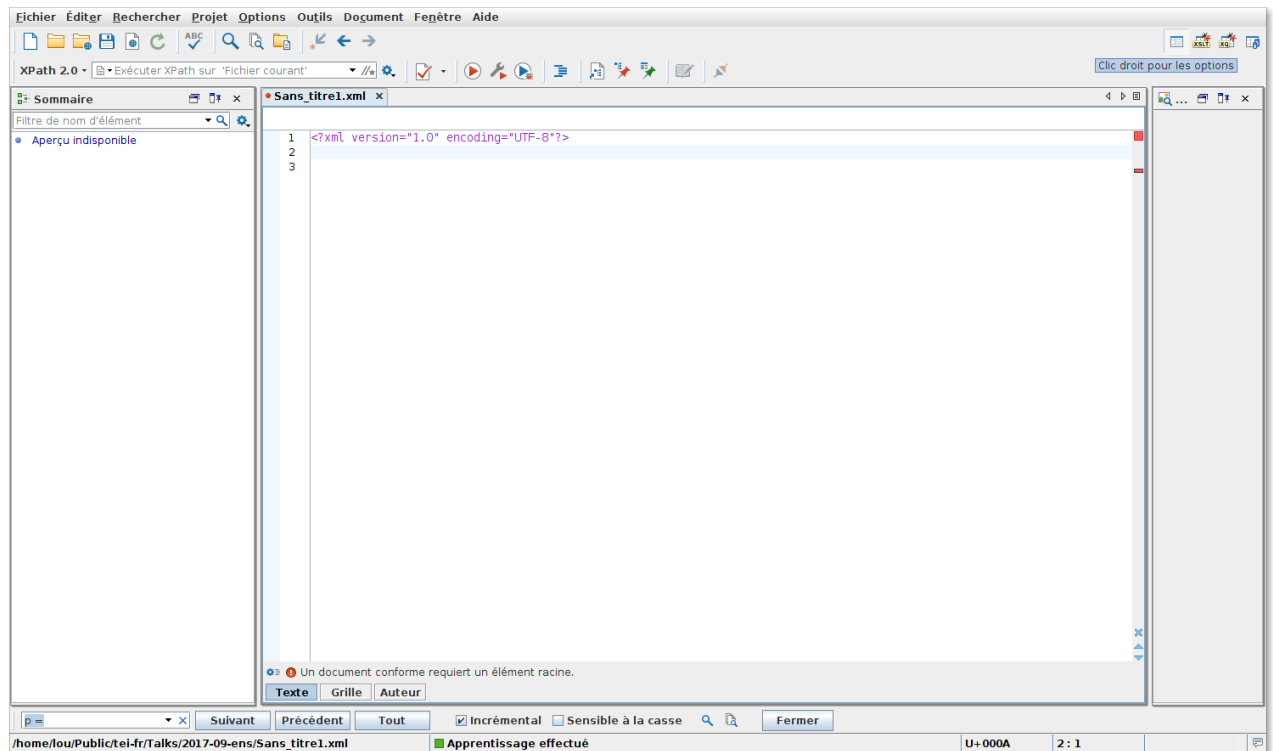
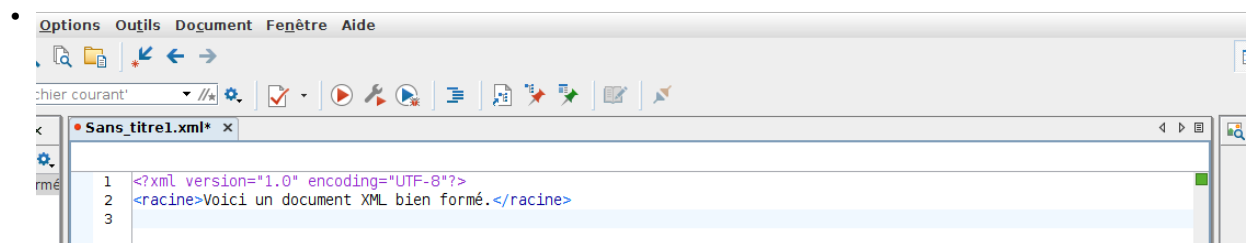


