```
In [1]: %cd '/Users/jwa/work/library-link-network-demo'
        from traitlets.config.manager import BaseJSONConfigManager
        path = "/Users/jwa/anaconda/etc/jupyter/nbconfig"
        cm = BaseJSONConfigManager(config dir=path)
        cm.update('livereveal', {
                 'theme': 'simple',
                 'transition': 'none',
                 'overview': True,
                 'center': True,
                 'scroll': True,
                 'hideAddressBar': True,
                 'slideNumber': True,
                 'start slideshow at': 'selected'
        })
        /Users/jwa/work/library-link-network-demo
Out[1]: {'center': True,
          'hideAddressBar': True,
          'overview': True,
          'scroll': True,
          'slideNumber': True,
          'start slideshow at': 'selected',
          'theme': 'simple',
          'transition': 'none'}
         %%javascript
```

Jupyter.keyboard manager.command shortcuts.remove shortcut('f');

publishing bibliographic data to the web

working with the Library Link Network

Jack Ammerman < jwa@bu.edu>



The internet has changed the very meaning of "research" *

*How Teens Do Research in the Digital World (Pew Internet and the American Life Project, Nov 2012)

http://www.pewinternet.org/files/old-

media//Files/Reports/2012/PIP TeacherSurveyReportWithMethodology110112.pdf

(http://www.pewinternet.org/files/old-

media//Files/Reports/2012/PIP_TeacherSurveyReportWithMethodology110112.pdf)

Advanced Placement & National Writing Teachers report how students search:

Google or other online search engine (94%).

Wikipedia or other online encyclopedia (75%).

YouTube or other social media sites (52%).

Their peers (42%).

Spark Notes, Cliff Notes, or other study guides (41%).

News sites of major news organizations (25%).

Print or electronic textbooks (18%)

Online databases such as EBSCO, JSTOR, or Grolier (17%).

A research librarian at their school or public library (16%).

Printed books other than textbooks (12%).

Student-oriented search engines such as Sweet Search (10%).

Google dependence: always returns to Google when confused; repeatedly asserts that "Google is my friend"; demonstrates the belief that Google has everything; uses Google as all-inclusive tool.

The Truth Is Out: How Students REALLY Search by Beth S. Bloom and Marta Deyrup (2012)

(DOI: 10.5703/1288284315103)

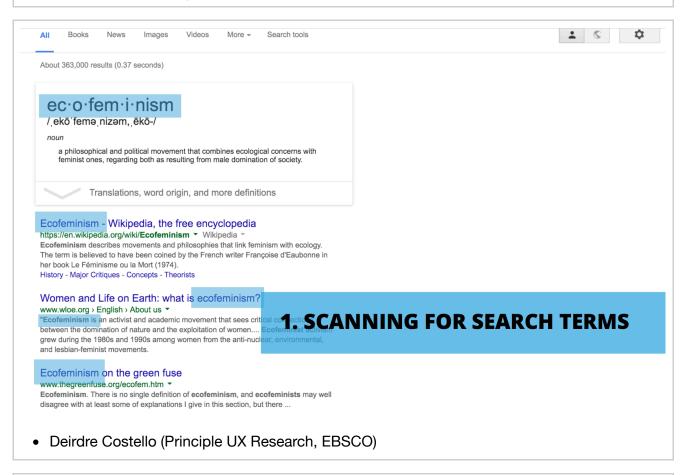
Students do use library databases – though they often start in Google

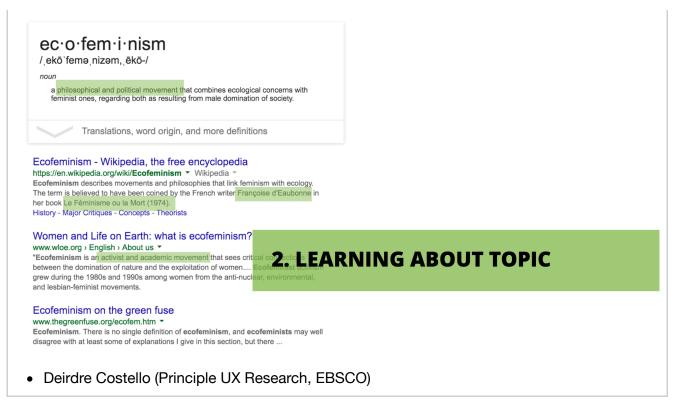
How college students *really* do research – findings from recent studies Emily Singley

