

Roma: TEI à la carte (2)

TEI@Oxford

2011-01

Roma: New

TEI Roma: generating validators for the TEI

These pages will help you design your own TEI validator, as a DTD, RELAXNG or W3C Schema.

Create a new or upload existing customization

- Build schema (Create a new customisation by adding elements and modules to the smallest recommended schema)
- Reduce schema (Create a new customization by removing elements and modules from the largest possible schema)
- Create customization from template
- Open existing customization

Start

Roma was written by Arno Mittelbach and is maintained by Sebastian Rahtz. Sanity check written by Iain Bernevig. Please direct queries to the [TEI@Oxford](#) project.

Roma: Customize



Roma: generating validators for the TEI

You are currently working on **TEI Absolutely Bare**

Set your parameters

[New](#) [Customize](#) [Language](#) [Modules](#) [Add Elements](#) [Change Classes](#) [Schema](#) [Documentation](#) [Save Customization](#) [Sanity Checker](#)

Set your parameters

Title

Filename

Namespace for new elements

Prefix for TEI pattern names in schema

Language English, Deutsch, Italiano, Español, Français, Portugues, Russian, Svenska, 日本語, 中文

Author name

Description

This customization creates a **TEI** schema with the bare minimum of tags to make a recognizable document. It combines the four basic modules, but removes most of the available elements and changes several attribute classes.

[Save](#)

Choisissez le fichier tei_cartes.odd dans votre dossier de travail.



Roma: Customize

Roma: Production des validateurs TEI - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.tei-c.org/Roma2/startroma.php?mode=customizeCustomization

Most Visited Latest Headlines Google Maps Adonis-related Travel TextAnalysis Shopping bit.Ly Sidebar Domestic

Roma: Production des validate...

TEI Roma: Production des validateurs TEI

Vous travaillez actuellement sur **TEI pour les cartes postales**

Personnaliser les paramètres

[Nouvelles](#) [Personnaliser](#) [Langage](#) [Modules](#) [Ajouter des éléments](#) [Modifier les classes](#) [Schéma](#) [Documentation](#) [Enregistrer](#) [Contrôleur de validité](#)

Personnalisation des paramètres en cours

Titre

Nom du fichier

Espace de nommage pour de nouveaux éléments


Préfixe pour des noms de modèle TEI dans un schéma

Langage English, Deutsch, Italiano, Español, Français, Portugues, Russian, Svenska, 日本語, 中文

Auteur

Description

Save



Roma: Schema

The screenshot shows a web browser window with the URL `http://tei.oucs.ox.ac.uk/Roma/startroma.php?mode=createSchema`. The page title is "Roma: Production des validateurs TEI". A blue banner at the top right says "Vous travaillez actuellement sur TEI Absolutely Bar". Below the title is the heading "Enfin je vous rends votre schéma...". A navigation menu contains buttons for "Nouvelle", "Personnaliser", "Langage", "Modules", "Ajouter des éléments", "Modifier les classes", "Schéma", "Documentation", "Enregistrer", and "Contrôleur de validité". The "Schéma" button is highlighted with a red border. A dropdown menu is open under "Schéma", showing options: "RELAX NG schema (compact syntax)", "RELAX NG schema (compact syntax)", "RELAX NG schema (XML syntax)", "ISO Schematron", "Schematron", "W3C schema (in ZIP archive)", and "DTD". A red "Generate" button is visible to the left of the dropdown. At the bottom, a footer note reads: "Roma was written by Arno Wapenaar and is maintained by Sebastian Rahtz. Sanity check written by Iain Bernevig. Documentation language en. Please direct queries to the TEI @ Oxford project. This is Roma version 4.0, last updated 2010-09-26. Using TEI P5 version 1.7.0. Last updated on July 1st 2010." The TEI logo is in the bottom right corner.

File Edit View History Bookmarks Tools Help

http://tei.oucs.ox.ac.uk/Roma/startroma.php?mode=createSchema

Most Visited Latest Headlines Google Maps Adonis-related Travel TextAnalysis Shopping bit.ly Sidebar Domestic

Problem loading page Roma: Production des valida...

TEI Roma: Production des validateurs TEI

Vous travaillez actuellement sur TEI Absolutely Bar

Enfin je vous rends votre schéma...

Nouvelle Personnaliser Langage Modules Ajouter des éléments Modifier les classes **Schéma** Documentation Enregistrer Contrôleur de validité

Creation du schéma en cours

Quel format préférez-vous?

