

# Bringing Up *EMMA*:

## Developing Writing Software with XML at the University of Georgia

Christy Desmet

Visit EMMA at <http://online.english.uga.edu>

Between January and December of 2002, a group of faculty and teaching assistants from the English Department got together weekly to discuss, develop, and then adopt for experimental classroom use a computer application.<sup>1</sup> EMMA, as we eventually named “her,” helps students, teachers, peer reviewers, and others to “mark up” writing in a way that makes particular features—ranging from argument and content to sentence structure—visible, and therefore, comprehensible in a new and striking way. EMMA (English Markup and Management Application) was created as part of a larger project in the English Department that centers on the use of markup languages to manage and disseminate information in a complete, compact, and efficient way. Already, developer Ron Balthazor has used XML markup to restructure the English Department Directory and to build useful databases for such purposes as faculty mentoring of undergraduate majors, storing and sorting out online student evaluations of First-year Composition classes, and generating Annual Reports for faculty. (Some of his efforts can be seen at <http://www.english.uga.edu> .)

Markup languages have begun to change the way we do business in Park Hall. But it is our hope that the EMMA project will also have a broad and lasting impact on the teaching and production of writing at all levels—which is, after all, the main job of an English Department. We want to use this new and flexible technology in the classroom, with and for the benefit of students. Nelson Hilton, Head of the Department of English and PI of the USG Learning Technologies Grant that has funded this ongoing project, remarks that “Students recognize already—or, at least, embody the fact—that writing today differs far from the undertaking it was prior to the information technology revolution; they enter the University assuming that research is done on the Internet and that text is processed on a computer. Documents are ‘engineered’ as much as ‘Englished,’ and, increasingly, composed of bits of multi-media content. Emerging technologies that utilize markup languages will enable a fundamental shift in the very nature of the production, examination, and distribution of text. The principles involved in such text production will be standard fare for these students when they graduate. Most text in the coming generation, even academic writing, will be electronic, and the degree to which it will be useful to the reader will in good measure be a function of its markup.”

---

<sup>1</sup>This essay is based on a group presentation made for the UGA Technology Expo in November 2002. The presenters were Nelson Hilton, Alexis Hart, Ron Balthazor, Bob Cummings, Angela Mitchell, and Christy Desmet. Although Christy wrote this report, the words and thoughts of these colleagues and others who developed and piloted the now adolescent EMMA (particularly Leigh Dillard and Robin Warren) are present everywhere in it. They are all co-authors in the fullest sense of that word.

So, one might ask at this point, what exactly is XML? What, indeed, is a markup language? Many people are familiar with HTML, the markup language used to display text, images, sound bites, and other information on the World Wide Web. XML (extensible markup language) works a lot like HTML, with one crucial difference. HTML tells a web browser how to *display* text. If a web designer writes the following HTML code--`<I>italics</I>`-- then a web browser knows to display that bit of marked text in *italics*. If the designer replaces the start and stop tags with `<B>` and `</B>`, the text is displayed in boldface as **italics**. In XML, however, the author describes not only the appearance of a text, but also its structure. Anyone—teacher, student, or administrator—can mark up text in a way that emphasizes pre-defined features of that text.



Take the case of a First-year Composition instructor who currently is working with students on essay structure, calling their attention to the development of a thesis statement and extolling the benefits of topic sentences for the arrangement of points in an argument. The teacher employs a DTD (or Document-Type Definition) that defines particular features (e.g., thesis, topic sentences, etc.) that, when marked in piece of text, will allow students to see (both literally and figuratively) the structure of their argument as a visual display.

At present, the student comes into the computer lab, picks up a CD with the latest version of EMMA and sits down at a computer.<sup>2</sup> After loading the program, the student opens the appropriate “assignment,” which has embedded within it the particular DTD for this exercise. None of the code—the inner workings of EMMA—is visible to the student, who sees only a split screen with a menu bar on top. On the left is a space for listing administrative information (so the document can be fully identified and tracked after uploading to the EMMA archive) and a blank section that says “Paste or compose your essay here.” The student either goes to the word processor of her choice and copies and pastes the essay into EMMA or composes directly in the editor. In the right-hand screen appears a menu of possible tags. In the case of our imaginary assignment, the student would first mark each paragraph of the essay with the `<paragraph>` tags, then use those for `<sentence>`. Next comes the essay’s `<thesis>` and the series of `<topic sentences>`. All the student must do is highlight the text she wants to mark and click on

---

<sup>2</sup> For those who care about these matters, the application we call EMMA is an open-source java editor called jEdit that we have modified for our purpose. The project then uses Cocoon (part of the Apache XML open-source development project), an XML publishing framework running on the Jakarta-Tomcat servlet engine. Cocoon is used to produce server-side XSL transformations of the XML documents. In the initial phase of the project, Cocoon produces a variety of HTML displays, but we continue to work toward using Cocoon as a complete document production and management environment that will include PDF production, a rich search capability, and multiple document concatenation. Currently we are using Filemaker Pro as the database for authentication and document tracking. (Thanks to Ron Balthazor for this description.)

the appropriate tag. When this work is done and the document has been checked for errors, the student uploads her essay and is able immediately to consider it in a variety of formats. One format just looks like a regular word-processing document; a second highlights the thesis and topic sentences in bright colors as they appear in the paper; a third collapses the intervening paragraphs to display the essay's skeletal structure, just the thesis and topic sentences in outline form.

