

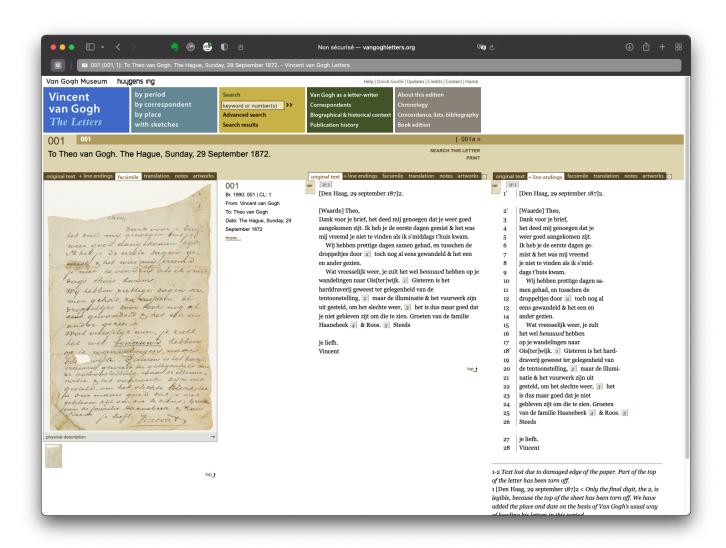
Université de Neuchâtel Master en littérature

Édition critique et génétique numérique

Élodie Paupe, chargée d'enseignement elodie.paupe@unine.ch



L'exemple des lettres de Vincent Van Gogh





Leo Jansen, Hans
Luijten, Nienke Bakker
(eds.) (2009), Vincent
van Gogh - The
Letters. Amsterdam &
The Hague: Van Gogh
Museum & Huygens
ING.

http://vangoghletters. org/vg/letters/let001/l etter.html.



Derrière le site web

- Une édition basée sur des fichiers XML encodés selon un schéma personnalisé qui respecte les stardards de la TEI.
- Un fichier XML-TEI par lettre qui contient:
 - des métadonnées
 - les liens vers les facsimilés
 - la transcription de la lettre
 - la traduction
 - les notes
- Transformation de ces fichiers XML-TEI en pages HTML pour permettre la lecture.



