

HISTORY 1993
Introduction to Digital History
Conference Course – Fall 2022

Time: Wed, 12-2:45pm
Location: Robinson Hall 107

Gabe Pizzorno
pizzorno@fas.harvard.edu
Office: Robinson Hall 210
Office Hours: Tuesdays, 10am-noon, and by appointment
Course Canvas site: <https://canvas.harvard.edu/courses/109350>

Note: Please plan to bring a laptop or equivalent to every class meeting.

Summary

This course trains students in a range of digital methods used for the acquisition, analysis, and visualization of data in the context of historical research. Beyond developing practical skills, students will learn how to critically evaluate the potential and limitations of new technologies, and how to integrate them into their work in a careful, theoretically informed way.

Format

The mastering of digital methods requires practice and the opportunity to learn from one's mistakes. Thus, this course is taught in a "workshop" style designed to foster collective learning in a safe environment where students can experiment and learn to embrace failure. Both through hands-on exercises and a semester-long research project, participants will have the opportunity to engage with problems associated with real and complex datasets.

Materials and Supplies

You will need a computer, of course, but there are no particular requirements as to what kind. The vast majority of the software covered in the course is either open source, or freely available with educational licensing. Any additional software, or other computing resources, that you may need for your research project will be provided. In order to contribute to class conversations and take full advantage of the discussion, it is important to read the texts assigned each week in advance of our meetings. These

readings will inform our approach to the more practical aspects of the course. All course readings will be posted online.

Assignments

Throughout the semester, you will be expected to work on a research project on a topic of your choosing. The specifics of this semester-long project will greatly depend on your own research interests. It could, for example, take the form of a small piece of research, either stand-alone or a component of a senior thesis or dissertation. In all cases, the overall requirement is that you engage with an actual dataset at every stage of the workflow covered in the course (i.e. the acquisition, manipulation, analysis, and presentation of data). The final output of this project will take the form of an online exhibition. Work towards this final goal will be incrementally realized through two multimedia essays, each contributing elements to the final project exhibition. The first multimedia essay will describe the nature of the dataset being used for the project and outline the methodology that will be employed. The second essay will provide historical and disciplinary context for the project, and explain how the research intersects with existing scholarship on the subject. The final online exhibition, bringing together these elements and presenting the outcome of the project, will be due after the end of the term. During the last week of the term, students will present their project to the class. These presentations are not intended to showcase the final results of the project, but rather to give everyone an opportunity to bring up their ideas and unresolved problems and get help and feedback from peers before the final submission.

Contents

Digital history, and digital scholarship more generally, focuses on the use of digital evidence, methods, and tools as a means to integrate knowledge within interdisciplinary, collaborative frameworks. Digital scholarship does not constitute a new field in itself. It can better be described as a methodological approach that emphasises new ways of exploring existing research questions, generating new ones, and expanding the dissemination of scholarship to new media and audiences, by taking advantage of the enormous potential of digital tools to transform the way in which we create, analyze, and share knowledge.

This course introduces participants to the different steps of a data-processing workflow that translates the general research process of the social sciences and the humanities (collecting sources, analyzing them, and presenting the results). This approach to teaching digital methods is designed to provide students with a solid intellectual framework that contextualizes the methodologies, techniques, and tools covered in class (as well as any learned in future) by relating them to each other and to the general process of engaging in digital scholarship. While the focus of the class is on historical

research, the methods covered are directly applicable to a range of disciplines in the humanities and social sciences.

During the semester, participants will build a toolbox of practical skills to aid in their research. The focus of digital scholarship, however, is not on tools, but on methodologies and their careful integration into disciplinary practice. While engaging in digital scholarship it is crucial to see past the instrumental aspects of tools and platforms and focus on how they can be adapted, in a theoretically informed manner, to the unique ways in which scholarly disciplines approach their subject. The course, then, goes beyond the teaching of particular technical skills, and aims to make students conversant across the spectrum of digital techniques so that they can critically evaluate the potential and limitations of new technologies, integrate them into sound research programmes, and fruitfully interact with experts to both acquire and refine their skills and produce results.

Collaboration and Teamwork

In line with digital scholarship's emphasis on collaboration and interdisciplinarity, participants in this course are keenly encouraged to engage in teamwork, including on their semester-long research projects. In the case of the latter, proposals for team projects should provide a clear plan to ensure that individual contributions are tracked and acknowledged accordingly.

