

# Culture du numérique et code

## XML – EXtensible Markup Language

Inspirés de:

<https://www.cours-gratuit.com/cours-xml/cours-xml-pour-debutant-pdf>

Tarak CHAARI - Dpt Télécommunications INSA

# Sommaire

- Fondements de la technologie XML
- Modélisation et validation de documents XML
- Présentation des documents XML

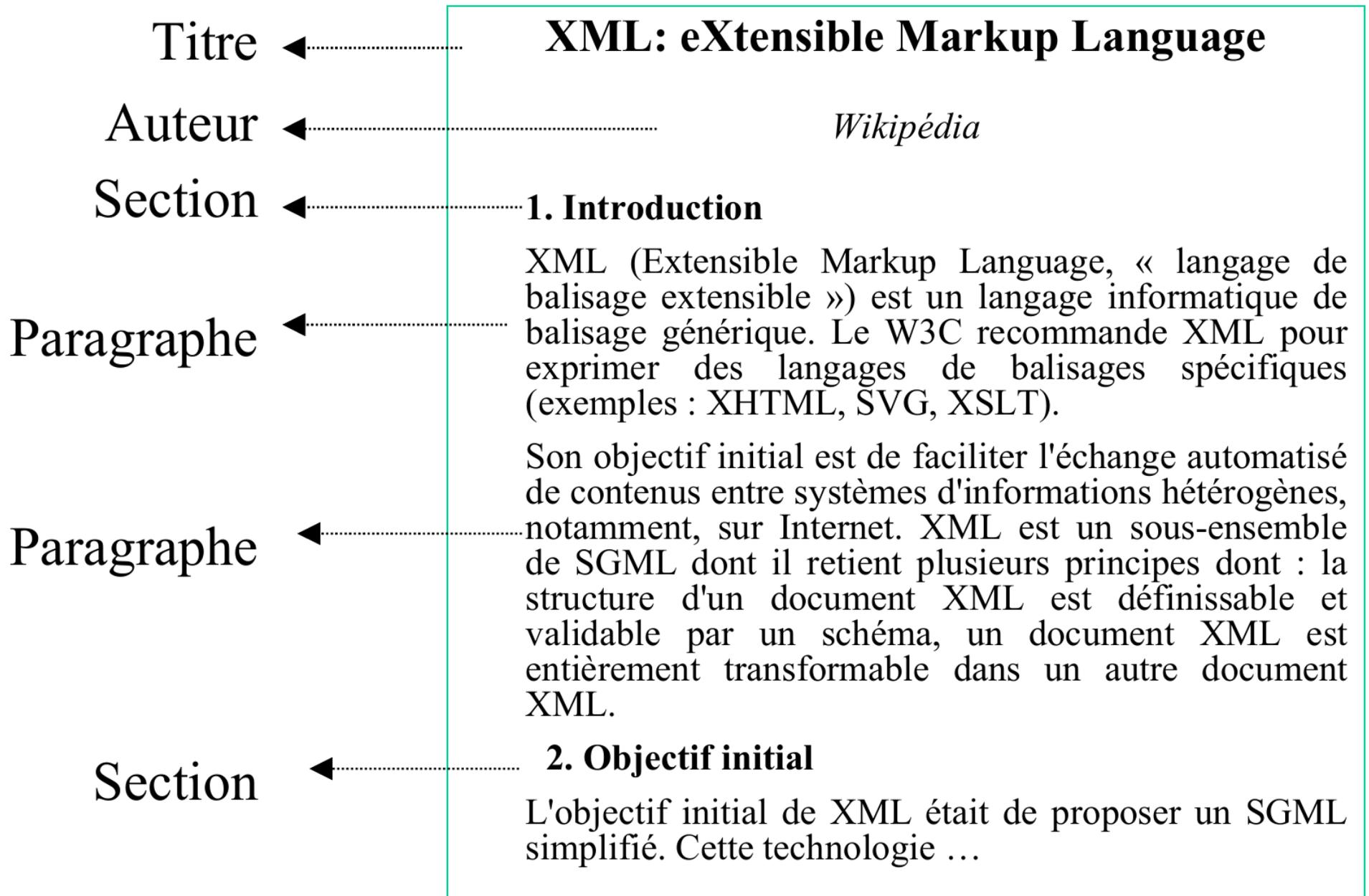
# Notion de document électronique

- Joue un rôle de médiateur entre les hommes dans le temps et l'espace (échange d'information)
- Moyen de communication de la pensée, de la connaissance, de l'information et du savoir
- « ensemble formé par un support et une information » (ISO TC-46)
- Se présente sous forme de: textes, tableaux, dessins, photos, ...
- Possède un contenu (structure logique) et un fond (structure physique)