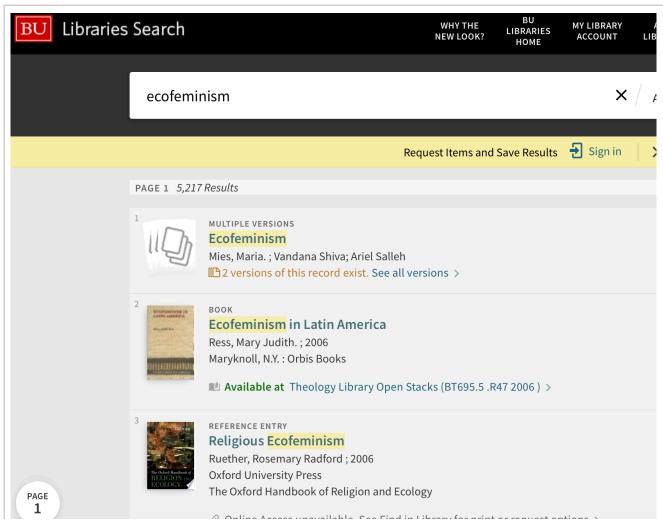
http://emilysingley.net/how-college-students-really-do-research-findings-from-recent-studies/ (http://emilysingley.net/how-college-students-really-do-research-findings-from-recent-studies/)

I always start with Google!

- Boston University Student









Google as "Pre-Search"

Could we make Google better?

Research Questions

Discovery

 How would inclusion of bibliographic data in traditional Internet search engines (Google, Bing, Yahoo, etc.) affect search results in those search engines?

Research Questions

Discovery

- How would inclusion of bibliographic data in traditional Internet search engines (Google, Bing, Yahoo, etc.) affect search results in those search engines?
- Would enabling users to discover our library resources in Google increase their use?

Research Questions

Discovery

- How would inclusion of bibliographic data affect search results for Internet search engines?
- Would enabling users to discover our library resources in Google increase their use?
- Would use of linked data enhance discovery of information in Primo? If so, in what ways? How would we measure the impact?

Research Questions

Discovery

- How would inclusion of bibliographic data affect search results for Internet search engines?
- Would enabling users to discover our library resources in Google increase their use?
- Would use of linked data enhance discovery of information in Primo? If so, in what ways? How would we measure the impact?
- What information from linked authority sources would be helpful to the user in a display of bibliographic search results? How would we display that information (ex: Google Knowledge Card)?

Examples of use Google's use of Linked Data

https://www.google.com/#q=thomas+edison (https://www.google.com/#q=thomas+edison)

https://www.google.com/#q=the+boss+baby (https://www.google.com/#q=the+boss+baby)

Research Questions

Non-discovery uses of bibliographic data

- How might bibliographic data be used beyond discovery of an information source?
 - Computational analysis?
 - Geographically and/or Bibliographically aware mobile apps?
 - Authority control?

Until bibliographic records are openly accessible on the Web, we won't know.

Practical Questions

Resources

- How to gain competence among BU Library staff in working with Bibframe or Schema?
- How to publish BU bibliographic records to the web with little internal technical support for the project?

Practical Questions

- Are there things we can do to make our bibliographic records better for the Web?
- If we enhance records with other authority sources, at what point should we do so?
 - In Alma before we export to the web?
 - Post Alma export but before publishing to the web?
 - How/should we enhance Primo records with linked data?

And finally,

- What is the right scale to address the various components of the effort?
 - What components are best addressed by library system vendors (like Ex Libris)?
 - When does working with a third party vendor (like Zepheira) make sense?
 - What parts of this effort are best addressed at the local level?
 - Can we use automated content generation and curation approaches at any of these levels?

