

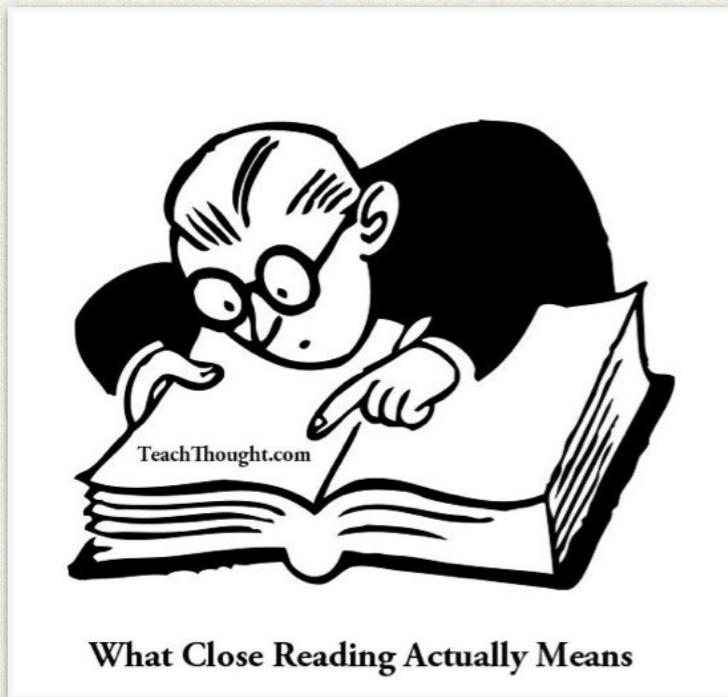
Methods in Digital and Computational Humanities

Module 1: Digital Materials & Collection

18 November, 2015

“micro”

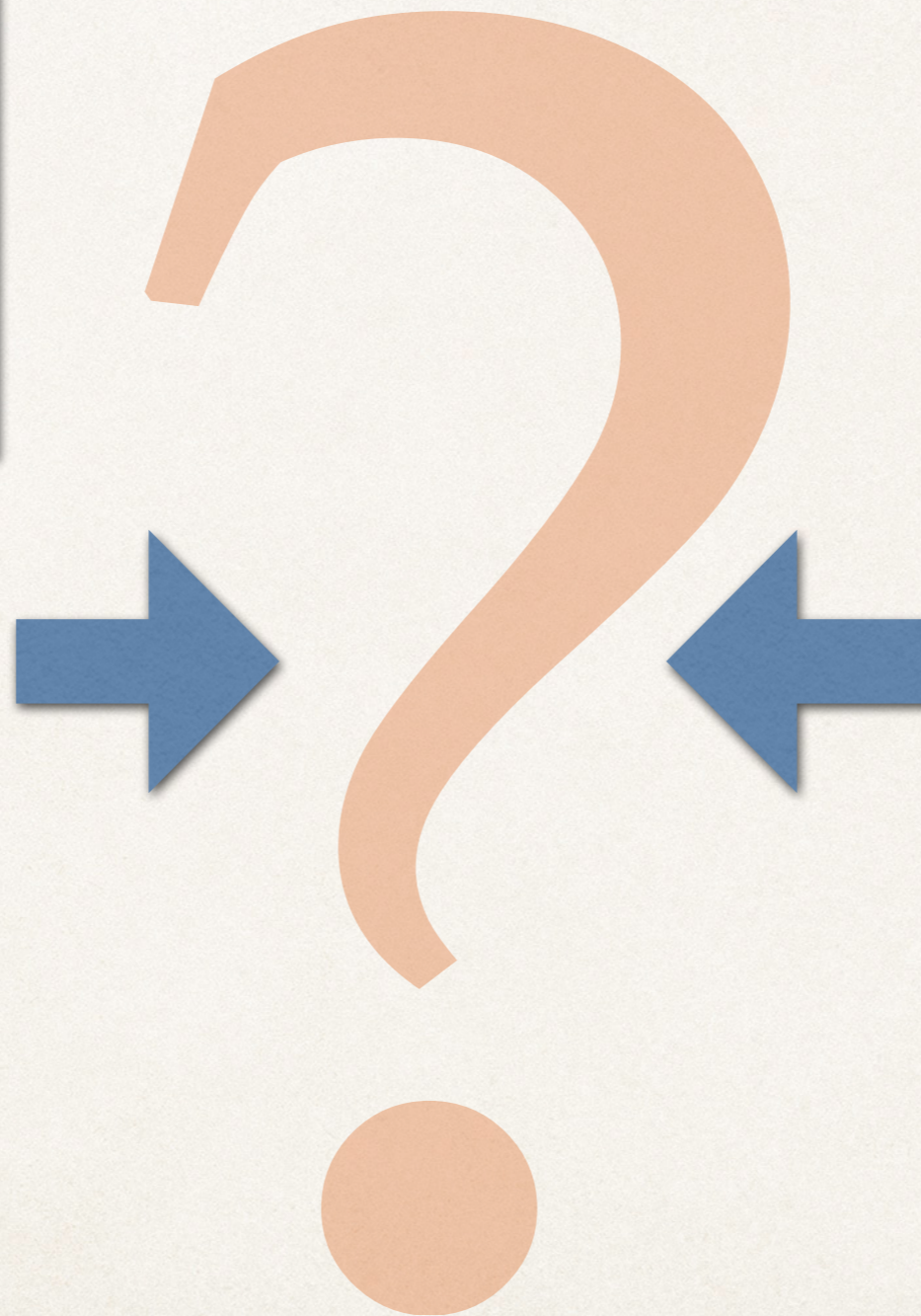
“macro”



