deux joueurs de part et d'autre d'un tablier appelé échiquier composé de soixante-quatre cases, 32 claires et 32 sombres nommées les cases blanches et les cases noires. Chaque joueur possède au départ un roi, une dame, deux tours, deux fous, deux cavaliers et huit pions.*] ²

¹Source: adapté de fr.wikipedia.org/wiki/Université_Jean-Monnet-Saint-Étienne

²Source: fr.wikipedia.org/wiki/Échecs

4 Étude de fichiers XML

De nombreuses applications utilisent le format XML pour stocker des données. Pour avoir une idée de ce que l'on peut représenter, nous allons étudier deux de ces formats SVG et MathML.

4.1 Scalable Vector Graphics (SVG)

SVG est un format d'image vectorielle (extension .svg) basé sur XML pour les images en deux dimensions.

- Télécharger le fichier [picture.svg](#) .
- Pour afficher l'image, il suffit d'ouvrir le fichier dans un navigateur ou avec un logiciel de visualisation d'images.
- Zoomer le plus possible, comment cela affecte-t-il la qualité de l'image?
- Ouvrir le fichier avec un éditeur de texte. Quelle est sa structure?
- A l'aide de l'éditeur de texte, ajouter un nouveau rectangle vert. Vérifier le résultat en ouvrant à nouveau le fichier dans le navigateur.

4.2 Mathematical Markup Language (MathML)

MathML est un format des équations mathématiques (extension .mml) basé sur XML pour l'affichage de symboles mathématiques, notamment sur Internet.

- Télécharger le fichier [equation.mml](#) .
- Pour afficher l'équation, il suffit d'ouvrir le fichier dans un navigateur (peut-être, il faut installer l'extension Native MathML).
- Ouvrir le fichier avec un éditeur de texte. Quelle est sa structure?
- Le document présente différents types d'éléments:
 - `<mrow>` pour des éléments à aligner horizontalement.
 - `<mi>` pour un identifiant,
 - `<mo>` pour un opérateur,
 - `<mn>` pour un nombre.

A quoi correspondent les éléments `<mfrac>`, `<sqrt>`, `<msup>` ?

- A l'aide de l'éditeur de texte, créer un document MathML qui affiche l'équation (la fonction sigmoïde):

$$S(x) = \frac{1}{1 + e^{-x}} = \frac{e^x}{e^x + 1} = 1 - S(-x) \quad (1)$$