

Seth Bernstein
Keene Flint Hall 209
M: 3:00pm – 6:00pm
Course Site: <https://ufl.instructure.com/courses/499574>

sethbernstein@ufl.edu
office hours:
Th: 11-1pm
(or by appointment)

Graduate Seminar:
Digital Methods in History

Course Description:

This is a course in the use of computational/digital methods for research and writing in History. It has three main goals:

1. Familiarize students with the basics of digital methods of analysis like text scraping, natural language processing, geographical information systems and (to a lesser extent) machine learning.
2. Learn methods of digital presentation of materials through visualizations and venues for online presentation (e.g., blogs or digital exhibits).
3. Analyze the benefits and problems of digital scholarship and the use of big data.

I have tried to balance the examples in this course so that all the sections (EUH/LAH/WOH) are covered. Nonetheless, much of my knowledge of these methods comes from experience working in my subfield of Soviet History. For better or worse, students will probably learn something about the USSR during this course but they will also learn something about Latin America and Africa.

This course does not demand technical expertise and will not make you into an expert on digital history.

Course Assignments:

Final Project (45% including pass/fail 10% for a two-page proposal): The final project is a flexible assignment that can take two forms:

1. An essay (~12 pages or 3500 words) that uses digital analysis (e.g., database analysis or GIS) as its main method and includes at least one significant computational output (e.g., chart, map, table).
2. A proposal (~12 pages or 3500 words) for a major digital project. The proposal should identify the entire project workflow: data acquisition; data processing/annotation; analysis in relevant historiographical framework.

The topic for the final essay should correspond to the course code. In other words, EUH section students need to do a European topic.

How Did They Make That (5%): Each student will present on a digital humanities project with a special emphasis on method. Miriam Posner's video "How Did They Make That?" (see week 1) and my version on my own project is the model.

Discussant Paper (15%): Imagine that you have been invited to be a discussant at a conference that includes the project from your "How Did They Make That?" assignment and two other relevant papers by different authors that you will need to find. Write a short paper (5 pages) with comments that highlight the dialogue between the project and those two papers.

Weekly (more or less) Technical Assignments (15%): Most classes will include a short technical assignment. I am grading these pass/fail based on good faith effort rather than results.

Brainstorming (5%): Following each class (by the Sunday before the next class) you should write a quick idea for how to implement an idea we talked about last week. E.g., if we talk about mapping, you might make a proposal to for a map project in your area of study.

Participation (15%): Students should come to our meetings prepared to discuss the readings/materials for the week. As part of this preparation, students should submit on Canvas two questions for discussion.

Texts:

Our one main text on the history of Stalinism is:

Oleg Khlevniuk, *Stalin: A New Biography of a Dictator* (New Haven: Yale University Press, 2015).

On digital history, many of the texts will come from the following sources:

Graham, Milligan and Weingart, *The Historian's Macroscope* (draft version available online here:

http://web.archive.org/web/20160315232134/http://www.themacroscope.org/?page_id=707)

Debates in Digital Humanities (a series of essay collections, available here:

<https://dhdebates.gc.cuny.edu/>)

The Programming Historian (various technical tutorials, available at

<https://programminghistorian.org/>)

Other materials will be available on our Canvas page as PDFs or links.

Required Software:

A specific aspect of this class is that you will need access to a laptop computer for our in-class activities and to complete your class projects. Most of the programs are available online with a Google account. I recommend signing in to Google via your UFL account, possibly via an

incognito tab to avoid a conflict with a non-UFL Google account. A non-exhaustive list of programs we will use include:

- Google Colabs (<https://colab.research.google.com/> Python platform. It might be worth downloading Python on its own but Colab is convenient because it means that everyone is working in the same programming environment.)
- Google Sheets (<https://sheets.google.com/> Spreadsheet)
- QGIS (<https://qgis.org/en/site/> Mapping)
- Voyant Tools (<https://voyant-tools.org/> Linguistic Analysis)

Academic Integrity: The University of Florida holds its students to the highest standards, and we encourage students to read the University of Florida Student Honor Code and Student Conduct Code (Regulation 4.040), so they are aware of our standards. Any violation of the Student Honor Code will result in a referral the Student Conduct and Conflict Resolution and may result in academic sanctions and further student conduct action. The two greatest threats to the academic integrity of the University of Florida are cheating and plagiarism. Students should be aware of their faculty's policy on collaboration, should understand how to properly cite sources, and should not give nor receive an improper academic advantage in any manner through any medium.

