

## TD1 Au vert pommier : Principes d'un document XML bien formé 1h

Léo et Marjorie Delcour gèrent le chenil familial « Au vert pommier » en Normandie, près de Caen. Ils élèvent plusieurs races de chiens. Leurs animaux, issus d'une lignée pure et réputée de champions de concours canins, peuvent être inscrits au LOF<sup>1</sup> et participer à des concours.

La gestion des lignées de race doit être rigoureuse afin de produire des animaux parfaits au sens du LOF<sup>2</sup> et d'éviter l'apparition de tares issues de croisées consanguines.

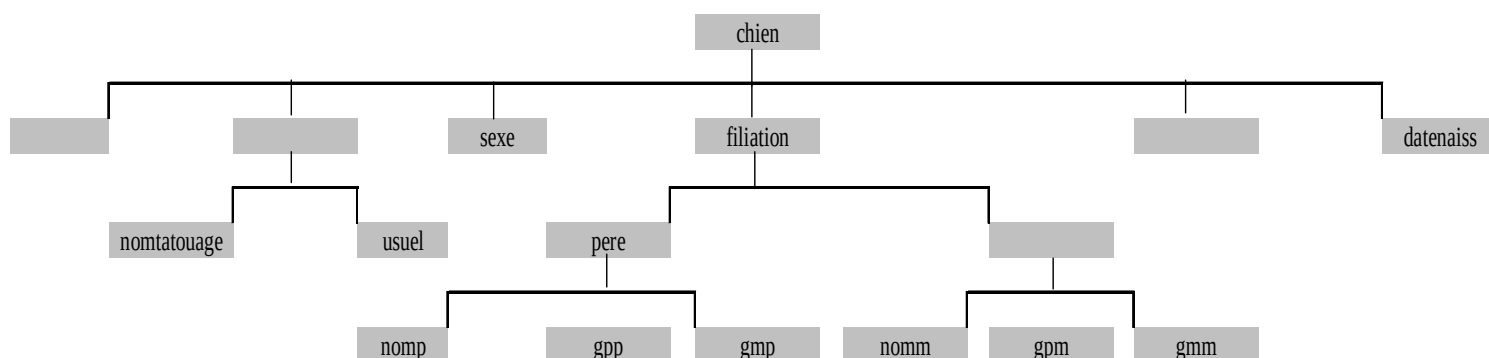
Des couples d'animaux produisent des petits qui sont tatoués lorsqu'ils ont un mois. Lors du tatouage, un carnet de santé est établi pour chaque animal par le vétérinaire, et l'éleveur donne un nom dont la première lettre est définie par ordre alphabétique chaque année. Ainsi, en 2006 c'était B et en 2007 les animaux doivent avoir un nom commençant par C.

L'acheteur de l'animal peut ensuite donner un nom d'usage différent lorsqu'il vient réserver son chien. (Les animaux sont mis à la vente dès leur naissance mais ne partent chez l'acheteur qu'à huit semaines révolues). Ce qui fait qu'un animal, jusqu'à ses quatre semaines, n'a pas de nom ni de tatouage, et de quatre à huit semaines, il a forcément un tatouage, un nom et parfois un second nom !

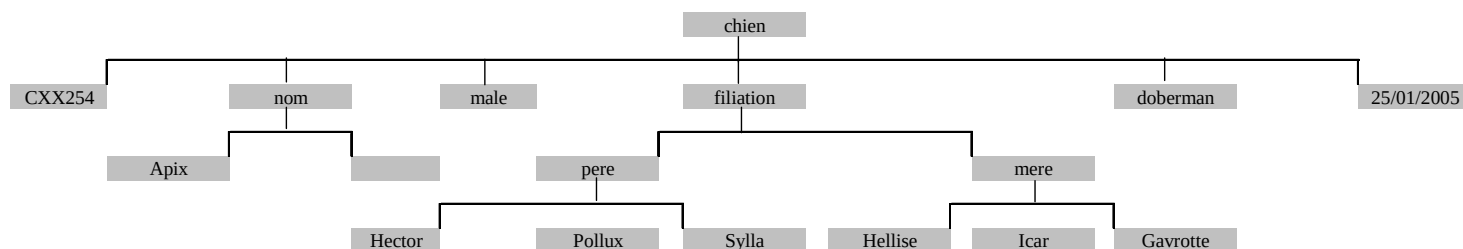
Marjorie recherche une nouvelle façon de présenter ses fichiers de lignées<sup>3</sup> et, à cet effet, décide d'étudier un nouveau mode de structuration de ses données.