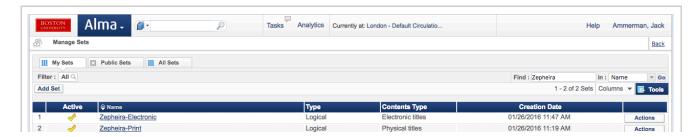
Began working with Zepheira in February 2016

- Bibframe training for staff
- · Publish bib records to the web

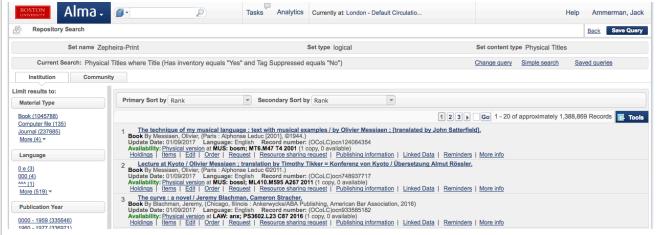
Publishing Bib Records to the Web

Our Workflow In Alma

Create one or more managed sets containing the records you want to publish



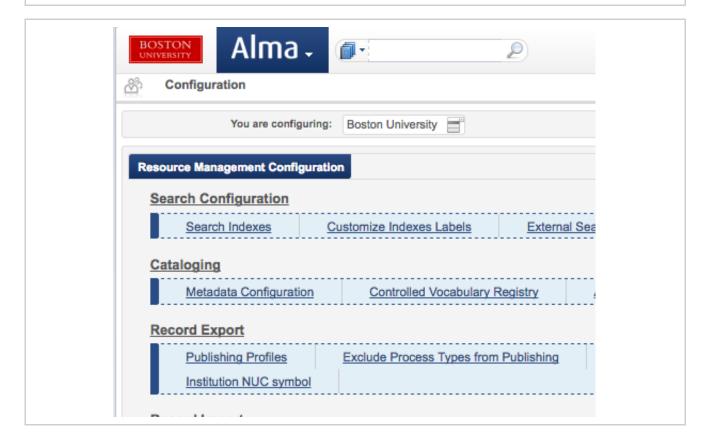


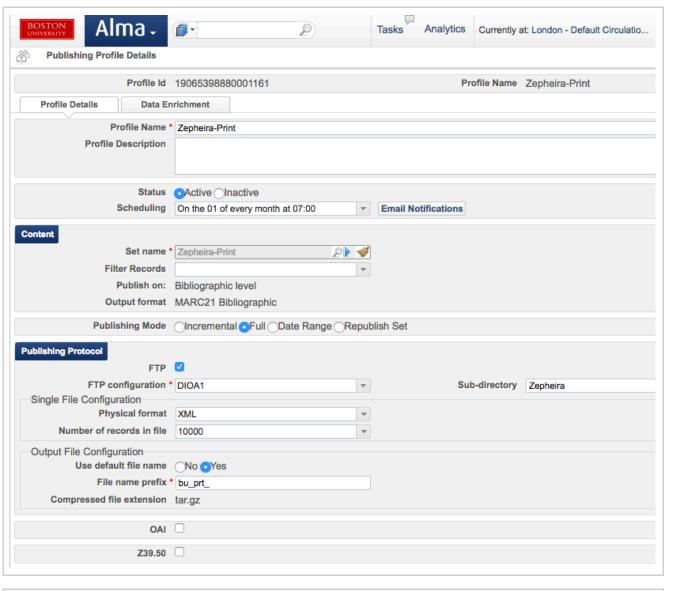


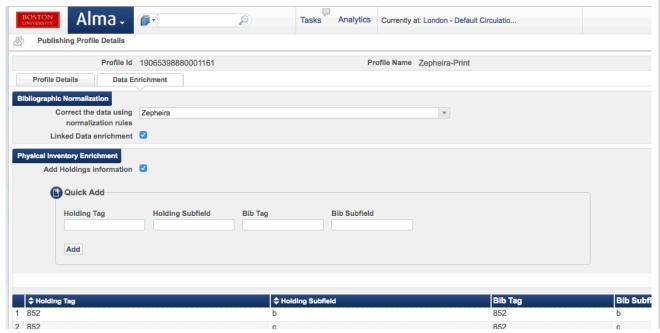
Publishing Bib Records to the Web

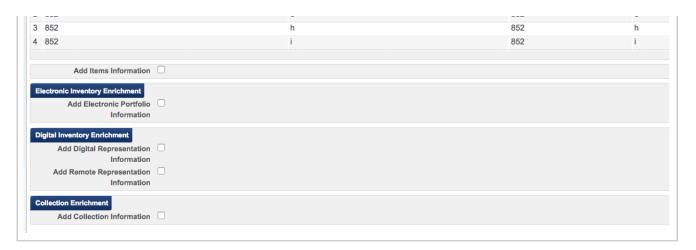
Our Workflow In Alma

- Create one or more managed sets containing the records you want to publish
- Create a General Publishing job(s) to publish the records to an FTP server









Normalization Rule Changes

- Change 856 tags to 956 for links that are licensed resources (proxied)
- Maintain 856 tags for open access resources
 - Internet Archive
 - HathiTrust
- Remove 7xx fields that have a CKB number in subfield 'w'
 - ex: removeField "770" if (exists "770.w.(CKB)*")
- Remove the "(uri)" prefix in subfield '0'
 - ex: ReplaceContents "600.0.(uri) " with ""

Pre-Process each record via script

- Copy the published marcxml records to a computing cluster for processing
- Strip unwanted data/subfields
 - ex: (uri)
 - Any subfield that goes through the Ex Libris resolver

Example:

https://open-na.hosted.exlibrisgroup.com/resolver/wikidata/oclc/909538078 (https://open-na.hosted.exlibrisgroup.com/resolver/wikidata/oclc/909538078)

Pre-Process each record via script

- Iterate through each field looking for any field with subfield 0 containing an LC authority number
 - ex: | http://id.loc.gov/authorities/names/n79091168
 (http://id.loc.gov/authorities/names/n79091168)
 - ex: http://id.loc.gov/authorities/subjects/sh92006372

(http://id.loc.gov/authorities/subjects/sh92006372)

- If an LC authority number is found, check a lookup table for additional data sources that are associated with the LC authority.