Lateness Policy: An important part of students' work is meeting deadlines. Late assignments will be docked three (3) percentage points per day overdue. Under extraordinary circumstances extensions will be granted but every effort should be made to avoid the need to take an extension.

Accommodations for Students with Disabilities: Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center by visiting <https://disability.ufl.edu/students/get-started/>. It is important for students to share their accommodation letter with their instructor and discuss their access needs, as early as possible in the semester.

Online Evaluations: Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at <http://gatorevals.aa.ufl.edu/students>. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at <http://gatorevals.aa.ufl.edu/public-results/>.

Grading Policy:

You can find University of Florida's grading policies here:

<http://catalog.ufl.edu/UGRD/academic-regulations/grades-gradingpolicies/>

Letter Grade	Percentage
--------------	------------

A	93-100
A-	90-92
B+	87-89
B	83-86
B-	80-82
C+	77-79
C	73-76
C-	70-72
D+	67-69
D	63-66
D-	60-62

Attendance and Makeup Policy

Attendance is mandatory, and unexcused absences will adversely affect your grade. Requirements for class attendance and make-up assignments in this course are consistent with university policies that can be found in the online catalog at Attendance Policies <https://catalog.ufl.edu/UGRD/academic-regulations/attendance-policies/>

Course Schedule:

Jan. 8: Introductions

Readings:

Miriam Posner, “How Did They Make That?” (<https://archive.org/details/howdidtheymakethat>)

Seth Bernstein and Irina Makhlova. “Aggregate Treason: A Quantitative Analysis of Collaborator Trials in Soviet Ukraine and Crimea.” *Soviet and Post-Soviet Review* 46, no. 1 (2019): 30-54. <https://doi.org/10.1163/18763324-20181345>.

Oleg Khlevniuk, *Stalin: A New Biography of a Dictator*

Assignments:

-Explore the Digital Humanities Working Group (DHWG) website:

<https://digitalhumanities.group.ufl.edu/dh-uf/>

-Explore the site of the National Endowment for Humanities Office of Digital Humanities

(<https://www.neh.gov/divisions/odh>)

Jan. 22: What is “Digital History/Humanities”? What is “Big Data”?

***Note: This week’s class will be unusual because we will spend half the class time with a visiting scholar, Cindy Ermus, either at her talk on January 24 at 3pm or at another meeting with graduate students. On January 22 we will meet from 3pm to 4:30pm.

Readings:

Graham et al, *Historian’s Macroscope*, front matter, chapters 1-2

(http://web.archive.org/web/20160315221828/http://www.themacroscope.org/?page_id=583)

David M. Berry, M. Beatrice Fazi, Ben Roberts, and Alban Webb, “No Signal without Symbol: Decoding the Digital Humanities” (<https://dhdebates.gc.cuny.edu/read/untitled-f2acf72c-a469-49d8-be35-67f9ac1e3a60/section/687d2d32-0ec6-4158-b776-ee5cb989ef8b#ch05>)

Marche, “Literature Is not Data” (<https://lareviewofbooks.org/article/literature-is-not-data-against-digital-humanities/>)

“Digital Humanities—Ways Forward; Future Challenges.” *Historical Reflections/Réflexions Historiques*, Winter 2023. (with article by Cindy Ermus)

Assignment (due in class):

-Prepare five-minute talk about yourself and what you think DH could do for you.

Jan. 29: Data Wrangling/Cleaning

Readings:

Rawson, Katie and Trevor Muñoz, “Against Cleaning”

(<https://dhdebates.gc.cuny.edu/read/untitled-f2acf72c-a469-49d8-be35-67f9ac1e3a60/section/07154de9-4903-428e-9c61-7a92a6f22e51#ch23>)

Posner, “Humanities Data: A Necessary Contradiction”

(<https://miriamposner.com/blog/humanities-data-a-necessary-contradiction/>)

Chatterjee, Choi. Imperial Subjects in the Soviet Union: M.N. Roy, Rabindranath Tagore, and Re-Thinking Freedom and Authoritarianism. *Journal of Contemporary History* 52, no. 4 (2017): 913-934. <https://doi.org/10.1177/0022009417716754>.

