

INLS 890: Digital Humanities: Explorations and Experimentation

<http://inls890spring2014.web.unc.edu/>

12:30 – 3:15 pm, Manning 304

Richard Marciano, Professor
School of Information and Library Science
e-mail: richard_marciano@unc.edu
skype: [richardjmarciano](#)
Office Hours:
Drop-in or by appointment

In this interdisciplinary graduate seminar, we will explore concepts and case studies in the digital humanities and experiment with software and tools. The class will consist of four main units, roughly one per month, with a hands-on assignment due at the end of each unit:

- January: *digitization / OCR*
- February: *GIS*
- March: *text analysis*
- April: *crowdsourcing*

Discussions of infrastructure and scalability will be incorporated into each of the units (as opposed to being a standalone theme).

Four sets of spatial and textual historical data will be provided and discussed at the start of the semester and teams of two to three students will be formed, allowing for reuse of data across all four units.

Suggested Textbooks

- Daniel J. Cohen and Roy Rosenzweig, *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web* (Philadelphia: University of Pennsylvania Press, 2006) – available free online at <http://chnm.gmu.edu/digitalhistory/>
- *A Companion to Digital Humanities*, edited by Susan Schreibman, Raymond George Siemens, and John M. Unsworth. Malden, MA: Blackwell Publishing, 2004. <http://www.digitalhumanities.org/companion/>.

Class Meetings

In the first part of each class, we will explore and discuss topics in each of the 4 class units (*digitization/OCR* in January, *GIS* in February, *text analysis* in March and *crowdsourcing* in April).

In the second part of each class, we will experiment with software and tools that relate to the current unit.

We will also welcome speakers and conduct field trips during the semester.

Evaluation & Grading

Attendance is expected at all class meetings. Your ability to participate, contribute, influence the class is essential.

Course grades will be determined as follows:

10 points: **Participation**.

10 points: **Presentation & Critique** of a map project.

Experimentation:

- 15 points: **Digitization/OCR** assignment, due end of January.
- 15 points: **GIS** assignment, due end of February.
- 25 points: 2 **Text Analysis** assignments, due end of March.
- 15 points: **Crowdsourcing** assignment, due end of April.

10 points: **Putting it all together** assignment, due end of course.

Letter grades will be: 100-95: H, 94-90: P+, 89-85: P, 84-80: P-, 79-70: L, < 69: F

Participation, the critique, and the final essay will be graded individually. For the experimentation assignments, you will be graded as a group.

Digital Humanities Blogs

There are many great DH blogs out there, among them:

- Digital Humanities Now: <http://digitalhumanitiesnow.org/>
- Digital Humanities Specialist: <https://dhs.stanford.edu/>
- GMU’s Center for History and New Media’s Digital Campus: <http://digitalcampus.tv/>
- Dan Cohen’s Digital Humanities Blog: <http://www.dancohen.org/>
- Fred Gibbs’ Blog, HistoryProef: <http://historyproef.org/>
- Digital Scholarship in the Humanities: Exploring the Digital Humanities: <http://digitalscholarship.wordpress.com/>
- Flowing Data Blog on Data Visualization: <http://flowingdata.com/>

Many DH centers maintain their own blogs and are easy to find via Google. The *New York Times*’ ongoing Humanities 2.0 series may be of interest to many of you, as well. There is also a social media LinkedIn “Digital Humanities / Humanities Computing” group.

