

Experimental customisation of the Versioning Machine

Elsa Pereira¹, Roman Bleier²

¹CLUL, University of Lisbon, Portugal

²ZIM, University of Graz, Austria

DOI: 10.36505/ExLing-2019/10/0041

Abstract

This paper deals with challenges in adapting the XML-TEI publishing framework Versioning Machine to compositional drafts of 20th-century literary works and describes the main customisations that have been implemented to suit a genetic edition of poetry by Pedro Homem de Mello. The case study emphasises that even minimal customisations require technical work that may go beyond an editor's skill.

Keywords: representation, drafts, TEI, Versioning Machine, customisation

Representation of compositional drafts

Early 20th-century literary works are often documented in a variety of authorial witnesses, including draft manuscripts with several layers of additions, deletions, and rearrangements. A scholarly edition with a genetic orientation – aimed at achieving an insightful overview of the compositional and revisional development of those works – needs to represent the overwriting layers in each draft, while signalling the successive variants that occur across different witnesses.

In digital scholarly editing, the process involves two main stages: data modelling and presentation. The first stage is achieved using descriptive markup to produce machine-readable transcriptions, and while graph and hypergraph data models such as TAG (Dekker & Birnbaum 2017) are emerging as an alternative to the XML paradigm, the standard maintained by the Text Encoding Initiative [TEI] prevails as one of the most robust schemas available. It allows for a representation of both the materiality and the textual dimension of manuscripts and makes it possible to combine intra- and inter-documentary variation, using chapters 11 and 12 of the TEI P5 guidelines. The second stage converts the encoded texts into a visual layer to be accessed by the reader and is achieved by using XML transformation and query languages (XSLT, XQuery, XPath), as well as HTML, CSS, and JavaScript, to build an interactive graphical user interface. Such a wide range of technologies requires extensive technical support, which is not always available to individual or discretely funded projects.

The TEI community is aware of this problem and has developed tools and publishing solutions for XML-TEI, shared as open-source. Among the light-

weight solutions available, mention should be made of the TEI Boilerplate (Walsh et al.), the JavaScript library CETELcean (Cayless & Viglianti 2018), EVT (Di Pietro & Rosselli Del Turco 2018), as well as the Versioning Machine [VM] (Schreibman et al. 2003; 2016) – one of the earliest and most widely known. Originally developed to trace the composition history of poems written by Thomas MacGreevy, the tool was used in several international projects and has been adapted for special publication needs over the years.¹

Customisation of the Versioning Machine

VM is specifically conceived to display multiple versions of text encoded with the TEI critical apparatus tag-set and additionally allows a limited number of elements to represent intra-documentary variation. By editing the tool's XSLT and CSS stylesheets,² we have achieved a wider range of presentational features, to suit an edition of print material and composition drafts by the Portuguese poet Pedro Homem de Mello (1904-1984). The alterations concern the display of three main categories: metadata, draft layers of revision, and inter-documentary variation.

The default VM setup provides a basic presentation of metadata, contained by the standard elements of the `tei:teiHeader`, to be displayed in the so-called “Bibliographic Panel”. However, an edition for the genetic orientation, focused “on the documents as sources of evidence of textual development and change through time” (Van Hulle & Shillingsburg 2015: 36), requires more detailed descriptions. For this reason, we included a “Compositional Synopsis”, containing the information from the `tei:creation` element (encoded within the *profile description*) and modified the “Original Source” section, by adding and structuring the content of `tei:msDesc` and `tei:biblStruct` (within the *source description*), as illustrated by example 1.³

Several alterations were made to suit the multi-layered transcription of complex documentary sources. Alternative readings, for instance, are not covered by the default VM setup, but it was necessary to include them for the project, since draft manuscripts occasionally have two or more phrasings at specific points of the text, without showing the author's preference for any of the options listed. In the XML-TEI this may be encoded using the `tei:seg` and `tei:add` elements with an `@type="alternative"`, for which we had to create a transformation rule, displaying one reading above the other, as seen in example 2.

Also important was rendering encoded gaps and significant spaces, as well as extending the display of authorial additions, through the processing of the values "top", "bottom", "marginLeft", "marginRight", "overleaf", and "opposite" of `@place`, since VM only deals with "above" and "below" by default (see example 3). However, revisions that involve several lines or

stanzas present a challenge. If the `tei:add` and `tei:del` elements are not allowed, due to conflicting XML hierarchies,⁴ the TEI guidelines suggest using *milestone* or *empty elements*: i.e. `tei:addSpan` and `tei:delSpan` with a `@spanTo` attribute pointing to the ID of a `tei:anchor` that indicates how far the addition or deletion goes. The default stylesheets of VM do not contain instructions to process these milestone-delimited ranges, so new transformation rules were required (see example 4).

