

TEI Simple for TEI Technical Council -- Executive Summary

Introduction

As The TEI Technical Council is already aware, the TEI Simple project aims to define a highly-constrained and prescriptive subset of the Text Encoding Initiative (TEI) Guidelines suited to the representation of early modern and modern books, a formally-defined set of processing rules which permit modern web applications to easily present and analyze the encoded texts, mapping to other ontologies, and processes to describe the encoding status and richness of a TEI digital text.

Making customisations of the TEI like this is nothing new, though it is desired that the TEI Council will take over future maintenance and updates of the TEI Simple customisation, similarly to TEI Lite, once the project has completed. The full and final prose accompanying the TEI Simple customisation has not yet been finished, but the initial documentation (showing some of the background and what elements are included in the customisation) is available at:

<http://htmlpreview.github.io/?https://github.com/TEIC/TEI-Simple/blob/master/teisimple.html>

Simple Processing Model

What makes the TEI Simple customisation different is that it also uses a new method for documenting processing models (SPM: Simple Processing Model). This documents one or more intended outputs in an implementation agnostic manner. TEI elements are generally descriptive of the interpreted semantics of the source text ('this is a title', 'this is a quotation'), where the interpretation is often based on a human interpretation of layout. The `rend`, `rendition` and `style` attributes are sometimes, but not necessarily, used to describe that layout. The resulting encoded text is very amenable to analysis, but for the common case of re-presenting the text on the web or in print as a digital edition, the TEI model by itself is incomplete. The Processing Model notation hopes to change that not only for TEI Simple, but for many other uses of the TEI as well. This enables developers with no knowledge of TEI semantics to develop front-end renderings of TEI materials.

The processing model notation allows each element to be assigned to a structural category (`<model>`), and given an outline rendition description (`<rendition>`). This allows a processor to know whether to handle the element or not, and broadly speaking how to display or otherwise process it. The model and rendition instructions are part of the TEI ODD, so can be changed in TEI customizations in the normal way. More information about the Processing Model, and how it works is available at:

<http://htmlpreview.github.io/?https://github.com/TEIC/TEI-Simple/blob/master/tei-pm.html>

An example of a processing model documentation for `<choice>` might be:

```
<model predicate="sic and corr" behaviour="alternate">
  <param name="default">corr</param>
  <param name="alternate">sic</param>
</model>
<model predicate="abbr and expan" behaviour="alternate">
  <param name="default">expan</param>
  <param name="alternate">abbr</param>
</model>
<model predicate="orig and reg" behaviour="alternate">
  <param name="default">reg</param>
```

```
<param name="alternate">orig</param>
</model>
```

Or for `<date>` based on a variety of known output formats:

```
<model output="print" predicate="text()" behaviour="inline"/>
<model output="print" predicate="@when and not(text())" behaviour="inline">
  <param name="content">@when</param>
</model>
<model predicate="@when" output="web" behaviour="alternate">
  <param name="default">.</param>
  <param name="alternate">@when</param>
</model>
```

It is important to understand the difference between the Processing Model (the notation for `<model>` etc, and the set of supporting functions), and the Simple Processing Model, which is the set of `<model>` elements for TEI Simple. Council does not have to agree with actual `<model>`s proposed for TEI Simple.

Changes for TEI Technical Council to Approve

In order to create the simple processing model (which has been debated by council members and others on the TEI Simple mailing list) it has been necessary to create a number of new elements and repurpose others.

New Elements:

The new elements proposed are:

<model>	processing model documentation of an intended processing model for this element
<modelSequence>	sequence of processing model a group of model elements documenting intended processing models for this element, to be acted upon in sequence
<modelGrp>	a group of model elements documenting intended processing models for this element
<param>	parameter of a model behaviour
<paramSpec>	specification of a parameter of a model behaviour
<paramList>	list of parameter specifications

http://htmlpreview.github.io/?https://github.com/TEIC/TEI-Simple/blob/master/tei-pm.html#index.xml-body.1_div.3 has more information on any of these elements including their attributes, class membership, examples, etc. or see <https://github.com/TEIC/TEI-Simple/blob/master/tei-pm.odd> for the underlying TEI ODD for these elements.

Changes to Existing TEI Elements:

The SPM needs a variety of changes to existing TEI elements where these

<elementSpec>	Allow <code><model></code> , <code><modelGrp></code> , or <code><modelSequence></code> in its content model
----------------------------	---

<valItem>	Allow <paramList> in its content model
<rendition>	Add model.oddDecl to its membership

Existing Implementations:

The Processing Model has been implemented multiple times in different languages:

- XSLT: Magdalena Turska and Sebastian Rahtz (University of Oxford)
- Java: Matthew Buckett (University of Oxford)
- XQuery & Exist-DB: Wolfgang Meier (eXist Solutions)

Options for TEI Council Decisions:

1. Reject the whole idea, even though there is evidence of community wanting it, and Council already agreeing it is worthwhile. Council members have had a chance to debate all aspects as they develop on the TEI Simple mailing list.
2. Accept concept but propose back significant modifications to the TEI Simple project (e.g. replace whole aspects with something different Council proposes), assisting project to produce guidelines-ready process and TEI ODDs for Council to review
3. Accept concept but propose back minor modifications to the TEI Simple Project (e.g. rename elements (modelSequence to modelSeq or paramList to listParam for example, or negotiate particular attribute datatypes, etc), instructing project to produce guidelines-ready process and TEI ODDs for Council to review
4. Accept concept and implementation as is with only very minor modifications to the TEI Simple project, instructing project to produce guidelines-ready process and TEI ODDs for Council to review

It should be noted that, if the general notation is accepted, the function library can be regarded as incomplete at this point. It is sufficiently rich for TEI Simple, but it is accepted that it can and will be extended in the future.