WORD LAB: An Experimental Space for Research Collaboration in Penn Libraries

Molly Des Jardin and Brian Vivier
University of Pennsylvania
What is WORD LAB?

• Based in Penn Libraries
• Research interest community
• Computational text analysis
• All disciplines – not just digital humanities
Why WORD LAB?

• Foster a scholarly community at Penn Libraries
• Lots of talk about DH, little doing. We do!
• Hands-on lab, learn by doing
• Push boundaries of what’s being done at Penn
Inspiration for WORD LAB

- Inspired by Stanford’s “Literary Lab” which meets together and produces research
- Franco Moretti: “We’ve got a room with a whiteboard.”
- We have one of those in the library!
- Administrator support to get started
- Started by Molly Des Jardin and Katie Rawson in Penn Libraries
Starting up WORD LAB

• Library support but minimal infrastructure. We have “a room with a whiteboard” with a regular meeting time, but no budget.
• Word of mouth to reach locally on campus, in all disciplines, world areas, and positions
• People from all over, mostly grad students and librarians
• Collaboratively decided what the group will do
What does WORD LAB do?

• WORD LAB discussion sessions: presenters
• OPEN LAB discussion time:
  – Collaborative self-directed learning
    • Creating a wiki for our programming tutorials
  – Journal article club
  – Free-form discussion
• Outreach:
  – Blog
  – Facebook
• Incubate spin-off research projects
Who is WORD LAB?

• Librarian Co-organizers
  – Molly Des Jardin, Japanese/Korean
  – Katie Rawson, Digital/English

• Core group
  – Grad students in Chinese, Middle Eastern studies
  – Humanities librarian
  – Recent Penn grad (Philadelphia U) in Chinese

• Occasional members
  – Annenberg School for Communication
  – Librarians from special collections center
Post Reach (03/07/2015 - 03/13/2015)

Matt Jockers will be visiting us next week, 3/17 at 1:30 PM for WORD LAB!
Reminder that Elias Saba will be presenting with a colleague on his project, Tahrir Documents. See you there!

WORD LAB discussion group - Molly Des Jardin
03/11/15 at 11:13am
21 people reached

WORD LAB open lab
03/09/15 at 3:02pm
26 people reached

WORD LAB open lab
03/03/15 at 3:01pm
16 people reached

Reach a new milestone
100 Likes

WORD LAB
Posted by Molly Des Jardin 1971 - March 11 at 11:13am

Matt Jockers will be visiting us next week, 3/17 at 1:30 PM for WORD LAB!
Reminder that Elias Saba will be presenting with a colleague on his project, Tahrir Documents. See you there!

Tahrir Documents
After that revolution ousted the monarchical system and put an end to those trailing behind it and other remnants, Egypt began a new era in a republican system, differing quite a bit from the system that preceded it, the military system. But was...
WORD LAB discussion group - Elia...
Tuesday, March 17, 2015 at 1:30pm
3 guests

WORD LAB open lab
Tuesday, March 24, 2015 at 1:30pm

WORD LAB open lab
Tuesday, April 7, 2015 at 1:30pm

Laura Gibson
January 20 at 10:43pm
A propos of today’s reading: another method for ranking (in this case, movies)

Dot Porter
December 15, 2014 at 12:45pm
Of interest?

Katie Rawson
October 13, 2014 at 10:33pm
We’d love to see a few Word Lab scholars present at the Keystones Digital Humanities Conference at Penn next summer!

WORD LAB via Jake Adelstein
Posted by Molly Des Jardin (?!) March 8 at 7:33am
This could be interesting for WORD LAB folks. Perhaps discuss it in the future? A lot of different techniques being used here.

Quantifying the Clickbait and Linkbait in BuzzFeed Article Titles
You probably do not know that the 3 most interesting things I found will blow your mind.
MINIMAXR.COM

WORD LAB created an event.
March 1 at 3:31pm

WORD LAB open lab
Tuesday, March 10 at 1:30pm
3420 Walnut St, Philadelphia, Pennsylvania 19104

12 people reached

WORD LAB created an event.
March 1 at 3:31pm

WORD LAB open lab
Tuesday, March 3 at 1:30pm
3420 Walnut St, Philadelphia, Pennsylvania 19104

41 people reached
natural and artificial languages, and scripts and gestures

March 3, 2015 ~ Molly Des Jardin ~ Leave a Comment ~ Edit

Is classical Chinese a natural language? Artificial language?

According to Wikipedia (our most reliable source, "naturally"), a natural language is "unpremeditated" when it is used. When someone is truly fluent in classical Chinese, what happens in their brains? Do they conceptualize poems in spoken language as a mediation layer and then "translate" it into classical Chinese/literary Sinitic, or do they think in that layer directly?

If they do think in it, is that enough to make it a natural language?

How about coding? Is pseudocode natural language (or, shorthand for a natural language), whereas programming languages are artificial? Does anyone think in Python, for example, or is it always mediated by a natural language (or pseudocode, or NL-then-pseudocode) stage beforehand?

What might edge cases teach us, such as sign language — never a spoken language, except for the unilateral hearing impaired? If
Main Page

NLTK Book Guide by WORD LAB

This is a guide to the book Natural Language Processing With Python, with solutions and explanations to the exercises developed by the Penn Libraries research community WORD LAB. To contribute, please create an account and start editing.

Chapter Guide

- Chapter 1: Language Processing and Python
- Chapter 2: Accessing Text Corpora and Lexical Resources
Activities

Current and upcoming activities in OPEN LAB:

- Natural Language Processing With Python
- Seven Databases in Seven Days (access via Safari Books Online at Penn)

Post your links and citations here in the comments for WORD LAB/OPEN LAB reading suggestions. Here are our past readings:

- Smith, David A; Cordell, Ryan; Dillon, Elizabeth Maddock. (2014) “Infectious texts: Modeling text reuse in nineteenth-century newspapers.” To be published in *Proceedings of the Workshop on Big Humanities.* [PDF]
Examples of WORD LAB

• Past presenters:
  – Hoyt Long on his research (Chicago, via Skype)
  – Molly Des Jardin on a web crawling project using Python to harvest Aozora Bunko texts
  – Brian Vivier’s project using the Chinese Biographical Database API
    • Led to Molly’s Python programming tutorial
  – Maddie Wilcox (recent Penn PhD) on her research questions in a workshop format
  – Others: American studies, linguistics, Middle East
What next for WORD LAB?

• Incubating a collaborative research project involving librarians, students, faculty on East Asian text analysis
• Building community and continuing outreach
• WORD LAB has exceeded all expectations in its turnout and success!
• Proven concept for collaboration in the library
• Build modest infrastructure, push for greater presence among DH on campus
• WORD BLAB podcast coming up