

Validation d'un document TEI

Un document TEI peut être instance d'un schéma personnalisant la TEI standard. Cependant, l'usage veut que :

- les éléments du namespace TEI soient utilisés avec leur sémantique ``normal''
- les éléments non TEI définis par un projet soient dans un autre namespace que le namespace TEI.

Conformance par rapport à tei_all

Parmi les critères de conformance à la TEI, il y a le fait que le document composé des nœuds du namespace TEI soit valide sur tei_all.

Cette consigne peut s'interpréter de deux façons :

1. si on enlève les nœuds hors du namespace TEI et tous leur descendants, alors on obtient un document valide sur tei_all.
2. si on enlève les nœuds hors du namespace tei et qu'on recolle leurs descendants qui sont dans le namespace tei, alors on obtient un document valide sur tei_all.

Les choses se compliquent encore si on considère les nœuds texte qui ne sont dans aucun namespace (les recopier a ou pas b ?)

D'une certaine façon, la question revient à se demander quel est le document TEI ``pur'' contenu dans le document de départ sachant que c'est ce document ``pur'' qui devra être valide sur tei_all.

La feuille de style ci-jointe prend en paramètres *recol* et *texte* à valeurs *vrai* ou *faux* et renvoie le document tei ``pur''. Je joins 4 documents supposés TEI pour jouer avec la feuille de style.

Code correspondant

La feuille de style ``extrairePureTEI.xsl" ci-jointe est supposée extraire le document à valider sur ``tei_all". Elle prend deux paramètres dont les valeurs par défaut sont ``vrai" et ``vrai". On peut l'appeler (sous un unix-like) par :

```
#xsltproc --stringparam recol "vrai" --stringparam texte "faux"
extrairePureTEI.xsl ex1.xml
```

ou bien

```
#saxon ex1.xml extrairePureTEI.xsl recol="vrai" texte="faux"
```

Le résultat est donc supposé être valide sur http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_all.rnc ou sur http://www.tei-c.org/release/xml/tei/custom/schema/xsd/tei_all.xsd selon qu'on préfère du relaxng ou du xsd.

Passage d'un fichier odd à un schéma

Actuellement, le passage se fait via Roma. Si on souhaite disposer d'une installation de Roma entièrement locale, on a besoin pour faire fonctionner Roma de :

- l'ensemble des fichiers TEI (recommandations). Remarque : de toute façon, on est supposé les archiver non ?
- les feuilles de style utilisées par Roma (incluses dans les feuilles de style TEI
- deux trois bricoles annexes telles que ``trang" si on souhaite passer des fichiers relaxNG que Roma engendre à des fichiers xsd par exemple...

L'ensemble des outils et feuilles de style est maintenu sur gitHub <https://github.com/TEIC/>.

Installer une copie de Roma (en ligne de commande)

Pour passer d'un fichier ODD à un schéma, on a besoin de deux projets GIT :

- `git clone https://github.com/TEIC/Stylesheets.git`
- `git clone https://github.com/TEIC/Roma.git`

A l'intérieur de Roma, on trouve le fichier shell ``roma2.sh" qui est l'outil qui servira à passer d'un fichier ODD à un schéma.

Ce qui manque maintenant est une copie des recommandations de la TEI (guidelines). Les guidelines sont développées sous subversion. Pour récupérer une copie locale :

```
svn checkout https://svn.code.sf.net/p/tei/code/trunk/P5
```

La commande précédente nous met donc une copie complète des guidelines dans le sous-repertoire odd.

On peut maintenant engendrer les schémas en tapant :

```
Roma/roma2.sh --xsl=Stylesheets --localsource=P5/p5subset.xml  
fichierOdd.odd Resultat
```

Dans le répertoire Resultat, on trouvera les schémas correspondant au fichier odd.

Remarque : on peut se contenter du rnc...(options --noxsd, --nodtd)