# Modèles de documents

- Structure logique
  - Décrit le contenu d'un document (information)
  - Par exemple: Chapitre, section, paragraphes, figures, notes...
- Structure physique
  - Décrit la forme et le support du document (formatage)
  - Par exemple: tomes, pages, cadres, pavés, fenêtres
- Caractéristiques communes des 2 structures:
  - Composées ou élémentaires (molécule / atomes)
  - Générique ou spécifique (date / le 29-11-2006)

# Exemple d'un document



# Exemple d'un document: structure logique

<Article>

<Titre> XML: eXtensible Markup Language </Titre>

<Auteur>Wikipédia</Auteur>

<Section titre = "Introduction">

<Paragraphe>XML (Extensible Markup Language, « langage de balisage extensible ») est un langage informatique de balisage générique. Le W3C recommande XML pour exprimer des langages de balisages spécifiques (exemples : XHTML, SVG, XSLT). ...

</Paragraphe>

<Paragraphe> Son objectif initial est de faciliter l'échange automatisé de contenus entre systèmes d'informations hétérogènes, notamment, sur Internet. XML est un sous-ensemble de SGML dont il retient plusieurs principes dont : la structure d'un document XML est définissable et validable par un schéma, un document XML est entièrement transformable dans un autre document XML. </Paragraphe>

</Section>

<Section titre= "Objectif initial"> ...

</Section>

</article>

# Exemple d'un document: structure physique

```
<Article>
```

```
<Titre police="Times" taille="24" position="centré" format="gras"/>
```

```
<Auteur police="Times" taille="20" position="centré"  
  format="italique"/>
```

```
<Section numero="1" police="Times" taille="18"  
  position="centré" format=" gras "/>
```

```
<Paragraphe police="Times" taille="18" position="justifié"/>
```

```
</article>
```

# WEB

- W3C - Fondé en 1994
- Consortium industriel international accueilli par différents sites
  - MIT/LCS aux Etats-Unis
  - INRIA en Europe
  - Keio University au Japon
- 448 membres industriels en septembre 2000
- Accroître le potentiel du WEB
  - Standards et Normes
  - Techniques, langages et architectures pour l'échange de documents sur le WEB

# Langages de représentation de documents

- SGML (Norme ISO 8879 en 1986, révisée en 1988 et 1994)
  - Méta-langage général
- HTML (Standard W3C depuis 1989)
  - Structure générale figée
- XML (Standard W3C depuis 1998)
  - Méta-langage simplifié compatible SGML

# XML

- XML= un nouveau langage d'échange basé sur le balisage
- XML= plus simple que SGML
- XML= plus ouvert que HTML
- XML = développé par XML Working Group dirigé par le W3C (depuis 1996)
- XML 1.0 = recommandation officielle du W3C depuis le 10 février 1998

# XML: objectifs (1/2)

- XML doit pouvoir être utilisé sans difficulté sur Internet
- XML doit soutenir une grande variété d'applications
- XML doit être compatible avec SGML et HTML
- Il doit être facile d'écrire des programmes traitant les documents XML

# XML: objectifs (2/2)

- Les documents XML doivent être lisibles par l'homme
- Les documents XML doivent être raisonnablement clairs
- La spécification de XML doit être disponible rapidement
- La conception de XML doit être formelle et concise
- Il doit être facile de créer des documents XML

# XML: forces

- Séparation de la structure et de la présentation
- Moins confus que HTML
- Plus simple que SGML
- Idéal pour l'échange de données semi-structurées
- Utilisable entre machines hétérogènes