Digital Humanities Syllabi and Tools

- DH Syllabi compiled by CUNY's DH Guide: http://commons.gc.cuny.edu/wiki/index.php/DH_Syllabi
- Digital Research Tools (DiRT): <http://digitalhumanities.org/centernet/resources/tools/>

Digital Humanities Journals

- DHQ: digital humanities quarterly: <http://digitalhumanities.org/dhq>

		Explorations		Experimentation	
		Monday – Part 1		Monday – Part 2	
Digitization / OCR	JAN.	Week 1:	1/13: #1 – Introduction	Introduction to VCL software / ABBYY	
		Week 2:	1/20: NO CLASS (MLK)	NO CLASS (Spring Break)	
		Week 3:	1/27: #2 – Digitization / OCR	Field Trip: digitalNC @ Wilson Library	
		Week 4:	2/03: #3 – Digitization at scale	ABBYY	
GIS	FEB.	Week 5:	2/10: #4 – Intro to GIS	ArcGIS intro	
		Week 6:	2/17: #5 – Recreating historical maps	Georeferencing	
		Week 7:	2/24: #6 – Presentation & critique of map projects	Tracing	
		Week 8:	3/03: #7 – Map services in the cloud	Geocoding	
Text Analysis	MAR.	Week 9:	3/10: NO CLASS (Spring Break)	NO CLASS (Spring Break)	
		Week 10:	3/17: #8 – Intro to text analysis	Voyeur	
		Week 11:	3/24: #9 – Linked open data	Text analysis	
		Week 12:	3/31: #10 – Big data	Text processing	
Crowdsourcing	APR.	Week 13:	4/07: #11 – Introduction to crowdsourcing	Toolkits	
		Week 14:	4/14: #12 – Hands-on Workshop	Hands-on	
		Week 15:	4/21: #13 – Crowdsourcing at scale	More hands-on exercises	
3D printing:		http://store.makerbot.com/replicator2			
Future of Higher Ed:		https://www.coursera.org/course/highered			

Schedule

#1: January 13

Part 1: Introduction

Readings:

DevDH.org, *Development for the Digital Humanities*, <http://devdh.org>

The Digital Humanities: A Gentle Introduction, Aug. 2013, Leslie Barnes & Jenn Riley, CDLA, UNC Libraries, <http://digitaldiscussions.web.unc.edu/session-august-2013-a-gentle-introduction-to-the-digital-humanities/>

A working Definition of Digital Humanities, 2013 blog: <http://www.scottbot.net/HIAL/?p=39749>

Digital Humanists: If You Want Tenure, Do Double the Work, https://chroniclevitae.com/news/249-digital-humanists-if-you-want-tenure-do-double-the-work?cid=at&utm_source=at&utm_medium=en

UCLA Center for Digital Humanities: Intro to Digital Humanities, <http://dh101.humanities.ucla.edu/>

A Companion to Digital Humanities, "1. The History of Humanities Computing," edited by Susan Schreibman, Raymond George Siemens, and John M. Unsworth. Malden, MA: Blackwell Publishing, 2004. <http://www.digitalhumanities.org/companion/>.

D. Cohen, and Roy Rosenzweig. "Introduction." *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*, 2006. <http://chnm.gmu.edu/digitalhistory/introduction/>.

P. Svensson, "The Landscape of Digital Humanities." *Digital Humanities Quarterly* 4, no. 1 (2010), <http://digitalhumanities.org/dhq/vol/4/1/000080/000080.html>

B. Bobley, "The Promise of Digital Humanities," Sep. 28, 2011, Inside Higher Ed, http://www.insidehighered.com/news/2011/09/28/national_endowment_for_the_humanities_celebrates_digital_humanities_projects

New York Times Series on Digital Humanities, http://topics.nytimes.com/top/features/books/series/humanities_20/index.html

Part 2: Introduction to VCL Software / ABBYY FineReader

Readings:

Virtual Computing Lab: <http://sils.unc.edu/it-services/remote-access/vcl> & <http://vcl.ncsu.edu>

Unit 1: Digitization & OCR

January 20: NO CLASS (MLK)

#2: January 27

Part 1: Digitization & OCR

Readings:

B. Bobley, "The Future of the Digital Humanities", 2/1/2009, HASTAC Scholars Discussion Forum, <http://hastac.org/forums/hastac-scholars-discussions/future-digital-humanities>

Univ. of Alberta, Day of DH, <http://ra.tapor.ualberta.ca/~dayofdh2010/>

Univ. of Alberta, What is DH?, http://tapor.ualberta.ca/taporwiki/index.php/How_do_you_define_Humanities_Computing_/Digital_Humanities%3F

