

# **Insérer des caractères spéciaux dans son encodage en TEI : DTD (*Document Type Definition*)**

# 1. Qu'est-ce qu'une DTD (Définition de type de document) ?

- Une sorte de « grammaire » du document d'encodage
- Parfois utilisée comme modèle de règles à appliquer à l'encodage (éléments et attributs utilisés, valeurs possibles, relations entre les éléments, etc.)<sup>1</sup>.
- Sert aussi à la déclaration et à l'affichage de glyphes particuliers (abréviations dans les textes anciens, symboles astronomiques, etc.).
- Elle peut être interne (déclarée à l'intérieur du document d'encodage) ou externe (dans un fichier à part qui est associé au document comprenant le texte encodé).

1. Cela revient à ce que nous faisons dans le schéma RNG que vous associez à l'encodage via la punaise rouge.

## 2. Comment ça marche ?

➤ Lorsque la DTD est interne, il faut la déclarer au début du document d'encodage de la manière suivante :

```
1 <?xml version="1.0" encoding="UTF-8"?>
2
3 <!-- DTD pour la déclaration des abréviations -->
4 <!DOCTYPE texte [
5 <!ENTITY q_que "<choice><abbr>&#59009;</abbr><expan>que</expan></choice>">
6 <!ENTITY a_tilde_n "<choice><abbr>&#227;</abbr><expan>an</expan></choice>">
7 <!ENTITY a_tilde_m "<choice><abbr>&#227;</abbr><expan>am</expan></choice>">
8 <!ENTITY z_et "<choice><abbr>z</abbr><expan>et</expan></choice>">
9 <!ENTITY co_com "<choice><abbr>c&#245;</abbr><expan>c<ex>om</ex></expan></choice>">
10 <!ENTITY e_tilde_n "<choice><abbr>&#275;</abbr><expan>en</expan></choice>">
11 <!ENTITY nre_nostre "<choice><abbr>nr&#771;</abbr><expan>n<ex>ost</ex>re</expan></choice>">
12 <!ENTITY e_tilde_m "<choice><abbr>&#275;</abbr><expan>em</expan></choice>">
13 <!ENTITY p_barre_pre "<choice><abbr>&#58981;</abbr><expan>pre</expan></choice>">
14 <!ENTITY qat_tilde "<choice><abbr>q&#227;t</abbr><expan>q<ex>uan</ex>t</expan></choice>">
15 <!ENTITY note_tyronienne_con "<choice><abbr>&#42863;</abbr><expan>con</expan></choice>">
16 <!ENTITY p_barre_per "<choice><abbr>&#42833;</abbr><expan>per</expan></choice>">
17 <!ENTITY mi_ser_icorde "<choice><abbr>mi&#59575;icorde</abbr><expan>mi<ex>ser</ex>icorde</expan></choice>">
18 <!ENTITY i_tilde_n "<choice><abbr>&#299;</abbr><expan>in</expan></choice>">
19 <!ENTITY q_barre_verticale "<choice><abbr>q&#781;</abbr><expan>qui</expan></choice>">
20 <!ENTITY p_barre_priere "<choice><abbr>&#58981;iere</abbr><expan><ex>pr</ex>iere</expan></choice>">
21 <!ENTITY o_tilde_n "<choice><abbr>&#245;</abbr><expan>on</expan></choice>">
22 <!ENTITY e_eur "<choice><abbr>&#7633;</abbr><expan><ex>ur</ex></expan></choice>">
23 <!ENTITY note_tyronienne_com "<choice><abbr>&#42863;</abbr><expan>com</expan></choice>">
24 <!ENTITY bn_bien "<choice><abbr>b&#781;n</abbr><expan>bien</expan></choice>">
25 <!ENTITY tre_terre "<choice><abbr>t&#781;re</abbr><expan>terre</expan></choice>">
26 <!ENTITY deps_depuis "<choice><abbr>de&#7765;s</abbr><expan>dep<ex>ui</ex>s</expan></choice>">
27 <!ENTITY tre_terre "<choice><abbr>&#59106;re</abbr><expan>terre</expan></choice>">
28 <!ENTITY note_tyronienne_us "<choice><abbr>&#42864;</abbr><expan>us</expan></choice>">
29 <!ENTITY seigneur "<choice>seigne<abbr>&#9676;&#771;</abbr><expan>seigne<ex>ur</ex></expan></choice>">
30 <!ENTITY vertu "<choice><abbr>&#59194;tu</abbr><expan><ex>ver</ex>tu</expan></choice>">
31 <!ENTITY ppoir "<choice><abbr>ppoir</abbr><expan>pouvoir</expan></choice>">
32 <!ENTITY erz_entiere "<choice><abbr>er&#732;z</abbr><expan>entiere</expan></choice>">
33 <!ENTITY dernier "<choice><abbr>der&#59364;</abbr><expan>der<ex>nier</ex></expan></choice>">
34 <!ENTITY po_pour "<choice><abbr>po&#61792;</abbr><expan>po<ex>ur</ex></expan></choice>">
35 <!ENTITY s_barré_ser "<choice><abbr>&#7836;</abbr><expan>ser</expan></choice>">
36 <!ENTITY double_p_lies "<choice><abbr>&#61143;</abbr><expan>pro</expan></choice>">
37 <!ENTITY mlt_moult "<choice><abbr>m&#58764;t</abbr><expan>moult</expan></choice>">
38 ]>
```