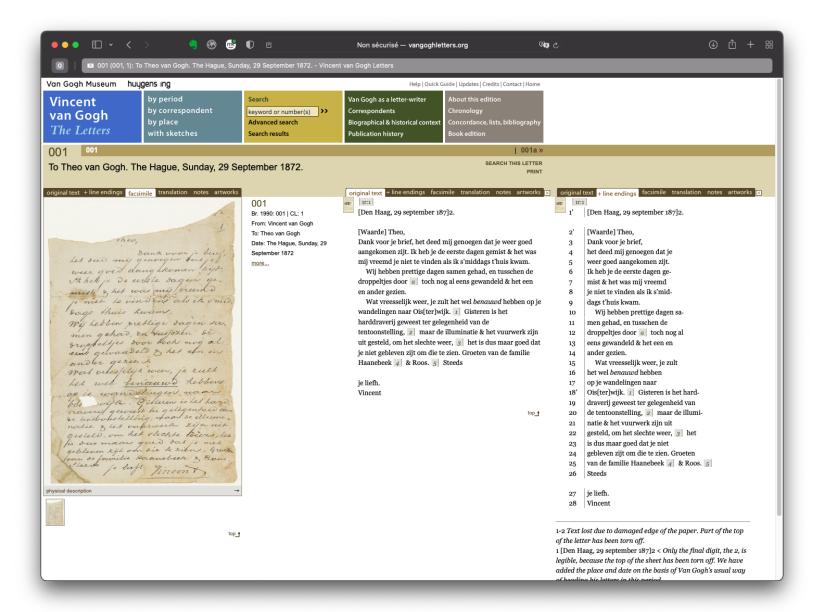
Fichier 1

```
DITA XSLT XQ I
                                                                                              - Q. | 🗸 - | ABC | 🕟 🔏 🕞 | 🖹 | 🧏 🦻 | 🖭 | 🗡
  XPath 2.0 ▼ 

Exécuter XPath sur 'Fichier courant'
   • Caligula_corr.xml X • let001.xml X
                                                                                                                                                                                                                                                                                                                                      4 ▷ 🗉
      1 <?xml version="1.0" encoding="UTF-8"?>
      <!-- Last Changed Date: 2020-01-30 22:42:22 +0100 (Thu, 30 Jan 2020) | SVN Revision: 2224 -->
       3 <!--Licensed under Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)-->
       5 VETEL xmlns:vg="http://www.vangoghletters.org/ns/" xmlns="http://www.tei-c.org/ns/1.0">
                    <teiHeader xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xml:lang="en"><fileDesc>
                          <titleStmt>
                                 <title>To Theo van Gogh. The Hague, Sunday, 29 September 1872.</title>
                                 <editor>Leo Jansen</editor>
                                 <editor>Hans Luijten</editor>
     11
                                  <editor>Nienke Bakker</editor></titleStmt>
     12
                           <publicationStmt>
     13
                                 <publisher>Van Goah Museum</publisher>
     14
     15
                                 <pubPlace>Amsterdam</pubPlace>
                                 <date type="first" when="2009">2009</date>
     16
     17
                                 <availability status="restricted">
                                        target="http://creativecommons.ora/licenses/by-nc-sa/4.0/ https://creativecommons.ora/licenses/by-nc-sa/4.0/leaalcode">
     19
                                              Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) 
     20
                                 <ptr target="http://vangoghletters.org/orig/let001"/></publicationStmt><sourceDesc><vg:letDesc><vg:letIdentifier>
     21
                                        <idno type="jlb">001</idno>
     22
                                        <idno type="collectedletters">1</idno>
     23
                                        <id><idno type="brieven1990">001</idno></vg:letIdentifier><vg:letHeading><author>Vincent van Gogh</author><vg:addressee>Theo van Gogh</vg:addressee><vg:placeLet>The Hague</vg:placeLet>Theo van Gogh</ra>
     24
                                              To Theo van Gogh. The Hague, Sunday, 29 September 1872.</vg:letContents>
     25 🗢
                                         <note type="sourceStatus" xml:id="sourceStatus">
     26
                                              Original manuscript</note>
     27 ▽
                                        <note type="location" xml:id="location">
                                              Amsterdam, Van Gogh Museum, inv. no. b1 V/1962</note>
     28
     29
                                         <note type="date" xml:id="date">
                                               Letter headed: '<hi rend="ital">[</hi>). The date is based on the harness racion
1 31 ▽
                     <facsimile xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
                          <surface n="b 0001 r" xml:id="surf-pb-1r"><graphic mimeType="image/tiff" url="VGM001000001_01.TIF" rend="norend"/><zone xml:id="zone-pb-1r-1">
     32 🗢
     33
                                 <graphic mimeType="image/tiff" url="VGM001000001_01_n.tif" rend="norend"/><graphic mimeType="image/jpeg" url="VGM001000001_01_nt.jpg" height="50px" rend="thumbnail"/><graphic mimeType="image/jpeg" url="vGM001000001_01_nt.jpg" height="50px" rend="thumbnail"/>
     34
                          <surface n="b 0001 v" xml:id="surf-pb-1v"><graphic mimeType="image/tiff" url="VGM001000001_02.TIF" rend="norend"/><zone xml:id="zone-pb-1v-1">
     35
                                 <qraphic mimeType="image/tiff" url="VGM001000001_02_n.tif" rend="norend"/><qraphic mimeType="image/jpeg" url="VGM001000001_02_n.tipg" height="50px" rend="thumbnail"/><qraphic mimeType="image/jpeg" url="VGM001000001_02_n.tipg" height="50px" rend="thumbnail"/><qraphic mimeType="image/jpeg" url="VGM001000001_02_n.tipg" height="50px" rend="thumbnail"/><qraphic mimeType="image/jpeg" url="VGM001000001_02_n.tipg" rend="thumbnail"/><qraphic mimeType="image/jpeg" url="thumbnail"/><qraphic mi
     37
                          </surface></facsimile>
     38
                    <text xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xml:lang="en"><body>
     39 🗢
                          <div xml:lang="nl" type="original">
                                 <pb f="1r" n="1" xml:id="pb-orig-1r-1" facs="#zone-pb-1r-1"/><lb n="1" xml:id="l-1"/>
     40
     41
                                 <ab><supplied>Den Haag, 29 september 187</supplied>2.</ab>
                                                                                                                                                   Texte Grille Auteur
/Users/elodiepaupe/Downloads/vangoghxml/let001.xml
                                                                                                  Validation réussie
                                                                                                                                                                                                                        ■ U+003C
                                                                                                                                                                                                                                                           5:1
                                                                                                                                                                                                                                                                                                                       38 nouveau(x.
```