Improving Access to Text, IMPACT, <http://www.impact-project.eu/>

IMLS Specifications for Projects that Develop Digital Products

http://www.ims.gov/applicants/projects_that_develop_digital_products.aspx

FADGI – Federal Agencies Digitization Guidelines Initiative <http://www.digitizationguidelines.gov/>

Becoming Digital (from Cohen & Rosenzweig book), <http://chnm.gmu.edu/digitalhistory/digitizing/>

Part 2: Wilson Library Visit: Nick Graham, Digital NC, <http://digitalnc.org> & Fred Stipe (Head of CDLA's Digital Production Center), overview & Scribe operation (large-format studio & Zeutschel/PhaseOne equipment)

#3: February 3

Part 1: Digitization at scale

Readings:

Hathi Trust Digital Library, <http://www.hathitrust.org/>

Google Books, <http://books.google.com/>

Google and Research Libraries Launch Massive Digitization Program (Dec. 20 , 2004)

<http://newsbreaks.infotoday.com/nbreader.asp?ArticleID=16307>

Digitizing Print Collections with the Internet Archives, <http://www.archive.org/scanning>

Ted Underwood: "The Task of Cleaning and Enriching Large Collections",

<http://www.youtube.com/watch?v=JaoaqfrnLE>

Addressing History:

- <http://addressinghistory.blogs.edina.ac.uk/2012/02/17/augmented-reality-view-of-addressinghistory-now-available-for-iphone-android-or-nokia-ovi-phone/>
- <http://addressinghistory.blogs.edina.ac.uk/>
- <http://addressinghistory.edina.ac.uk/>
- <http://www.slideshare.net/LIFESHAREProject/addressing-history>

Part 2: ABBYY

Unit 2: GIS

#4: February 10

Part 1: An Introduction to GIS

Readings:

Ian Gregory, "Using Geographical Information Systems to Explore Space and Time in the Humanities" in *The Virtual Representation of the Past*, eds. Mark Greengrass and Lorna Hughes

Ian Gregory, "A Place in History: A Guide to Using GIS in Historical Research", 2nd Edition, Ian N. Gregory,

<http://www.ccsr.ac.uk/methods/publications/jg-gis.pdf>

Michael Goodchild, "Combining Space and Time: New Potential for Temporal GIS" in *Placing History: How Maps, Spatial Data, and GIS Are Changing Historical Scholarship*, ed. Anne Kelly Knowles

White, Richard. What is Spatial History? Spatial History Lab: Stanford University, 2010.

<http://www.stanford.edu/group/spatialhistory/cgi-bin/site/pub.php?id=29>.

Part 2: Introduction to ArcGIS

#5: February 17

Part 1: Recreating historical maps

Kurt Schlichting, Peter Tuckel, Richard Maisel, "Residential Segregation and the Beginning of the Great Migration of African Americans to Hartford, Connecticut: A GIS-Based Analysis," *Historical Methods* 39, number 3 (summer 2006), 132-143.

Logan, John R., Jason Jindrich, Hyoungjin Shin, and Weiwei Zhang, "Mapping America in 1880: The Urban Transition Historical GIS Project," *Historical Methods* 44, number 1 (2011), 49-60, 60.

Part 2: Hands-on: Georeferencing

#6: February 24

Part 1: Presentation and critique of map projects

Student presentations

Part 2: Hands-on: Tracing

#7: March 3

Part 1: Map services in the cloud

Readings:

NYPL Labs:

NYPL Map Warper: <http://maps.nypl.org/warper/>

NYPL Map Polygon & Feature Extractor: <http://www.gislounge.com/automating-extracting-gis-data-scanned-maps/>

Terrahub exercise: <http://geocompute3.renci.org/>

“Big Historical Data’ Feature Extraction”, Richard Marciano et al., in Journal of Map & Geography Libraries, published online Apr. 17, 2013, “Working Digitally with Historical Maps” special issue. Vol. 9, Issue 1-2, Jan. 2013, pp. 69-80. See: [big-historical-data](#)

ESRI:

Geocommons: <http://geocommons.com>, Share open data and maps

GeoIQ: <http://www.geoiq.com/>, “Provide geospatial data management, visualization and analysis platform providing collaborative, browser-based data analysis tools for use by both technical and non-technical users.”