## 2. Comment ça marche ?

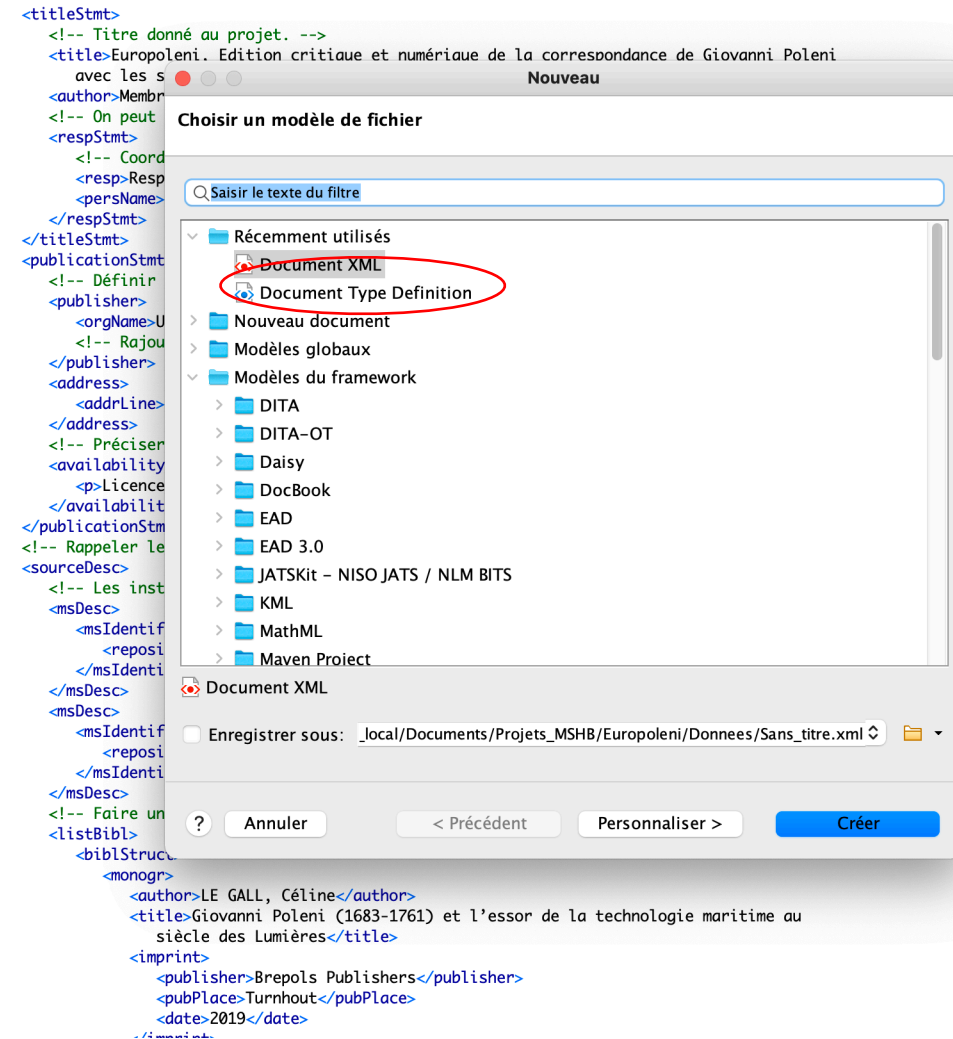
On remarque la structure suivante :

```
<!DOCTYPE Nom_donné_à_la_DTD [  
<!ENTITY Nom_donné_au_caractère "code_associé_au_caractère">
```

