

Séance 5 : Initiation à l'édition scientifique

Les enjeux et les métiers de l'édition

Les éditions ont pour but de rendre accessible et compréhensible un texte pour un lectorat donné.

Il existe différentes formes d'édition (liste non exhaustive) :

- Les éditions génétiques
 - [Madame Bovary : l'histoire du texte à travers ses brouillons](#)
- Les éditions "paléographiques"
 - [Le Didascalicon d'Hugues de Saint-Victor](#)
- Les éditions critiques
 - [La Queste del saint Graal](#)

Les enjeux d'une édition "graphétique"

Geste: un corpus de chansons de geste,

<http://dev.chartes.psl.eu/elec/geste/>

Corpus test : La séquence de sainte Eulalie

- Lecture interactive : <https://bibliotheque.ville-valenciennes.fr/iguana/www.main.cls?v=6bb63866-fa45-4563-a75e-626d3cccb43c>
- Voir le manuscrit sur Gallica : <https://gallica.bnf.fr/ark:/12148/btv1b84526286/f288.item>
- Obtenir le texte sur wikisource : https://fr.wikisource.org/wiki/Séquence_de_sainte_Eulalie

Pour aller plus loin, voir l'article : *Paléographie statistique pour décrire, identifier, dater... Normaliser pour coopérer et aller plus loin ?* de

Dominique Stutzmann, en ligne : <https://halshs.archives-ouvertes.fr/halshs-00596970/document>

Définition des enjeux

- Identifier et délimiter les phénomènes à signaler;
- Définir les futurs usages de l'édition;

Les éléments que l'on veut reproduire

- Mise en page;
- Mise en valeur typographique;
- Graphie originale du texte.

Les éléments que l'on veut apporter :

- Résolutions des abréviations;
- Régularisations des graphies;
- Corrections (si nécessaires).

Les solutions TEI

Le set pour la **structuration des textes**

Quelques balises courantes

- `<lg>` contient un groupe de vers fonctionnant comme une unité formelle, par exemple une strophe, un refrain, un paragraphe en vers;
- `<l>` contient un seul vers;
- `<lb>` marque le début d'une nouvelle ligne;
- `<pb>` marque le début d'un nouveau folio; @facs permet de lier l'encodage à une image du folio;
- `<cb>` marque le début d'une nouvelle colonne.

Exercice 1

À l'aide des documents sur la *Séquence de sainte Eulalie* :

- Structurer le texte selon ses caractéristiques littéraires;
- Encoder la structuration matérielle du texte dans le manuscrit;
- Lier le texte à l'image correspondante dans Gallica.

la description des sources primaires

Quelques balises courantes

- Développements d'abréviation : `<choice><abbr></abbr><expn></expn></choice>` ;
- Régularisations : `<choice><orig></orig><reg></reg></choice>` ;
- Corrections : `<choice><sic></sic><corr></corr></choice>` .

NB : Dans une édition numérique, il est conseillé d'utiliser les caractères UTF-8. Il existe des [fontes spécialisées](#) dans la représentation des manuscrits :

- Palemonas MUFI
- Junicode

Pour les caractères qui n'existent pas dans toutes les fontes, passer par l'entité hexadécimale XML.

Astuce : convertisseur pour les caractères <https://r12a.github.io/app-conversion>

Traitement des variations graphiques dans un fichier TEI

Le module gaiji

```
<encodingDesc>
  <charDecl>
    <char xml:id='q1'>
      <charName>Q SEMICOLON</charName>
      <charProp>
        <localName>entity</localName>
        <value>&#61868;</value>
      </charProp>
    </char>
  </charDecl>
</encodingDesc>
```

[...]

```
<l>Niule cose non la pouret om<g ref="#q1">que</g> pleier.</l>
```

Déclaration d'entités

Exemple :

```
<!DOCTYPE TEI [  
  <!ENTITY pbarre-pre  
    "<choice><abbr>&#58981;</abbr>  
    <expn>pre</expn></choice>">  
]>
```

La déclaration se fait dans une DTD interne, soit dans une DTD externe.

Exercice 2

- Reprendre le fichier XML de la *Séquence de sainte Eulalie*;
- Encoder les graphies originales du manuscrit et leur normalisation.

