# Teaching Digital Editing and Manuscript Studies: A Project-Based Short Course Approach

## Anya Adair, Katherine Hindley, and Joe Stadolnik

This article describes the objectives, structure, and teaching plan of a two-day workshop offering basic skills in paleography, codicology, and digital text encoding. This short course provides students with a supported experience of the full editing process. The teaching model here is project-based and collaborative: participants jointly produce a basic digital edition of a manuscript artifact (in this example case, a medieval manuscript roll). Combining manuscript studies with digital text encoding allows for substantive training in introducing both areas of study, while also helping to overcome barriers to access—especially for students who are newcomers to paleography and digital scholarly editing. The peer-teaching model of this course addresses several limitations in the digital editing training commonly available to undergraduate and graduate students. The model is adaptable to semester-long graduate or undergraduate courses, and to situations where access to physical manuscripts is limited or impossible.

## Introduction

The short course "Digital Editing and the Medieval Manuscript Roll" (DEMMR)[1] makes concentrated use of digitized manuscripts, and adapts freely available digital tools and resources to create a low-cost and flexible workshop suitable for both experienced students and complete beginners in the field of medieval studies. The course design responds to a central challenge in beginners' study of both textual editing and digital humanities: students must master a number of

Anya Adair is Assistant Professor in Law and Humanities at the University of Hong Kong.

Katherine Hindley is Assistant Professor of Medieval Literature at Nanyang Technological University in Singapore.

Joe Stadolnik is a Postdoctoral Researcher and Instructor in the Stevanovich Institute on the Formation of Knowledge at the University of Chicago.

technical skills for which training is difficult to access. The DEMMR learning model brings together a group of students with varying levels of prior knowledge in these areas to work collaboratively to extend their skills with structured teaching support. The course opens a door for newcomers to manuscript study and digital humanities, even as it acknowledges that a short course cannot address every detail of these scholarly fields. A vital aspect of this accessible introductory course of study is that it reveals a pathway to the further advanced study necessary for mastery of the field in a way that is accessible and welcoming to a diverse population of beginning students.[2]

DEMMR is a flexible course model that can be modified for a range of student audiences. It has been successfully taught to groups of graduate students, with participation from early-career scholars and undergraduates (usually in classes of 12 to 20). The course has most frequently been offered in the U.S., and has drawn participants from a variety of academic disciplines and institutions across North America. DEMMR has also been offered in London to graduate students, and in Australia to both graduate students and to a mixed group of under-graduates, novice postgraduates, and nonacademic participants from the local community. The model has also been taught as a unit within a semester-long undergraduate course in Singapore. The focus of the learning model on collaborative and supported learning makes it an adaptable program attractive to a range of student audiences and classroom situations.

**Digital Editing and the Medieval Manuscript Roll**

The creation of a digital edition of a medieval manuscript is the energizing goal of this course. To achieve this end, it covers the fundamentals of medieval manuscript cataloging and transcription, collaborative editing, text encoding, and the use and marking up of digital images. Participants work together to produce an edition of a medieval manuscript, complete with digital facsimile, edited text, and catalog data. They are introduced to the Extensible Markup Language (XML), the digital editing Schema of the Text Encoding Initiative (TEI Schema), and the XML editing software Oxygen. XML is a general coding language used for marking, or "tagging," data. In theory, an XML document can tag any information in a text—marking and making searchable, for example, all nouns, all people, or all uses of the color blue. The TEI Schema provides a set of guidelines for textual scholars specifying which XML tags should be used to mark

which textual features, thus offering a common scholarly standard for creating machine-readable text for academic purposes.

The course takes a hands-on approach that requires students to apply their technical skills as soon as—or while—they learn them; in this way, it seeks both to train participants in new skills and to provide them with the real experience of digital manuscript editing. The course uses the following minimal resources:

- A website for participants to post responses before and during the course.[3]

- Readings on digital textual editing assigned before the course.

- Handouts for use during the course (e.g., sample manuscript catalog entries and guidelines for transcription).