- RELAX NG schema (compact syntax)
- RELAX NG schema (compact syntax)
- RELAX NG schema (XML syntax)
- ISO Schematron
- Schematron
- W3C schema (in ZIP archive)
- DTD

Generate


Roma was written by Arno Wapenaar and is maintained by Sebastian Rahtz. Sanity check written by Iain Bernevig. Documentation language en. Please direct queries to the TEI @ Oxford project. This is Roma version 4.0, last updated 2010-09-26. Using TEI P5 version 1.7.0. Last updated on July 1st 2010.

Roma: Documentation

The screenshot shows a web browser window with the URL `http://tei.oucs.ox.ac.uk/Roma/startroma.php?mode=createDocumentation`. The page title is "Roma: Production des validateurs TEI". A blue notification bar at the top right says "Vous travaillez actuellement sur TEI Absolutely Bar". Below the title is a navigation menu with buttons: "Nouvelle", "Personnaliser", "Langage", "Modules", "Ajouter des éléments", "Modifier les classes", "Schéma", "Documentation" (highlighted with a red border), "Enregistrer", and "Contrôleur de validité".

The main content area is titled "Création de la documentation en cours". It contains a form with the question "Quel format préférez-vous?" and a dropdown menu with the following options: "HTML web page" (selected), "HTML web page", "PDF", "TEI Lite", and "TEI ODD". A red "Generate" button is located below the dropdown.

At the bottom of the page, there is a footer text: "Roma was written by Arno Mittelbach and is maintained by Sebastian Rahtz. Sanity check written by Ioan Bernevig. Documentation language en. Please direct queries to the TEI@Oxford project. This is Roma version 4.0, last updated 2010-09-26. Using TEI P5 version 1.7.0. Last updated on July 1st 2010."



Qu'est-ce qu'on vient de faire?

On a traité un fichier ODD déjà existant, qui contient de prose discursive et en plus une spécification de schéma comme ceci:


```
<schemaSpec ident="tei_cartes">
  <moduleRef key="tei"/>
  <moduleRef key="textstructure" include="TEI teiHeader text body div"/>
  <moduleRef
    key="core"
    include="list item p address addressLine del add reg lb hi figure
    graphic title"/>
  <moduleRef key="namesdates" include="persName placeName"/>
  <specGrpRef target="#classmods"/>
  <specGrpRef target="#hdrIncludes"/>
</schemaSpec>
```


Testons le schema

- Dans Oxygen, créez un nouveau fichier XML
- En bas de l'écran il y a un bouton Personnaliser. Cliquez-le.
-

Nouveau ✕

Customize editor

 **Nom:** Document XML
Extension: xml

URL du schéma: ▼  ▼

Type de schéma: Schéma XML ▼

ID publique:

Espace de nommage:

Préfixe:

Exercice

En vous servant de cette schéma, essayez de compléter le balisage des fichiers cartes-1.xml, cartes-2.xml, et cartes-3.xml disponibles dans votre dossier Travaux; vous y trouverez aussi des fichiers images des rectos et versos des cartes concernées.

Exploration de TEI par Roma

Roma est également un outil apte à l'exploration des possibilités du système TEI.

- Sélectionner l'onglet [Modules] pour voir les modules disponibles ou déjà pris
- Sélectionner le nom d'un module pour voir les éléments qu'il contient, et pour en faire le choix.
- (On peut par défaut supprimer tous les éléments d'un module, ou bien en sélectionner tous)
- Sélectionner le nom d'un élément pour voir sa documentation complète

Roma: Modules

[Nouvelle](#) [Personnaliser](#) [Langage](#) [Modules](#) [Ajouter des éléments](#) [Modifier les classes](#) [Schéma](#) [Documentation](#) [Enregistrer](#) [Contrôleur de validité](#)

Liste des modules TEI

	Titre du module	Description brève	Modification
ajouter	analysis	Mécanismes analytiques simples	
ajouter	certainty	Degré de certitude et responsabilité	
ajouter	core	Éléments disponibles pour tous les documents TEI	modifié
ajouter	corpus	Corpus linguistiques	
ajouter	dictionaries	Dictionnaires	
ajouter	drama	Théâtre	
ajouter	figures	Tableaux, formules et graphiques	
ajouter	gaji	Représentation des caractères et des glyphes non standard	
ajouter	header	En-tête TEI	modifié
ajouter	iso-fs	Structures de traits	
ajouter	linking	Liens, segmentation et alignement	
ajouter	msdescription	Description de manuscrits	
ajouter	namesdates	Noms, dates, personnes et lieux	
ajouter	nets	Graphes, réseaux et arbres	
ajouter	spoken	Transcriptions de la parole	
ajouter	tagdocs	Éléments de déclaration d'un modèle	
ajouter	textcrit	Apparat critique	
ajouter	textstructure	Structure textuelle par défaut	modifié
ajouter	transcr	Représentation de sources primaires	
ajouter	verse	Poésie	

