

History GR8975  
**What is a Book in the 21st Century?**  
**Working with Historical Texts in a Digital Environment**

Spring 2017  
Wednesdays, 4:10pm-6pm  
Friday Labs, 2-4pm

**INSTRUCTORS:** Terry Catapano (CU Libraries) and Pamela Smith (History), with guest lectures by Steven Feiner (Computer Science)

**Course Instructors**

Prof. Pamela Smith

Prof. Terry Catapano  
Special Collections Analyst, Columbia Libraries (DLIST)

**Project Manager**

Naomi Rosenkranz

**Course Assistants**

Atif Ahmed  
Mehul Kumar  
Varsha Maragi  
Jeffrey Wayno

This course will introduce graduate students to techniques of working in digital environments. The course is intended mainly for humanities and social science students who are novices with little or no experience in using digital platforms, but we also welcome students from all disciplines, as well as those who might be familiar with constructing websites or blogs, or even with creating minimal editions. Through hands-on assignments (with plenty of assistance), you will master a variety of skills that constitute literacy in digital humanities, and, by the end of the semester, you will be able to take your newfound digital literacy with you as you pursue your own study, research, and future work.

Throughout the course, your skills will be built by implementing them to collectively create a small scale digital edition, which will be festively launched at the end of the semester. This digital edition will draw on collaboration with and research done by the [Making and Knowing Project](http://www.makingandknowing.org/) (<http://www.makingandknowing.org/>) on an anonymous sixteenth-century French compilation of artistic and technical recipes (BnF Ms. Fr. 640).

The Project's existing English translation of this manuscript will constitute the "data" with which students in this course will work to create their small scale edition.

This rare French manuscript resulted from the compilation of craft knowledge over time, followed by its subsequent "disassembly" in a late sixteenth-century workshop by an author-compiler-practitioner who experimented on techniques contained in the manuscript's "recipes." While the course will focus on this intriguing manuscript and the research that has been carried out on it, the skills you will learn over the course of the semester are widely applicable to other types of Digital Humanities projects, and, indeed, in many fields outside of traditional academic study.

The [Making and Knowing Project](#), directed by Professor Smith, has produced the transcription and English translation of this manuscript, "disassembling" Ms. Fr. 640 through research seminars and workshops, involving multidisciplinary teams of students and scholars. The Project is now engaged in creating a complete critical digital edition, which represents a reassembly of this manuscript in a 21<sup>st</sup>-century form. In this course, you will be an active participant in the Project's exploration of the technologies that allow not just a *reading* of the text but an *interaction* with the content itself. This is in direct resonance with the ways that this sixteenth-century recipe collection can only be transformed from *text* to *knowledge* when the techniques contained within it are *practiced*, whether in the sixteenth century or in the Making and Knowing Laboratory reconstructions today. Through this exploration, the course aims to foster reflection on the constraints of the codex as a framework and vehicle for the production of knowledge, and to re-think the technology of the book and what it means to read a text. To this end, the course also includes collaboration with Professor Steven Feiner's Computer Graphics and User Interfaces Lab ([CGUI](#), <http://graphics.cs.columbia.edu/home/home/>).

*This course is one component of the [History in Action Initiative](#) of the Columbia Department of History. The American Historical Association (AHA) and the Andrew W. Mellon Foundation are collaborating to re-think career education for history PhD candidates at four selected universities (Columbia, Chicago, New Mexico, and UCLA) and to continue, expand, and enhance the AHA's "[Career Diversity and the History PhD](#)" initiative. The long-term goal is to establish a new norm: that doctoral graduates in history and the humanities will be equipped with the skills to pursue a wide spectrum of career opportunities and communicate their research to a broad audience.*

**ASSESSMENT:**

Participation, initiative, effort: 10%

Weekly assignments and field notes: 30%

Final edition project: 60%

## **SCHEDULE:**

**Please note: You will encounter many unfamiliar and possibly intimidating terms in the following syllabus, but FEAR NOT! Learning a new craft involves not just “how to do” it, but also “how to talk” about it. Hands-on techniques are in general difficult to put into words, so this practitioners’ jargon is often necessary.**

**Please see here for a short and easy to read version of the class schedule and syllabus that includes the digital skills introduced in each class.**

**Be sure to bring your computer to every class.**

### **Week 1: Jan 18 - Introduction**

Get to know your many collaborators in this class!