1. Création d'un document XML

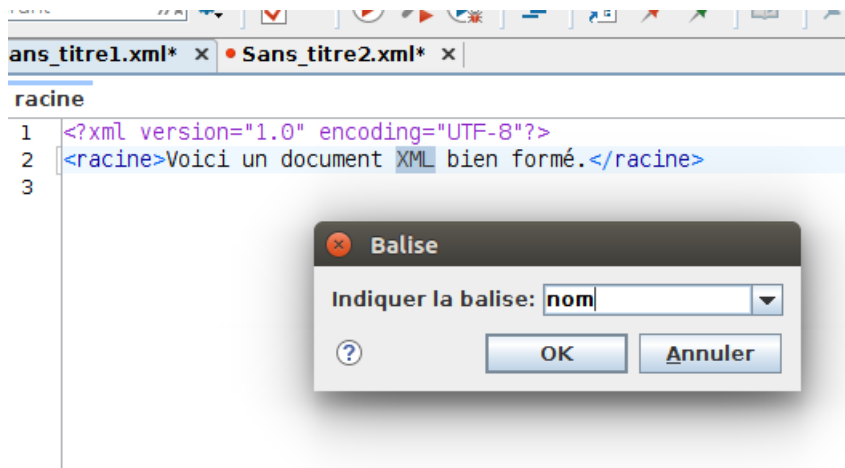
- Démarrez oXygen.
- Cliquez sur l'icône **Nouveau** en haut à gauche (ou sélectionnez Fichier Nouveau du menu, ou tapez CTRL - n) pour ouvrir la boîte de dialogue Nouveau
- Dans ce dialogue, sélectionner **Nouveau Document - Document XML** et cliquer le bouton **Créer**.
- Vous êtes maintenant face à l'espace de travail typique d'oXygen. Par défaut, il y a plusieurs petites fenêtres, autour d'une grande blanc où vous allez éditer votre document.



- Le petit carreau rouge en haut à droite signale un problème. Regardez bien le message en bas : 'Un document conforme requiert un élément racine' OK. Tapez la balise `<racine>` : Oxygen va compléter cela avec une balise fermante `/racine` aussitôt que vous tapez `>`. Votre document est maintenant valide : le petit carreau rouge devient vert. Fournissez un peu plus d'intérêt en ajoutant du contenu à cet élément, par ex, tapez 'Ceci est un document XML bien formé.'



- Oxygen dispose de plusieurs méthodes pour le balisage d'un document. Essayons un deuxième. Sélectionner le mot XML avec le souris et tapez CTRL-E. Tapez un nom pour la balise souhaitée: nous proposons `nom`



Cliquer OK

- Les éléments s'imbriquent: nous avons un `<racine>` qui contient un morceau de texte, suivi d'un `<nom>`, suivi d'un autre morceau de texte. Maintenant sélectionnez tout le contenu du `<racine>` et le baliser `<phr>` (pour phrase)

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <racine><phr>Voici un document <nom>XML</nom> bien formé.</phr></racine>
3
```

- Une troisième manière de baliser un document est de le diviser en morceaux automatiquement. Mettez le curseur à la fin de l'élément `<phr>` que vous venez de créer (entre `formé.` et `/phr`) et tapez ALT-MAJ-d. Qu'est ce qui arrive? Vous pouvez maintenant ajouter une deuxième phrase à votre document.

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <racine><phr>Voici un document <nom>XML</nom> bien formé.</phr><phr>Une deuxième phrase.</phr></racine>
3
```

- Il y a également plusieurs méthodes pour ajouter des attributs. Mettez d'abord le curseur à l'intérieur de la première balise `<phr>` juste avant le `>` et tapez un blanc, suivi de `n="`. Oxygen ajoute un `"` et le curseur reste immobile. Tapez le chiffre 1.

```
version="1.0" enco
:><phr n="1">Voici
```

- Numérotez la 2ème phrase "2" de la même manière.
- Oxygen fait de son mieux pour maintenir le document en bon état. Expérimentez cela. Si vous proposez un nouveau élément, en tapant une balise ouvrante (`<wibble>` par exemple), il va vous proposer tout de suite la balise fermante correspondante. Si vous changez le nom de la balise ouvrante, il va changer tout de suite le nom de la balise fermante qui lui correspond. Si vous changez le nom d'une balise fermante, par contre, Oxygen signale une erreur. Pour annuler vos modifications, il suffit de taper CTRL-z
- Finalement, on va faire le doc un peu plus joli. Tapez CTRL-MAJ-P ou sélectionnez.