- The Oxygen XML Editor application for each participant.[4]

- A digitized manuscript image (either a digitization of a manuscript housed at a local institutional library, or facsimile images found on one of the growing number of repositories of open-access manuscripts).

- PowerPoint slides developed by instructors (slides used in DEMMR, and especially those relating to the TEI Schema, are in part based upon slides generously made available under a Creative Commons Attribution License as part of the TEI @ Oxford Summer School.)[5]

Instructors divide the chosen manuscript into sections and distribute these among the students, so that each one has a manageable selection for transcription and description; instructors also put students in pairs, chosen to balance any differences in prior knowledge. Before the course, participants are required to read the "Guidelines" on the website of the Text Encoding Initiative (http://www.tei-c.org/guidelines/) as well as two articles dealing with issues of text encoding. We have found Lou Burnard's "On the Hermeneutic Implications of Text Encoding" and Amanda Gailey's "A Case for Heavy Editing"[6] to be good introductory texts, in that they provide a range of justifications for the work of text encoding with which students can immediately engage. Participants post online a brief response to the two articles, which introduce both the benefits and limitations of text encoding and the TEI Schema.

The contact hours of the course are divided into eight taught sessions of 75 to 90 minutes each (see Table 1). In a two-day workshop, this means two morning classes and two afternoon classes each day. The first two sessions are ideally undertaken in a manuscript library with the physical manuscript, although they have been successfully taught without this access (in this variation, more time is given to discussing the benefits and challenges of working from digital artifacts, and the pedagogical focus shifts slightly towards theories of digital humanities). In the first session, items from the collection, including the focus manuscript at the center of the workshop, are shown and described, and students are encouraged to consider the meanings inherent in material form and textual layout. When no manuscript library can be visited, manuscripts may be selected from a freely accessible online collection, such as Parker Library on the Web.[7] The second session covers catalog conventions (and their limitations for objects in roll format, in the case of DEMMR). Students work in pairs to research and draft an assigned section of the catalog entry. The third and fourth sessions introduce text encoding and XML, the Oxygen XML Editor, and the TEI Schema. In the fifth session, some challenges of encoding nonstandard texts (such as rolls) are introduced; students then collaborate to determine the editorial conventions they will use for the manuscript in response to those challenges. In the sixth session, students encode the links between digital manuscript facsimile images and sections of transcribed text they have prepared. In the seventh and eighth sessions, the technical rules of TEI encoding for catalog information are minimally set out. Student pairs are then allowed a substantial amount of time to complete their assigned section of the catalog entry and their encoded, transcribed texts using the Oxygen XML Editor, according to rules they have agreed upon in session five.

By the end of the eighth session, most students have completed their transcriptions and encoding, and have submitted their work to the instructors. They may alternatively submit their markup text via email within one month of the completion of the course. The instructors compile these files into a single XML document, which is developed into a digital edition using a modified version of the freely available software of the Edition Visualization Technology (https://github.com/evt-project/evt-viewer).

Table 1. Syllabus Overview of Eight-Session DEMMR Course

| Session No. | Instruction Topics | Participant Tasks |
|---|---|---|
| 1 | • Introduction to manuscripts: show and tell.<br>• Introduction to focus MS: distribute sections and introduce partnerships.<br>• "Rules" of MS transcription. | *Think:* What are the challenges these MSS pose to editors? Why and how should they be digitized?<br>*Think:* Using the transcription handout as a guide, consider how you will transcribe your MS. |
| 2 | • How to catalog an MS or early printed book.<br>• Practical session: provide guided description of our MS. | *Think:* What are the shortcomings of a traditional MS catalog? Discuss examples on the handout.<br>*Do:* Get to know your MS, and develop the text for your part of the MS catalog entry. |
| 3 | • Introduction to text encoding.<br>• Introduction to computer processing, markup language, HTML, and XML. | *Think:* What is "text?" Why encode?<br>*Learn:* What is a coding "language?" What is HTML? What is XML?<br>*Do:* In pairs, race to write a short piece of XML by hand. |
| 4 | • How to use Oxygen.<br>• Overview of TEI, its structure and common elements.<br>• Using Oxygen for TEI P5. | *Do:* Follow along to set up a well-formed XML document in Oxygen.<br>*Think:* How does TEI constrain the XML tags and document structure?<br>*Do:* Follow along to set up a validated (TEI P5) XML document in Oxygen. |
| 5 | • Lecture on some issues of TEI for nonstandard texts: ideas for solutions.<br>• Practical session: make editorial decisions for our MS. | *Think:* How can we use and adapt TEI P5 hierarchies to represent our MS?<br>*Do:* As a group, discuss and decide upon your editorial policy. |