#### **To prepare in advance of the class on Jan 18:**

- **To do:**
  - If you do not already have one, please create a GitHub account.
- **Read and Explore:**
  - The [Making and Knowing Project website](#)
  - The Making and Knowing Project [Flickr account](#), the Project’s photo repository from lab reconstructions
  - Computer Graphics and User Interfaces Lab ([CGUI website](#))
  - [Digital competencies](#) (and the larger site: <http://allaboardhe.org/>)
    - See also the [Short Guide to the Digital Humanities](#)

#### **In class on Jan 18:**

- Introduction
  - Aims and overview of the course (Smith and Catapano)
  - The Making and Knowing Project and BnF. Ms. Fr. 640 (Smith)
  - Digital literacy - resources and digital competencies (Jessica Brodsky)
  - Overview of and Digital Editions and Editing (Catapano)
  - Computer Graphics and User Interfaces Lab (Feiner)

#### **Homework assignment Jan 18:**

- Reading (for lab on Jan 20):
  - Milligan and Baker: [Introduction to the Bash Command Line](#)
- Assignment 1 (due Jan 25):

- Begin to familiarize yourself with your assigned folios from the [course GitHub](#) and read through them.
- Complete Digital Competencies Evaluation #1 and permission and contribution forms, and bring them to class.
- Reading (for Jan 25):
  - “User Story” [https://en.wikipedia.org/wiki/User\\_story](https://en.wikipedia.org/wiki/User_story)
  - G. Thomas Tanselle. *A Rationale of Textual Criticism* (Philadelphia: University of Pennsylvania Press, 1992). Available at Book Culture.

### **Lab 1: Jan 20 - workshop with Dennis Tenen**

- Introduction to Command Line, and help with GitHub
  - Be sure to have read Milligan and Baker: [Introduction to the Bash Command Line](#)
  - [D. Tenen GitHub notes and tutorials](#)
  - For Reference: [Cunix/unix tutorial](#)
    - [More info about using PuTTY](#) (for Windows users)
- Practical help with GitHub

### **Week 2: Jan 25 - General introduction to text editing and scholarship**

What is a “book”? How does it organize text and content? What aims does it achieve? Who does it reach? What is Scholarly Editing and Textual Criticism? What are the rationale, purposes, scope, and features of scholarly editions?

#### **In class on Jan 25:**

- Discussion: What *is* a book? Digital Humanities projects, scholarly editions, user stories
- Introduction of the [Casebooks Project](#) by director, Prof. Lauren Kassell, Cambridge University

#### **Homework assignment Jan 25:**

- Assignment 2 (due Feb 1):
  - Read about User Stories: [https://en.wikipedia.org/wiki/User\\_story](https://en.wikipedia.org/wiki/User_story)
  - From class on January 25, think about the [Casebooks Project](#) and the [Making and Knowing Project](#), derive 3-5 user stories related to our proposed online edition of BnF Ms. Fr. 640, based on your reading of your folios. In class, we will discuss the user stories and create a document that we will collaboratively add to GitHub.

- If you can't come to Monday's class, there will be office hours for troubleshooting help announced later this week.

Reading (for Feb 1):

- Identify and read the annotations relevant to your folios.

## **Lab 2: Jan 27 - Wiki, GD, and GitHub workshop**

- Dennis Tenen, Introduction to GitHub
- Photos of today's lecture notes from Dennis Tenen
  - [Whiteboard](#)
  - [Workflow](#)

## **Week 3: Feb 1 - Data and Project Management**

How do we think about the social, intellectual, and physical infrastructure of producing a "book" or a "digital project"? What is distinctive about digital projects? What is the range of concerns for a digital editing project?

### **In class on Feb 1:**

- Class discussion and exercise:
  - User stories
  - The whole class will:
    - Discuss and refine user stories and collaboratively contribute to a shared document on the class Github repository
    - From user stories create "feature requests" in the issue tracker
- Lecture and discussion:
  - Project Management
    - "Agile" development and management
    - Collaboration and Communication
    - Release Management
    - Technical Debt and Digital Obsolescence
  - Data Management
    - Identifiers
    - Metadata
    - Tracking
    - Preservation and Sustainability
    - Licensing: Creative Commons and open access
    - Optimising for Re-use

### **Homework assignment Feb 1:**

- Begin taking rudimentary field notes that record the process of doing your homework.
- Metadata: For 5 of your folios (r and v), create a metadata table
  - Remember, please read and use the folios in GitHub. At the same time, have a look at the manuscript pages, including looking at the HD images of your folios.
- If you come up with additional metadata fields, please create a new issue in GitHub. Also use the issue tracker if you come across problems while filling in the table.
- Reading (for Feb 8):
  - Federal Agencies Digitization Guidelines Initiative “[Technical Guidelines for Digitizing Cultural Heritage Materials](#)”

### **Lab 3: Feb 3 - Metadata**

What are the categories by which we organize our “content,” our “materials,” our “digital assets”?

