Possible Data Sources

Initially from the whiteboard, 1/30/14

Bibliographic Data

MARC

- Harvard HOLLIS
- LibraryCloud enhanced MARC
- SUL catalog
- CUL catalog

Thesis records (a good opportunity to link to persons)

Person Data

use case: when does something in SearchWorks have a Stanford author?

Faculty profile data

- CAP
- VIVO
- Harvard profiles (Faculty Finder?)

Usage data

- LibraryCloud – usage
- Cornell circulation data
- Stanford circulation data
- Supplementary online circ data?
  - Hathi?
  - Open Library?
  - Project Gutenberg?
  - LibraryThing book mentions in their online conversations (They were willing to give LibraryCloud that info)

Other collections

Archival collections

- EAD
  - At the component level? (Harvard has a project doing this.)

Repositories

- Cornell eCommons
- Harvard DRS
  - One source at Harvard says that all items in DRS are also in the HOLLIS catalog, but we think that's not quite right
  - There may or may not be useful metadata in DRS about the items (even if they're in HOLLIS), but DRS metadata tends to be quite operational: info about page scans for use by page-turning software, etc.

Digital Collections

- Arbitrary OAI-PMH
- Possible Harvard collections to pull in:
  - Nuremberg
Emily Dickinson
Colonial North America (which also has EAD component metadata)

Image databases
• IIIF
• VIA/Olivia/Shared Shelf
• HWP

Arbitrary OAI-PMH
• plug in any AO

Curated information and annotations

Annotations
• Cornell’s CULLR
• Bloglinks
• DMS (Stanford digital manuscripts)
• DBpedia
• LibGuides (Cornell and Stanford at least)

Subjects & Authorities

Subjects
• NYTines subjects (entity resolution)
• BBC
• Europeana
• DPLA
• Freebase
• LC Classification Codes (LCC)
  • LibraryCloud has a working API to serve these up
    • http://hiswebtest.law.harvard.edu/dev/david/subjectBrowser/subjectBrowser.html

Authorities
• taxonomies
• FAST
• researcher ids
  • VIAF
  • Social Networking in Archival Context
    • encoding of EACCPF
  • id.loc.gov
  • ISNI
  • ORCID
• OCLC researcher networking group – roughly 30 sources
• organizations
  • ISNI/Ringold
• journals
• geographic locations
  • Geonames
  • DBpedia
  • Geopolitical
• Getty placenames thesaurus
• events