Another specificity of draft manuscripts is the repetition of revised stanzas or lines to clarify the wording. The TEI guidelines suggest encoding such clarifications with the `tei:retrace` element, which was introduced in 2011, as part of an encoding model for genetic editing. VM was designed for the critical apparatus tag-set and, therefore, does not support this element by default. A template rule had to be implemented to process and display `tei:retrace` and its `@rend` attributes, as shown in example 5.

For the same reason, VM does not deal with displacements of text (indicated by arrows or other symbols in the author's drafts), and new transformation rules were necessary to process these occurrences. In the project's encoding model, `tei:div` containing displaced text is specified by an `@type` with the value "displacedFragment", a `tei:metamark` indicates the place of insertion, and both elements are linked using an ID. Example 6 illustrates the rendering achieved by the modified stylesheets.

Besides the presentation of internal layers of revision, customisations were finally made to display the intricacies of inter-documentary variation in works with multiple drafts. According to the TEI P5 guidelines, apparatus entries may nest, using parallel segmentation with optional location-based referencing. While the VM is compatible with this encoding method, the default setup obscures the display of subvariants, because of the indistinct yellow highlight applied to all `tei:rdg` elements. We changed this by applying a different background colour whenever `tei:app` elements nest with different `@loc` ID, as illustrated by example 7.

Conclusions

By looking at one specific software application that was developed for the publication of TEI data – the Versioning Machine – this paper explored the customisations necessary to adapt the software to an edition's needs and encoding model. In this case, a series of alterations had to be made to combine intra- and inter-documentary variation in poems where multiple drafts coexist.

The TEI tries to suit a diverse community of humanities scholars, which requires great flexibility to support different textual models and editing strategies. This circumstance means that the standard is not particularly well suited for the development of generic publication solutions. While some of

the changes described in this paper could be achieved with minimal changes to the VM source code, it still requires an understanding of web technologies to implement them, and not every textual scholar with basic XML knowledge has these skills or the resources to pay for them. Hence, the 1S/1P/1DH (one scholar, one project, one digital humanist) paradigm described by Robinson (2013) inevitably persists. The question that remains is how this can be overcome in the future.

Notes

1. A selection of projects is listed on the website: <http://v-machine.org/vm-in-use>. The Wandering Jew's Chronicle is a good example of special customisation.
2. See `vmachine.xml` and `vmachine.css` in the "src" folder of our VM instance: https://www.dropbox.com/sh/etrhahpwahu83j1/AAAE7nWx9woNrdF5Hu_aU41na?dl=0
3. Examples mentioned in this paper are in the "samples" folder of our VM instance.
4. See chapter 20 "Non-hierarchical Structures" of the TEI guidelines.

Acknowledgements

The authors wish to thank the VM team for developing the Versioning Machine and Gael Vaamonde for his help during customising experiments.

References

- Cayless, H., Vigiante, R. 2018. CETEIcean: TEI in the Browser. In Proceedings of Balisage: The Markup Conference 2018. Balisage Series on Markup Technologies 21. <https://doi.org/10.4242/BalisageVol21.Cayless01>.
- Dekker, R.H., Birnbaum, D. J. 2017. It's more than just overlap: Text As Graph. In Proceedings of Balisage: The Markup Conference 2017. Balisage Series on Markup Technologies 19. <https://doi.org/10.4242/BalisageVol19.Dekker01>.
- Di Pietro, C., Rosselli Del Turco, R. 2018. Between Innovation and Conservation: The Narrow Path of User Interface Design for Digital Scholarly Editions. In Bleier, R. et al. (eds.) 2018, *Digital Scholarly Editions as Interfaces*, 133-163. Norderstedt: Books on Demand.
- Robinson, P. 2013. Why digital humanists should get out of textual scholarship. <http://scholarlydigitaleditions.blogspot.com/2013/07/why-digital-humanists-should-get-out-of.html>.
- Schreibman, S., Kumar, A., McDonald, J. 2003. The Versioning Machine. *Literary and Linguistic Computing* 18-1, 101-107. <https://doi.org/10.1093/lc/18.1.101>.
- Schreibman, S., et al. 2016. Versioning Machine 5.0. <http://v-machine.org>.
- Text Encoding Initiative. 2019. P5: Guidelines for Electronic Text Encoding and Interchange. <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>.
- Van Hulle, D., Shillingsburg, P. 2015. Orientations to text, revisited. *Studies in Bibliography* 59, 27-44.
- Walsh, J. et al. TEI Boilerplate. <http://dcl.ils.indiana.edu/teibp/index.html>.