

## SVG Visualization of TEI Texts

*Wendell Piez* ([wapiez@mulberrytech.com](mailto:wapiez@mulberrytech.com))  
*Mulberry Technologies, Inc.*

### SVG Visualization of TEI Texts

One of the more interesting benefits of XML technology for text processing has been the 'network effects' we get from using different XML technologies together. For example, XSLT proves to be suitable for a great range of tasks beyond simply the routine formatting of texts for display in a browser or on the page (the job for which it was designed): the investment we make in learning XSLT to generate reading versions of our XML texts also pays off many times over in enabling us to perform other kinds of tasks such as extra-schema validation, heuristic analytics of the markup or the text itself, and even (up to a point) querying. Likewise, it proves easy to produce a wide range of different kinds of output to represent the results of these operations. An XML application such as SVG proves to be a straightforward target for a transformation from XML data. The resulting SVG graphics can be anything. For example, graphs and bar charts of information captured in numerical data sets and represented in XML are easy to create using XSLT/SVG. But so are more arcane kinds of depictions of source datasets or their features, including using SVG as a display format for 'maps' of a document's structure.

This basic architecture, XML + XSLT -> SVG, has been demonstrated repeatedly in both the commercial and academic sectors in recent years (see Bibliography; several applications by the author demonstrating the use of XSLT to create SVG graphical depictions of various kinds are included (Piez 2000, 2002, 2003a, 2003b). There is nothing particularly innovative at this point (late 2004) about this inexpensive and powerful method of creating graphics. What has been explored perhaps less deeply is what can be done with stylesheets generating graphical depictions of specifically *literary* works, leveraging descriptive tagging of the 'pure' kind (that is, tagging that has been designed to reflect documents' logical organization, without any particular renditions in mind). Not only are the structures and features of such works of intrinsic interest to students of literature; they can also serve as a diverse and heterogeneous testbed for prototyping techniques of rendition and visualization that could be used on other sources or indeed, on other kinds of XML data. These techniques would be widely applicable both to works of narrative or discursive prose and to more highly structured literary texts such as verse and drama.

Earlier demonstrations of this approach make it clear that we are now, with the maturation of XML technologies and the increasing support of SVG in readily available tools (the *Mozilla* development team has lately been implementing SVG for their browser, and Adobe continues work on the technology as well), in a position where we can perform these operations on a larger scale. One of the features of the architecture is that a family of documents marked up consistently with the same tag set (say, TEI) should be processable with the same stylesheet. The marginal effort required to create a graphic depiction of a new text, consequently, is negligible when that text's tagging conforms to a known and supported usage pattern (preferably valid to a known DTD). In theory, it should be possible to generate an entire library of graphics to represent a library of texts, all with a single stylesheet.

The poster I am proposing for ACH/ALLC 2005 will present the results of a set of experiments testing these ideas, applying stylesheets (both extant and new) on a variety of texts from the *Women Writer's Project* at Brown University (with their kind permission and collaboration). This will have the twofold purpose of exploring what kinds of visual representation of these structures are most revealing, as well as testing to what extent single stylesheets or small families of stylesheets can be used across a document repository, to draw interesting and revealing comparisons among texts. (It is quite possible that per-document "tuning" of the presentation logic will be necessary, through a customization layer, for best results; but until we have tried the technique on a range of texts, we will not know the extent to which stylesheet reuse is practical. This extent may also vary between different stylesheets used to create different sorts of graphics.)

Stylesheets developed for this poster will also be contributed to the *WVO* (*Women Writers Online*) project, and made available to the wider TEI community.

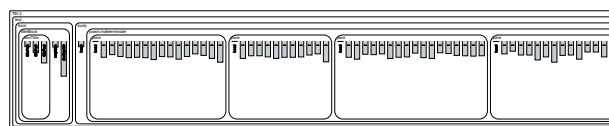


Figure 1: Aphra Behn, "A Pindaric Poem to the Reverend Doctor Burnet" (1689). An example of a free verse form.

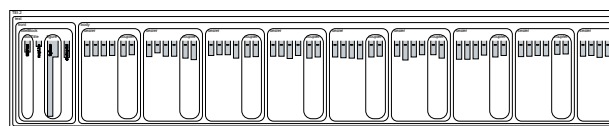


Figure 2: Catherine Clive. "The Case of Mrs. Clive" (1744). An example of a work in prose.

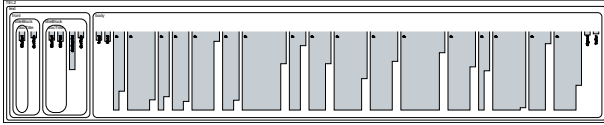


Figure 3: Mary Sidney, Countess of Pembroke. "The Doleful Lay of the Fair Clorinda" (1595). An example showing a regular verse form (sestets containing couplets).

## Bibliography

Birnbaum, David J. "Analyzing and visualizing the structure of medieval encyclopedic works with XML-related technologies." Paper delivered at the Extreme Markup Languages 2003, Montreal. August 2003.

Cagle, Kurt. *SVG Programming: The Graphical Web*. Berkeley, CA: Apress, 2002.

Eisenberg, J. David. *SVG Essentials*. Sebastopol, CA, USA: O'Reilly, 2002.

Mangano, Sal. *The XSLT Cookbook*. Sebastopol, CA, USA: O'Reilly, 2002.

Mansfield, Philip A., and Darryl W. Fuller. "Graphical Stylesheets: Using XSLT to Generate SVG." Presented at XML 2001. 2001. On line at <http://www.idealliance.org/papers/xml2001/papers/html/05-05-02.html>

Piez, Wendell. *The Sonneteer: A demonstration of structured form*. . Accessed 2005-04-13. <http://sonneteer.xmlshoestring.com>.

Piez, Wendell. "SVG By Way of XSLT." Tutorial delivered at Extreme Markup Languages 2001, Montreal. August 2001.

Piez, Wendell. "Visualizing XML document structure using XSLT and SVG." *interChange, the journal of ISUG (the International SGML Users' Group)* (December 2003): n. pag. On line at <http://www.xmlshoestring.com/xml49/visualizingxml>

Piez, Wendell. "XSL: Characteristics, Status and Potentials for the Humanities." Presented at ALLC/ACH 2000, Glasgow. July 2000. On line at <http://www.idealliance.org/papers/xml2001/papers/html/05-05-02.html>

Tennison, Jeni. *Beginning XSLT*. Birmingham, UK: Wrox Press, 2002.