Lecture 24: Digital humanities, then wrapup

- collecting, curating, analyzing, and presenting data that comes from traditional humanities fields like history, literature, art, music, …

- examples of DH projects here and elsewhere
- examples of undergrad DH work
- one example of my own
- potential topics
- issues

- Disclaimer: I'm a dilettante, not an expert

Center for Digital Humanities

The Center for Digital Humanities at Princeton is an interdisciplinary research center that brings together faculty, staff, students, and community partners working at the intersection of the humanities and technology. We create and apply digital tools to humanistic questions and critically engage with the promises and risks that technology poses to society.

Princeton Prosody Archive

Compelling users to rethink the past through a collection of historical prosodic

Shakespeare and Company Project

Recreating the world of the Lost Generation in interwar Paris.

Princeton Ethiopian Miracles of Mary Project

Folklore about How the Virgin Mary Helps Believers in Ethiopian Literature and Art

New Languages for NLP: Building Linguistic Diversity in the Digital Humanities

An NEH-funded workshop to help scholars use NLP for new languages

Princeton Geniza Project

Accessing the medieval Islamic world through digital tools

Princeton Prosody Archive

(Meredith Martin, English)

Of | man's | first diso|bedience | and the | fruit Of | that
for|bidden | tree | whose | mortal | taste Brought | death |
into the | world | and | all our | woe | Sing | Heavenly | Muse.

"Of man's first disobedience and the fruit"

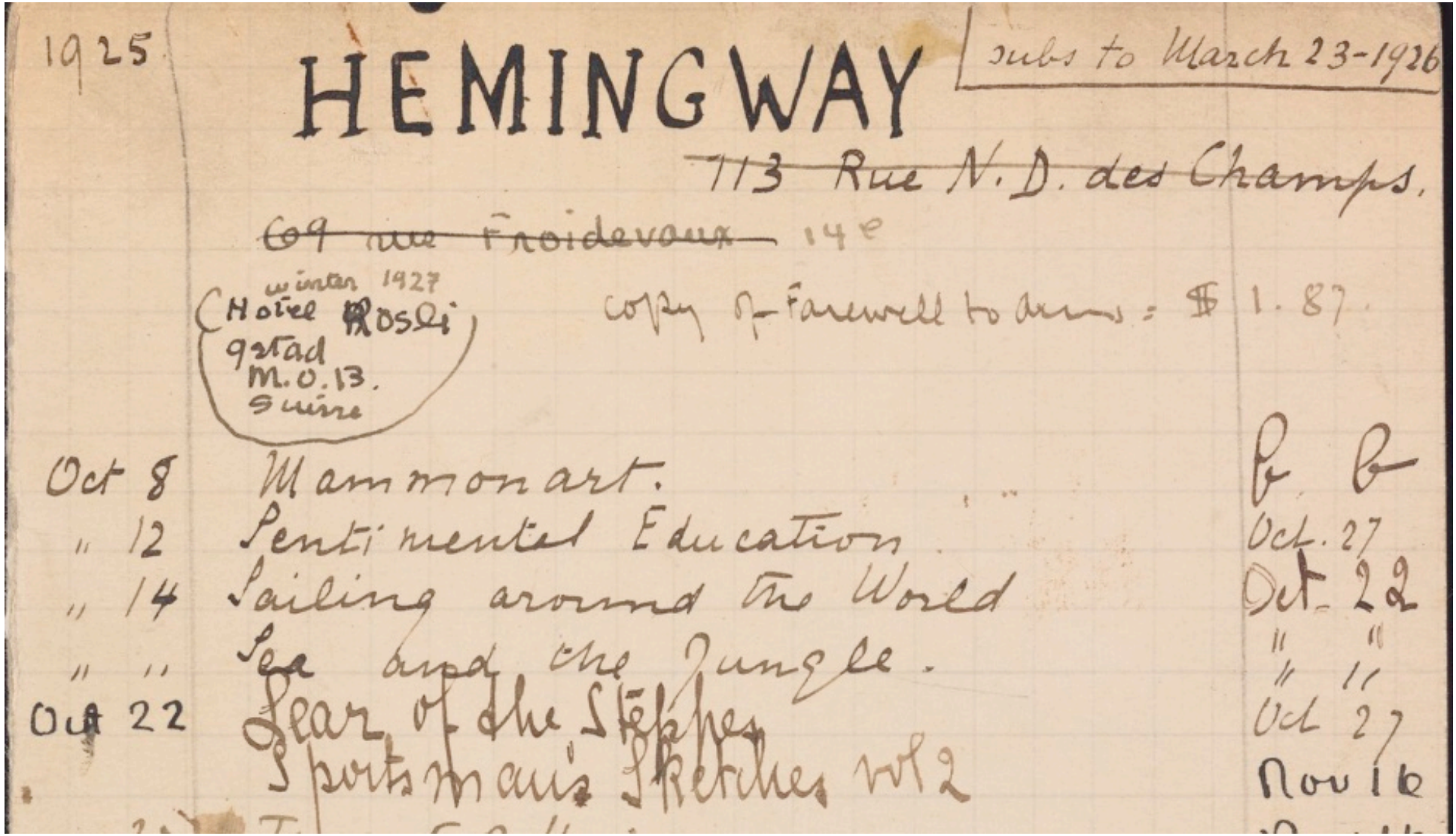
Ov mánz férst disɷbédiens, and ðɛ frút	0-1-,½-0,01,00-0-,0-1- (1)
Ov ðát fɐrbíd'n tré, huz mə́rtal tást	0-½-,01,0-½-,0-1,0-1- (2)
Brə́t dét intu ðɛ wúrld, and él ɤr wɷ́	½-1-,00-,0-1-,0-½-,0-1- (3)
Wiɗ lós ov 'Ɛd'n, til wun gráter mán	0-½-,0-1,0-0-,0-1,0-1- (4)