Lier son encodage à des zones d'un fac-similé

Ex : *Le Didascalicon* d'Hugues de Saint-Victor :

<http://theleme.enc.sorbonne.fr/dossiers/vue100.php>

Ex : Lecture interactive de la *Séquence de sainte Eulalie*

<https://bibliotheque.ville-valenciennes.fr/iguana/www.main.cls?v=6bb63866-fa45-4563-a75e-626d3cccb43c>

Comment faire ?

TEI guidelines [11.1 Digital Facsimiles](#)

- Déclarer un élément `<facsimile>` dans l'élément racine TEI;
- Déclarer en élément `<graphic>`, enfant de `<facsimile>`
 - `<graphic url='adresse_image' width='largeur' height='longueur' />`

- Déclarer des éléments `<zone>` , enfants de `<facsimile>` , pour chacune des parties de l'image à identifier
 - `<zone xml:id="<!-- identifiant -->" ulx="<!-- coordonnée -->" uly="<!-- coordonnée -->" lrx="<!-- coordonnée -->" lry=" <!--coordonnée -->" />`

@	définition
<i>ulx</i>	donne la valeur x de l'abscisse du coin supérieur gauche d'un rectangle.
<i>uly</i>	donne la valeur y de l'ordonnée du coin supérieur gauche d'un rectangle.
<i>lrx</i>	donne la valeur x de l'abscisse du coin inférieur droit d'un rectangle.
<i>lry</i>	donne la valeur y de l'ordonnée du coin inférieur droit d'un rectangle.

Astuce pour les coordonnées : [TElzonner](#)

Exercice 3

- Reprendre le fichier XML de la *Séquence de sainte Eulalie*;
- Récupérer l'image du folio où se trouve le texte;
- Lier l'encodage à des zones du folio.

Décrire les sources manuscrites

EAD (Encoded Archival Description)

EAD est une DTD (un schéma d'utilisation) de XML.

- 2000, encodage en XML EAD du catalogue général des manuscrits des bibliothèques publiques de France.
- 2010, la DeMarch recommandation est publiée.
 - Attention : DeMArch est une règle de description indépendante du format informatique.
- 2012, deux tiers des manuscrits et documents d'archives conservés par les bibliothèques françaises sont décrits en EAD (« BnF archives et manuscrits », Calames, le sous-domaine « Manuscrits » du Catalogue collectif de France CCFr).

Quelques exemples

- [Inventaire du fonds du Théâtre de Poche-Montparnasse \(1942-2011\)](#)
 - [Voir le fichier XML](#)
- [Catalogue du fonds général, Ms 1 à Ms 721](#)
 - [Voir le fichier XML](#)

Pour résumer

Extrait de Florent Palluault, « Informatiser des descriptions complexes : l'utilisation de l'EAD et de la TEI pour les manuscrits et les livres anciens en France », *IFLA 2012 (Helsinki)*, en ligne : <https://www.ifla.org/past-wlic/2012/212-palluault-fr.pdf>

"La nécessité d'une arborescence est moindre pour ces documents que pour des ensembles d'archives, et la sémantique EAD n'est pas aussi développée que celle de la TEI-P5, utilisée pour les manuscrits médiévaux dans plusieurs pays d'Europe."

Au-delà des manuscrits et archives, **l'EAD est pertinente pour tout ensemble de documents hiérarchisés, quelle que soit leur nature. Elle permet de visualiser aisément la structure de fonds ou de collections composites** comme celles du Département des Arts du Spectacle de la BnF, qui comprennent fréquemment scénarios manuscrits, programmes imprimés, costumes, objets de décor, etc.

Naissance du msDesc

- 1999-2001 : projet européen MASTER (Manuscript Access through Standards for Electronic Records);
- 2001, la DTD Master, personnalisation de la TEI qui complète par l'ajout d'un élément msDescrption, et d'autres éléments spécialisés qui représentent un enrichissement des possibilités pour la description des manuscrits.
- 2007-2009 : lancement du projet ENRICH (European Networking Resources and Information concerning Cultural Heritage)

Les éléments TEI de description des sources : `<msDesc>`

10.2 Manuscript Description Element

[Voir le fichier TEI](#)

Mise en pratique

- Reprendre le fichier de la séquence de sainte Eulalie
- Compléter les éléments du `<msDesc>` à l'aide des renseignements accessibles sur le site de la bibliothèque de Valencienne et sur Gallica.