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <racine>
3   <phr n="1">Voici un document <nom>XML</nom> bien formé.</phr>
4   <phr n="2">Une deuxième phrase.</phr>
5 </racine>
6
```

Petit rappel. Pour baliser un document :

- je tape une balise ouvrante directement au bon endroit et oxyGen fournit sa balise fermante correspondante
- je sélectionne un morceau de text, je tape CTRL-E, et je sélectionne la balise qu'il me faut

- je mets mon curseur à l'intérieur d'un élément, (i.e. entre une balise ouvrante et une balise fermante) je tape ALT-MAJ-D; oxyGen divise l'élément en deux
- j'ajoute les attributs à l'intérieur d'une balise ouvrante

2. Création d'un document TEI

- Démarrez oXygen.
- Cliquez encore 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
- Cette fois, nous allons utiliser un schéma TEI pour contrôler la disponibilité des balises. Dans ce dialog rechercher la mention **Modèles du framework**, sélectionnez **TEI P5**, et ensuite **Bare** de la longue liste des schemas connus par oXygen.
- oXygen affiche un document TEI conforme à ce schema que vous allez compléter.

2.1. Ajout des métadonnées

Notez d'abord dans la fenêtre **Sommaire** un aperçu de la structure d'un document TEI. Il 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>`.



Tout schéma TEI exige trois types de métadonnées : un titre, une information sur l'édition, et une information sur l'origine du document. Pour cet exercice, nous vous proposons

- Entre les balises `<title>` et `/title>` ajouter un titre pour votre document: **Exercice de balisage guidée** par oXygen par exemple.
- Entre les balises `<p>` et `/p>` suivant la balise `<publicationStmt>` ajouter la mention 'Exercice d'encodage non publié'
- Entre les balises `<p>` et `/p>` suivant la balise `<sourceDesc>` ajouter cette mention 'Aucune source pré-existante'

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

- Mettez le curseur après la balise fermante `</title>` et avant la balise fermante `</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 `<titleSmt>`, d'insérer un commentaire XML etc.

- Tapez sur ENTRÉE pour accepter la première proposition. oXygen insère 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. Faites bouger le curseur d'un caractère vers la droite et tapez votre nom, comme auteur du fichier.

2.2. Ajout du texte au document

- Avec la souris, sélectionnez le contenu du `<body>`, i.e. la séquence `<p>Some text here</p>`, y compris les balises..
- 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 ENTRÉE 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.

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-dou 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.

2.3. L'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 CTRL-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 ces options.
- L'option **Toutes les balises** affiche les balises en icônes; l'option **Aucune balise** les fait disparaître totalement.