Sylvia Beach

(1887-1962)

Mapping Expatriate Paris

The Sylvia Beach Lending Library Project



1925

HEMINGWAY

subs to March 23-1926

113 Rue N. D. des Champs.

~~69 rue Froidevaux~~ 14e

winter 1927
Hotel Rosli
grad
M.O.13.
Suisse

copy of Farewell to Arms = ₣ 1.87

Oct 8	Mammon art.	P. P.
" 12	Sentimental Education	Oct. 27
" 14	Sailing around the World	Oct. 22
" "	Sea and the Jungle.	" "
		" "
Oct 22	Year of the Stiffes	Oct 27
	Sportsman's Sketches vol 2	Nov 16

Some existing tools…

- Voyant: textual analysis
- Mallet: classification, topic modeling
- Gephi, Cytoscape: network graph visualization
- Palladio, RAW: visualizing network maps, graphs, timelines
- OpenRefine: cleaning up CSV data
- Juxta: compare textual variants

- D3.js: Javascript graphics
- Leaflet, Mapbox: map displays
- NLTK: natural language processing
- Scikit-Learn: machine learning
- R: statistics, plotting

Hardware

- logical/functional/architectural structure
  - bus connects processor, primary memory, disks, other devices
  - caching
  - CPU cycle: fetch-decode-execute; kinds of instructions
    - toy machine as an example
    - different processor families are incompatible at the instruction level
  - von Neumann: architecture; Turing: equivalence of all machines
- physical implementation; sizes and capacities
  - chips; Moore's law, exponential growth
- analog vs digital
- representation of information
  - bits, bytes, numbers, characters, instructions
  - powers of 2; binary and hexadecimal numbers
  - interpretation determined by context
- it's all bits at the bottom

Software

- algorithms: sequence of defined steps that eventually stops
  - complexity: how number of steps is related to amount of data
    - linear: searching, counting, …
    - quadratic: simple sorting
    - logarithmic: binary search (logarithm = number of bits needed to store)
    - n log n: quicksort
    - exponential: towers of Hanoi, traveling salesman problem, …
- programs and programming languages:
  - evolution, language levels: machine, assembly, higher-level
  - translation/compilation; interpretation
  - a program can simulate a machine or another program
- basic programming, enough to figure out what some code is doing
  - variables, constants, expressions, statements, loops & branches (if-else, while), functions, libraries, components
- operating systems: run programs, manage file system & devices
  - file systems: logical: directories and files; physical: disk blocks
- application programs, interfaces to operating system, APIs

Communications

- local area networks, Ethernet, wireless, broadcast media
- Internet: IP addresses, names & DNS, routing; packets
  - bandwidth
- protocols: IP, TCP, higher-level; layering
  - synthesis of reliable services out of unreliable ones

- Web: URLs, HTTP, HTML, browser
  - caching
- security & privacy: viruses, cookies, spyware, …
  - active content: Javascript, plugins, addons

- cryptography
  - secret key; public key; digital signatures; secure hashes
- compression; error detection & correction

- wireless, cell phones, GPS, …

Real world issues

- **legal**
  - intellectual property: trademarks, patents, copyrights, licenses
  - jurisdiction, especially international
- **social**
  - privacy, security
- **economic**
  - open source vs proprietary
  - who owns what
- **political**
  - policy issues
  - balancing individual, commercial and societal rights and concerns

Things to take away

- **some skills, some specific technical knowledge**
  - how computers and communications work today
  - what's ephemeral, what's likely to still be true in the future
- **improved numeracy / quantitative reasoning**
  - what makes sense, what can't possibly make sense, and why
    - plausible estimates, engineering judgment, enlightened skepticism
- **another way of thinking**
  - how do things work?
  - how *might* something work?
  - you can often figure it out
- **some appreciation of tradeoffs & alternatives**
  - you never get something for nothing
- **some historical perspective**
  - everything derives from what came before

- **informed opinions about the role of technology**

Final exam (watch the web page for updates)

- Exam will be emailed to you early on Saturday Dec 17
  - must be returned by Thu Dec 22, 5 PM EST in person / email / pony express
  - Q/A session 4:30 Wednesday Dec 14
- similar to midterm but twice as long
- open book, as with midterm:

  open notes, book, problem sets, labs, old exams, …
- see instructions on web site

- I'm usually looking for something brief that shows that you understand or can reason
- if you're writing or calculating a lot, you're likely on the wrong track
- questions try to test understanding of basic ideas
  - meant to be simple and straightforward, if you understand
  - not meant to be tricky or rely on obscure facts
- think about plausibility and where I'm likely coming from
- if it still seems ambiguous, say "I'm assuming this..." and carry on