- We will start this assignment in the lab, and students will finish at home:
  - Create master table of metadata elements collaboratively, to be filled in for homework: What are the considerations we may need to have for creating a digital edition? What should our metadata be?
  - Here is the Schema/template we came up with in class:
  - Use the template to create table of metadata for your assigned folios and add to GitHub.
    - Add to issue tracker as you come across issues while filling in the table

### **Week 4: Feb 8 - Digital Representation Fundamentals**

Representing a representation: How are images represented digitally? How are they viewed, processed, and referenced? What are their advantages and limitations?

#### **In class on Feb 8:**

- Discussion:
  - Reconciliation and resolution of metadata issues
- Lecture and discussion:
  - Digital image fundamentals
- Tool:
  - Viewshare

### **Homework assignment Feb 8:**

- Add a few more items to your metadata table
- Attempt to create a viewshare account, at viewshare.org
- Assignment 4 (to be started during Lab 4 and due Feb 15):
  - Using the new metadata table, enter the data for your entries
  - Update your field notes
- Reading (for Feb 15):
  - Read <http://www.regular-expressions.info/quickstart.html>
  - Re-read and re-do the exercises in Milligan and Baker: [Introduction to the Bash Command Line](#)
  - Take a look at “Unix for Poets”  
<https://web.stanford.edu/class/cs124/lec/124-UnixForPoets.pdf>, we will be doing the exercises together in class
  - If you’re interested, here is a video on regular expressions:  
<https://www.youtube.com/watch?v=DRR9fOXkfRE>

### **Lab 4: Feb 10 - Using metadata in Viewshare**

- Review metadata table - create a composite table
- Create presentation metadata in Viewshare
- Start Assignment 4 (if needed, finish for homework)

### **Week 5: Feb 15 - Text Fundamentals**

What is digital text? What can it do that printed type on paper cannot? How may digital or “electronic” text be “processed”? What sorts of study and inquiry does text “processing” facilitate? How does the way digital text is “prepared” affect its possible uses?

### **In class on Feb 15:**

- Metadata and Viewshare conclusion
- GitHub and command line hands-on
- Lecture and discussion:
  - Text fundamentals, representation, and encoding
  - What can you do with digital text?
- Regular expression exercises: see: <http://dh.obdurodon.org/#regex>
- Tool:
  - Linux command line text utilities

### **Homework assignment Feb 15:**

- Assignment 5 (due Feb 22):

- Continue with last week work
- Play around with your cloned Viewshare “View”
- Try Unix for Poets exercises on your command line
- Update your field notes

Reading (for Feb 22)

Be sure to have read and worked with previous assignments:

- <http://www.regular-expressions.info/quickstart.html>
- Re-read and re-do the exercises in Milligan and Baker: [Introduction to the Bash Command Line](#)
- Take a look at “Unix for Poets” <https://web.stanford.edu/class/cs124/lec/124-UnixForPoets.pdf>, we will be doing the exercises together in class
- If you’re interested, here is a video on regular expressions: <https://www.youtube.com/watch?v=DRR9fOXkfRE>

### Lab 5: Feb 17 - GitHub lab

- Git/Github and command line hands-on

### Week 6: Feb 22 - More on Text Processing, Command Line, Utilities, and Version Control in GitHub

The mess of digital reproduction: how to maintain control of content, issue, edition, “release”? How can digital tools accommodate textual “instability”?

#### In class on Feb 22:

- Hands-on Exercises:
  - Text Processing, Command Line, Utilities, Version control and representation of textual variance in traditional critical editions

#### Homework assignment Feb 22:

- Assignment 6 (due Mar 1):
  - Update your field notes
- Reading (for Mar 1):
  - Tenen, Dennis and Wythoff, Grant: [Sustainable Authorship in Plain Text using Pandoc and Markdown](#)

### Lab 6: Feb 24 - SPEAKER - STUDIO@BUTLER 2-3PM

“Digital Amati: Structure and Interpretation of Classical Stringed Instruments” by Harry Mairson, professor of computer science at Brandeis University. Professor Mairson is also an amateur violoncello maker, has conducted research on type systems in programming languages and their relation to problems in logic and complexity theory. In this lecture, he introduces the Digital Amati Project which explores the structure, interpretation, and making of stringed instruments, and how modern software can be used to represent historical practices of instrument design. The lecture discusses digital humanities tools, and the creative work done with them, and will be of interest to historians, musicologists, practitioners of digital humanities, and makers. Link to event page can be found [here](#). The event is hosted by the Making and Knowing Project and co-sponsored by History in Action in the Columbia University History Department.

