Digital Critical Editions: Coding and/as Scholarship

Although there is no shortage of Latin texts on the internet, most websites provide only the text, not the other features that make a critical edition a vital, original contribution to scholarship (e.g., an introduction, a critical apparatus, analysis of the primary and secondary sources for the text, etc.). This is due in no small part to the technical challenges involved in displaying those features on a screen. The Digital Latin Library's series *The Library of Digital Latin Texts* (LDLT) has developed a data model for representing all of the information in a critical edition, and it will soon launch several applications for presenting that information in a variety of formats, including HTML, PDF, and in tools for visual data analysis. Texts submitted to the LDLT will undergo peer review and, if selected for publication, they will bear the imprimatur of one of the learned societies affiliated with the project: the Society for Classical Studies, the Medieval Academy of America, and the Renaissance Society of America. Accordingly, there should be plenty of incentive to publish new critical editions in the LDLT, but potential editors may balk at the prospect of having to learn how to work with XML to prepare their texts for publication. This paper aims to demonstrate a technique developed by a professor and an undergraduate researcher that will automate the encoding, leaving editors free to focus on textual scholarship. At the same time, it aims to make the point that developing scripts for processing textual scholarship is itself an act of scholarship and deserves to be evaluated as such.

We will begin by demonstrating some of the things that can be done with a text encoded according to the DLL's guidelines, as opposed to a traditional printed edition or a digital text with more basic encoding. We will then reveal the encoded text itself and briefly discuss the process for encoding a text manually, that is, by working directly with XML. Although some

might enjoy working in this manner, many scholars do not have the time or technical background to learn XML, and the encoding process is tedious even for those that do. That is why we have developed a script to facilitate their participation in this new way of publishing textual scholarship without working directly with XML

We will argue that creating a data model for critical editions and building an application that converts the complex information contained in a critical edition into machine-readable code is an act of scholarship. We will support this argument by demonstrating how our application works in real time. We will begin with a plain text file of a critical edition's text and another text file of the edition's critical apparatus. Using a series of scripts written in the Python programming language, we will convert the two files into a fully encoded, valid XML document that complies with the DLL's encoding guidelines. The end product will reflect the scholarly judgment and critical decision-making that went into the creation and development of the application.

We will conclude with remarks about the potential that this kind of work presents for involving undergraduate students in research. The field of digital humanities scholarship is opening up great opportunities for applying the critical thinking, reasoning, and language skills that we promote as strengths of Classical education. As projects like the DLL develop, and as coding becomes an ever more common part of the curriculum at all stages of education, faculty members willing to venture into digital scholarship will have new ways of exploring Classical texts with their students. But it is vital to begin by establishing that this sort of work is scholarship and that involving undergraduate students has far-reaching pedagogical value.