- Add additional subfield 0 subfields when a source is found in the lookup table

Exported from Alma:

Enhanced:

```
<datafield tag="700" ind1="0" ind2=" ">
        <subfield code="a">Augustine,</subfield>
        <subfield code="c">Saint, Bishop of Hippo.</subfield>
        <subfield code="0">| http://id.loc.gov/authorities/names/
n80126290</subfield>
        <subfield code="0">https://viaf.org/viaf/66806872</subfie</pre>
1d>
        <subfield code="0">https://www.wikidata.org/wiki/Q8018</s</pre>
ubfield>
        <subfield code="0">https://en.wikipedia.org/wiki/Augustin
e of Hippo</subfield>
        <subfield code="0">https://thesaurus.cerl.org/record/cnp0
1302049</subfield>
        <subfield code="0">| http://vocab.getty.edu/page/ulan/500
104317</subfield>
        <subfield code="0">| http://plato.stanford.edu/entries/au
gustine</subfield>
        <subfield code="0">| http://id.worldcat.org/fast/61554</s</pre>
ubfield>
        <subfield code="0">| http://www.imdb.com/nm1240835</subfi</pre>
eld>
    </datafield>
```

Exported from Alma:

Enhanced by script:

```
<datafield tag="700" ind1="1" ind2=" ">
        <subfield code="a">Souter, Alexander,</subfield>
        <subfield code="d">1873-1949.</subfield>
        <subfield code="0">| http://id.loc.gov/authorities/names/
n88638867</subfield>
        <subfield code="0">https://viaf.org/viaf/1159</subfield>
        <subfield code="0">https://www.wikidata.org/wiki/Q4720124
</subfield>
        <subfield code="0">https://en.wikipedia.org/wiki/Alexande
r Souter</subfield>
        <subfield code="0">| http://www.npg.org.uk/collections/se
arch/person/mp141707</subfield>
        <subfield code="0">| http://id.worldcat.org/fast/1485105<</pre>
/subfield>
        <subfield code="0">| http://www.oxforddnb.com/index/10107
5166</subfield>
    </datafield>
```