## **Week 7: March 1 - Text Markup: Introduction and Overview**

Digital text: How it works in practice, part 1. Approaches for preparing textual data to represent implicit “formal” or “structural” features

### **In class on Mar 1:**

- Go over homework
- 4:30-5:15pm: Alex Gil: Introduction to structured text, basic markup technologies
  - Markdown and Jekyll platform
  - HTML/XHTML/HTML5
- 5:15-5:30: hands-on markdown.
  - READ here:  
<https://guides.github.com/features/mastering-markdown/#syntax>
- 5:30-6pm: Review

### **Homework assignment Mar 1:**

- Assignment 7 (due Mar 8):
  - Read about markdown
  - Create files of the entries in your folios. Then commit AND push to the shared repository in GitHub
    - if you need a text editor, you can download Atom here:  
<https://atom.io/>
  - Practice markdown of your entry files.
  - Consider the affordances and limitations of markdown and bring your ideas to the next class about what more you want from markup of your folios
  - Update your field notes
- Reading (for Mar 8):

- DeRose SJ, Durand DG, Mylonas E, Renear AH. “What is text, really?” SIGDOC Asterisk J. Comput. Doc. 1997 Aug; 21(3):1-24. [doi: 10.1145/264842.264843](https://doi.org/10.1145/264842.264843)
- “What is XML and why should humanists care? An even gentler introduction to XML” <http://dh.obdurodon.org/what-is-xml.xhtml>

**NO LAB Week 7: March 3 - Terry Office Hours**

## **Week 8: Mar 8 - Text Markup Continued: Semantic Markup**

Digital Text: How it works in practice, part 2. “Text Encoding” or “Markup” for preparing textual data to represent both “formal” and “semantic” textual features.

### **In class on Mar 8:**

- Lecture and discussion:
  - Introduction to XML
  - Text Encoding Initiative (TEI) and customized markup
- Exercise:
  - Begin determination of our possible markup tag set
- Tool:
  - XML-aware text editor, eg. Sublime or Atom

### **Homework assignment Mar 8:**

- Assignment 8: Due Monday March 20:
  - Create an xml file with the name in the form:
    - [entry\_id].xml
    - For example: p104r\_4.xml
  - For entry files, follow the sample entry markup file
  - For marginal blocks, follow the sample markup
  - For the marginal block entry files’ filenames use the pattern:
    - [p+folio number]\_mb[sequence #].xml
    - For example: p001r\_mb1.xml
    - Determine the sequence number from the order that they appear in the entry files now. If you have individual files for marginal notes, just assign a different number to make the file names unique (don’t worry about “sequence”).
    - Find features that are not currently marked up and apply your own.

- Due Wednesday March 22: Compare and review your partner's tag set and, using the issue tracker, comment on your partner's tag set. Be prepared to present to the rest of the class on Mar 22 during class.
- Update your field notes
- Digital Competencies Evaluation #2 [Please print this form, fill it out and bring it to class on March 22 to turn in]
- Prepare for NYU class visit.
- Reading:
  - John Lavagnino, "Electronic Textual Editing: When not to use TEI" [http://www.tei-c.org/About/Archive\\_new/ETE/Preview/lavagnino.xml](http://www.tei-c.org/About/Archive_new/ETE/Preview/lavagnino.xml)

### **Lab 6: Mar 10 - Markup - Fayerweather 513**

- Text editor and markup troubleshooting and help
  - Working in teams (form teams)
  - Assign peer reviewers
  - Choose presentation/presentation for March 24

### **Spring Break: Mar 15 - NO CLASS**

- Working on Assignment 8 [DUE Monday Mar 20 to your peer reviewers]

### **Week 9: Mar 22 - Text Markup Continued: Establishing Consensus**

Digital Text: How it works in practice in collaborative projects (part 3). How to decide what to tag and what not to tag. The role of the "schema" in formally defining (i.e., for a computer) a "document type" or "tag set"

#### **In class on Mar 22:**

- Present your comparison and review of your own and partner group's tag set. Discussion of different markup
- Begin to form a consensus markup strategy
- Begin to formalize this consensus markup in a schema
- Begin applying the consensus markup - troubleshooting, understanding, reporting, diagnosing, and fixing errors

#### **Homework assignment Mar 22:**

- Continue to think about semantic issues. Come up with things you think might be significant to the user?
- Assignment 9 (due Mar 29):
  - Apply consensus markup (in its current form) to your folios

- Review partner group's markup (through issue tracker)
  - Identify Issues:
    1. Bugs
    2. Commentary
- Update your field notes
- Reading (for Mar 29):
  - David Birnbaum: Digital humanities course XSLT materials
  - <http://dh.obdurodon.org/#xslt>
    - <http://dh.obdurodon.org/xslt-basics.xhtml>
    - <http://dh.obdurodon.org/xslt-basics-2.xhtml>

### **Lab 7: Mar 24 - NYU Student Visit**

- Meeting with graduate digital humanities students from NYU
- Optional Event: March 28th at 6PM at NYU Center for the Humanities. "Digital Humanities Meets Art Galleries" This event is free. More info can be found [here](#). The NYU Center for the Humanities website can be found [here](#).

