

Linked Data for Libraries (LD4L) Use Cases

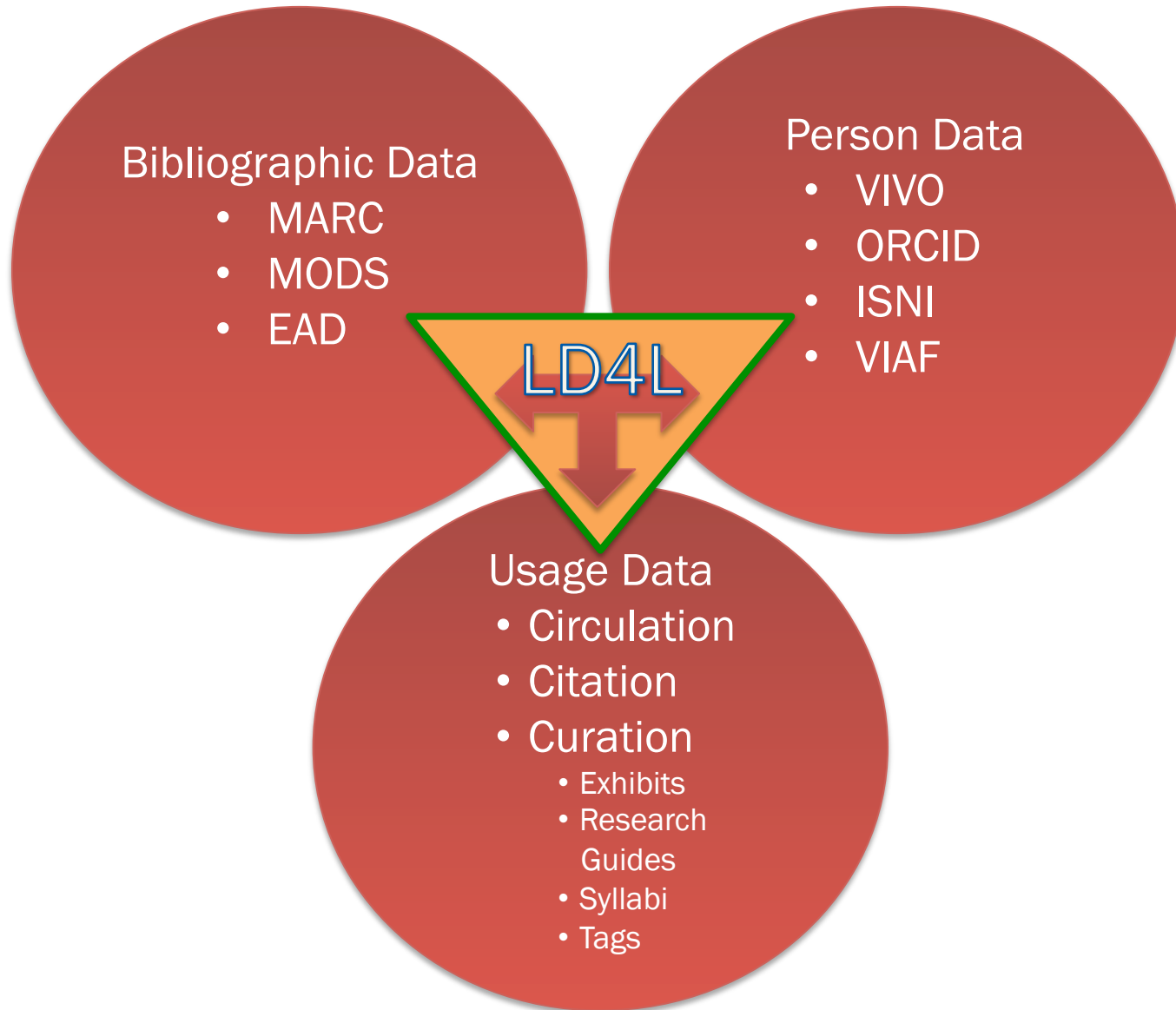
Tom Cramer (Stanford)
Simeon Warner (Cornell)

Monday, 22 February 2015

Process

- Brainstorming and refinement spring and summer 2014
- Want manageable set of uses cases that
 1. illustrates benefits of Linked Data
 2. has enough data available to demonstrate
 3. shows benefits of cross-institutional approach
- Started engineering work based on use cases summer 2014
- Want feedback on value, importance/interest, feasibility

LD4L Data Sources



Stories as the Basis of Use Cases

As a _____, I want to _____, so that I can
<realize this benefit>.