| Session No. | Instruction Topics | Participant Tasks |
|---|---|---|
| 6 | • Lecture on digital images and linking encoded text with coordinates on an image.<br>• Practical session: insert links between <surface> areas and <div> sections. | *Think:* How should image and text be digitally linked? For whom is the edition being created?<br>*Do:* Create some image links using pixel coordinates in your XML template. |
| 7 | • Lecture on TEI metadata: <teiHeader> and <msDesc>.<br>• Practical session: build header section of XML document. | *Learn:* What information goes in the header of a TEI P5 document?<br>*Do:* Encode the catalog information developed earlier into the header. |
| 8 | • Practical session: transcribe MS, and build text section of XML document. | *Do:* Encode your transcription, and complete your header, image links, and text markup. |

## Project-Based Learning with Manuscripts

A single manuscript object (available in digital form) serves as the focus of each course. For the first DEMMR course, we chose the Beinecke Rare Book & Manuscript Library's Takamiya MS 35, a thirteenth-century chronicle of the kings of England with a fifteenth-century continuation. The manuscript is a parchment chronicle roll just over eleven feet long and presents a range of challenges for the potential editor: a complex layout moving between one and two columns, a later continuation, an incomplete text, multiple languages, and visually elaborate genealogical diagrams. Takamiya MS 35 was selected deliberately for these varied and frustrating features. The textual issues of language, lacunae, and multiple dates of production expose students to a representative range of editorial challenges. The complexities of layout focus attention on a crux of pre-modern textual editing: how to represent in the format of a modern edition the visual elements and organizational principles of the medieval manuscript artifact. Subsequent courses have selected similarly confounding texts.[8] Such artifacts challenge student-editors to clarify the visual and textual complexities for a modern readership, thereby forcing a careful consideration of audience, transcription, layout, and, above all, the purpose that a modern edition should serve. In creating detailed catalog information and a complete transcription, participants

experience a genuine sense of undertaking important work in their field, under constraints of time, resources, and available technology, all of which anticipate the real challenges of major editorial and digital projects. Carefully chosen partnerships that pair students whose skills are complementary (e.g., a student with some digital experience with a student who can transcribe Latin) allow even novices to participate in this editorial process. Challenged to produce and follow a comprehensive editorial plan that digitally captures the complexities of the manuscript object, students are inspired to master the technicalities of the systems of markup as they discover the possibilities, and fight against the limitations, of digital markup systems.

These workshops focus on manuscript materials that record previously unedited texts or that bear on broader research projects for pedagogical reasons: student enthusiasm is generated by the feeling of engagement with a real research project. Thus, authenticity is at the center of our project-based learning model.[9] The problems with which participants are presented and which they themselves discover and solve are genuine ones. Accordingly, instructors ensure that the scaffolding created by their preparation and lesson planning does not invalidate the authenticity of the project experience. For example, course instructors of DEMMR transcribe the focus manuscript before the course, so that students unfamiliar with paleography may be given a crib if they are still struggling with their transcriptions on day two. However, students are made aware that this transcription is rough and may be inaccurate: final decisions rest with them.

## Active Learning and Teaching with Markers and Poster Paper

Despite the short length of the course and the quantity of material to be covered, its design deliberately avoids the lecture model. Instead, every session contains several active learning tasks in which participants solve the problems they encounter. One consequence of this model is that we cannot familiarize participants with all the features of XML or the TEI, or train them in a full range of paleographic skills. Our aim instead is to increase participants' confidence with the tools relevant to the focus manuscript, thereby helping them gain an understanding and sense of control over the processes by which they can increase their own knowledge after the course.