# XML: utilités (1/2)

- XML est un Méta-langage universel pour représenter les données échangées sur le Web qui permet au développeur de délivrer du contenu depuis les applications à d'autres applications ou aux navigateurs
- XML standardise la manière dont l'information est :
  - échangée
  - présentée
  - archivée
  - retrouvée
  - transformée
  - cryptée

# XML: utilités (2/2)

- Définir vos propres langages d'échange
  - Commande, facture, bordereau de livraison, etc.
- Modéliser des documents et des messages
  - Modèle logique de données
  - Eléments typés agrégés (DTD, XML Schema)
- Publier des informations
  - Neutre du point de vue format
  - Mise en forme avec des feuilles de style
- Archiver des données
  - Auto-description des archives (recherche d'information)

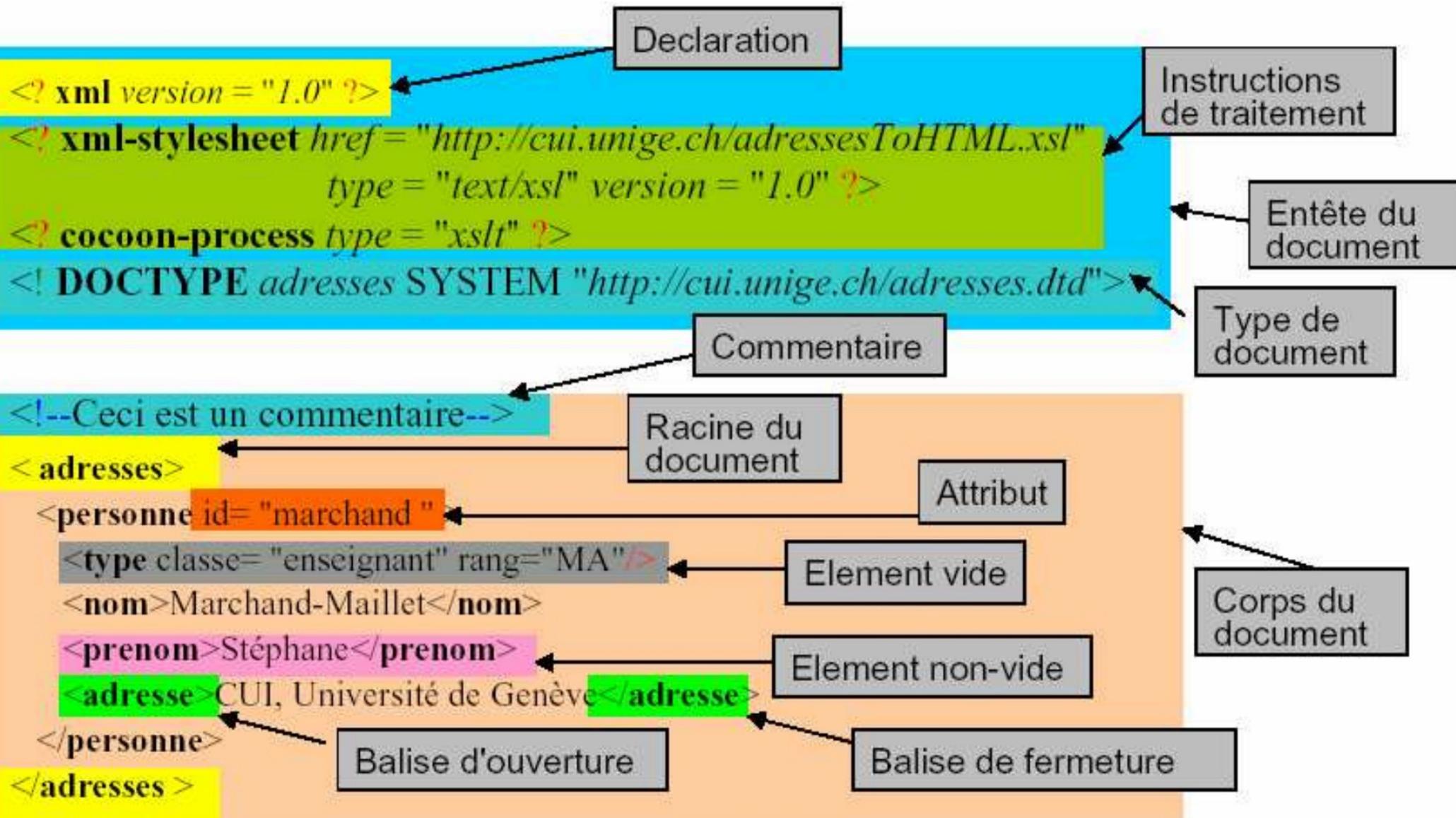
# XML: concepts

- **Balise (ou tag ou label)**
  - Marque de début et fin permettant de repérer un élément textuel
  - Forme: <balise> de début, </balise> de fin
- **Élément de données**
  - Texte encadré par une balise de début et une de fin
  - Les éléments de données peuvent être imbriqués

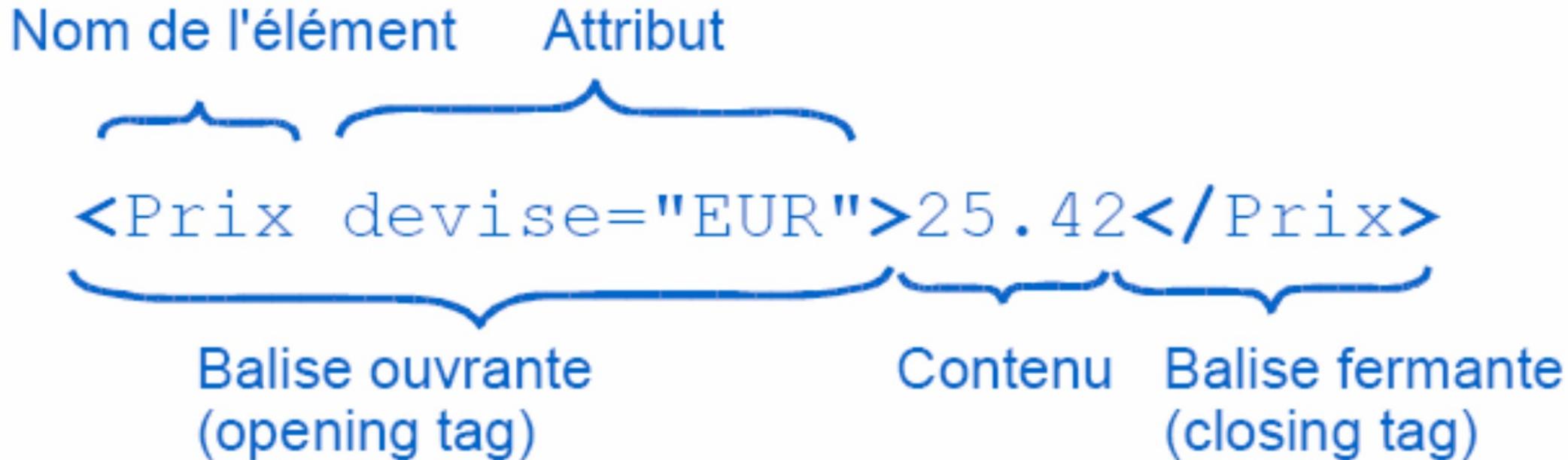
```
<producteur>
  <adresse>
    <rue>A. Einstein</rue>
    <ville>Villeurbanne</ville>
  </adresse>
</producteur>
```
- **Attribut**
  - Doublet nom="valeur" qualifiant une balise

```
<producteur no="160017" region="Rhône">
```