WikiData Property	Linked Data Source
P1649	KMDb person ID
P727	Europeana ID
P830	Encyclopedia of Life ID
P724	Internet Archive ID
P760	DPLA ID
P846	Global Biodiversity Information Facility ID
P214	VIAF ID
P1816	National Portrait Gallery (London) person ID
P245	ULAN ID
P243	OCLC control number
P1667	TGN ID
P957	ISBN-10
P1795	Smithsonian Amer Art Mus: person thesaurus id
P345	IMDb ID
P496	ORCID

P2671	Google Knowledge Graph identifier
P1871	CERL ID
P486	MeSH ID
P1566	GeoNames ID

WikiData Property	Linked Data Source
P2163	FAST-ID
P213	ISNI
P1157	US Congress Bio ID
P672	MeSH Code
P402	OpenStreetMap Relation identifier
P2036	African Plant Database
P1184	handle
P1417	Encyclopaedia Britannica Online ID
wp	Wikipedia
wd	WikiData
P3241	Catholic Encyclopedia ID
P212	ISBN-13
P1960	Google Scholar ID
P1415	Oxford Biography Index Number
P236	ISSN
P1669	CONA ID
P1014	AAT ID
P1230	JSTOR journal code
P1144	LCOC LCCN (bibliographic)
P3123	Stanford Encyclopedia of Philosophy ID
P244	LCAuth ID
P1156	Scopus Source ID
P675	Google Books ID

Let's look at link.bu.edu

http://link.bu.edu (http://link.bu.edu)

Try a few searches

The C.P. Snow controversy. Binkowski, Joanna

Joannis Tideman Dissertatio de deo Platonis, sive, Diatribe in Platonis philosophiae locum, qui est de deo, divinaque natura : nunc primum edita et indice aucta. Tideman, J.

Maimonides Rambam : spomenica povodom osamstogodišnjice od njegovog rođenja, 1135-1935. Daniel Ozmo

Lion of the Senate: When Ted Kennedy Rallied the Democrats in a GOP Congress, Nick Littlefield and David Nexon

link.bu.edu

- Not a search interface
- Designed for computers, but with humans in mind
- · Embedded with rich metadata using schemas understood by the Web
- Resource pages for concepts (subjects)
- · Designed to give visibility to the Library

Google Analytics

Recently released functionality

Resource Navigator

- a staff tool for data cleanup
- http://navigate.library.link/link.bu.edu/ (http://navigate.library.link/link.bu.edu/)

Recently released functionality

Resource Navigator

- a staff tool for data cleanup
- http://navigate.library.link/link.bu.edu/ (http://navigate.library.link/link.bu.edu/)

Embed content

• http://link.bu.edu/resource/nv6k8YI8TvQ/ (http://link.bu.edu/resource/nv6k8YI8TvQ/ (http://link.bu.edu/resource/nv6k8YI8TvQ/)

```
In [2]:
```

```
%%html
<h1>Let's comsume some content</h1>
<h2>Authority Links associated with Kennedy, Edward M., (Edward Moore), 1
<br/>
<div id="R-nv6k8YI8TvQ-this" class="lln-embed">
<iframe width="100%" height="485px" src="http://link.bu.edu/resource/nv6k<//iframe>
</div>
```

Let's comsume some content

Authority Links associated with Kennedy, Edward M., (Edward Moore), 1932-2009

(/concept/) Kennedy, Edward M., (Edward Moore), 1932-2009

Authority link (http://bibfra.me/vocab/lite/authorityLink)

