

**Christopher W. BLACKWELL**

**XML and Undergraduate Writing: Self-describing Data for Valid Arguments**

This presentation suggests ways in which Extensible Markup Language (XML), a technology that is now accessible to humanists, is a useful tool for teaching undergraduates to write. It offers a scheme by which beginning students can write simple research papers using XML, which the instructor can then transform in various ways for printing, online display, and analysis. It also briefly covers freely available tools to assist instructors in following a similar scheme. Additionally, all materials discussed will be distributed to any interested party for their use or adaptation.

XML is a standard for defining systems of markup to describe particular sets of data. Unlike procedural markup languages (like HTML) that describe how parts of a document should look, descriptive markup languages (such as XML) describe the component parts of a document. Moreover, an XML document can be required to conform to a specific Document Type Description (DTD) that specifies what kind of data can appear in the document, in what order, and in what combinations. Those who introduce students to the writing of academic papers are familiar with a similar process: defining the structure of a text, the parts that it must have, the elements that must appear in each part, in what order, and in what combination. Using a DTD forces students to be explicit about elements of their paper. Because XML documents are not subject to formatting, students cannot use italics or underlining. So, for example, where they might use italics when writing in a word-processor, in XML they must specify "emphasis" or "bookTitle" (the only two options in this simple DTD). A student writing a paper in XML can use a web-based program to validate her paper, that is, to check the structure of her paper against the requirements of the DTD.

In this presentation, I discuss a simple DTD that prescribes the structure of a simple research paper. The DTD defines a text (the student's paper) that has an author, a title, and a date, followed by an introductory paragraph, one or more body paragraphs, and a concluding paragraph. The introductory paragraph must contain one (and only one) statement of thesis. The body paragraphs can consist of unstructured text or support (that is, the evidence on which the students' argument are built). Each instance of support must be identified explicitly, and must have a citation attached. The students must identify explicitly citations, too, as being either primary or secondary. The concluding paragraph cannot include support; this prohibition forces students to conclude in their own words. The concluding paragraph must include a restatement of the thesis, marked explicitly as such.

Texts marked up in XML are not pretty, nor are they always easy for humans to read. But XML documents can be transformed into other formats by means of a related technological standard, Extensible Style-sheet Language (XSL). The presentation offers a few simple XSL transformations that allow XML papers to be used in various ways. One style sheet transforms XML documents for printing, with all the support elements appearing as endnotes. Another formats XML documents into HTML for display on the Web. A third transforms the students' paper more radically. It first produces a version of

the essays with all support removed. This allows the instructor and students to focus on the students' own intellectual contribution. This style sheet will also print initial statements of thesis and the final restatements side-by-side, showing the extent to which the arguments are unified and consistent. Finally, it sorts and lists all citations, distinguishing primary sources and secondary sources.