

Making your own schema with Roma

Sebastian Rahtz and Lou Burnard

1 Before you start

In this exercise, you will use **Roma**, a web tool available from the TEI web site and also included on your TEI Knoppix CD, to make an XML schema. If you are not using the TEI Knoppix CD, you will need access to the net and a web browser such as IE, Mozilla, or Opera. Once you have your schema you will also need an XML-aware editor to use it (Emacs or oXygen in our setup).

Our goal is to make a schema which we can use to mark up the Punch document. We don't need all of TEI Lite, much less the full TEI, but we do need bits of various modules. We'll also have to tinker with some of the modules, to make our schema more helpful with daily editing.

2 Making your own schema

1. Open the Roma application. If you are online, you can point your favourite web browser at <http://tei.oucs.ox.ac.uk/Roma/>, or, if you are working from the CD, at <http://localhost/Roma/> (on the TEI Knoppix CD the recommended web browser is Mozilla Firefox, available on the toolbar; but you can also use Konqueror or even Lynx if you prefer).
2. The Roma start screen allows you to create a new customization, or to upload an existing customization for further work. We will start from scratch, which is the default option. Press the `Submit` button at bottom right of the screen to continue.

The next and subsequent screens show you a row of tabs for acting on your customization (`Save`, `Customize`, `New`, and `Help`), and a row of tabs for each of the major stages or tasks making up a customization (`Modules`, `Add elements`, `Change classes`, `Language`, `Schema`, and `Documentation`). We won't explore all of these in this exercise. By default the `Customize your Customization` screen is displayed. This allows you to specify a file name and other details for the schema, and also to change the interface language if you wish. For now, accept the defaults. Go to the `Modules` tab to proceed.

The modules screen shows two lists: on the left are all available TEI modules; on the right are the modules currently selected for your schema. You can add modules from the list on the left, and remove modules from the list on the right, by clicking the appropriate word next to the module you wish to operate on.

1. For this exercise, we will need the following extra module:

- `figures`

Click the word `add` next to each of these modules.

2. The modules chosen contain many more elements than we need, so we will now remove some of them, simplifying the view in the XML editor. Click the name of a module in the *List of selected modules* (the right-most column) to see a list of the elements this module defines.

Each element listed has a name, a radio button indicating whether it is to be included or excluded, a tag name, a description, and a link to a further screen where its attributes are specified. You can toggle inclusion or exclusion of all elements in the list by clicking the appropriate column heading. You can click on `Exclude` to remove all elements from the module.

Now work down the list clicking the radio button to restore or add the elements needed for this exercise. Remember to press the `Submit` button when you have finished with each module. Press the `Modules` or `back` links to go back to the list of modules.

from the core module delete `<add>`, `<analytic>`, `<cb>`, `<cit>`, `<corr>`, `<dateRange>`, ``, `<distinct>`, `<expan>`, `<gap>`, `<gloss>`, `<headItem>`, `<headLabel>`, `<index>`, `<meeting>`, `<mentioned>`, `<milestone>`, `<orig>`, `<pb>`, `<postBox>`, `<postCode>`, `<reg>`, `<series>`, `<sic>`, `<street>`, `<time>`, `<timeRange>`, and `<unclear>`

from the textstructure module delete `<argument>`, `<byline>`, `<closer>`, `<dateline>`, `<epigraph>`, `<imprimatur>`, `<opener>`, `<titlePage>`, `<trailer>` and `<salute>`

from the figures module delete `<formula>``<table>`, `<row>` and `<cell>`

We are now ready to generate a schema. Click the *Schema* tab, and then press `Submit`, taking the default option of a Relax NG compact schema. Your browser will ask whether you want to save or open the generated file: you should save it into your Desktop directory (the default). Look at the result, if you feel strong, or experiment with other options of the web application.

Complete this stage by going to the *Save* tab and saving your work. Do not close the web browser, we'll use it again shortly.

3 Using your schema in oXygen

You can use oXygen and the file `punch-ok.xml` to check that you've made your schema correctly. Proceed as follows, remembering the Punch file is in the `samples` directory.

- Open the file `punch-ok.xml` with oXygen.
- Go to the menu `Document`, then `XML Document` and then `Associate Schema`. Choose the `RelaxNG` tab, and locate your schema file (using the middle folder icon on the right to browse).
- If all goes well, oXygen will insert some code to mark the schema location, and attempt to validate the file. Try inserting some new elements, and you should see a much-reduced collection.

4 Enhancing your schema

Now let us return to the Roma, and make the schema more constrained, to make sure we get just what we want in our documents. The example we have chosen is to constrain the allowed values of the `type` attribute on `<div>`, and to make it compulsory.

Go back to Roma. If you have closed the browser, you can restart Roma and load the session you saved earlier. Go to the *Modules* tab and click on `textstructurea` in the right-hand column. Find `<div>` and click on `Change attributes` on the right-hand side. This

will show you all the attributes of `<div>`. Click on `type`, and you will be able to change its properties:

- Change the `Is it optional` radio button to make it compulsory
- Change the radio button for `Closed list?` to make it a closed list
- In the box for `List of values`, type



```
cartoon, verse, drama
```

(ie a list of possible values, separated by commas).

Now save the schema as before and reload the file in oXygen. There should be a validation error, because the last `<div>` in the Punch file has no `type` attribute.