How Did They Make That?

EUH:

Mapping the Republic of Letters (<http://republicofletters.stanford.edu/publications/>)

Old Bailey Online (<https://www.oldbaileyonline.org/>)

Google N-Gram Viewer (<https://books.google.com/ngrams/>)

NKVD Database: <https://nkvd.memo.ru/> (in Russian)

List of Repressed in USSR: <https://lists.memo.ru/> (in Russian)

WOH/LAH:

Slave Voyages: <https://www.slavevoyages.org/>

Tutorial: Approaches to Data Acquisition

Indian Marxists at Marxists.org: (<https://www.marxists.org/archive/roy/index.htm>,
<https://www.marxists.org/archive/tagore/index.htm>)

Web Scraping with Python

Excel Cleaning

No Assignment

Feb. 5: Databases/Historians See Like a State

Readings:

James Scott, *Seeing Like a State*, introduction-102.

Teo Ballvé, *The Frontier Effect: State Formation and Violence in Colombia*. Ithaca: Cornell University Press, 2020. Excerpts on Canvas.

How Did They Make That?

EUH:

Mapping the Republic of Letters (<http://republicofletters.stanford.edu/publications/>)

Old Bailey Online (<https://www.oldbaileyonline.org/>)

Google N-Gram Viewer (<https://books.google.com/ngrams/>)

WOH/LAH:

Slave Voyages: <https://www.slavevoyages.org/>

Tutorial: Database Work

Database from the Truth Commission in Colombia (https://nd-lucy-cdt-archive.s3.amazonaws.com/data_resource/Informes+y+Casos/Relacion_Casos_informes_anonimizado_datos.xlsx)

Cross Tabulation: Excel/Python

Basic Statistics

Assignment

Recall the article on collaborators from the first week and how the team from that project transformed Soviet police files into a database with clean (and not so clean) data. Find a dataset in your field (as defined by the course section WOH, LAH, EUH)—a set of novels, diaries, archival documents. Design a database that quantifies important aspects of that data, at least 5 fields. Provide 5 entries. Write a paragraph about the choices necessary in processing this data.

Feb. 12: Working with Text I: Concepts and Basics

Readings:

Historian's Macroscope, chapter 3

Franco Moretti, "Style, Inc.: Reflections on Seven Thousand Titles (British Novels, 1740–1850)"

Matthew Connelly, Raymond Hicks, Robert Jervis, and Arthur Spirling, "New Evidence and New Methods for Analyzing the Iranian Revolution as an Intelligence Failure," *Intelligence and National Security* (2021). <http://www.matthewconnelly.net/wp-content/uploads/2023/02/New-evidence-and-new-methods-for-analyzing-the-iranian-revolution-as-an-intelligence-failure.pdf>.

How Did They Make That?

EUH:

Tolstoy Digital: <https://tolstoy.ru/projects/tolstoy-digital/> (Russian)

Google N-Gram Viewer (<https://books.google.com/ngrams/>)

WOH:

Bitter Aloe: <https://bitteraloeproject.createuky.net/>

LAH:

Comision de la verdad: <https://www.comisiondelaverdad.co/> (Spanish)

Tutorial: Basic Text Counting

Foreign Relations of the United States, 1977-1980 Volume 11, Part 1

(<https://history.state.gov/historicaldocuments/frus1977-80v11p1>)

Easy tools (Voyant)

Loading text in Python

Assignment (due by class):

Using Excel or Python, count the number of incidents registered in Antioquia cross tabulated with the type of organization recording the data (autor_tipo_organizacion). For instance, you will find a total number of times a civil society organization (Organizaciones de sociedad civil) registered incidents in the region. The result should be a table. In a short paragraph, discuss what these results might mean in light of Ballve's book on Antioquia. Does the data fit with the narrative? Does it provide other areas for investigation?