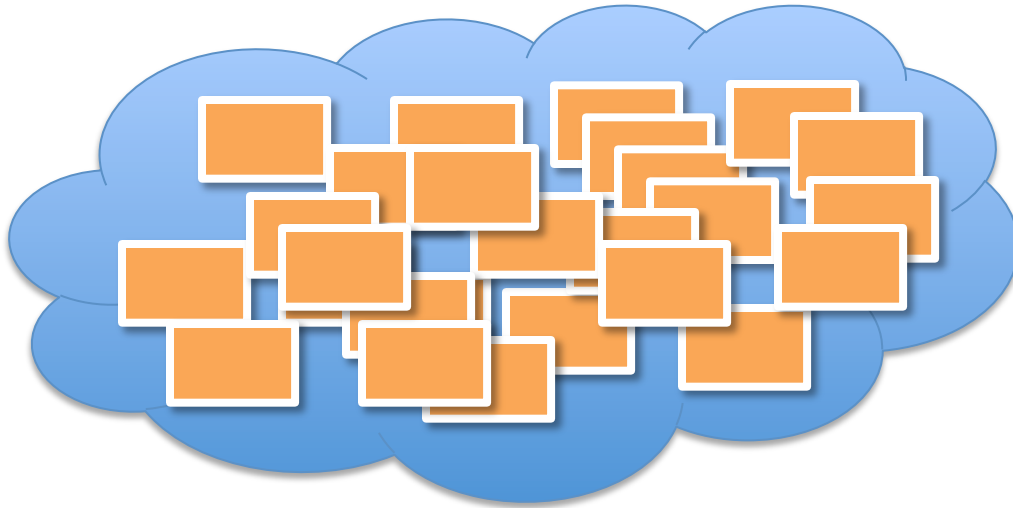
Potential Demonstrations:

- A. Demo 1
- B. Demo 2
- C. Demo 3

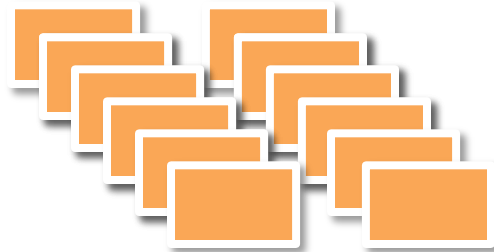
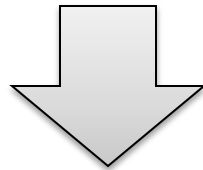
Data Sources Needed:

Ontology Requirements:

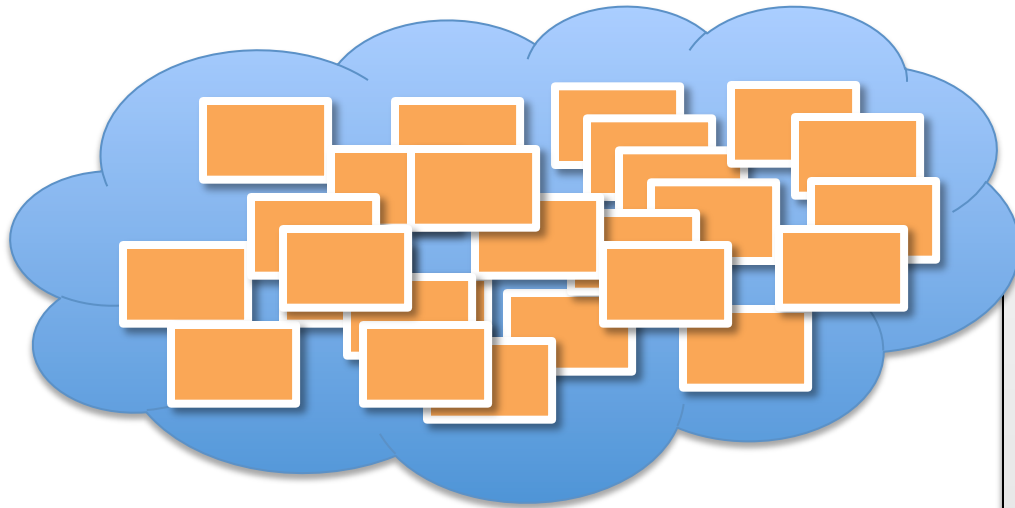
Engineering Work:



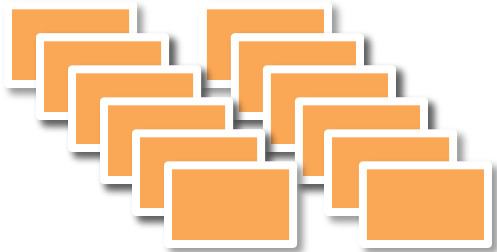
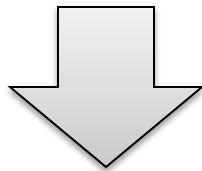
42 Raw Use Cases



12 Refined Use Cases
in 6 clusters...



42 Raw Use Cases



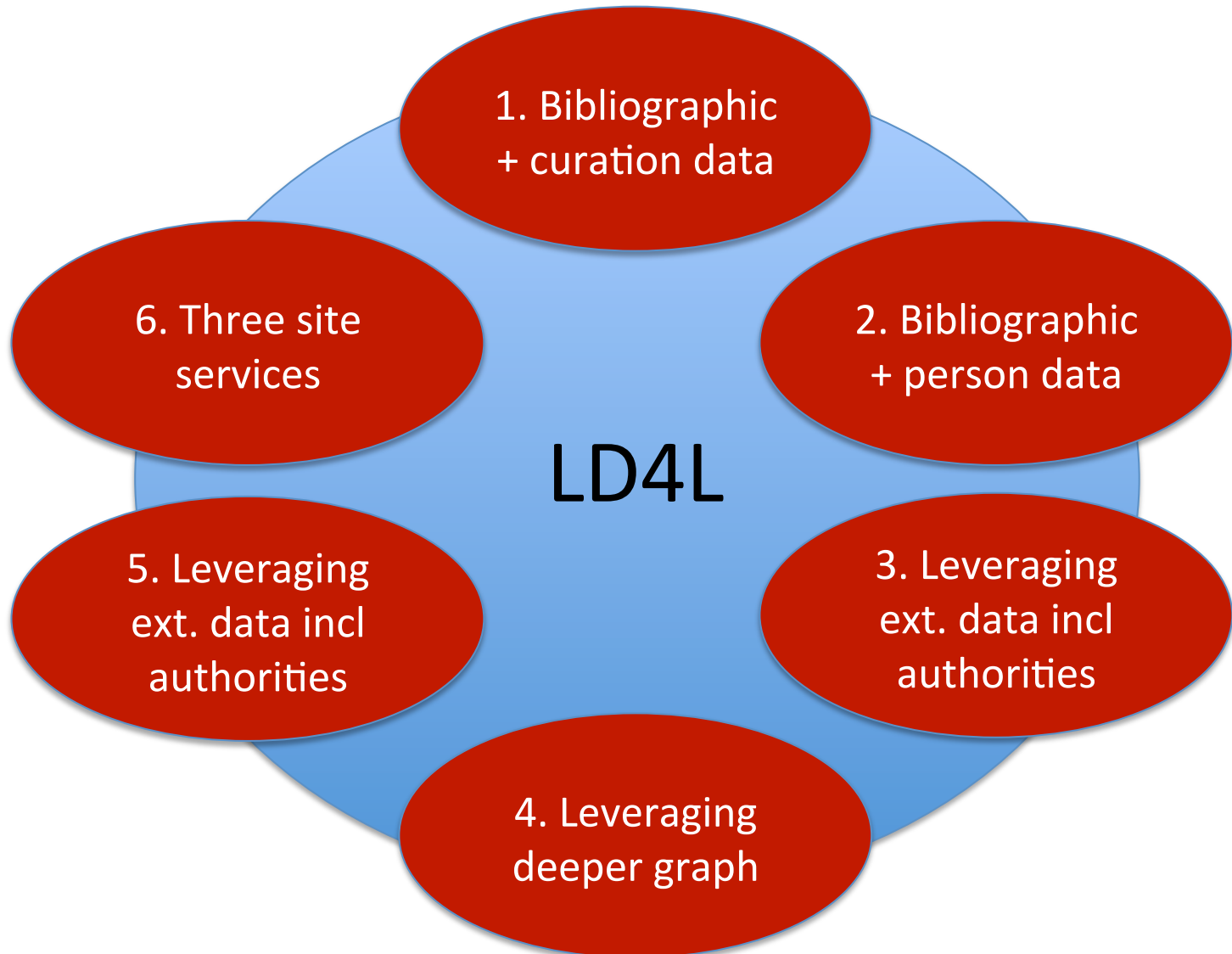
12 Refined Use Cases
in 6 clusters...