Liste des Modules sélectionnés

[supprimer](#) [core](#)
[tei](#)

[supprimer](#) [header](#)
[supprimer](#) [textstructure](#)

Roma: Change Module



Vous travaillez actuellement sur **TEI Absolutely Bare**

Modification d'un module

[Nouvelle](#) [Personnaliser](#) [Langage](#) **[Modules](#)** [Ajouter des éléments](#) [Modifier les classes](#) [Schéma](#) [Documentation](#) [Enregistrer](#) [Contrôleur de validité](#)

[back](#)

Liste des éléments compris dans ce module core					
	Inclure	Exclure	Nom	Description	Attributs
abbr	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="abbr"/>	? (abréviation) contient une abréviation quelconque.	Changer les attributs
add	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="add"/>	? (ajout) contient des lettres, des mots ou des phrases insérées dans le texte par un auteur, un copiste, un annotateur ou un correcteur.	Changer les attributs
addrLine	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="addrLine"/>	? (ligne d'adresse) contient une ligne d'adresse postale.	Changer les attributs
address	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="address"/>	? contient une adresse postale ou d'un autre type, par exemple l'adresse d'un éditeur, d'un organisme ou d'une personne.	Changer les attributs
analytic	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="analytic"/>	? (niveau analytique) contient des éléments descriptifs qui décrivent la bibliographie d'une ressource (par exemple un poème ou un article de revue) publiée à l'intérieur d'une monographie ou d'une ressource et non publiée de façon indépendante.	Changer les attributs
author	<input checked="" type="radio"/>	<input type="radio"/>	<input type="text" value="author"/>	? (auteur) dans une référence bibliographique contient le nom de la (des) personne(s) physique(s) ou du collectif, auteur(s) d'une oeuvre ; la première mention de responsabilité comme seul élément bibliographique.	Changer les attributs
bibl	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="bibl"/>	? (référence bibliographique.) contient une référence bibliographique faiblement structurée dans laquelle les sous-composants peuvent ou non être explicitement balisés.	Changer les attributs
biblScope	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="biblScope"/>	? (extension d'une référence bibliographique.) définit l'extension d'une référence bibliographique, comme par exemple une liste de numéros de pages, ou le nom d'une subdivision d'une oeuvre plus étendue.	Changer les attributs
biblStruct	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="biblStruct"/>	? (référence bibliographique structurée) contient une référence bibliographique dans laquelle seuls des sous-éléments bibliographiques apparaissent et cela,	Changer les attributs



Autre possibilités de personnalisations

Pour notre schema...

- on veut contraindre les valeurs légaux de l'attribut *@type* sur `<div>`
- on veut ajouter un nouveau élément `< saintName >` pour encoder les noms des saints (pourquoi pas?)
- on veut insister que chaque `<text>` contienne un `<div type="recto">` suivi d'un `<div type="verso">`

L'avantage ODD

Toutes ces contraintes peuvent être exprimées dans notre ODD, et on peut en générer un schéma pour contrôler qu'elles soient respectées

- Une schéma peut être générée en
 - ISO RELAX NG language
 - W3C Schema Language
 - XML DTD language
- Les modèles de contenu s'expriment avec une syntaxe RELAX NG
- Les types de données (datatypes) sont définis selon le standard W3C datatypes
- Il y a quelques contraintes (e.g. alternation, namespaces) qui ne sont pas exprimables en DTD — donc RELAX NG schema est recommandé
- Il y a quelques contraintes qui ne sont pas exprimables en aucune langue de schéma: pour celles-ci, on se sert d'une autre langue ISO, qui s'appelle *schematron*

