

TD 1 – XML – DTD

XML permet de structurer une information. Il est donc nécessaire, avant d'envisager d'utiliser ce format, de se familiariser avec cette structuration.

Exercice 1

Le paragraphe suivant contient de l'information "en vrac". Réorganisez-la de manière à mettre en évidence sa structure logique, sans forcément passer par une mise en forme XML.

« Une bouteille d'eau **Kirène** de 150 cl contient par litre 71 mg d'ions positifs calcium, et 5,5 mg d'ions positifs magnésium. On y trouve également des ions négatifs comme des chlorures à 20 mg par litre et des nitrates avec 1 mg par litre. Elle est recueillie à **Thiès**, dans le département de **Mbour**. Son code barre est 3274080005003 et son pH est de 7,45. Comme la bouteille est sale, quelques autres matériaux comme du fer s'y trouvent en suspension.

Une seconde bouteille d'eau **Kirène** a été, elle, recueillie à la source de **Toglou** dans la commune de **Diass**. La concentration en ions calcium est de 98 mg/l, et en ions magnésium de 4 mg/l. Il y a 3,6 mg/l d'ions chlorure et 2 mg/l de nitrates, pour un pH de 7,4. Le code barre de cette bouteille de 50 cl est 3268840001008.

Une bouteille de même contenance est de marque **Séo**, et a été puisée à **Bambey**. **Séo**, bien connu pour ses sources donnant un pH neutre de 7. Elle comprend 11,5 mg/l d'ions calcium, 8,0 mg/l d'ions magnésium, 13,5 mg/l d'ions chlorures et 6,3 mg/l d'ions nitrates. Elle contient également des particules de silice. Son code barre est 3057640117008. ». **Séo** est dans la région de Diourbel.

Exercice 2

Dans cet exercice, nous considérons un maraîcher qui a décidé, en 2021, de passer à l'informatique et d'archiver tout son stock à l'aide d'un document XML. Donnez une DTD correspondant au document XML utilisé par le maraîcher. Le document donné doit être valide par rapport à votre DTD, et de plus elle doit vérifier les critères suivants :

- Le nombre d'éléments **fruit** et le nombre d'éléments **légume** sont quelconques.
- Les éléments **fruit** et **légume** peuvent être mélangés.
- Les éléments **producteur**, **origine** et **qntity** sont obligatoires, les autres sont optionnels.
- L'ordre des éléments est fixe.
- L'attribut **région** de l'élément **origine** est optionnel.
- Les attributs **type**, **prix** et **calibre** des éléments **fruit** et **légume** sont obligatoires
- L'attribut **calibre** peut uniquement prendre les valeurs **1**, **2** ou **3**.
- L'élément **bio** est vide.

Exercice 3

Dans le passage au LMD l'UT décide d'uniformiser la présentation des enseignements suivant un canevas commun nommé syllabus. Un cours est décrit par 3 blocs : description du cours et renseignements sur le(s) enseignants et les étudiants inscrits au cours.

La description du cours donne l'intitulé du cours, celui de l'unité d'enseignement à laquelle appartient le cours, la charge de travail (crédits, coefficient et volume horaire correspondant), le semestre du cours, le niveau et les prérequis.

Les renseignements sur le (s) enseignant (s) donne le nombre d'enseignements qui animent le cours. Pour chaque enseignement on donne les prénom(s) et nom, son adresse de bureau, son courriel et ses heures de bureau.

- 1) Donnez un exemple de document XML qui décrit le cours Technologies XML et qui respectent les spécifications ci-dessus.
- 2) Donnez une DTD et le schéma XML correspondant à cette description.

Les étudiants sont identifiés par des numéros et suivent plusieurs cours. Un étudiant est décrit par ses nom et prénoms, par sa classe (son niveau de L1 à D3), sa situation financière (boursier ou non), sociale (logé ou pas) et son adresse familiale. Chaque élément étudiant contient un élément notes qui à son tour contient des références à des cours et les notes de

l'étudiant correspondantes au projet et à l'examen. Dans ce cas nous ne considérons que 2 notes : projet et examen TechXML. Enfin les étudiants sont consignés dans un fragment à insérer dans le document principal.

- 1) Donnez un exemple de document XML qui décrit votre classe L2 qui respectent les spécifications ci-dessus et contient une liste de 5 étudiants.
- 2) Donnez le schéma XML et une DTD correspondant à cette description.

Exercice 4

On veut représenter dans un format XML des séries de bandes-dessinées et leurs auteurs et illustrateurs. On veut respecter les contraintes suivantes :

- pour une série de BD, on doit pouvoir donner son titre ;
- une série peut être soit une BD unique, soit plusieurs BDs ;
- on veut pouvoir donner, pour chaque BD : son (ou ses) auteur(s), son illustrateur, son résumé, sa date de sortie, et éventuellement sa couverture (une image) et son sous-titre ;
- si la BD fait partie d'une série, on précisera en plus son numéro dans la série ;
- différentes BDs d'une même série peuvent être écrites et/ou dessinées par des personnes différentes ;
- une même personne peut être à la fois auteur et illustrateur ;
- pour un auteur/illustrateur on peut donner son nom et son prénom.

Vous utiliserez un système de références pour séparer les auteurs/illustrateurs des bandes-dessinées. Il ne devra pas y avoir de redondance sur les auteurs/illustrateurs. Le résumé d'une BD doit autoriser du contenu mixte pour mettre en lumière les noms des personnages.

1. Fournir une DTD décrivant un tel document.
2. Fournir un XML Schéma décrivant ce document. Pour les références, vous utiliserez le système de clés offert par XML Schéma. Les données pouvant être typées (dates, longueurs de chaînes, etc.) doivent l'être.
3. Fournir un schéma Relax NG. Vous utiliserez la définition de patterns offerte par le format Relax NG. Les données pouvant être typées doivent l'être.

Exercice 5

On veut encoder un arbre généalogique dans un format XML. La racine est gene, elle a pour fils des éléments personne. On autorise une généalogie vide.

L'élément personne a un attribut facultatif **id** de type *ID* et un attribut obligatoire **sexe** pouvant prendre au choix les valeurs *m* et *f*. Il a un fils nom, obligatoire, contenant uniquement du texte et un fils parents, facultatif, vide. L'ordre est important. L'élément parents possède un attribut ref, de type IDREFS, obligatoire.

1. Écrivez la DTD correspondant à cette description.
2. Insérez-là en tant que DTD interne d'un document XML décrivant la généalogie suivante: Fati a deux parents, Marieme et Badou. Babou a deux parents, Sidi et Niania l'arbre XML, vous pouvez utiliser p en place de personne.

Exercice 6

Voilà un document XML avec DTD interne.

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE texte [
  <!ELEMENT texte (#PCDATA) >
]>
<text>
  Thiès est une commune sénégalaise de la region ouest.
<text>
```

Transformez ce document pour avoir le mot "Thiès" dans une entité générale que vous appellerez ville. Indiquez juste les lignes que vous ajoutez ou modifiez.