### **Week 10: March 29 - Text Markup Continued: Establishing Consensus**

#### **In class on March 29:**

- Protocols and Guidelines
- Continue refining consensus markup: What problems have we encountered, and how might we overcome them? Think back to our user stories.

### **Lab 8: March 31 - Transformations, Representations, and Interfaces to Digital Resources Part 1**

Digital text: How it really works.

#### **In Lab on March 31:**

- oXygen setup and intro (XML editing and XSLT frameworks)
- Review and troubleshooting of your marked-up folios
- Lecture and discussion:
  - Transformation of XML - XSLT

#### **Homework assignment Mar 31:**

- Assignment 10 (due Apr 5):
  - Complete markup of as many entries as possible
  - Update your field notes
- Reading (for Apr 5):
  - Birnbaum DH course XSLT materials

- “Losing the Thread,” *Aeon*, 2016:  
<https://aeon.co/essays/how-textiles-repeatedly-revolutionised-human-technology>

## **Week 11: April 5 - Transformations, Representations, and Interfaces to Digital Resources Part 2**

Moving from preparation of digital textual data to “processing” and “application”, particularly “transformation” or “conversion” into appropriate formats for publishing in an online edition.

### **In class on Apr 5:**

- Transformation of XML - XSLT
- Publishing platform via Jekyll using Ed (<http://elotroalex.github.io/ed/documentation/>); see the Mini Lazarillo Project (<https://minilazarillo.github.io>)
- For more information on the technologies used for publishing our edition, see:
  - Ed (Jekyll theme):
    - <http://elotroalex.github.io/ed/documentation/>
    - mini lazarillo (uses Ed theme): <https://minilazarillo.github.io/>
  - Jekyll:
    - <http://jekyllrb.com>
    - "How (and Why) to Generate a Static Website Using Jekyll" <http://www.chronicle.com/blogs/profhacker/jekyll1/60913>
  - kramdown (markdown dialect used by Ed):
    - <https://kramdown.gettalong.org/quickref.html>
  - HTML/CSS/JavaScript:
    - Jump in at the appropriate level at Mozilla Developer Network 's "Learn Web Development" <https://developer.mozilla.org/en-US/docs/Learn>
  - Liquid (template language used by Jekyll): <https://shopify.github.io/liquid/>
  - And if you want to dive really deep, lunr and elasticlunr (client side search engine):
    - <https://lunrjs.com/>
    - <http://elasticlunr.com/>

### **Homework assignment Apr 5:**

- Assignment 11 (due Apr 12):
  - Transform marked-up pages to Jekyll markdown using Oxygen XML Editor

- Revisit feature requests from Week 3 and evaluate status of requests
- Update your field notes

### **Lab 8: Apr 7 - XSLT workshop**

- Hands-on XSLT session: We will discuss progress on markup, tackle problems, and refine our markup strategy accordingly

### **Week 12: April 12 - Transformations, Representations, and Interfaces to Digital Resources Part 3**

#### **In class on Apr 12:**

- Continued: Transformation of XML - XSLT
- Continue the process of refining our markup
- Troubleshoot problems with Oxygen XML Editor
- Update field notes

### **Week 13: Apr 19**

- Continue markup in light of ongoing challenges; refer back to user stories to help determine what we should be marking up, and why
- Begin discussion of what functionality we want our online edition to have
- Update field notes

### **Lab 9: Apr 21 - Computer Graphics and User Interface Lab**

- Presentation by Computer Graphics and User Interface Lab

### **Week 14: Apr 26 - Review and Conclusion**

Finalizing markup while preparing for the first launch of the digital edition

#### **In class on Apr 26:**

- Finalize markup strategy and discuss timeline to complete all entry files
- Discuss in more depth functionality required in our digital edition

### **Week 15: Digital Bytes and Bites**

#### **May 3**

- Working session to hone our markup and complete work for the launch of 1.1

#### **May 5**

- Working session to hone our markup and complete work for the launch of 1.1

## **Week 16: May 10 Digital Launch**

- Digital launch to be held