

## Séance 2 : Spécifications XML

Les spécifications XML définissent un modèle pour les documents XML qui réglemente l'imbrication des balises, les noms d'éléments autorisés, la façon dont les attributs sont rattachés aux éléments.

Les documents XML qui respectent, en plus de la syntaxe XML, les règles d'un modèle de document sont dits **valides**.

# La syntaxe DTD

## Déclaration d'un élément :

```
<!ELEMENT nom_element (contenu)>
```

- modèle de contenu :
  - #PCDATA : données textuelles
  - EMPTY
  - ANY
- Exemple : 

```
<!ELEMENT titre (#PCDATA)>
```

## Déclaration d'un sous-élément :

```
<!ELEMENT nom_element (nom_sousElement)>
```

- Déclaration d'une séquence de sous-éléments : 

```
<!ELEMENT  
nom_element (nom_sousElement1, nom_sousElement2)>
```

*Attention : l'ordre de déclaration des sous-éléments est significatif.*

- Nombre d'éléments autorisés :
  - \* ? zéro ou un ;
  - \* \* zéro ou plusieurs éléments ;
  - \* + un ou plusieurs éléments.
- Exemple : `<!ELEMENT texte (chapitre+)>`

NB : si vous voulez déclarer des sous-éléments et du texte dans un élément, il faut déclarer `#PCDATA` avant les autres éléments.

## Déclaration des attributs :

- `<!ATTList nom_element nom_attribut type_Données type_Attribut>`
  - Les données d'attributs :
    - `CDATA` n'importe quelle chaîne de caractères;
    - `ID` doit contenir une valeur unique dans le document XML. Un seul attribut de ce type par élément est accepté.
    - `IDREF` fait référence à un attribut de type ID d'un élément du document. Les attributs Idref sont généralement utilisés pour établir des relations entre les éléments.

- Les types d'attributs:
  - \* `#IMPLIED` optionnel ;
  - \* `#REQUIRED` obligatoire ;
  - \* `#FIXED` à valeur fixe.
- Exemple : `<!ATTLIST chapitre n CDATA #REQUIRED >`

# Déclarer sa DTD dans son document XML

- DTD interne : `<!DOCTYPE ElementRacine [<!ELEMENT nom_element (contenu)>]>`
- DTD externe (à déclarer avant l'élément racine) :  
`<!DOCTYPE ElementRacine SYSTEM "nomDTD.dtd">`

## Exercice

À l'aide du lien suivant, encoder la lettre LXXXIX des *Liaisons dangereuses* :

<https://gallica.bnf.fr/ark:/12148/bpt6k9738925r/f23.item>

- Déterminez les différentes zones de texte et les informations que l'encodage pourrait permettre d'explicitier;
- Proposer un encodage XML de la lettre et sa dtd associée;