Roma: selecting attributes

Liste des attributs: div

Ajouter des attributs

Changer un attribut	Inclure	Exclure	Nom	Description	Supprimer
org	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="org"/>	précise l'organisation du contenu de la division	
sample	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="sample"/>	indique si cette division est un échantillon de la source originale et dans cas, de quelle partie.	
part	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="part"/>	précise si la division est vraiment fragmentée ou non par quelques autres éléments structurels, par exemple une prise de parole qui est partagée en au moins deux strophes de vers.	
type	<input checked="" type="radio"/>	<input type="radio"/>	<input type="text" value="type"/>	caractérise l'élément en utilisant n'importe quel système ou typologie de classification approprié.	
subtype	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="subtype"/>	donne une sous-catégorisation de l'élément, si c'est nécessaire.	
decls	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="decls"/>	identifie un ou plusieurs éléments déclarables dans l'en-tête TEI, qui sont destinés à s'appliquer à l'élément portant cet attribut et à son contenu.	
xml:id	<input checked="" type="radio"/>	<input type="radio"/>	<input type="text" value="xml:id"/>	fournit un identifiant unique pour l'élément qui porte l'attribut	
n	<input checked="" type="radio"/>	<input type="radio"/>	<input type="text" value="n"/>	donne un nombre (ou une autre étiquette) pour un élément, qui n'est pas nécessairement unique dans le document TEI.	
xml:lang	<input checked="" type="radio"/>	<input type="radio"/>	<input type="text" value="xml:lang"/>	indique la langue du contenu de l'élément en utilisant les codes du RFC 3066	
rend	<input checked="" type="radio"/>	<input type="radio"/>	<input type="text" value="rend"/>	indique comment l'élément en question a été rendu ou présenté dans le texte source	
rendition	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="rendition"/>	pointe vers une description du rendu ou de la présentation utilisés pour cet élément dans le texte source	
xml:base	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="xml:base"/>	donne une référence URI de base au moyen de laquelle les applications peuvent résoudre des références d'URI relatives en références d'URI absolues	
xml:space	<input type="radio"/>	<input checked="" type="radio"/>	<input type="text" value="xml:space"/>	signale que les applications doivent préserver l'espace blanc	

Save

Roma: constraining attribute values

Création d'attribut en cours

Nom d'attribut type

Nome de class

Facultatif? yes
 no

Contenu Text

Valeur par défaut

Liste fermée yes
 no

Liste de valeurs

Description

characterizes the element in some sense, using any convenient classification scheme or typology.

Save

Qu'est ce qu'on vient de faire?

Notre fichier ODD contient maintenant :

```
<elementSpec ident="div" module="textstructure" mode="change">
  <attList>
    <attDef ident="type" mode="change" usage="req">
      <valList type="closed" mode="replace">
        <valItem ident="recto"/>
        <valItem ident="verso"/>
        <valItem ident="obliteration"/>
        <valItem ident="message"/>
        <valItem ident="destinataire"/>
      <!-- ... -->
    </valList>
  </attDef>
</attList>
</elementSpec>
```

Nota: on peut ajouter de documentation directement dans l'ODD:

```
<valItem ident="obliteration">
  <desc>contient
    le texte d'une obliteration quelconque sur une carte postale</desc>
</valItem>
```

Recréer le schema, et voir l'effet de ces changements avec Oxygen...

Ajout d'un élément nouveau < saintName >

Il ressemble a quels autres éléments? C'est une espece de nom, évidemment

Il peut contenir quoi? Il peut contenir que de texte.

Ou est-ce qu'il peut apparaitre? A l'interieur des titres, des phrases, des paragraphes, etc. mais non pas entre les deux.

Conclusion:

- on le fait membre de la classe `model.nameLike`
- son contenu est definie par `macro.xText`

Roma: Definition d'un nouveau élément



Vous travaillez actuellement sur **TEI Absolutely Bare**

Création d'un élément

[Nouvelle](#) [Personnaliser](#) [Langage](#) [Modules](#) [Ajouter des éléments](#) [Modifier les classes](#) [Schéma](#) [Documentation](#) [Enregistrer](#) [Contrôleur de validité](#)

[go back to list](#)

Création d'un élément en cours:

Nom

Namespace

Description

Classes structurales

- | | |
|--|--|
| <input type="checkbox"/> model.addrPart | <input type="checkbox"/> model.addressLike |
| <input type="checkbox"/> model.applicationLike | <input type="checkbox"/> model.biblLike |
| <input type="checkbox"/> model.biblPart | <input type="checkbox"/> model.castItemPart |
| <input type="checkbox"/> model.catDescPart | <input type="checkbox"/> model.choicePart |
| <input type="checkbox"/> model.common | <input type="checkbox"/> model.dateLike |
| <input type="checkbox"/> model.dimLike | <input type="checkbox"/> model.div1Like |
| <input type="checkbox"/> model.div2Like | <input type="checkbox"/> model.div3Like |
| <input type="checkbox"/> model.div4Like | <input type="checkbox"/> model.div5Like |
| <input type="checkbox"/> model.div6Like | <input type="checkbox"/> model.div7Like |
| <input type="checkbox"/> model.divBottom | <input type="checkbox"/> model.divBottomPart |