Feb. 19: Working with Text II: Entities

Readings:

Salvatore, Ricardo. *Disciplinary Conquest: U.S. scholars in South America, 1900-1945*. Raleigh: Duke University Press, 2016. PAGES. <https://library.oapen.org/handle/20.500.12657/30113>.

How Did They Make That?

EUH:

Tolstoy Digital: <https://tolstoy.ru/projects/tolstoy-digital/> (Russian)

Google N-Gram Viewer (<https://books.google.com/ngrams/>)

WOH:

Bitter Aloe: <https://bitteraloeproject.createuky.net/>

Islamic West Africa: <https://islam.zmo.de/s/westafrica/page/home>

LAH:

Comision de la verdad: <https://www.comisiondelaverdad.co/> (Spanish)

Tutorial: Deciphering an Entire Field

Working with Constellate

Tokenizing with NLTK

Assignment (due by class)

Load a set of texts from your section into Voyant Tools. I have some if you are struggling to find any. Look at the word bubble. Are there words that should really be stopwords? Add them to the stopwords. Find the five most frequent words overall and the five most distinctive words for a few of the documents that seem notable (either surprising or perhaps understandable but interesting). In a short paragraph, write about what questions you might investigate by looking closely at the document.

Feb. 26: Working with Text III

***Note: One of the “working with text” weeks will involve a meeting over Zoom with Dr. Steven Davis of the University of Kentucky and his digital humanities class. We will discuss his project on the South African Truth and Reconciliation Commission. Date TBD. That week there will just be the tutorial, with little or no reading.

Readings:

Salvatore, Ricardo. *Disciplinary Conquest: U.S. scholars in South America, 1900-1945*. Raleigh: Duke University Press, 2016. PAGES. <https://library.oapen.org/handle/20.500.12657/30113>.

How Did They Make That?

EUH:

Tolstoy Digital: <https://tolstoy.ru/projects/tolstoy-digital/> (Russian)

Google N-Gram Viewer (<https://books.google.com/ngrams/>)

WOH:

Bitter Aloe: <https://bitteraloeproject.createuky.net/>

Islamic West Africa: <https://islam.zmo.de/s/westafrika/page/home>

LAH:

Comision de la verdad: <https://www.comisiondelaverdad.co/> (Spanish)

Tutorial: Deciphering an Entire Field

Continuing work with Constellate, this time with authors!

Entity Recognition with spaCy

Assignment

Find and download a journal in your section field from Constellate. Try to tokenize (remove stopwords, make uniform) the following: a single article text; a single article's title; an entire year of texts; an entire year of titles; the entire run of texts; the entire run of titles. Bonus points (like in Mario, not extra credit – this is a graduate course) for making a database by year. Do you notice any terms that raise questions for you?

Mar. 4: Mapping I: Basics

Reading:

Giesecking, “Where Are We? The Method of Mapping with GIS in Digital Humanities” *American Quarterly* (<https://jgieseking.org/new-publication-where-are-we-the-method-of-mapping-with-gis-in-digital-humanities/>)

Gobat, Michel. “The Invention of Latin America: A Transnational History of Anti-Imperialism, Democracy, and Race.” *The American Historical Review* 118, no. 5 (2013): 1345–75.
<http://www.jstor.org/stable/23784580>.

Mattern, “Gaps in the Map: Why We’re Mapping Everything, and Why not Everything Can, or Should, Be Mapped”
(<https://web.archive.org/web/20151208035349/http://www.wordsinspace.net/wordpress/2015/09/18/gaps-in-the-map-why-were-mapping-everything-and-why-not-everything-can-or-should-be-mapped/>)

How Did They Make That?

WOH:

World History Gazetteer: <https://whgazetteer.org/>

Islamic West Africa: <https://islam.zmo.de/s/westafrika/page/home>

EUH:

Imperiia Project: <https://imperiia.scalar.fas.harvard.edu/imperiia/index>

Ta Storona: <https://tastorona.su> (Russian)

LAH:

Power of Attorney: Native People, Legal Culture, and Social Networks:

<https://www.powerofattorneynative.com/>

Digging into Early Colonial History: <https://www.lancaster.ac.uk/digging-ecm/>

Tutorial:

-Basic map production

-Geocoding with Python

Assignment:

Using the same Constellate journal from last week, find the top 25 GPE words that appear with spaCy in an article, entire year, entire journal. Does spaCy provide good results?