Puis on fait autant de balises <!ENTITY > que nécessaire.

## 2. Comment ça marche ?

➤ Lorsque la DTD est externe, on crée un document à part au format .dtd.



## 2. Comment ça marche ?

Le document créé doit adopter la syntaxe suivante :

```
1 <?xml version="1.0" encoding="UTF-8"?>
2
3 <!ENTITY mercure "&#9791;">
4 <!ENTITY soleil "&#9737;">
```

Il contient simplement la déclaration des caractères spéciaux.

## 2. Comment ça marche ?

Pour l'associer au document contenant le texte encodé, il faut ajouter la ligne suivante :

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <?xml-model href="http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_all.rng" type="application/xml"
3   schematypens="http://purl.oclc.org/dsdl/schematron"?>
4 <?xml-model href="out/ODD_Europoleni.rng" type="application/xml" schematypens="http://relaxng.org/ns/structure/1.0"?>
5 <?xml-model href="out/ODD_Europoleni.rng" type="application/xml" schematypens="http://purl.oclc.org/dsdl/schematron"?>
6 <?xml-stylesheet type="text/css" href="Customisation_mode_auteur.css"?>
7
8 <!DOCTYPE symbole SYSTEM "symboles.dtd">
9
10 <!-- Le <teiCorpus> englobe l'ensemble du corpus : les métadonnées du projet ; les différentes lettres encodées avec leurs m
11 <teiCorpus xmlns="http://www.tei-c.org/ns/1.0">
12   <!-- Métadonnées du projet : à compléter et enrichir, notamment les index et les références bibliographiques. -->
13   <teiHeader>
14     <fileDesc>
```

## 2. Comment ça marche ?

Cela revient à la forme suivante :

```
<!DOCTYPE élément_racine SYSTEM "Nom_du_fichier.dtd">
```

ATTENTION : Si le fichier contenant la DTD n'est pas dans le même répertoire que le fichier contenant l'encodage, il faut ajouter le chemin de fichier correct pour que le lien se fasse.



## 2. Comment ça marche ?

- Pour appeler les caractères définis préalablement dans la DTD au sein de l'encodage, il suffit de taper :

&mercure;

Ce qui revient à la syntaxe :

&nom\_entité;

### 3. Résultat

*Définir des valeurs d'attribut fixes pour @type et respecter codes du RFC 3066 pour @xml-lang.*

type de lettre transcription

langue it › Test affichage des symboles astronomiques. ‹ | ☿ ‹ et › ☉ ‹ › Exemple de formules

scientifiques avec MathML. ‹ › Type de notation : mathml ›  $E = mc^2$  ‹‹ › Type de notation : mathml

›  $\frac{1}{10} + \frac{1}{100} + \frac{1}{1000} + \frac{1}{10000} + \dots$  ‹‹ ‹‹ › Cf. premier élément <TEI>. ‹

Numéro de page : ??

# Pour aller plus loin

- Pinche (Ariane), Cours TNAH\_XML-TEI, séance 2 ([https://github.com/ArianePinche/coursTNAH\\_XML-TEI/blob/master/seance02/DTD.md](https://github.com/ArianePinche/coursTNAH_XML-TEI/blob/master/seance02/DTD.md))
- CHAGON (Gilles), Définir le format d'un document XML : DTD et Schémas, (<https://gilles-chagnon.developpez.com/cours/xml/dtd-et-schemas/?page=dtd>)
- XML - Déclaration d'entités dans les DTD, (<https://web.maths.unsw.edu.au/~lafaye/CCM/xml/xmltdtd2.htm>)