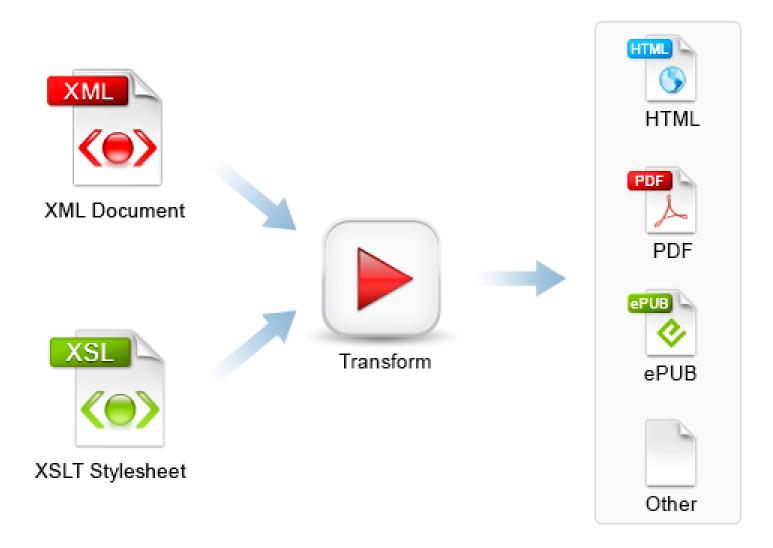
👼 let001.xml [/Users/elodiepaupe/Downloads/vangoghxml/let001.xml] - <oXygen/> XML Editor (Usage académique seulement)







https://www.oxygenxml.com /xml_editor/xslt_transformat ion.html





Une solution pour la visualisation: http://v-machine.org

- une chaîne de traitement complète
 - normes d'édition numérique XML-TEI à respecter
 - fichier XSLT qui permet la transformation d'un document correctement formé en page HTML
 - fichiers CSS et JSON qui permet la visualisation dynamique de l'édition
- pour réaliser l'édition, il faut un éditeur de texte capable d'exécuter une feuille de style XSLT: Oxygen est recommandé.



XML-TEI



Qu'est-ce que le XML?

X Extensible

M markup

L Language

• Un (méta)langage de balisage

• 1996: premiers travaux

• 1998: XML 1.0

• Objectif: permettre l'interopérabilité



Autrement dit:

- Pas un langage de programmation: un document XML n'exécute aucune fonction.
- Un langage qui sert à ajouter des métadonnées sur un contenu qui n'est pas mis en forme = un langage informatique de représentation des données.
- Un langage ouvert (= extensible) qui peut s'adapter aux besoins de l'utilisateur.



Concept de base

Transformer un texte brut (= une chaîne de caractères) en base de données à l'aide d'un langage informatique (= XML) qui isole des portions de texte (= sous-chaînes de caractères) en fonction des informations que l'éditeur souhaitent mettre en évidence.



Structuration d'un texte brut

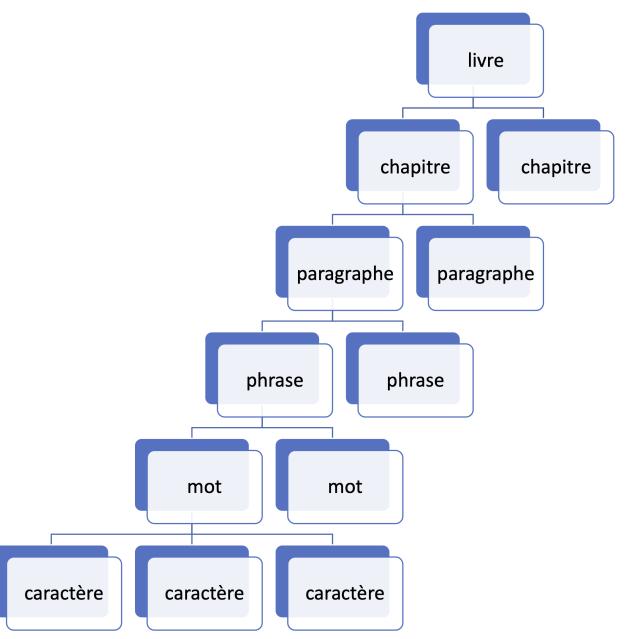
Il faut penser le texte comme une architecture:

- livre
 - chapitres
 - paragraphes
 - phrases
 - mots
 - caractères











```
vre>
    <chapitre>
        <paragraphe>
            <phrase>
                <mots>
                    <caractère>
                    </caractère>
                    <caractère>
                    </caractère>
                    <caractère>
                    </caractère>
                </mot>
                <mot>popule</mot>
            </phrase>
            <phrase> </phrase>
        </paragraphe>
        <paragraphe> </paragraphe>
    </chapitre>
    <chapitre> </chapitre>
</livre>
```



TEI

- Problème du XML: pas de standards communs
- Pour l'encodage de texte, développement d'une pratique commune à travers

T = Text

E = Encoding

I = *Initiative*



• Version standardisée du XML pour représenter un texte sous une forme numérique:

The Text Encoding Initiative (TEI) is a consortium which collectively develops and maintains a standard for the representation of texts in digital form. Its chief deliverable is a set of Guidelines which specify encoding methods for machine-readable texts, chiefly in the humanities, social sciences and linguistics.

• Recommandations émises par un consortium:

The TEI Consortium is a nonprofit membership organization composed of academic institutions, research projects, and individual scholars from around the world.

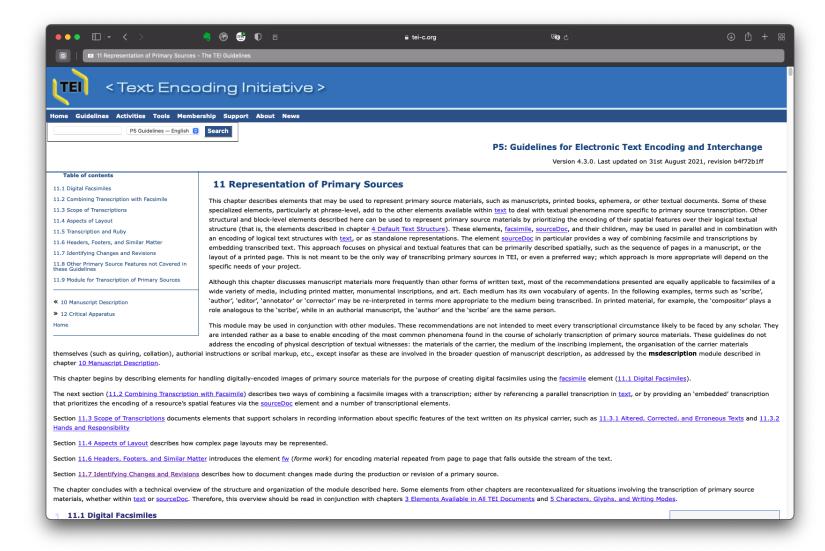
Source: https://tei-c.org



XML et TEI

- La TEI est donc une normalisation du langage XML utilisée par une communauté.
- Cette normalisation implique d'utiliser les balises définies par les guidelines établies et continuellement mises à jours.
- Il existe des balises standardisées pour encoder la génétique d'un texte.







Encoder en XML-TEI



Fonctionnement de base

Ce langage de balisage fonctionne de manière simple

<élément attribut="valeur">données</élément>



- 1. Les données sont la chaîne de caractères (= le texte brut).
- 2. On appelle "balise" les <élément> qui sont toujours écrits entre chevrons.
- 3. Une <balise> s'ouvre et doit impérativement être fermée </balise>
- 4. Des balises ne peuvent pas se chevaucher, c'est le principe des poupées russes:
 - Correct: <baliseA> <baliseB> données </baliseB> </baliseA>
 - o Fautif: <baliseA> <baliseB> données </baliseA> </baliseB>
- 5. Une <balise/> peut être auto-fermante: elle ne contient donc pas de données.
- 6. Un <élément> peut porter un @attribut (noté avec un @)
- 7. L' @attribut a une "valeur" (entre guillemets)
 - < <élément @="valeur de l'attribut">

14 octobre 2021

23



Structure minimale d'un fichier XML-TEI

Pour être dit "bien formé", un fichier XML-TEI doit respecter la syntaxe XML et les *guidelines* de la TEI.