Mar. 18: Mapping II: Teaching and Visualizing

Reading:

Bell, Wilson T. “Was the Gulag an Archipelago? De-Convoyed Prisoners and Porous Borders in the Camps of Western Siberia.” *Russian Review* 72, no. 1 (2013): 116–41.
<http://www.jstor.org/stable/23355602>.

Westerveld, L. and A. K. Knowles. Loosening the Grid: Topology as the Basis for a More Inclusive GIS. *International Journal of Geographical Information Science* 35, no. 10 (2021): 2108-2127. <https://doi.org/10.1080/13658816.2020.1856854>.

“I Was There.” <https://visionscarto.net/i-was-there>.

How Did They Make That?

WOH:

World History Gazeteer: <https://whgazetteer.org/>

EUH:

Imperii Project: <https://imperii.scalar.fas.harvard.edu/imperii/index>

Ta Storona: <https://tastorona.su> (Russian)

LAH:

Power of Attorney: Native People, Legal Culture, and Social Networks:

<https://www.powerofattorneynative.com/>

Digging into Early Colonial History: <https://www.lancaster.ac.uk/digging-ecm/>

Tutorial:

Time Mapping

StoryMaps

Assignment:

Using the geographical entities retrieved for the previous assignment, find the coordinates with geocoding and map them on ArcGIS Online. Make the map public and share it in a link or embedding on Canvas. What areas for investigation do you see from the map? What problems with the geocoding or spaCy entity recognition do you see?

Mar. 25: Machine Learning/AI I: Mostly Text

***Note: This week we will be meeting with Jeffrey Binder and reading excerpts from his book. Depending on the amount of engagement, I will move class entirely to our meeting(s) with Binder. Depending on where the class is with projects, we may also spend the class as a workshop/catching up session this week.

Reading:

Lara Putnam, "The Transnational and the Text-Searchable: Digitized Sources and the Shadows They Cast" *American Historical Review*.

<https://academic.oup.com/ahr/article/121/2/377/2581842>

Artificial Intelligence and the Practice of History. *American Historical Review* 128, no. 3 (2023): 1345-1389. (entire forum) <https://academic.oup.com/ahr/issue/128/3>.

Binder, Jeffrey M. *Language and the Rise of the Algorithm*. Chicago: University of Chicago Press, 2020. Excerpts.

How Did They Make That?

EUH:

The Sphere (CorDeep): <https://sphaera.mpiwg-berlin.mpg.de/>

LAH:

Unlocking the Colonial Archive: <https://unlockingarchives.com/>

Assignment:

Make a heat map or time map and make a quick presentation about it in a story map.

Tutorial:

-Breaking ChatGPT

-OCR in Python

Apr. 1: Machine Learning/AI I: Mostly Text II

Reading:

Putnam, “The Transnational and the Text-Searchable: Digitized Sources and the Shadows They Cast” *American Historical Review*.

Artificial Intelligence and the Practice of History. *American Historical Review* 128, no. 3 (2023): 1345-1389. (entire forum)

How Did They Make That?

EUH:

The Sphere (CorDeep): <https://sphaera.mpiwg-berlin.mpg.de/>

LAH:

Unlocking the Colonial Archive: <https://unlockingarchives.com/>

Tutorial:

-Breaking ChatGPT

-OCR in Python

Assignment:

Use OCR to process one of your archival documents. Try to process it with spaCy. Are the results totally unusable? What are the uses and what are the limitations?

Apr. 8: Machine Learning III: Mostly Images

Reading:

Taylor Arnold, Lauren Tilton. Distant Viewing: Analyzing Large Visual Corpora, *Digital Scholarship in the Humanities* 34, Issue Supplement_1 (2019): i3–i16.

Haley Di Pressi, Stephanie Gorman, Miriam Posner, Raphael Sasayama, and Tori Schmitt, with contributions from Roderic Crooks, Megan Driscoll, Amy Earhart, Spencer Keralis, Tiffany Naiman, and Todd Presner. A Student Collaborators' Bill of Rights.