- https://en.wikipedia.org/wiki/Ted_Kennedy (https://en.wikipedia.org/wiki/Ted_Kennedy)
- http://www.npg.org.uk/collections/search/person/mp70784 (http://www.npg.org.uk/collections/search/person/mp70784)
- http://id.loc.gov/authorities/names/n79091168 (http://id.loc.gov/authorities/names/n79091168)
- http://id.worldcat.org/fast/01730723 (http://id.worldcat.org/fast/01730723)
- http://id.worldcat.org/fast/1730723 (http://id.worldcat.org/fast/1730723)
- http://bioguide.congress.gov/scripts/biodisplay.pl?index=K000105 (http://bioguide.congress.gov/scripts/biodisplay.pl?index=K000105)
- https://viaf.org/viaf/68969344 (https://viaf.org/viaf/68969344)
- http://www.imdb.com/nm0448356 (http://www.imdb.com/nm0448356)
- http://id.loc.gov/authorities/n79091168 (http://id.loc.gov/authorities/n79091168)
- https://www.wikidata.org/wiki/Q134549

/https://www.wikidata.org/wiki/0121510)

```
In [5]: %%html
```

```
%%ntml
<h2>Items from BU Libraries about Kennedy, Edward M., (Edward Moore), 193
</br>
<div id="R-nv6k8YI8TvQ-related-by-concept" class="lln-embed">
<iframe width="100%" height="429px" src="http://link.bu.edu/resource/nv6k
</iframe></div>
<div class="citation" vocab="http://schema.org/">
<i class="fa fa-lg fa-external-link-square fa-fw"></i> Data from
<span resource="http://link.bu.edu/resource/nv6k8YI8TvQ/" typeof="Intangi
<span property="name http://bibfra.me/vocab/lite/label">
<a href="http://link.bu.edu/resource/nv6k8YI8TvQ/">Kennedy, Edward M., (E
<span property="offeredBy" typeof="Library ll:Library" resource="http://l
<span property="name http://bibfra.me/vocab/lite/label">
<a property="name http://bibfra.me/vocab/lite/label">
<a property="url" href="http://link.bu.edu/">Boston University Libraries</a>
</span></span></span></span></div>
```

Items from BU Libraries about Kennedy, Edward M., (Edward Moore), 1932-2009

2 18 Items that are about the Concept ❖ Kennedy, Edward M., (Edward Moore), 1932-2009

Death at Chappaquiddick, by Thomas L. Tedrow and Richard L. Tedrow (http://buprimo.hosted.exlibrisgroup.com/primo-explore/search? institution=BOSU&vid=BU&lang=en_US&query=any,contains,99188787730001161)

(http://link.bu.ed/portal/Death-at-Chappaquiddick-by-Thomas-L.-Tedrow-and/Ni1TQr3zVVA/)

Edward Kennedy: the myth of leadership, Murray B. Levin and T.A. Repak (http://buprimo.hosted.exlibrisgroup.com/primo-explore/search? institution=BOSU&vid=BU&lang=en_US&query=any,contains,99186489210001161)

(http://link.bu.ed/portal/Edward-Kennedy--the-myth-of-leadership-Murray/xAwtBnlz2LM/)

Edward Kennedy and the Camelot legacy, James MacGregor Burns (http://buprimo.hosted.exlibrisgroup.com/primo-explore/search? institution=BOSU&vid=BU&lang=en_US&query=any,contains,99181761510001161)

(http://link.bu.ed/portal/Edward-Kennedy-and-the-Camelot-legacy-James/b5dg5_LNdfo/)

Data from Kennedy, Edward M., (Edward Moore), 1932-2009 (http://link.bu.edu/resource/nv6k8YI8TvQ/) - Boston University Libraries (http://link.bu.edu/)

Primo as a consumer of Linked Data?

http://buprimo.hosted.exlibrisgroup.com/primo-explore/search?vid=jwa&lang=en_US (http://buprimo.hosted.exlibrisgroup.com/primo-explore/search?vid=jwa&lang=en_US)

Thanks!

Jack Ammerman Boston University Libraries jwa@bu.edu

Github: https://github.com/jwacooks/library-link-network-demo

(https://github.com/jwacooks/library-link-network-demo)

Link Network: http://link.bu.edu (http://link.bu.edu)

BU Libraries: http://www.bu.edu/library (http://www.bu.edu/library)

Goals for Linked Data

- Publish bib records to the web to allow discovery by Internet Search Engines
- Allow for computational analysis of bibliographic data using a wide variety of tools
- Use linked authority sources to enrich our bib records and address authority control
- link to other data sources (DPLA, LinkedJazz, etc.)