Active learning in the course ranges from simple breakout tasks ("30 seconds: turn to the person beside you, and tell them the main

differences between HTML and XML")[10] to more involved activities. For example, session three (which introduces XML) places the course partnerships in competition with one another: each races to be the first to create a well-formed XML document encoding a few lines of medieval poetry. Crucially, teams write out their code by hand, with poster paper and Sharpie markers, painstakingly re-creating and correctly applying the symbols and syntax of XML. Their under-standing of the rules of XML is immediately exposed, and under the (deliberately) rather panicked conditions of a race, the way in which tiny errors in placement can render the whole code dysfunctional is brought home. As teams take turns to stand up with their completed handwritten code, the group judges their success ("Is there only one root element? Is there an attribute-value pair? Are the quotation marks correctly placed?"). This fifteen-minute activity regularly surprises both instructors and students with its success: the group is energized to put what they have just learned into practice, and instructors can immediately see which aspects of XML need more explanation. But perhaps the chief benefit lies in the avoidance of technological issues: there is no program to load (or fail to load) on each student's computer, and no need to lead the group slowly through opening and using an XML editor. They are coding before they open their laptops.

Throughout the course, the teaching approach privileges participants' own attempts to build knowledge and solve problems. This has two crucial effects for the course as a whole: first, it empowers participants to work through the complexities of the variously technical and arcane skills required in digital manuscript editing; and second, it shifts the focus away from the instructors' knowledge of technical minutiae and onto the group's ability to discover the information needed to move forward with their growing edition. The sessions become shared problem-solving endeavors, in which instructors work to develop the participants' own understanding of the issues at stake. This "owned understanding" is central, for example, in the session that requires collaborative editorial decision-making.

**Fighting for Collaborative Cross-Disciplinary Learning**

The TEI Schema offers a number of pedagogic benefits beyond its importance to textual scholarship. In order to understand TEI guidelines, one must learn XML, a markup language that is reassuringly simple and accessible even to novice students and can be

mastered quickly. This provides a huge boost to students' confidence with digital tools, since it demystifies the process of learning coding languages and gives a vital first insight into computer processes. The limitations of TEI are as important to the course as its capabilities. One major issue relates again to the choice of focus text: the TEI system (as with any XML validation schema) is hierarchical—it best describes a codex text like the Victorian novel, with its set and neatly nested hierarchy of book, chapter, paragraph, sentence, and word. The meandering genealogies and rich illustration of pre-modern manuscripts (particularly those with a complex format like Takamiya MS 35) challenge the TEI's fundamental system and expose its limitations.

There is an important lesson here for digital manuscript scholars and editors. What the TEI Schema can and cannot mark about a medieval text, and which of its features should be used, is the centerpiece of session five. In the first four sessions, students have been introduced to their manuscript and guided through gathering information using rulers and magnifying glasses. Based on this work, they prepare a detailed catalog entry about size, ink colors, decoration, damage, and stitching. They are familiar with the manuscript, and their own descriptions of its features have given them some sense of ownership: it has become *their* manuscript. They have also been introduced to the basics of XML and the TEI, enough to be aware of the limitations and rigidities of these systems of data encoding. The group is thus primed for some real editorial decision-making. To facilitate this session, instructors provide a large roll of paper, numerous Sharpies, and sticky labels colored red, yellow, and green. On the long roll paper, participants together write the elements of the text they think should (or should not) be marked up. Some students may insist that all abbreviations be noted, expanded, and encoded twice using the <choice> tag, so that original text and editorial intervention can both be displayed. Others may feel strongly about the encoding of all colors of ink in the manuscript. Still others may point out that tagging all place-names will be useful as a future search feature. As students fill the paper with their ideas and preferences, instructors prompt them with questions, seeking to open up deeper inquiry: How should proper names be treated? Should we modernize spelling? What happens if we mark that one prominent scribal punctuation mark early in the manuscript—do we have to mark all punctuation thereafter? Students are next given sticky labels in "traffic light" colors, which they use to vote on one another's editorial choices.