Grading

The final grade for the class will be determined as follows:

Participation	45%
Project proposal (Fri, 16 September)	5%
Essay 1: Dataset (Fri, 14 October)	15%
Essay 2: Context (Fri, 4 November)	15%
Project Exhibition (Tue, 13 December)	20%

Schedule and Readings

Week 1 (Sep 1): Digital Scholarship and its Tools

Introduction to the course and showcase of digital scholarship tools and methods.

Readings:

Cohen, Daniel J., and Roy Rosenzweig. 'Promises and Perils of Digital History'. *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*, University of Pennsylvania Press, 2005.

Week 2 (Sep 8) 2: Understanding Data

The nature of data, data formats, types, and structures. Finding and using online data sources and local repositories. Data collection strategies and digitization. Introduction to scraping.

Readings:

Marche, Stephen. 'Literature Is Not Data: Against Digital Humanities'. *Los Angeles Review of Books*, 28 Oct. 2012.

Owens, Trevor. 'Defining Data for Humanists: Text, Artifact, Information or Evidence?' *Journal of Digital Humanities*, vol. 1, no. 1, 2011.

Schöch, Christof. 'Big? Smart? Clean? Messy? Data in the Humanities'. *Journal of Digital Humanities*, vol. 2, no. 3, 2013.

Renear, Allen H. 'Text Encoding'. *Companion to Digital Humanities*, edited by Susan Schreibman et al., Hardcover, Blackwell Publishing Professional, 2004.

Week 3 (Sep 15): Cleaning Data

Basic techniques for data manipulation. Introduction to Open Refine.

Readings:

Thaller, Manfred. 'The Need for a Theory of Historical Computing'. *History and Computing II*, edited by Peter Denley et al., Manchester University Press, 1989, pp. 1–11.

Doulton, David, and Arno Kitts. 'The Storing and Processing of Historical Data'. *History and Computing II*, edited by Peter Denley et al., Manchester University Press, 1989, pp. 81–87.

Gibbs, Fred, and Trevor Owens. 'The Hermeneutics of Data and Historical Writing'. *Writing History in the Digital Age*, edited by Jack Dougherty and Kristen Nawrotzki, University of Michigan Press, 2013, pp. 159–70.

Schmidt, Benjamin M. 'Theory First'. *Journal of Digital Humanities*, vol. 1, no. 1, 2011.

Week 4 (Sep 22): Working with Data

Advanced techniques for the acquisition, cleaning, and manipulation of datasets.

Readings:

Knox, Doug. 'Understanding Regular Expressions'. *Programming Historian*, June 2013.

van Hooland, Seth, Ruben Verborgh, and Max De Wilde. 'Cleaning Data with OpenRefine', *Programming Historian*, August 2013.

Week 5 (Sep 29): Advanced Data Manipulation

Batch processing, scripting, and programming basics. Introduction to Python.

Readings:

Karsdorp, Folgert and Maarten van Gompel. 'Python Programming for the Humanities', Chapter 1: Getting Started.

Wolff, Robert S. 'The Historian's Craft, Popular Memory, and Wikipedia'. *Writing History in the Digital Age*, edited by Jack Dougherty and Kristen Nawrotzki, University of Michigan Press, 2013, pp. 64–74.

Week 6 (Oct 6): Online publishing

Online publication platforms and key issues: copyright, intellectual property, information overload, and data curation. Introduction to Omeka.

Readings:

Cummings, Alex Sayf, and Jonathan Jarret. 'Only Typing? Informal Writing, Blogging, and the Academy'. *Writing History in the Digital Age*, edited by Jack Dougherty and Kristen Nawrotzki, University of Michigan Press, 2013, pp. 246–58.

Wolff, Robert S. 'The Historian's Craft, Popular Memory, and Wikipedia'. *Writing History in the Digital Age*, edited by Jack Dougherty and Kristen Nawrotzki, University of Michigan Press, 2013, pp. 64–74.

Hitchcock, Tim. 'Academic History Writing and Its Disconnects'. *Journal of Digital Humanities*, vol. 1, no. 1, 2011.

Kalfatovic, Martin R. 'Chapter 1: Online Exhibitions Versus Digital Collections', *Creating a Winning Online Exhibition: A Guide for Libraries, Archives, and Museums*. American Library Association, 2002.

Week 7 (Oct 13): Exploring Datasets

Gaining insight into complex datasets. Principles of exploratory and inferential data analysis. Introduction to data modelling.

Readings:

Harris, Jacob. 'Word Clouds Considered Harmful'. *Nieman Lab*, 13 Oct. 2011.

