

TP1 : oXygen

TEI@Oxford

octobre 2012

1 À la découverte d'oXygen

Dans ce premier exercice, nous allons nous servir d'oXygen pour :

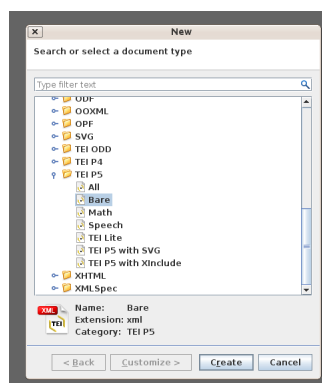
- la création d'un nouveau document XML
- le balisage progressif d'un document
- la validation continue d'un document
- l'affichage du document sans balise

2 Le document né-numérique

La TEI ne sert pas uniquement à l'encodage des documents anciens, déjà existants sur papier ou pierre. Elle sert aussi à la création de documents nouveaux *ex nihilo*....

3 Création d'un document

- Démarrez oXygen.
- Cliquez sur l'icône Nouveau en haut à gauche (ou sélectionnez Nouveau du menu Fichier, ou tapez CTRL-N) pour ouvrir la boîte de dialogue Nouveau
- Choisissez Cadre des modèles, ensuite TEI P5, puis TEI-Bare, pour sélectionner un schéma basique TEI.



- Cliquez sur le bouton Créer en bas. oXygen affiche un document TEI minimal que vous allez compléter.

Le schéma TEI bare contient un minimum de balises, mais il permet quand-même de décrire un document complet TEI.

On est maintenant dans l'espace de travail typique d'oXygen. Par défaut, il y a plusieurs petites fenêtres, autour d'une grande blanc où nous allons éditer notre document. À gauche, par exemple, vous voyez dans la fenêtre Sommaire un aperçu de la structure du document.

Notez qu'un document TEI est divisé en deux parties. En haut, vous avez les métadonnées, entourées de la balise `<teiHeader>`. En bas se trouve le texte, entouré de la balise `<text>`.



En bas de l'écran, vous voyez qu'il est possible de modifier l'affichage du document en cliquant sur les boutons *Texte*, *Grille*, ou *Auteur*. Nous, on n'a pas peur des balises, donc on va commencer à l'éditer en mode *Texte*.

D'abord, on va fournir les métadonnées essentielles à tout nouveau document : son titre, sa distribution, et ses origines. oXygen vous propose des morceaux de texte qu'il faut changer puis compléter...

- Remplacez le mot 'Title' (entre les balises `<title>` et `</title>`) par un titre pour votre document. À vous de choisir ! Quelque chose comme "À la découverte d'oXygen" suffirait.
- Remplacez les deux autres morceaux de texte proposés par des mentions plus appropriées. Par exemple, pour publication, on pourrait mettre 'Brouillon inédit' ou 'Distribué sur mon blog' ; pour `<sourceDesc>`, on vous propose 'Œuvre originale, guidé par une formation TEI'.

oXygen par défaut cherchera à maintenir votre document dans un état de validité. Qu'est-ce qui arrive si vous l'endommagez ?

- Effacez quelques lettres à l'intérieur de n'importe quelle balise : par exemple changez `<publicationStmt>` en `<publicat>`.
- oXygen fait de son mieux. D'abord, il change la balise fermante correspondante (qui devient instantanément `</publicat>`): le document reste donc bien formé. Mais il n'est plus valide : par conséquent les balises fautives sont soulignées en rouge ; il y a un message d'erreur en bas de l'écran ; et à droite il y a quelques taches rouges.
- Cliquez sur les taches rouges pour essayer de comprendre la manière dont une erreur peut en entraîner une autre.
- Cliquez sur CTRL-Z (ou sélectionnez **Undo** du menu **Edit**) pour revenir dans un état de grâce... vous verrez que le petit carreau rouge en haut à droite devient vert après correction de toutes les erreurs.