By writing their names on a red sticky label and placing that label beside a requested editorial choice, they announce their opposition; green marks wholehearted approval, and yellow asks for further discussion or explanation. Instructors jump into the conversation to encourage engagement with editorial decisions, which tends to produce lively discussions of the worth of one editorial method above another. As the allotted 75 or 90 minutes begins to run out, the discussion becomes more urgent: students will have to abide by the decisions reached, which will impact their own time and effort, as well as the product of their combined work.

In this session, the benefits of cross-disciplinary collaboration become apparent. Our courses have drawn participants from many different humanities fields, and elements of a manuscript that are treated as peripheral to the interests of one discipline are frequently revealed in this session to be absolutely central to another. From participants' discussions about editorial policy, then, emerges a dialogue among fields of study. This interaction, in addition to the partnerships into which all participants are placed, makes the course a truly cooperative one, in which successful progress to the completion of the project goal depends upon interdisciplinary collaboration, both in the construction of the edition itself and in the immediate context of the sessions.[11] Experience has taught us that these deliberations over collaborative editing discussions are among the most productive and successful elements of the course. But it takes very engaged and proactive instruction to manage this session. The instructor must know which of the groups' ideas absolutely will not work within the constraints of the technology, thereby steering discussion towards consensus by the close of the session in order that the handwritten notes and votes might be translated into a series of collective rules to be followed by the group. Following this session, these collectively agreed-upon rules are posted on the course website by an instructor and become a reference for students in later editing sessions.

## Learning through Teaching: Peer Instruction in the Graduate Context

From its pilot program and in the ensuing regular workshops, DEMMR teaching has been conducted entirely by volunteer graduate students. These instructors are generally students who have taken the course in the past, or who have some experience with manuscript studies or digital editing. Each instructor is responsible for a small

amount of the course content, either individually or as a member of a partnership, in which experienced instructors are paired with newcomers. Instructors are encouraged to manage their own study of the necessary skills using free online courses like the XML tutorials on www.w3schools.com. Although the course can be adapted for the undergraduate classroom, or as a semester-long graduate-level course, there are many benefits to this peer-teaching model, in which graduate participants become instructors in subsequent courses. These include the low cost of the course and a welcoming atmosphere in which new participants are from the outset invited to think of themselves as future instructors. It is in keeping with the active, student-focused and project-focused model of the course that peer teaching be the primary mode of instruction. It is not, ultimately, the instructors' expertise that drives learning; rather, it is the instructors' ability to present the challenges of the course as inviting opportunities for participants to work with pre-modern manuscript materials and to do so using a suite of new digital skills and tools. In this way, teaching a DEMMR workshop becomes part of the learning model of the course itself. The move from student to teacher reinforces the skills learned while offering valuable teaching experience in manuscript studies and digital editing. Graduate instructors themselves participate in a complex collaborative project: checking one another's transcriptions, developing handouts on codicology and paleography, and working to help participants solve the many issues that emerge in the process of encoding manuscripts using the TEI Schema. DEMMR is a perfect demonstration of the dictum that "you learn by teaching."

**Adapting the Course**

The course as a whole adopts a readily generalizable position on pedagogy and the study of manuscript and text. As Burnard has argued, the process of encoding a text requires that its implicit features be labeled and made explicit: "Before a text can be encoded, it must first be decoded."[12] This process of articulating meaning can be put to practical use in the manuscript studies and textual scholarship classroom. The assumptions of the textual community whose shared understanding of language and manuscript culture worked to "create" the original text must be identified and articulated by students, and the features of the text prioritized, before it can be encoded. This text is an abstraction, a product of the process of making explicit and assigning a value to each of the visual elements of the manuscript at

the center of the "edition." As a pedagogical exercise, this process is as much about method as it is about content; as Matthew K. Gold puts it, "what sets the digital humanities apart from other disciplines is its commitment to building things as a way of knowing."[13] The central building project of the DEMMR course as a "way of knowing" is replicable across different fields and student levels. Its benefits are many:

- It forces close interaction with the material selected for encoding, and privileges active rather than passive responses to text.

- It necessitates an evaluation of the pros and cons of new technologies in textual editing.