Au minimum, un fichier XML-TEI se compose:

- d'un élément racine <TEI> qui contient l'entier du fichier
- d'un <teiHeader> qui contient les métadonnées du fichier dans une balise <fileDesc>
- d'un corps de texte compris dans la balise <body> elle-même comprise dans uen balise <text>

14 octobre 2021

24



```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
     <fileDesc>
        <titleStmt>
           <title>Mon édition numérique</title>
        </titleStmt>
        <publicationStmt>
           Élodie Paupe, Université de Neuchâtel. Cours d'édition, semestre d'automne 2020
        </publicationStmt>
        <sourceDesc>
           Le(s) manuscrit(s) de l'édition
        </sourceDesc>
     </fileDesc>
 </teiHeader>
 <text>
   <body>
       Texte
   </body>
 </text>
</TEI>
```



Texte en vers

Guidelines TEI pour la versification



Strophe/groupe de vers 1g et vers 1



Saut de ligne 1b



Encoder la création

Guidelines: Representations of Primary Sources



Suppressions et ajouts

Les étapes de l'écriture (rature, d'ajouts, repentences...) sont encodées avec les balises suivantes

- permet d'encoder un passage effacé
- <add> permet d'encoder un passage ajouté

Mon cher<add>vieil</add> ami



Liste des témoins

Pour être interprétés correctement par les fichiers de Versioning Machine, les différentes versions du texte qui sont assemblées dans l'édition génétique sont listées grâce à l'élément <witList> dans le du document XML.

Chaque témoin est identifié avec un @xml:id unique vers lequel pointent les balises <rdg wit="#...">.



Niveaux de réécriture

On utilise donc l'élément <app> pour introduire une variation et l'élément <rdg> pour renseigner chaque variante de réécriture. L'attribut @wit identifie le témoin.



? Le Pont Mirabeau Le Pont Mirabeau 1 Sous le pont mirabeau coule la Seine. 1 Sous le pont Mirabeau coule la Seine 2 Et nos amours faut-il qu'il m'en souvienne? 2 Et nos amours Faut-il qu'il m'en souvienne 3 La joie venait toujours après la peine. 3 La joie venait toujours après la peine. he NAF 4 Vienne la nuit, sonne l'heure, 5 Les jours s'en vont, je demeure. 4 Vienne la nuit sonne l'heure 5 Les jours s'en vont je demeure 6 Les mains Vient le dans les mains, et restons face à face, 6 Les mains dans les mains restons face à face Pont 7 Tandis que sous le pont de nos bras passe 7 Tandis que sous & Des éternels regarde l'ande si lacce I a nont da noe brae nacea



Versioning machine 5.0

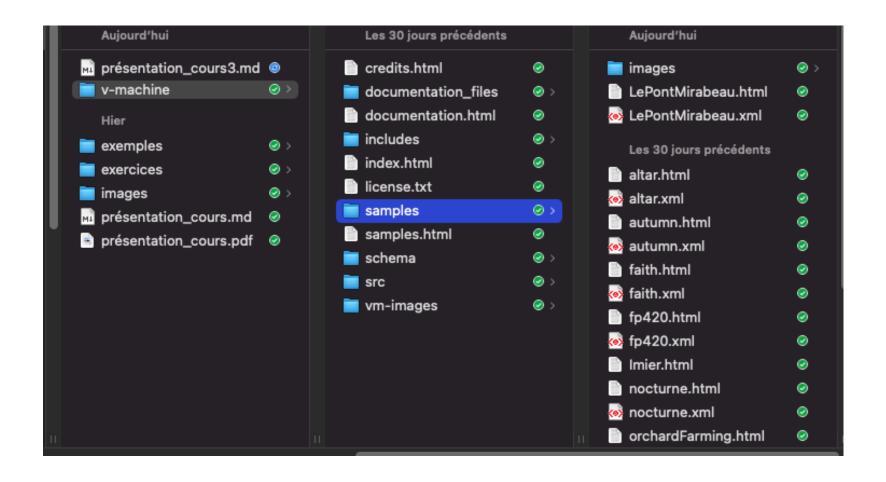


• Les normes d'encodage présentées sont décrites dans documentation.html et le schéma vmachine.rng contrôle que votre document soit bien formé. Pour cela, il faut le lier à votre document XML en ajoutant les lignes suivantes au début du fichier:

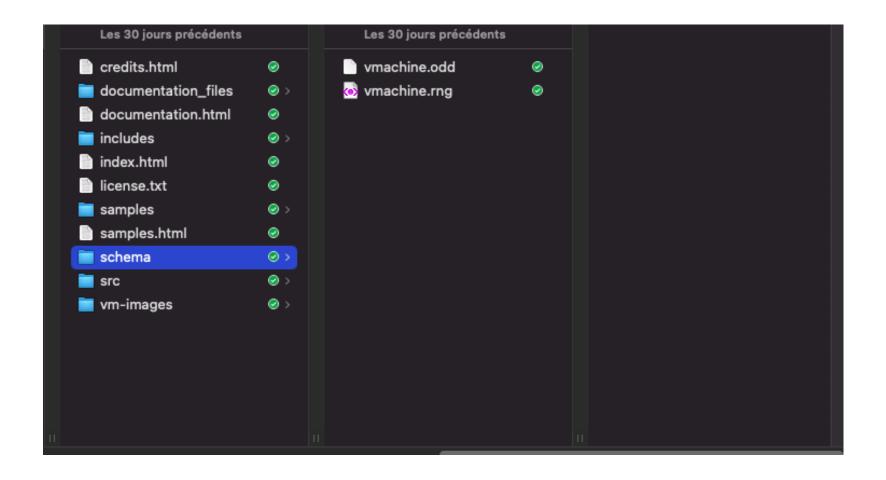
```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="../schema/vmachine.rng" type="application/xml" schematypens="http://relaxng.org/ns/structure/1.0"?>
<?xml-model href="../schema/vmachine.rng" type="application/xml" schematypens="http://purl.oclc.org/dsdl/schematron"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
...
</TEI>
```

Sous @href se trouve le chemin relatif dans l'arborescence de votre ordinateur qui lit votre document XML au schéma .rng.











• Une fois que vous avez un fichier XML bien formé, vous devez le transformer en fichier HTML pour qu'il puisse être visualisé. Pour cela, vous allez utiliser le fichier XSLT fourni dans la documentation de Versioning Machine 5.0; vmachine.xsl

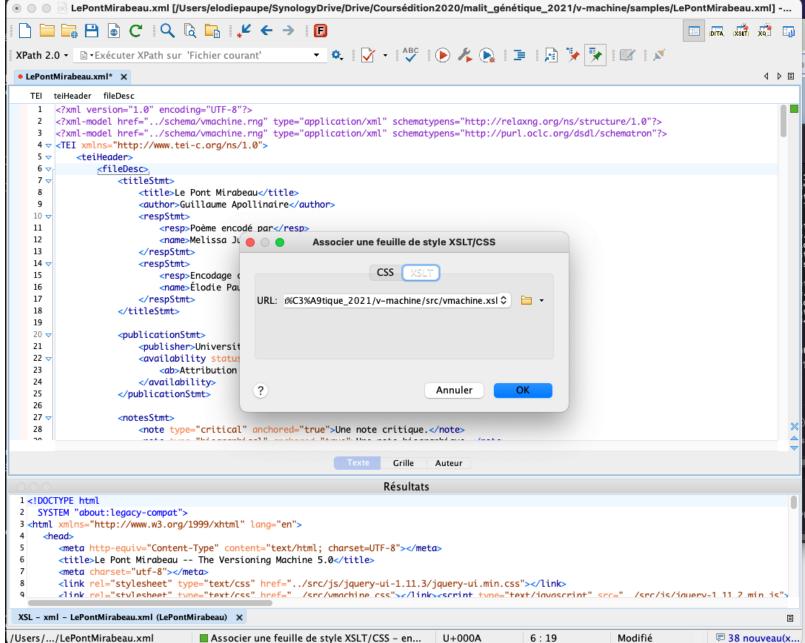
Voyons comment faire grâce à Oxygen, pas à pas.



1. Lier le fichier XSL au fichier XML dans Oxygen en cliquant sur l'épingle verte et en sélectionnant le fichier.

La déclaration <?xml-stylesheet type="text/xsl" href="../src/vmachine.xsl"?> s'ajoute en haut de votre document, sous les déclarations de schéma.