Les informations à manipuler sont le numéro de tatouage, les noms du chien, le sexe, le nom de ses 2 parents et 4 grands parents (le pedigree remonte à ce niveau), la race, la date de naissance.<sup>4</sup>

1. Complétez le tableau suivant qui permettra de représenter l'arborescence des informations à gérer.



Illustrons cette représentation par un exemple pour un chien particulier : Apix



<sup>1</sup> LOF : Livre des Origines Français qui permet de confirmer les origines d'un animal

<sup>2</sup> Tout chiot dont les deux parents ont un pedigree est dans un premier temps inscrit au LOF. Ce n'est qu'à l'issue de la confirmation à l'âge adulte (à partir de 15 mois et sans limite d'âge maximum) par un juge habilité qu'il reçoit ou non son pedigree. Chaque pays a son livre des origines et si un chien n'est inscrit à aucun livre, il ne peut être déclaré que comme genre (ou type) Labrador ou Golden retriever ou caniche ...

<sup>3</sup> Dans ce cas d'école, nous ne prendrons en compte que des éléments de base et non pas toutes les informations nécessaires à une déclaration au LOF !

<sup>4</sup> les informations sur les acheteurs de l'animal ne sont pas considérées à ce niveau.

- 2) Quelle est la race de ce chien ?
- 3) Quel est le nom de son grand-père paternel ?
- 4) Cet animal a un nom de tatouage mais pas de nom usuel, qu'est-ce que cela nous apprend sur lui ?

La structuration des informations sur un chien comporte différents niveaux de détails : Au premier niveau, nous avons une première structuration : Tatouage, nom, sexe, filiation etc. (ligne 2 de l'arbre). Le niveau suivant (ligne 3) est un niveau de détail de certaines informations du premier niveau. Etc.

Cette représentation structurée des informations relatives à un chien, peut s'écrire sous la forme d'un fichier XML. Examinons le script XML suivant :

<u>commentaires XML</u>	<u>fichier chien.xml</u>	<u>Niveau de l'arborescence</u>
déclaration xml (prologue) →	<?xml version="1.0" ?>	
commentaires XML →	<!--AU VERT POMMIER -->	
élément racine →	<listechiens>	
indentation primaire →	<chien>	←1er niveau
Indentation secondaire →	<tatouage></tatouage>	←2d niveau
	<nom></nom>	
	<sexe></sexe>	
	<filiation></filiation>	
	<race></race>	
	<datenaissance></datenaissance>	
fermeture de balise →	</chien>	
fermeture de la balise de la racine →	</listechiens>	

*Les niveaux suivants ne sont pas représentés ici*

- 5) Comment a-t-on appelé la première ligne du script XML ? À quoi correspond-elle ?
- 6) En observant ce document, comment peut-on définir la balise appelée « élément racine » ?

