

Master 2 - GL

Fiche TP02 : Introduction à XML

Mini projet : modélisation

Une entreprise veut gérer la procédure de recrutement en ligne via une application de gestion des CV (CVTHEQUE). Dans un premier temps, la modélisation des données (les CV) sera basée sur XML, afin de récupérer les données en XML.

Travail demandé :

1. Expliquer les avantages et les inconvénients d'une telle approche (modélisation XML) ?
2. Proposez une DTD (fichier cv.dtd), pour la validation de document XML, selon les règles et les contraintes suivantes (vous devez spécifier les types de données (texte, ...) :

A. Le CV contient 09 sections : 1- Entête (nom, prénom, date et lieu de naissance, adresse e-mail, image), 2- Postes occupés (1 ou plusieurs, exemple : étudiant, développeur, gestionnaire, ...etc), 3- Diplômes obtenus (1 ou plusieurs), 4- Stages (zéro ou plusieurs), 5- Compétences techniques (1 ou plusieurs, exemple : java , ...), 6- Langues (1 ou plusieurs), 7- Loisirs (1 ou plusieurs), 8- Niveau (débutant, intermédiaire, expert), 9- Salaire demandé.

B. Vous devez prendre en considération l'unicité des CV (chaque CV doit avoir un code unique).

3. Générer automatiquement un fichier Cv.xml à base du schéma DTD précédente, et saisir vos données (je dois trouver un nombre de CV équivalent au membre du groupe de projet).
4. Lire et afficher le contenu du document XML à l'aide de :
 - o Un parseur DOM (ParserDOM.java).
 - o Un parseur SAX (ParserSAX.java).
5. Proposez une solution pour convertir ces CV sous format DOC, PDF et HTML.
6. Afin de stocker vos documents XML, on a plusieurs possibilités : expliquer brièvement comment stocker vos documents XML dans un SGBD relationnelle (Exemple : SQL Server), et dans SGBD XML native (proposer un SGBD), **détailler les choix et expliquer brièvement.**

Rappel: *Les espaces des noms* est un mécanisme qui permet de résoudre l'ambiguïté au niveau des noms des éléments et attributs (différentes balises qui porteraient le même nom),

- Les espaces de noms favorisent la modularité des documents XML et les sources multiples de ces derniers,
- Les espaces de nom permettent, indirectement, de rendre plus lisible un document XML contenant des informations de sources diverses.
- Trois formes de type d'espace de nom sont possibles : implicite, explicite et hybride.
- Quand on déclare plusieurs espaces de noms, l'espace de nom par défaut est celui sans préfixe.
- La portée d'une déclaration est limitée au sous-arbre dans lequel la déclaration a été faite.
- Les espaces de noms peuvent être appliqués aux éléments ainsi qu'aux attributs.

The diagram shows an XML snippet with several annotations pointing to specific parts:

```
<?xml version="1.0" encoding="UTF-8" ?>
<disc:discothèque xmlns:disc="namespaceURI">
  <disc:disque disc:nb="1">
    <disc:titre>Just Jazz</disc:titre>
    <disc:artiste>Romano</disc:artiste>
    <disc:disque>
  </disc:discothèque>
```

Annotations on the right side:

- déclaration (points to `xmlns:disc="namespaceURI"`)
- URI (espace de noms) (points to `namespaceURI`)
- préfixe (points to `disc:` in `disc:titre`)
- partie locale (points to `titre` in `disc:titre`)
- nom qualifié (points to `disc:titre`)