- Notez en plus que votre document est maintenant stylisé ; il y a des changements de police ; les retours de ligne sont traités comme des blancs ; etc.
- Dans le mode auteur, on peut diviser un élément simplement en tapant **ENTRÉE** deux fois. Essayez cela.

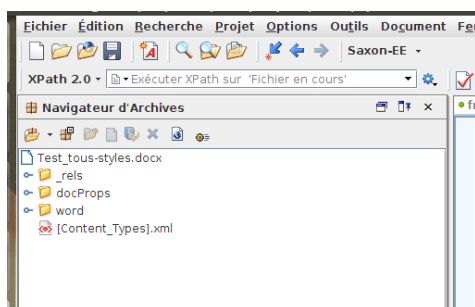
3. Importation d'un document bureautique

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 donnent l'apparence d'être du texte brut. 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.

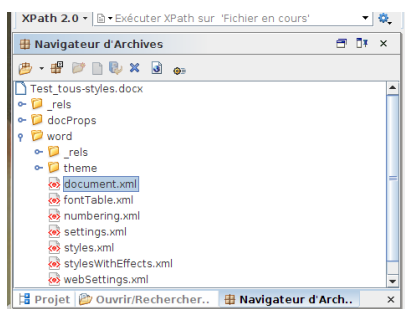
Dans votre dossier Travaux il y a un fichier **lacigale.docx** : cliquez là-dessus pour l'ouvrir. Sans doute vous le reconnaitrez. On va se servir d'oXygen pour le transformer en XML TEI.

Un fichier Word docx est une archive comprimée qui contient plusieurs autres fichiers, y compris des fichiers XML qu'on peut gérer et transformer avec oXygen comme n'importe quel autre fichier XML. La conversion prend en compte les styles Word existant dans le fichier, et donc peut être personnalisée.


- Chargez oXygen. Sélectionnez **Ouvrir** dans le menu **Fichier** (ou tapez **CTRL - O**, ou cliquez l'icône dossier) et sélectionnez le fichier **lacigale.docx** dans votre dossier **Travaux** .
- Une fenêtre intitulée **Navigateur d'Archives** s'ouvre à gauche de l'écran principal. La structure de l'archive docx y est affichée.



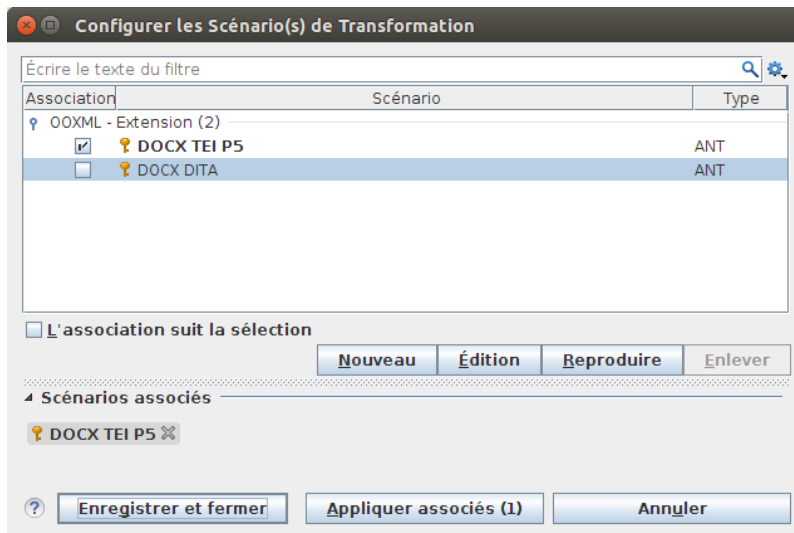
- Cliquez sur la petite clef bleue à côté du dossier **word** pour voir ce qu'il contient




- Sélectionnez le fichier **document.xml** et double-cliquez pour l'ouvrir (peut prendre un peu de temps si le fichier est grand)
- Voilà : c'est un document XML, plein de balises, pour la plupart appartenant à un espace de noms définie par Microsoft. Convertissons-le !
- Avec ce fichier **document.xml** ouvert dans l'écran principal, sélectionnez **Transformation -> Configurer les Scénario(s) de Transformation** du menu **Document**. Ou tapez **CTRL - MAJ - C**. Ou cliquez sur la petit

icone clef à molette (celui-ci : )

- oXygen vous présente les options de transformations disponibles pour ce document, parmi lesquelles une transformation nommée DOCX TEI P5 devrait figurer.



- Cochez la case à côté et cliquez sur le bouton **Appliquer associés** au dessous. Une nouvelle fenêtre s'ouvre en bas pour l'affichage des messages concernant le processus de transformation: vous pouvez les ignorer sans problème.
- Dans l'écran principal, un document TEI-conforme s'affiche. Sélectionnez **Source -> Format and Indent** du menu **Document** menu, ou tapez CTRL - MAJ - p ou cliquez sur le bouton **Indent** (celui-ci : ) pour regarder sa structure
- L'attribut *@rend* indique le nom du style Word utilisé dans cette portion du document original. Une transformation XSLT plus complexe pourrait donc s'en servir pour effectuer une conversion plus intelligente.