En plus, oXygen ne vous permettra d'ajouter que des balises prévues dans le schéma actuellement sélectionné pour le document.

- Mettez le curseur à la fin de la 7ème ou au début de la 8ème ligne, entre `</title>` et `</titleSmt>`;
- tapez `<` et arrêtez un instant ;
- le carreau vert redevient rouge, parce que vous n'avez pas encore entré le nom de l'élément souhaité.

- oXygen vous propose une liste de balises disponibles à cet emplacement dans le document : notamment vous avez le droit d'insérer ici un élément `<author>` ou encore un `<title>`, de terminer le `<titleStmt>`, d'insérer un commentaire XML etc.
- Tapez sur ENTRER pour accepter la première proposition. oXygen insert dans votre document les balises `<author>` et `</author>`.
- Le curseur reste à l'intérieur de la balise ouvrante, au cas où vous désireriez spécifier ses attributs. Faire bouger le curseur d'un caractère à droite et tapez votre nom, comme auteur du fichier.

4 Ajout de texte au document

Enfin, nous sommes prêts ... commencez la création de votre document. À vous de décider de son contenu et de son balisage.

D'abord, quels composants proposeriez-vous pour votre document ? Un document contenant un seul paragraphe n'est guère probable ! Avec le schéma *bare* vous avez le droit de signaler :

- des titres ... à baliser avec la balise `<head>`
- des chapitres ... à baliser avec la balise `<div>`
- des listes ... à baliser avec la balise `<list>`
- des paragraphes ... à baliser avec la balise `<p>`

Allons-y !

Dans le reste de ce petit tutoriel, nous allons vous montrer quelques manipulations utiles offertes par ce logiciel.

D'abord, une question importante. Pour le contenu, OK, il faut le créer – mais est-ce qu'il faut toujours taper soi-même les balises ? Non : oXygen est là pour vous aider. Vous avez déjà vu comment il peut vous proposer la liste des balises disponibles à un certain emplacement du document. Il peut aussi vous proposer comment baliser un morceau de texte sélectionné.

- Avec la souris, sélectionnez le contenu du `<body>`, i.e. la séquence `<p>Some text here</p>`.
- Tapez CTRL-E (ou sélectionnez XML Refactoring et ensuite Entourer des balises dans le menu Document).
- oXygen vous propose toutes les balises qui pourraient éventuellement être insérées à cet endroit précis du document : dans ce cas, il est conseillé de sélectionner `<div>`. Puis, cliquez sur Accepter
- L'élément `<div>` va contenir le premier chapitre de votre document. Insérez donc un titre au début de ce `<div>`, en vous servant de la balise `<head>`
- Il serait aussi utile de numéroter vos chapitres d'une manière explicite : pour cela, on vous propose l'attribut `n` qui est disponible sur n'importe quel élément TEI, y compris `<div>`.
- Mettez le curseur à l'intérieur de la balise `<div>` juste avant le `>` et tapez un seul espace. oXygen vous propose un menu des attributs disponibles.
- Tapez sur RETOUR pour sélectionner `n`. oXygen vous informe qu'il faut dans ce cas ajouter une valeur, et le curseur reste entre les guillemets.
- Tapez (par exemple) `1`, et votre document redeviendra valide.

Votre document devrait maintenant ressembler à ceci :

```
<TEXT>
  <body>
    <div n="1">
      <head>Chapitre premiere</head>
      <p>Some text here</p>
    </div>
  </body>
```

À vous d'ajouter maintenant du texte à ce paragraphe : un peu plus intéressant que 'some text here' de préférence ! Décrivez-nous vos premières réactions avec oXygen ; partagez vos idées sur la situation politique actuelle ; écrivez ce que vous voulez.

Sans doute aurez-vous besoin de plus d'un seul paragraphe. Comment faire pour terminer celui-ci et commencer un nouveau ? Bien-sûr on peut simplement taper les balises `</p><p>` mais il y a une autre manière plus simple.