<https://humtech.ucla.edu/news/a-student-collaborators-bill-of-rights/>.

Alexander Reid. Graduate Education and the Ethics of the Digital Humanities. *Debates in Digital Humanities*. Minneapolis: University of Minnesota, 2012.

<https://dhdebates.gc.cuny.edu/read/40de72d8-f153-43fa-836b-a41d241e949c/section/b4f365e7-eefa-4546-9857-dfd35b316183#ch20>.

Jones, Polly. From Stalinism to Post-Stalinism: De-Mythologising Stalin, 1953-56. *Totalitarian Movements and Political Religions* 4, no. 1 (2003): 127-148.

<https://doi.org/10.1080/14690760412331326108>.

Tutorial:

Object recognition in YOLOv8

Training a model to recognize Stalin

Assignment:

Use OCR to process one of your archival documents.

****Discussant Paper Due

Apr. 15: Digital Museums and Visualization

Reading:

Michael Friendly. The Golden Age of Statistical Graphics. *Statistical Science* 28, no. 4 (2008): 502-535. (Canvas)

Allison Hedley. Florence Nightingale and Victorian Data Visualisation. *Significance*, April 2020, 26-30. (Canvas)

Biedermann, Bernadette. Virtual Museums as an Extended Museum Experience: Challenges and Impacts for Museology, Digital Humanities, Museums and Visitors – In Times of (Coronavirus) Crisis. *Digital Humanities Quarterly* 15, no. 3.

<http://digitalhumanities.org:8081/dhq/vol/15/3/000568/000568.html>

Tutorial:

What makes a good visualization/website?

Timeline.js

Flourish.studio

Omeka

Apr. 22: Professional Issues/Wrapping Up

Readings:

Read H-Net's job listings that give Digital Humanities as a category: https://www.h-net.org/jobs/job_browse.php?category_id=29.

American Historical Association. "Guidelines for the Professional Evaluation of Digital Scholarship by Historians." <https://www.historians.org/teaching-and-learning/digital-history-resources/evaluation-of-digital-scholarship-in-history/guidelines-for-the-professional-evaluation-of-digital-scholarship-by-historians>

LaDale Winling. "Getting Tenure with Digital History: How One Scholar Made His Case." *Perspectives on History*. April 8, 2019. <https://www.historians.org/research-and-publications/perspectives-on-history/april-2019/getting-tenure-with-digital-history-how-one-scholar-made-his-case>

Lauren Tilton. "On Tenure in Digital History." *Perspectives on History*. May 20, 2019. <https://www.historians.org/research-and-publications/perspectives-on-history/may-2019/on-tenure-in-digital-history>.

***Presentations

***Final discussion

Final Project/Essay Due on May 2, 5pm

Recommended Readings/Viewings:

Further Reading:

2019: Underwood, “Digital Humanities as a Semi-Normal Thing”

(<https://dhdebates.gc.cuny.edu/read/untitled-f2acf72c-a469-49d8-be35-67f9ac1e3a60/section/ac5fc1c4-abc4-4a04-8a4b-fa5fe763220e#ch10>)

Weingart, “What’s under the Big Tent?” (<https://www.digitalstudies.org/article/id/7312/>)

Schmidt, “Words Alone” (<http://journalofdigitalhumanities.org/2-1/words-alone-by-benjamin-m-schmidt/>)

Moretti, “The Slaughterhouse of Literature”

The Spatial Humanities (eds. Bodenhamer, Corrigan, Harris) (any essay)

Wainwright, “Teaching History Theory through Video Games,” *The History Teacher*

Schmidt, “The History of Looking at Data Visualizations”

(<http://sappingattention.blogspot.com/2017/04/the-history-of-looking-at-data.html>)

Marsh, “Omeka in the classroom: The challenges of teaching material culture in a digital world”

Teaching Resources:

Miriam Posner, Digital Humanities course syllabus: <http://miriamposner.com/classes/dh201w21/>

Programming:

Programming Historian (<https://programminghistorian.org/>)