Behrens, John T. 'Principles and Procedures of Exploratory Data Analysis'. *Psychological Methods*, 1997, pp. 131–160.

Michel, J. B., et al. 'Quantitative Analysis of Culture Using Millions of Digitized Books'. *Science*, vol. 331, no. 6014, Jan. 2011, pp. 176–82.

Thomas, William G. 'Computing and the Historical Imagination'. *Companion to Digital Humanities*, edited by Susan Schreibman et al., Hardcover, Blackwell Publishing Professional, 2004.

Week 8 (Oct 20): Visualisation

Highlighting the key aspects of a dataset in visually engaging ways to enable discovery and support research.

Readings:

Theibault, John. 'Visualizations and Historical Arguments'. *Writing History in the Digital Age*, edited by Jack Dougherty and Kristen Nawrotzki, University of Michigan Press, 2013, pp. 173–85.

Manovich, Lev. 'What Is Visualisation?' *Visual Studies*, vol. 26, no. 1, Mar. 2011, pp. 36–49.

Tufte, Edward R. 'Escaping Flatland'. *Envisioning Information*, Graphics Press, 1990, pp. 12–35.

Jessop, M. 'Digital Visualization as a Scholarly Activity'. *Literary and Linguistic Computing*, vol. 23, no. 3, Sept. 2008, pp. 281–93.

Week 9 (Oct 27): Time and Networks

Using time and relationships as organizing elements for data analysis. Introduction to timelines and network analysis.

Readings:

Moretti, Franco. *Graphs, Maps, Trees: Abstract Models for a Literary History*. Verso, 2005, pp. 1–64.

Weingart, Scott B. 'Demystifying Networks, Parts I & II'. *Journal of Digital Humanities*, vol. 1, no. 1, 2011.

Meeks, Elijah. *More Networks in the Humanities or Did Books Have DNA?* 2011.
Rosenberg, Daniel, and Anthony Grafton. *Cartographies of Time: A History of the Timeline*. Princeton Architectural Press, 2012, pp. 10–25, 210–237.

Week 10 (Nov 3): Space and Geography

Acquiring, manipulating, and analyzing spatial data.

Readings:

White, Richard. 'What Is Spatial History?' *Spatial History Lab*, Jan. 2010.

Ayers, Edward L. 'Turning toward Place, Space, and Time'. *The Spatial Humanities: GIS and the Future of Humanities Scholarship*, edited by David J. Bodenhamer et al., Indiana University Press, 2010, pp. 1–13.

Bodenhamer, David J. 'The Potential of Spatial Humanities'. *The Spatial Humanities: GIS and the Future of Humanities Scholarship*, edited by David J. Bodenhamer et al., Indiana University Press, 2010, pp. 14–30.

Bembeneck, Emily. 'Spatial Storytelling'. *Play the Past*, 27 June 2012.

Week 11 (Nov 10): Text and Words

Quantitative corpus linguistics and textual analytics. Introduction to topic modelling.

Readings:

Underwood, Ted. 'Seven Ways Humanists Are Using Computers to Understand Text.' *The Stone and the Shell*, 4 June 2015.

Jockers, Matthew Lee. 'Chapter 4: Macroanalysis', *Macroanalysis: Digital Methods and Literary History*. University of Illinois Press, 2013, pp. 24-32.

Blei, David M. 'Topic Modeling and Digital Humanities'. *Journal of Digital Humanities*, vol. 2, no. 1, 2012.

Brett, Megan R. 'Topic Modeling: A Basic Introduction'. *Journal of Digital Humanities*, vol. 2, no. 1, 2012.

Week 12 (Nov 17): Presenting Data

Communicating information clearly and effectively. Issues of perception, audience, and mediation. Style and aesthetics.

Readings:

Drucker, Johanna. 'Image, Interpretation, and Interface'. *Graphesis: Visual Forms of Knowledge Production*, Harvard University Press, 2014, pp. 17–63.

Schlatter, Tania, and Deborah A. Levinson. 'Introduction', *Visual Usability: Principles and Practices for Designing Digital Applications*. Morgan Kaufmann Publishers is an imprint of Elsevier, 2013, pp. xi-xix.

Cohen, Daniel J., and Roy Rosenzweig. 'Designing for the History Web', *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*, University of Pennsylvania Press, 2005.

Presner, Todd. 'How to Evaluate Digital Scholarship'. *Journal of Digital Humanities*, Dec. 2012.

Week 13 (Dec 1): Final Project Presentations

Presentation of final projects.