LD4L Use Case Clusters

1. Bibliographic + curation data
2. Bibliographic + person data
3. Leveraging external data including authorities
4. Leveraging the deeper graph (via queries or patterns)
5. Leveraging usage data
6. Three-site services, e.g. cross-site search

Clusters



CLUSTER 1. BIBLIOGRAPHIC & CURATION DATA

UC1.1: Build a virtual collection

“As a faculty member or librarian, I want to create a virtual collection or exhibit containing information resources from multiple collections across multiple universities either by direct selection or by a set of resource characteristics, so that I can share a focused collection with a <class, set of researchers, set of students in a disciplinary area>.”

UC1.1: Build a virtual collection

- Individual selection or resources
- Collection may be exhibit, reading list, etc.
- Includes description, ordering
- Shared or private
- **Enable across institutions**

- Well developed, uses annotations
- Demo and discussion after coffee (Lynette Rayle/
Naomi Dushay/Simeon Warner/Rob Sanderson)

UC1.2: Tag scholarly information resources to support reuse

“As a librarian, I would like to be able to tag scholarly information resources from one or multiple institutions into curated lists, so that I can feed these these lists into subject guides, course reserves, or reference collections. I'd like these lists to be portable (into Drupal, into LibGuides, into Spotlight! or Omeka, into Sakai, e.g.) and durable. I'd like these lists/tags to selectively feed back into the discovery environment without having to modify the catalog records.”

UC1.2: Tag scholarly information resources to support reuse

- Scale $O(100,000)$ items
- Use for e.g. virtual subject library
- Selection by rules and patterns
- Collaborative curation
- Feed data into discovery environment
- Same model as 1.1, partially developed

Search for Engineering Lib: x

https://engineering.library.cornell.edu/cullrseek/search/%2A?rqt=Biomedical,%20Biological%20and%20Chemical%20Engine...

Engineering Library

Looking for course reserves? Search the course reserves page.

Sort by

- Relevance
- Title
- Year

Format

- book (13776)
- journal (1046)
- ebook (413)
- ejournal (375)
- manuscript (86)
- video recording (82)
- database (23)
- computer file (19)
- sound recording (4)
- map (1)

Content

- thesis (1809)
- package (5)
- handbook (4)
- encyclopedia (2)
- directory (1)
- standard (1)

Search for Engineering Library Resources

"use quotes around exact phrases"

Start new search

Biomedical, Biological and Chemical Engineering

Search results

Total found: 15825

CRC handbook of chemistry and physics	Searchable ready-reference resource of chemical and physical data, updated with new research. Coverage includes physical constants of organic compounds, properties of the elements and inorganic compounds, thermochemistry, electrochemistry and kinetics, fluid properties, biochemistry, analytical chemistry, molecular structure and spectroscopy, atomic, molecular and optical physics, nuclear and particle physics, properties of solids, polymer properties, geophysics, astronomy, and acoustics, practical laboratory data, health and safety information. Searchable by text, and by substance/property.	database	connect
PubMed	Provides access to over 16 million citations from MEDLINE and other life science journals for biomedical articles back to the 1950s. Includes links to many sites providing full text articles and other related resources.	database	connect
Marquis who's who	Marquis Who's Who on the Web, a 'high-speed', searchable online database, features comprehensive profiles on over 1.4 million of the most accomplished individuals from all fields of endeavor including: government, business, science and technology, the arts, entertainment, and sports. Search by name, gender, occupation, geography, hobbies and interests, ...	database	connect