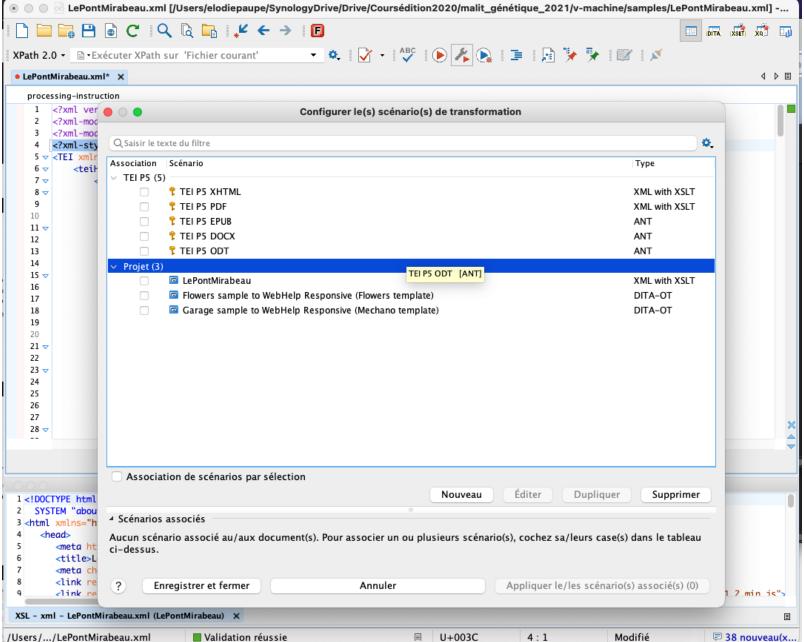






2. Créer le scénario de transformation en pressant sur la clé à outil et sélectionner "nouveau" --> "XML transformation with XSLT".

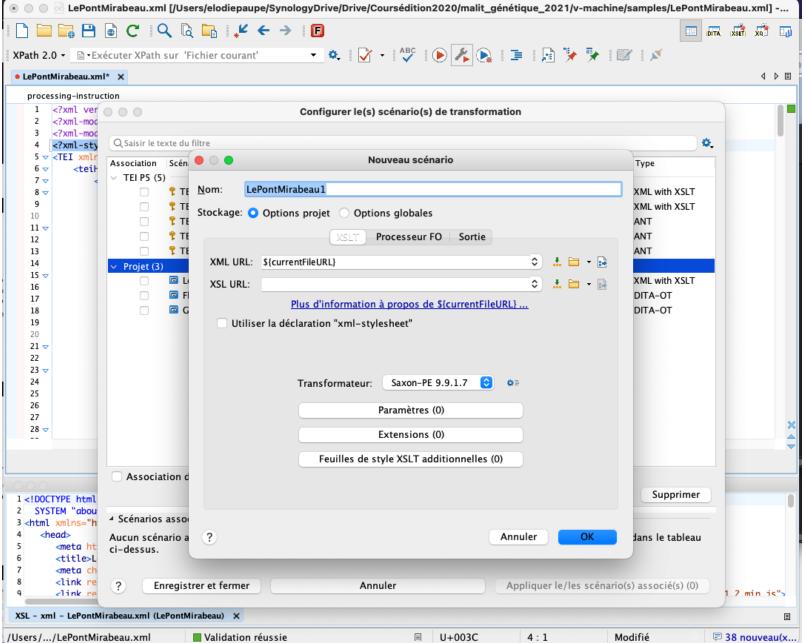






- 3. Donner un nom au scénario.
- 4. Sous XSLT, renseigner le chemin vers le fichier XML (celui que vous souhaitez transformer) et vers le fichier XLS (vmachine.xsl)