5 Division des éléments

Supposons que vous avez maintenant quelques centaines de mots dans votre seul paragraphe, et que vous voulez le diviser en plusieurs.

- Mettez le curseur au point de chaque division souhaitée
- Tapez ALT + MAJ + D ou sélectionnez XML Refactoring et ensuite Élément de division (sic) dans le menu Document.

Vous pouvez faire pareil pour diviser le chapitre unique, mais attention à choisir un endroit valide pour effectuer la division : *entre* deux éléments `<p>`, et non pas dans l'enceinte d'un `<p>`.

Dans un de vos chapitres, n'omettez pas d'ajouter une liste, balisée avec un élément `<list>`, contenant une séquence de `<item>` éléments, précédée d'un `<head>` facultatif.

6 Affichage du texte balisé

Est-ce que la quantité de balises dans votre document commence à vous inquiéter un peu ?

On peut mettre en ordre le balisage que l'on a fait, en cliquant sur le bouton d'indentation, ou en tapant ALT+MAJ+P, ou en sélectionnant Source->Indenter le document sur le menu Document. Mais peut être préféreriez-vous voir disparaître les balises ?

- Au fond de la fenêtre vous voyez trois onglets: Texte , Grille, et Auteur. Sélectionnez le dernier. L'affichage de votre document change, et un nouveau menu TEI P5 devient disponible. Sur ce menu, sélectionnez Mode d'affichage des balises et regardez les possibilités offertes par ses options.
- L'option Toutes les balises affiche les balises en icones; l'option Aucune balises les fait disparaître totalement.
- Notez en plus que votre document est maintenant stylé ; il y a des changements de police ; les retours de chariots sont traités comme des blancs ; etc.
- Dans mode auteur, on peut diviser un élément simplement en tapant RETOUR deux fois. Expérimentez cela.

Plus tard, nous verrons comment vous pourrez contrôler l'affichage de votre document avec une *feuille de style*.

7 Importation d'un document bureautique

La plupart des documents numériques que vous rencontrerez dans le monde actuel semble innocent de toute balise XML. Par exemple, les documents préparés à l'aide des outils bureautiques tels Microsoft Word ou Open Office ont l'apparence d'être de texte pure. Mais effectivement, ce n'est qu'une apparence: derrière tout cela il existe un format XML, que l'on peut manipuler, et donc qu'on peut transformer dans un format XML TEI simple.

Dans votre dossier Travaux il y a un fichier `travprat1.odt` : ce document est créé avec Open Office. Cliquez là-dessus pour l'ouvrir. Sans doute vous le reconnaîtrez. On va se servir d'un outil online TEI-Roma pour le transformer en XML TEI.

Nous pourrions exporter un fichier bureautique au format « plein text ». Nous pourrions aussi faire un copier-coller. Mais dans chaque cas, nous perdrons le formatage qui distingue (par exemple) les titres des sections.

Si le fichier bureautique est bien stylé, on peut le transformer en XML TEI, sans perte d'information. L'outil OxGarage nous aidera.

- Ouvrez votre navigateur Internet, et allez sur le site <http://www.tei-c.org/ege-webclient/>
- Cliquez sur **Text Documents**. Dans la liste des formats de saisie proposée, cochez la case pour Open Office.
- Une liste des formats cibles proposés apparaît. Cochez la case pour TEI P5 XML.
- En haut de l'écran un bouton **Browse** apparaît. Cliquez sur ceci, et naviguez jusqu'au fichier `travprat1.odt` dans votre dossier Travaux.
- Cliquez sur le grand bouton **Convert** et patientez.
- Le site vous renvoie (après un bref délai) un fichier `travprat.xml`. Enregistrez-le dans votre dossier Travaux.
- Ouvrez ce fichier avec Oxygen, et considérez si la conversion a été effectuée correctement.

Si vous disposez d'autres documents Word ou ODT, n'hésitez pas d'essayer de les transformer de la même manière!