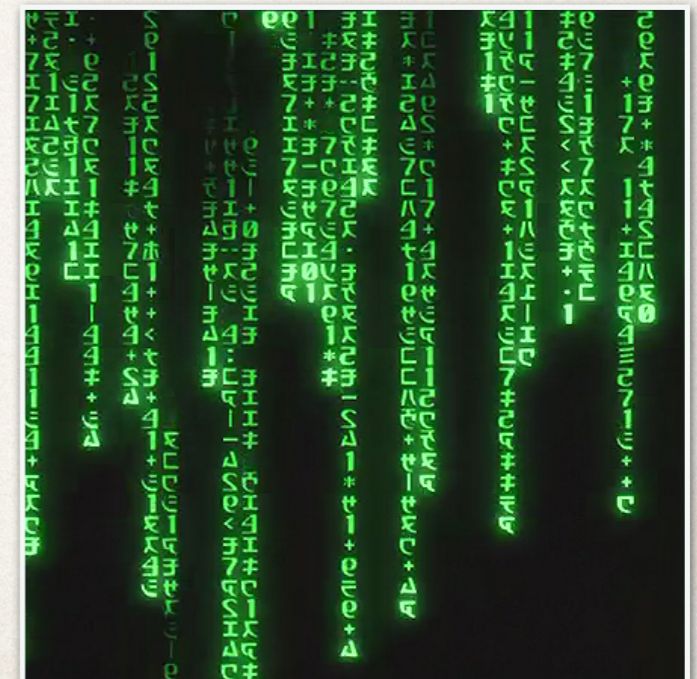
<http://www.teachthought.com/wp-content/uploads/2013/05/what-close-reading-actually-means.jpg>



<https://bookspeakvolumes.files.wordpress.com/2013/02/doorstop.jpg>



Argonne National Laboratory's Flickr page



<http://blog.onlyice.net/work/2012-12-Matrix-Green-Text/matrix-green-screen.png>

Connecting micro to macro

Three challenges for using micro-level data in macro-level analyses:

1. **Heterogeneity of data:** little common structure, where structure exists at all.
2. **Loose coupling between data and inferences:** citation practices vary, and are usually too vague to identify specific evidences for historical claims.
3. **Lack of specificity:** computers are literalists; historical data are rarely encoded in a way that facilitates precise interpretation.

This Module

- ❖ Collecting documents;
- ❖ Curating metadata;
 - ❖ Names;
 - ❖ Dates.
- ❖ Extracting useable text;
- ❖ Semantic web: encoding information in ways that computers can understand.

What is data?

What is worth sharing? (Why share?)

- ❖ Evidence
- ❖ Reuse by other researchers
- ❖ Funding agency requirements
- ❖ Journal requirements
- ❖ Moral imperative?

Flavors of Data

- ❖ Quantitative data vs “thick data”
- ❖ “Big data” (macro) vs micro (& meso?)

Where can/should you share it?

- ❖ Personal website?
- ❖ Institutional repository
- ❖ Field/domain-specific repository
 - ❖ UC Irvine Network Data Repository: <http://networkdata.ics.uci.edu/index.php>
 - ❖ ASU Digital HPS Repository: <http://hpsrepository.asu.edu>
- ❖ Big catch-all repositories:
 - ❖ Dryad: <http://datadryad.org/>
 - ❖ Harvard Dataverse Network: <http://thedata.harvard.edu/dvn/>
 - ❖ UK Data Archive: <http://www.data-archive.ac.uk/deposit/who>
 - ❖ Cultural Policy and the Arts National Data Archive: <http://www.cpanda.org/cpanda/home>
 - ❖ Faculty of 1000 — <http://f1000.com/> — what about posters and presentations?
- ❖ Code repositories
 - ❖ Sourceforge, GitHub

What should other people be able to do with it?

- ❖ CC0 - does it make sense? — <https://creativecommons.org/about/cc0>
- ❖ Other CC licenses — <https://creativecommons.org/choose/>
- ❖ Open Source Software licenses
 - ❖ <http://choosealicense.com/>