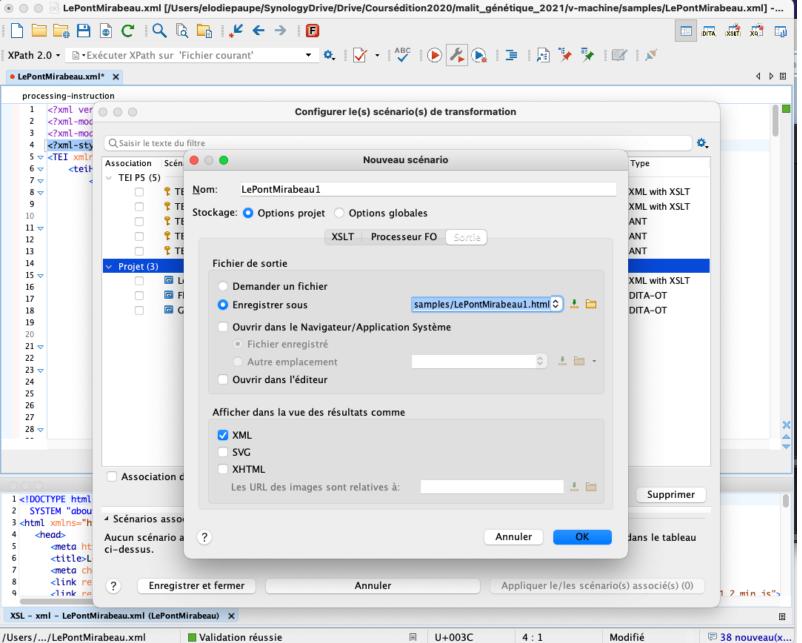


5. Sous sortie, choisir l'endroit où enregistrer le fichier et lui donner un nom.

Pour simplifier les choses, nous allons enregistrer le fichier dans le dossier "sample" de Versioning Machine 5.0 et nommer le fichier avec une extension en .html .

Cela signifie que nos deux fichiers XML et HTML sont enregistrés au même endroit.



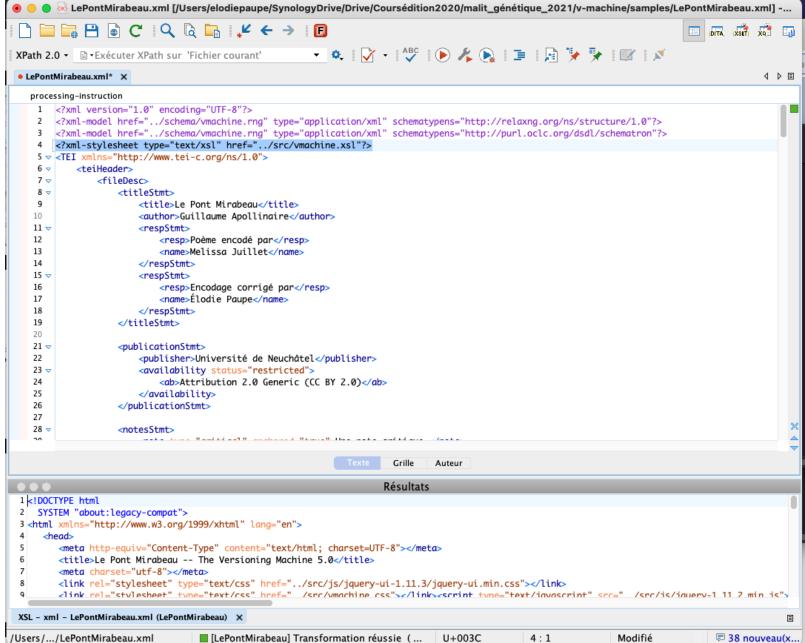


14 octobre 2021 Validation réussie U+003C 4:1 Modifié 38 nouveau(x...)



- 6. Cliquer sur "ok" puis sur "Appliquer le/les scénario(s) associé(s)"
- 7. Un fichier .html apparaît dans le bas de l'écran et se trouve enregistré dans le dossier "sample".







- 8. Ouvrir le document dans Google Chrome.
- 9. Ouvrir le document dans Oxygen.



Sources



Burnard, Lou, *Qu'est-ce que la Text Encoding Initiative?* [en ligne], Marseille: OpenEdition Press, 2015. Disponible sur Internet : http://books.openedition.org/oep/1298. DOI : https://doi.org/10.4000/books.oep.1237.

Versioning Machine 5.0, by Lara Vetter, Maryland Institute for Technology in the Humanities, Summer 2003; Jarom McDonald, Maryland Institute for Technology in the Humanities, January 2005; Sean Daugherty, University of Maryland Libraries, November 2006, March 2007, and July 2007; Tanya Clement, Digital Cultures and Creativity, University of Maryland and Susan Schreibman, Digital Humanities Observatory, June 2010; and Roman Bleier and Joshua D. Savage, An Foras Feasa at the Maynooth University Institute for the Humanities, 2015.