ArcGIS Online: <http://www.esri.com/software/arcgis/arcgisonline>

Kurt Schlichting, Peter Tuckel, Richard Maisel, “Residential Segregation and the Beginning of the Great Migration of African Americans to Hartford, Connecticut: A GIS-Based Analysis,” Historical Methods 39, number 3 (summer 2006), 132-143. See: [schlichting2006_paper](#)

Logan, John R., Jason Jindrich, Hyoungjin Shin, and Weiwei Zhang, “Mapping America in 1880: The Urban Transition Historical GIS Project,” Historical Methods 44, number 1 (2011), 49-60, 60. See: <http://www.archives.gov/applied-research/events/1940-census.html>

Google Fusion Tables:

<http://www.google.com/drive/apps.html#fusiontables> AND <https://support.google.com/fusiontables/answer/184641>

GIS Cloud: <http://www.giscloud.com>

Urban Mapping: <http://www.urbanmapping.com/>

Part 2: Hands-on: Geocoding

Unit 3: Text Analysis

March 10: NO CLASS (Spring Break)

#8: March 17

Part 1: Introduction to text analysis

Readings:

Blue Velvet “Re-dressing New Orleans in Katrina’s wake”, D. Goldberg

<http://www.vectorsjournal.org/projects/index.php?project=82>

Explore the TEI initiative: <http://www.tei-c.org/index.xml>

“What is TEI”, Scott Hamlin:

http://www.colleges.org/techcenter/Archives/workshops/xml05/what_is_tei/slideswhat_is_tei-slides0.html

John Unsworth: “Monk” project. <http://www3.isrl.illinois.edu/~unsworth/monk.hdc.ppt>

Explore TAPoR tools: <http://portal.tapor.ca/portal/portal> and <http://entry.tapor.ca/>

Part 2: Text analysis using OpenRefine and Voyeur

Readings:

The Stone and the Shell (historical questions raised by a quantitative approach to language), by Ted Underwood, <http://tedunderwood.com/2012/08/14/where-to-start-with-text-mining/>

“Doing Things with Text”, Lisa Spiro, Jan. 2013, <http://digitalscholarship.files.wordpress.com/2013/01/doing-things-with-text-workshop-rice-2013.pdf>

Exercises: https://docs.google.com/document/d/1p1-K-MAndX_trL33ZI4of1m3VXkVhkmkmY10G3wWius/edit#

Open Refine: <http://openrefine.org/>

Google Refine tutorial: <http://davidhuynh.net/spaces/nicar2011/tutorial.pdf>

GREL functions: <https://github.com/OpenRefine/OpenRefine/wiki/GREL-Functions>

Book: Using OpenRefine

#9: March 24

Part 1: Linked open data

Readings:

Linked Open Data, What is it? <http://vimeo.com/36752317>

What is Europeana doing with semantic web and linked open data?, in LODLAM blog, June 18, 2013, Linked Open Data, <http://www.diglib.org/community/groups/linkedata/>

A. Renear, “Text Encoding”, Chapter 17, A companion to Digital Humanities, 2004, http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-3-5&toc.depth=1&toc.id=ss1-3-5&brand=9781405103213_brand

Part 2: Text Analysis

Voyant Tools: <http://voyant-tools.org/>

–<http://hermeneuti.ca/voyeur>

–<https://blogs.nd.edu/emorgan/2014/01/voyant-tools/>

–<http://docs.voyant-tools.org/start/>

Stylometry:

Patrick Juola:

- PAPER: “The Professor Who Declared, I’s J.K. Rowling, Steve Kolowich, The Chronicle of Higher Education, July 29, 20-13, <https://chronicle.com/article/The-Professor-Who-Declared/140595/>
- Java Graphical Authorship Attribution Program: http://evllabs.com/jgaap/w/index.php/Main_Page
- PAPER: <http://www.mathcs.duq.edu/~juola/papers.d/jsb-aut.pdf>, Patrick Juola, John Sofko, and Patrick Brennan. (2006). “A Prototype for Authorship Attribution Studies,” Literary and Linguistic Computing 21:169-178

Jose Nilo Binongo:

- PAPER: <http://dh.obdurodon.org/Binongo-Chance.pdf>, Jose Nilo G. Binongo, “Who Wrote the 15th Book of Oz? An Application of Multivariate Analysis to Authorship Attribution,” Published online, Taylor & Francis, , 9-17

#10: March 31

Part 1: Big data

Readings:

Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, et al. "Quantitative Analysis of Culture Using Millions of Digitized Books." *Science* 331, no. 6014 (January 14, 2011): 176 -182.

<http://www.sciencemag.org/content/331/6014/176.abstract>.

Window on Culture" *NYT* 16 December 2010:

<http://www.nytimes.com/2010/12/17/books/17words.html?scp=2&sq=patricia%20cohen%20digital&st=cse>

JISC Digging into Data, <http://diggingintodata.org>

Part 2: Text processing

Unit 4: Crowdsourcing

#11: April 7

Part 1: Crowdsourcing in the cultural realm

Readings:

Some examples (among many...):

- Brave New World: Crowdsourcing Maps —<http://news.dice.com/2012/11/19/brave-new-world-crowdsourcing-maps/>
- **WebWise 2012: Session 2 – Sharing Public History Work: Crowdsourcing Data and Sources:** <http://www.tvworldwide.com/events/webwise/120229/default.cfm>, **21:30 to 44:45** Crowdsourcing the Museum
- Crowdsourcing and big data analytics aid search for Flight 370, March 26, 2014,<http://www.isgtw.org/visualization/crowdsourcing-and-big-data-analytics-aid-search-flight-370>
- What's on the Menu, <http://menus.nypl.org/>
- FamilySearch.org: <https://familysearch.org/indexing/>

Incentivizing Volunteer Workforces for Crowd-sourced Projects, at SAA 2013 New Orleans, <http://bighumanities.net/events/saa-2013-crowdsourcing/>

Crowd-Sourcing Scoping Study: Engaging the Crowd with Humanities Research by Stuart Dunn and Mark Hedges from the Centre for e-Research, Department of Digital Humanities King's College London. <http://stuardunn.files.wordpress.com/2013/04/crowdsourcing-connected-communities.pdf>

L. Carletti, G. Giannachi, D. Price, D. McAuley, "Digital Humanities and Crowdsourcing: An Exploration, in MW2013: Museums and the Web 2013, April 17-20, 2013, <http://mw2013.museumsandtheweb.com/paper/digital-humanities-and-crowdsourcing-an-exploration-4/>

J. Oomen, L. Aroyo, Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges, In: 5th International Conference on Communities & Technologies. Brisbane, Australia – 29 June – 2 July 2011, [CLICK HERE](#)

AND ALSO:

- J. Oomen, L. Aroyo, Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges, <http://www.cs.vu.nl/~marieke/OomenAroyoCT2011.pdf>
- H. Rose, Crowdsourcing: How and Why Should Libraries Do It?," <http://www.dlib.org/dlib/march10/holley/03holley.html>, March/April 2010, Vol. 16, No. 3 / 4

- Crowdsourcing and Curating Collective Memory, Legends, and Local History with Facebook Groups: <https://theunquietlibrarian.wordpress.com/2011/08/03/crowdsourcing-and-curating-collective-memory-legends-and-local-history-with-facebook-groups/#com-head>
- Explore Civil War Diaries Transcription Project: <http://digital.lib.uiowa.edu/cwd/transcripts.html>
- Explore the Australian Newspapers Digitisation Program, esp. OCRing and Crowdsourcing: http://www.nla.gov.au/ndp/project_details/