7) Modifiez le fichier chien.xml afin de pouvoir y représenter les 3<sup>ème</sup> et 4<sup>ème</sup> niveaux d'arborescence

```
1  <?xml version="1.0" ?>
2  <!--AU VERT POMMIER -->
3  <listechiens>
4      <chien>
5          <tatouage></tatouage>
6          <nom>
7
8
9          </nom>
10         <sexe></sexe>
11         <filiation>
12
13
14
15
16
17
18
19
20
21
22         </filiation>
23         <race></race>
24         <datenaissance></datenaissance>
25     </chien>
    ...
</listechiens>
```

Marjorie vous donne un extrait de fichier XML pour Apix :

```

1  <?xml version="1.0" encoding="iso-8859-1">
2  <!--AU VERT POMMIER -->
3  <listechiens>
4      <chien>
5          <tatouage>.....</tatouage>
6          <nom>
7              <nomtatouage>Apix</nomtatouage>
8              <usuel></usuel>
9          </nom>
10         <sexe>mâle</sexe>
11         <filiation>
12             <pere>
13                 <nomp>.....</nomp>
14                 <gpp>.....</gpp>
15                 <gmp>.....</gmp>
16             <pere>
17                 <mere>
18                     .....Hellise.....
19                     .....Icar.....
20                     .....Gavrotte.....
21                 </mere>
22                 .....
23             <race> Dobermann </race>
24             <datenaissance></datenaissance>
25             .....
26             .....
```

- 8) Complétez les zones grisées en fonction des questions précédentes.
- 9) Saisissez ce script XML dans un éditeur de texte puis affichez-le dans un navigateur.
  - a) Remarquez que le navigateur indente les éléments selon l'arborescence même s'ils ne l'ont pas été dans l'éditeur de texte.
  - b) Lorsqu'on ferme tous les éléments développés de l'arborescence dans le navigateur, il ne reste plus qu'une seule balise avec le signe + devant, à quoi correspond t-elle ?
  - c) Apix n'ayant pas de nom usuel, remarquez l'affichage particulier de la balise correspondante. Les balises vides <usuel></usuel> s'affichent : <usuel />

Léo à vu le travail de Marjorie mais il lui rappelle que les déclarations au LOF demandent un degré d'information plus fin en ce qui concerne les races<sup>5</sup> et la « destination <sup>6</sup> » de l'animal.

- 10) Proposez une amélioration du script XML afin de prendre en compte ces besoins spécifiques.
- 11) L'affixe est une dénomination qui s'ajoute au nom du chien et qui permet de savoir de quel éleveur l'animal provient. Pensez-vous que cet affixe doit apparaître dans le fichier xml ?
- 12) Ouvrez le fichier *auvertpommier.xml* que vous a fourni votre professeur.
- Combien d'animaux sont enregistrés dans ce document ?
  - Comment se présente la structuration des informations sur ce document, par rapport aux questions précédentes ?
  - Modifiez le prologue, dans un éditeur, en effaçant l'encoding : **encoding="iso-8859-1"**. Sauvegarder puis actualiser l'affichage dans le navigateur. En déduire l'utilité de cet encoding.
  - Corrigez l'erreur précédente en remettant le bon encoding puis effacer l'une des balises `</chien>`, que se passe-t-il ?
  - Corrigez le problème précédent et modifiez maintenant la première balise, `<chien>` en mettant `<CHIEN>` que se passe-t-il ?
  - Modifiez la première balise `<chien>` en mettant `<1erchien>` et la première balise `</chien>` avec `</1erchien>` que se passe-t-il ?

Nous voyons donc que le XML impose quelques règles strictes de construction :

- Un nom de balise ne peut pas commencer par un chiffre mais il peut commencer par toute lettre alphabétique, un « \_ » ou « : »
- Le xml est sensible à la casse (majuscules/minuscules)
- A toute balise ouvrante, correspond une balise fermante, les balises portant un contenu vide peuvent être écrites de la forme `<balise />`, sinon le navigateur les interprétera de la sorte.

Un document xml respectant ces critères est dit *bien formé (well formed)*.

<sup>5</sup> Plusieurs types : Grande race, petite race, race moyenne ou naine ou toy. Exemple : Race = Caniche, Type = Toy

<sup>6</sup> La destination d'un chien concerne le but de son propriétaire à l'achat : chien de compagnie, chien de chasse, chien de garde, chien d'attaque etc.