Now the real work begins. Teacher and students watch as different students' work is projected on the screen. "What do we think of Dan's thesis?" the teacher asks. "A little general," says one student. "He argues against mandatory sentencing in drug cases, but I'm not sure whether he is talking about all court cases involving drugs or whether he makes a distinction between different kinds of crimes or situations." "The thesis introduces the topic but doesn't exactly make an argument," says another. "It is awfully short," Dan himself comments, looking critically at the projection screen. "Based on the information provided by your topic sentences, the first three paragraphs seem to be about marijuana, Dan. Are you mostly thinking about this drug?" the teacher asks. "I think so," says Dan. "I'm thinking of marijuana because it is a less destructive drug than ecstasy or heroin."

The scenario I have sketched out here is a rather traditionalist one. Teachers have always used highlighters, crayons—any tool at hand—to make students "look" critically at their writing. But marked-up text is different. First of all, one set of markings can be displayed in an infinite number of ways, and more important, can be displayed in a way that jolts writers out of the visual complacency encouraged by the smooth surface of word-processed text. Second, the writer's conceptual decisions about how to mark text persist through multiple versions, allowing students, peer review groups, and finally, teachers to track the progress of an argument through successive revisions. Third, the DTD that informs any rhetorical exercise in EMMA has a logical structure that legislates what elements are marked and how they are "nested" within one another, so that DTD writers and students alike see particular features of text as part of a rhetorical system. In the example given above, thesis sentences and topic sentences no longer exist in a conceptual vacuum, as fetish objects earnestly desired by teachers for mysterious reasons of their own, but are placed in the logical hierarchy of academic writing, whose structure begins at the level of the essay and moves down through that of the paragraph to the micro-level of the sentence. Finally, we need not even define writing in the textbook terms I have used here. One might, like structuralist critic Roland Barthes, mark a text in terms of semiotic units, or "semes." One might mark in terms of "text" and "hypertext," or even "meta-text." The only limit is the human imagination.



There are very many things one can do with XML markup to slice and dice writing into different conceptual units and so see it in a new light. One of the beauties of markup is to make macro the micro-level effects that constitute style. Angela Mitchell, for instance, asks students to take a biographical essay and mark the devices by which writers develop their ethical (i.e. "you can trust me") personae;

Bob Cummings has students trace the development of different argumentative themes as they take their essays through successive drafts, so that the students can see just how much or little their thought has developed and so function as partners in, rather than consumers of, the evaluation process. EMMA can even blur the line between reading and writing. *Macbeth*, for instance, is a play in which the line between truth and falsehood is remarkably flexible. The witches who persuade Macbeth to murder his king “equivocate”—that is, they speak lies in the form of truth. Imagine members of a Shakespeare class (as I hope mine will next semester) marking up key speeches in *Macbeth*, labeling portions of text as lie, truth, or equivocation. Most likely, each person will mark the text differently, and assessing those differences can help us begin to understand, and perhaps even experience, the spiritual dilemma of Macbeth as he attempts to “read” the witches’ promise that he will wear the Scottish crown.

EMMA is not a technological babysitter, a “teacher-proof” piece of software. The best thing about the endeavor, in fact, is how it can bring together class members into a conversation about writing. During infancy, EMMA was “cranky,” and so we devoted most of our attention to working out its “bugs.” But as we moved through the first phase of the software development, we found ourselves talking to students more about the structures they were being asked to mark and identify. These conversations can be very simple (“Is this my main verb?” “I can’t remember what words are prepositions”), but they can become complex and even philosophical in nature. Most important, since EMMA is a tool that can be employed by teachers, students, and peer reviewers alike, students can finally learn the answer to that all important question, “What does the teacher want?”, by assuming themselves the authority to define, characterize, and evaluate their writing.

In the upcoming semesters, we plan to develop more pedagogical uses for EMMA. We also plan to extend the “management” aspect of our application’s acronym. The papers submitted through EMMA can be organized as a database that faculty can search in evaluating programs or that students themselves can use to create writing portfolios for their job searches. The possibilities are limited only by the imagination of individuals, server space (more ample every year), and the ongoing kindness of the Committee for Applied Instructional Technologies, the Office of Instructional Support and Development, and the Franklin College of Arts & Sciences—without whom, EMMA would not have flourished and grown into the promising young application whose best years, we hope, lie still ahead.

