

```

<?xml version="1.0"?>
<!DOCTYPE encyclopedie SYSTEM "encyclopedie.dtd">
<encyclopedie>
  <personne datenaissance="1942-01-08" sexe="H">
    <nom>HAWKING</nom>
    <prenom>Stephen</prenom>
    <publication>Une brève histoire du temps</publication>
  </personne>
  <personne datenaissance="1932-07-13" sexe="H">
    <nom>REEVES</nom>
    <prenom>Hubert</prenom>
    <publication>L'Univers expliqué à mes petits-enfants</publication>
    <publication>Patience dans l'azur, L'évolution cosmique</publication>
    <publication>Poussières d'étoiles</publication>
  </personne>
  <personne datenaissance="1879-03-14" sexe="H">
    <nom>EINSTEN</nom>
    <prenom>Albert</prenom>
    <publication>Des ondes gravitationnelles</publication>
    <publication>Sur la théorie quantique du rayonnement</publication>
  </personne>
  <personne datenaissance="1867-11-07" sexe="F">
    <nom>CURIE</nom>
    <prenom>Marie</prenom>
    <publication>traité de radioactivité</publication>
  </personne>
</encyclopedie>

```

Voici ce que pourrait-être sa DTD:

```

<?xml version="1.0"?>
<!ELEMENT encyclopedie      (personne*)>
<!ELEMENT personne (nom,prenom,publication+)>
<!ATTLIST personne sexe (H | F) "H">
<!ELEMENT nom      (#PCDATA)>
<!ELEMENT prenom   (#PCDATA)>
<!ELEMENT publication (#PCDATA)>

```

- Pour chacun des éléments, on définit sa composition par **<!ELEMENT nom_element (structure)>**
- Pour chacun des attributs, on définit sa composition par **<!ATTLIST nom_element nom_attribut (structure)>**
- L'attribut **sexe** de la balise **<personne>** est, ici, défini comme étant un choix imposé entre "H" et "F". Les valeurs possibles sont séparées par des caractères '|'. Cet attribut n'étant pas indiqué comme obligatoire, s'il n'est pas précisé alors il prendra par défaut la valeur "H" (la valeur précisée entre guillemets après la liste)
- **personne*** signifie que dans un élément **<encyclopedia>** on peut trouver **de 0 à plusieurs éléments** de type **<personne>** à l'exclusion de tout autre élément
- **nom,prenom,publication+** signifie que dans un élément **personne** on doit trouver **un et un seul (*)** élément **<nom>**, un et un seul élément **<prenom>** et **un ou plusieurs (+)** éléments **<publication>** et **dans cet ordre** (comme imposé par la virgule)
- **<nom>**, **<prenom>** et **<publication>** sont des éléments qui peuvent contenir du texte comme l'indique **#PCDATA**. On parle de contenu mixte
- Deux autres types pour les éléments existent, **EMPTY** et **ANY**, **empty** signifie que l'élément est vide (comme **
** en HTML), **ANY** indique que l'élément peut comporter tout autre élément déclaré dans la DTD.

La syntaxe

Reprenons les différents indicateurs d'occurrences:

- pas de symbole veut dire 1 et 1 seul élément
- + veut dire 1 ou plusieurs éléments
- * veut dire 0 ou plusieurs éléments
- ? veut dire 0 ou 1 éléments

Ces symboles sont placés juste après les noms des éléments.

Et voici les différents connecteurs possibles:

- , les différents éléments séparés par une virgule doivent apparaître dans l'ordre donné
- | signifie qu'un seul des deux éléments séparés par ce symbole doit apparaître dans le fichier XML

Ces symboles sont placés entre deux éléments.

- On peut composer les différents connecteurs en utilisant les parenthèses. Toutefois, il faut veiller à ne pas définir une grammaire trop ambiguë par exemple ((nom,prenom)|(nom,surnom)) n'est pas correct, il est préférable d'écrire **(nom,(prenom|surnom))**.
- On peut autoriser plusieurs éléments répétés, par exemple **(adresse,telephone)*** ce qui permet de créer des couples (adresse,telephone).

Les attributs (déclarés par ATTLIST) possèdent différents contrôles:

- **CDATA** pour une chaîne quelconque de caractères
- **ID** pour un nom unique dans le document (identifiant)
- **IDREF** pour un numéro d'identifiant du document (création d'un lien dans le document)
- **ENTITY** entité externe
- **ENTITIES** liste d'ENTITY séparées par des espaces
- **NMTOKEN** pour un mot sans espace
- **NMTOKENS** pour une liste de NMTOKEN séparés par des espaces

Les valeurs des attributs peuvent être précisées par la DTD.

- **#REQUIRED** signifie que la valeur est **obligatoire** dans le document
- **#IMPLIED** signifie que s'il n'y a pas de valeur dans le document, c'est à l'application de donner la sienne par défaut
- **#FIXED** signifie que la valeur doit être celle de la DTD
- une valeur par défaut que va prendre l'attribut si aucune valeur n'est donnée dans le fichier XML