# XML: exemple



# XML: Elements/Attributs



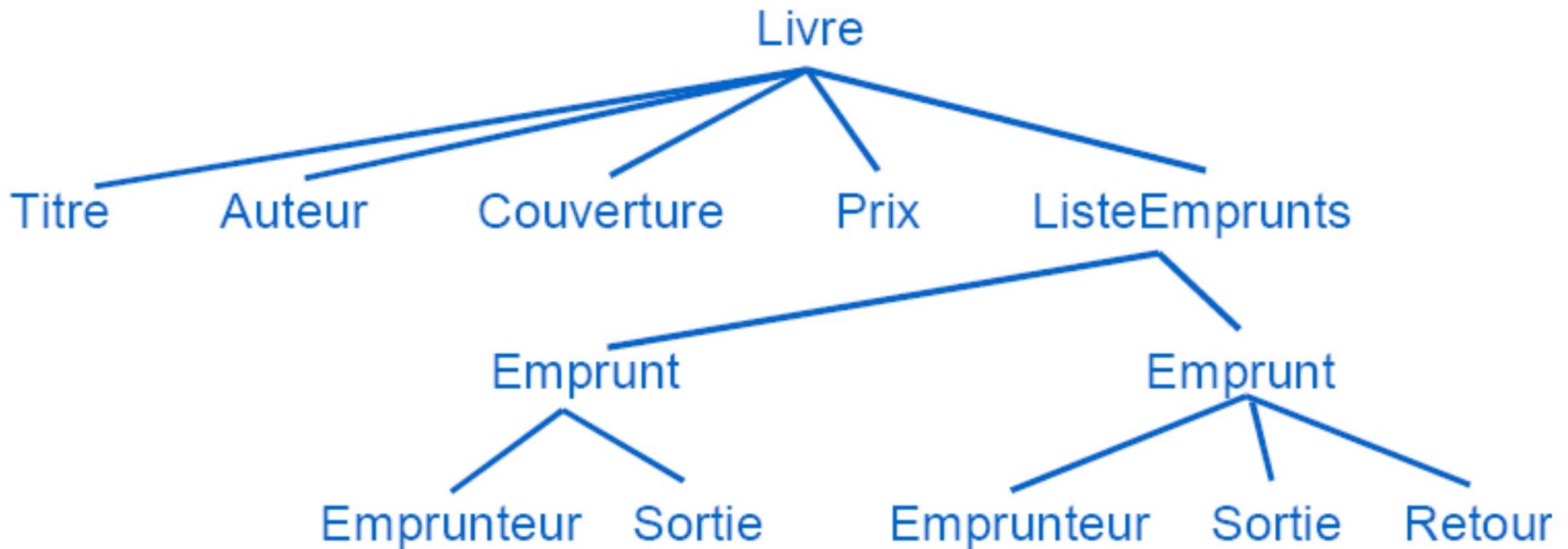
- Quand utiliser les attributs?
  - Valeur unique de type simple (information monovaluée)

# XML: exemple

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<?xml-stylesheet href="bib.css" ?>
<?cocoon-process type="xslt" ?>
<!DOCTYPE Livre SYSTEM "Livre.dtd">

<Livre>
  <Titre>Les réseaux</Titre>
  <Auteur>A. Tanenbaum</Auteur>
  <Couverture imgsrc="/imgs/res-tan.jpg"/>
  <Prix devise="EUR">25.42</Prix>
  <ListeEmprunts>
    <Emprunt>
      <Emprunteur>François Duchemin</Emprunteur>
      <Sortie>25/09/2000</Sortie>
      <Retour>02/10/2000</Retour>
    </Emprunt>
    <Emprunt>
      <Emprunteur>Hervé Delarue</Emprunteur>
      <Sortie>05/10/2000</Sortie>
    </Emprunt>
  </ListeEmprunts>
</Livre>
```

# XML: structure arborescente d'un document



# Exercice: structurer ce document en XML

Une bouteille d'eau Cristaline 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 **St-Cyr la Source**, dans le département du Loiret. 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 Cristaline a été, elle, recueillie à la source d'**Aurèle** dans les Alpes Maritimes. 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 Volvic, et a été puisée à... **Volvic**, 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.

# Exercice: structurer ce document en XML

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
_ <eau>
_ <bouteille>
  <marque>Cristaline</marque>
_ <composition>
  <ion type="positif">calcium 71mg/l</ion>
  <ion type="positif">magnésium 5,5mg/l</ion>
  <ion type="negatif">chlorure 20mg/l</ion>
  <ion type="negatif">nitrate 1mg/l</ion>
  <autre type="metal">fer</autre>
  </composition>
_ <source>
  <ville>St-Cyr la Source</ville>
  <departement>Loiret</departement>
  </source>
  <code_barre>3274080005003</code_barre>
  <contenance unit="cl">150</contenance>
  <ph>7,45</ph>
  </bouteille>
```

**... COMPLETER EN TP**