Part 2: Toolkits

#12: April 14

CrowdConf agenda, Oct. 2013, <http://www.crowdconf.com/agenda-2013>

Hector Garcia Molina from Stanford: "Using Crowdsourcing for Data

Analytics", <http://ci.drexel.edu/bigdata/bigdata2013/Using%20CrowdSourcing%20for%20Data%20Analytics.pdf>

Crowd business models (from <http://www.resultsfromcrowds.com>):

Matt Lease at U. Texas iSchool, "Crowdsourcing News, Events, and Resources", <http://ir.ischool.utexas.edu/crowd/>

Crowd Research, <http://crowdresearch.org>

List of crowdsourcing projects, http://en.wikipedia.org/wiki/List_of_crowdsourcing_projects

Crowdsourcing For Dummies, David Alan Grier, <http://www.dummies.com/how-to/content/crowdsourcing-for-dummies-cheat-sheet-uk-edition.html>

Part 2: Hands-on

"Top 6 Crowdsourcing Websites", <http://www.ichitect.com/best-crowdsourcing/>

Richard's list: [software](#)

Crowdflower: <http://www.crowdflower.com/>

- "CrowdFlower Lawsuit Could Change Crowd Labor Industry Forever", Seth Weinstein, Jul. 29, 2013, <http://tinywork.wordpress.com/2013/07/29/crowdflower/>
- "Should Creative Crowdsourcing Participants Get Working Contracts and Salaries", Yannig Roth, Jan. 7, 2014, <http://yannigroth.com/2014/01/07/should-creative-crowdsourcing-participants-get-working-contracts-and-salaries/>

Amazon Mechanical Turk: <http://www.mturk.com>

- "Conducting behavioral research on Amazon's Mechanical Turk", Winter Mason, Siddarth Suri, in Behav. Res., 2011, [CLICK HERE](#)
- "Who are the Crowdworkers? Shifting Demographics in Mechanical Turk", J. Ross, L. Irani, S. Silberman, A. Zaldivar, B. Tomlinson, in CHI 2010, April 10-15, 2010, [CLICK HERE](#)
- "Demographics of Mechanical Turk", P. Ipeirotis, [CLICK HERE](#)
- "Amazon Mechanical Turk: The Digital Sweatshop", E. Cushing, Jan/Feb 2013, <http://www.utne.com/science-and-technology/amazon-mechanical-turk-zm0z13jzfzlin.aspx#axzz2yoznRjVx>

#13: April 21 (last class)

Part 1: Crowdsourcing at scale

Readings:

D. Ferriero, April 12, 2010 blog, "Cultivating Citizen Archivists", April 12, 2012, <http://blogs.archives.gov/aotus/?p=144>

D. Hilgers, Citizensourcing: Applying the Concept of Open Innovation to the Public Sector, The International Journal of Public Participation, Vol. 4, No. 1, Jan. 2010,

http://www.iap2.org/associations/4748/files/Journal_10January_Vol4_No1_6_Hilgers%26Ihl_Citizens_ourcing.pdf

- G. Gugliota, NY Times, Mar. 28, 2011, "Deciphering Old Texts, One Woozy, Curvy Word at a Time",
http://www.nytimes.com/2011/03/29/science/29recaptcha.html?_r=2&scp=1&sq=recaptcha&st=cse
- K. Smith-Yoshimura, C. Shein, "Social Metadata for Libraries, Archives, and Museums, Part 1: Site Reviews," An OCLC Research Report, <http://www.oclc.org/research/publications/library/2011/2011-02r.htm>
- K. Smith-Yoshimura, C. Jean Godby, H. Kofler, K. Varnum, E. Yakel, "Social Metadata for Libraries, Archives, and Museums, Part 2: Survey Analysis," An OCLC Research Report, <http://www.oclc.org/research/publications/library/2011/2011-03r.htm>

Part 2: Hands-on