Definition de modele de contenu

- Le contenu d'un élément TEI est défini en langage RELAXNG
- La plupart des éléments TEI définissent leur contenu par référence aux *classes* d'élément plutôt que par référence à des éléments spécifiques.
- Il existe des modèles prédéfinis très utiles, par exemple:
 - `macro.paraContent` le contenu des éléments qui ressemblent aux paragraphes
 - `macro.phraseSeq` séquence de caractères mélangé d'éléments de la classe `model.phraseLike`
 - `macro.xText` séquence de caractères mélangé avec l'élément `<g>` (qui sert à encoder les caractères non-Unicode)

Roma: Defining a new élément 2

File Edit View History Bookmarks Tools Help

http://tei.oucs.ox.ac.uk/Roma/startroma.php?mode=listAddedElem Google

Most Visited Latest Headlines Google Maps Adonis-related Travel TextAnalysis Shopping

Problem loading page Roma: Production des valida...

- att.damaged
- att.dateable
- att.dateable.iso
- att.dateable.w3c
- att.declarable
- att.declaring
- att.dimensions
- att.divLike
- att.docStatus
- att.duration
- att.duration.iso
- att.duration.w3c
- att.editLike
- att.enjamb
- att.enjamb.w3c
- att.handFeatures
- att.internetMedia
- att.lexicographic
- att.metrical
- att.naming
- att.placement
- att.pointing.group
- att.ranging
- att.readFrom
- att.scoping
- att.sourced
- att.tableDecoration
- att.timed
- att.translatable

data.numeric
data.outputMeasurement
data.pattern
data.pointer
data.probability
data.sex
data.temporal.iso
data.temporal.w3c
data.truthValue
data.version
data.word
data.xTruthValue
macro.anyXML
macro.limitedContent
macro.paraContent
macro.phraseSeq
macro.phraseSeq.limited
macro.schemaPattern
macro.specialPara
macro.xtext

Contenu

User content

```
<content xmlns:rng="http://relaxng.org/ns/structure/1.0">
</content>
```

Save

Qu'est-ce qu'on vient de faire?

On a ajouté une spécification pour notre nouvel élément:

```
<elementSpec
  ident="saintName"
  ns="http://www.example.org/ns/nonTEI"
  mode="add">
  <desc>contains the name of a saint.</desc>
  <classes>
    <memberOf key="model.nameLike"/>
    <memberOf key="att.typed"/>
  </classes>
  <content>
    <rng:ref name="macro.xtext"/>
  </content>
</elementSpec>
```

Nota: ce nouveau élément n'est *pas* un élément TEI! Il appartient donc à une autre espace de noms.

Autre contraintes

- On peut contraindre le contenu d'un élément, ou la valeur d'un attribut en se servant des *datatype* (par exemple, pour insister que le contenu de l'élément `<date>` soit vraiment une date)
- La TEI prédéfinit plusieurs *macros* pour répondre à cette possibilité. Par exemple
 - `data.word` a single word or token
 - `data.name` an XML Name
 - `data.enumerated` a single XML name taken from a documented list
 - `data.temporal.w3c` a W3C date
 - `data.truthValue` a truth value (true/false)
 - `data.language` a human language
 - `data.sex` human or animal sex
- Ou on peut créer ses propres contraintes avec le langage Schematron

Contraintes Schematron

- Une spécification d'élément peut contenir un élément `<constraintSpec>` contenant des règles exprimés en langage ISO Schematron

```
<elementSpec ident="body" module="teisttructure" mode="change"
  xmlns:s="http://purl.oclc.org/dsdl/schematron">
  <constraintSpec mode="add" ident="bodyLim" scheme="isoschematron">
    <constraint>
      <s:assert
        test="tei:div/@type='recto' and tei:div/@type='verso' and
count(tei:div)=2">le body doit contenir un recto et un verso et rien de plus
      </s:assert>
    </constraint>
  </constraintSpec>
</elementSpec>
```

A noter...

- Pour ajouter de telles règles, il faut éditer le fichier ODD. Roma ne vous aide pas.
- L'implantation de telles règles n'est pas forcément disponible avec tout processeur XML.