- It requires a thoughtful defense of final positions adopted on issues of textual meaning and editorial structure.

- It makes collaboration and cooperation central to the success of the learning endeavor.

- It allows students to experience fully the real work of the discipline (and the interdisciplinary areas) within which they aspire to work.

This level of rigor and empowerment is both desirable and possible across a range of disciplines and interests in the digital humanities field. The peer-teaching/team-teaching approach is a highly flexible one and fits well with the requirements for authentic learning. We have found the experience of working together to design the DEMMR course enjoyable and instructive, and have been privileged to share the enthusiasm and extend the skills of our participants. We are very grateful for the insights into learning and teaching, as well as into manuscript studies and digital tools, which their generous participation has given us.

**Notes**

[2]Barriers to entry (particularly with regard to the digital-technological field) can be especially pernicious for women and scholars of color. On the benefits of bringing digital humanities into the classroom for women in particular, see Kara Kennedy, "A Long-Belated Welcome: Accepting Digital Humanities Methods into Non-DH Classrooms," *DHQ: Digital Humanities Quarterly* 11, no. 3 (2017): 6–20.

[3]Working within an institutional context, this site has been successfully set up on university CoursePress or CampusPress sites. Working outside the institutional context, organizers have created their own free WordPress sites to host the course website: these require little to no web design experience to create.

[4]The Oxygen Editor is available for download on a 30-day free trial with full functionality at www.oxygenxml.com.

[5]Slides used as part of the development of the DEMMR course were those made by James Cummings. See http://tei.oucs.ox.ac.uk/ for a list of Creative Commons teaching resources relating to the TEI.

[6]Lou Burnard, "On the Hermeneutic Implications of Text Encoding," in *New Media and the Humanities: Research and Applications*, ed. D. Fiormonte and J. Usher (Oxford: Humanities Computing Unit, 2001); Amanda Gailey, "A Case for Heavy Editing: The Example of Race and Children's Literature in the Gilded Age," in *The American Literature Scholar in the Digital Age*, ed. Amy E. Earhart and Andrew Jewell (Grand Rapids: University of Michigan Press, 2011).

[7]Parker Library on the Web: Manuscripts in the Parker Library at Corpus Christi College, Cambridge, https://parker.stanford.edu.

[8]Such as New Haven's Beinecke Library MSS 410 and Osborn a14, and London's Wellcome Collection MS 410.

[9]On the contextual instruction of project-based learning, see Phyllis C. Blumenfeld, Elliot Soloway, Ronald W. Marx, Joseph S. Krajcik, Mark Guzdial, and Annemarie Palincsar, "Motivating

Project-Based Learning: Sustaining the Doing, Supporting the Learning," *Educational Psychologist* 26, no. 3-4 (1991): 369–98. On the importance of authenticity as an element of this model, see the summary in Laura Helle, Päivi Tynjälä, and Erkki Olkinuora. "Project-Based Learning in Post-Secondary Education—Theory, Practice and Rubber Sling Shots," *Higher Education* 51, no. 2 (2006): 287–314, esp. 290–94.

[10]On the success of these simple models in raising retention levels, particularly in STEM subjects, see Scott Freeman, Sarah L. Eddy, Miles McDonough, Michelle K. Smith, Nnadozie Okoroafor, Hannah Jordt, and Mary Pat Wenderoth, "Active Learning Increases Student Performance in Science, Engineering, and Mathematics," *PNAS: Proceedings of the National Academy of Sciences of the United States of America* 111, no. 23 (2014): 8410–15.

[11]On the central elements of cooperative work such as this (including positive interdependence, individual accountability, face-to-face promotive interaction, and group processing), see D. Johnson, R. Johnson, and K. Smith, *Active Learning: Cooperation in the College Classroom* (Edina, MN: Interaction Book Company, 1998).

[12]Burnard, 43.

[13]Matthew K. Gold, quoted in "Day of DH: Defining the Digital Humanities," in *Debates in the Digital Humanities*, ed. Matthew K. Gold (Minneapolis: University of Minnesota Press, 2012), 67–74, at 69.