Cornell: provide virtual library “views/collections” as part of main Blacklight discovery system

CLUSTER 2. BIBLIOGRAPHIC & PERSON DATA

UC2.1: See and search on works by people to discover more works, and better understand people

“As a researcher, I'd like to see / search on works <by, about, cited by, collected, taught> by University faculty <in an OPAC, profiles system>, to discover works of interest based on connection to people, and to understand people based on their relation to works.”

“As a researcher, I'd like to see a list of works from the most prolific authors in my field at my institution and at other institutions.”

UC2.1: See and search on works by people to discover more works, and better understand people

- Profile + catalog data
- Perhaps between institutions: multiple catalogs + multiple profile systems
- Poor journal data suggests likely can do more in non-journal disciplines
- Afternoon demo and discussions today (Rebecca Younes/Steven Folsom/Darren Weber/Jon Corson-Rikert/Phil Schreur)

CLUSTER 3. LEVERAGING EXTERNAL DATA INCLUDING AUTHORITIES

UC3.1: Search with Geographic Data for Record Enrichment and Pivoting

“As a researcher, I'd like to see the geographic context of my search results, and be able to pivot, extend or refine a search with a single click, in order to better assess found resources, find related resources, and filter or expand search results to broaden or narrow a search on the fly.”

UC3.2: Search with Subject Data for Record Enrichment and Pivoting

“As a researcher, I'd like to see the subject area contexts for my search results and be able to pivot, extend or refine a search ...”

UC3.3: Search with Person Data for Record Enrichment and Pivoting

“As a researcher, I'd like to see the person contexts for my search results and be able to pivot, extend or refine a search ...”

UC3.x: Search with XYZ for Record Enrichment and Pivoting

XYZ = geographic / subject / person data

- External data and entity linking
- Things not strings
- Different data types have different challenges and data sources

- Included in afternoon brainstorming discussions

CLUSTER 4. LEVERAGING THE DEEPER GRAPH (VIA QUERIES OR PATTERNS)

UC4.1: Identifying related works

“As a scholar, I would like to find all the images associated with various instances of a work sorted by time, so that I can see how the depictions of or illustrations in a work have changed over time.”

“(GloPAD specific): As a scholar, I would like to find all costume photographs and scene illustrations for various stagings and performances of the plays of a particular author or the operas of a particular composer, so that I can see how the visual look of performances of the plays or operas have changed over time.”

UC4.1: Identifying related works

- use complex graph relationships via queries or patterns (rather than direct connections)
- allow discovery that would not be possible without the semantics of different relationship
- restricted by available data

- Tuesday after-lunch demonstration and discussion (Paolo Ciccarese/Steven Folsom)

UC4.2: Leverage the deeper graph to surface more relevant works

*“As a researcher, I would like to see resources in response to a search where the **relevance ranking of the results reflects the “importance”** of the works, based on how they have been used or selected by others, so that I can **find important resources that might otherwise be “hidden”** in a large set of results.”*

UC4.2: Leverage the deeper graph to surface more relevant works

- calculate scholarly importance
- include items not well represented in commodity services
- do better than Amazon?

CLUSTER 5. LEVERAGING USAGE DATA

UC5.1: Research guided by community usage

*“As a researcher, I want to **find what is being used (read, annotated, bought by libraries, etc.)** by the scholarly communities not only at my institution but at others, and to find sources used elsewhere but not by my community”*

UC5.1: Research guided by community usage

- need to understand community of user
 - direct: auth and from identity?
 - inferred as part of discovery?
- cross-institutional – need usage data that can be shared and merged

UC5.2: Be guided in collection building by usage

“As a librarian, I would like help building my collection by seeing what is being used by students and faculty.”

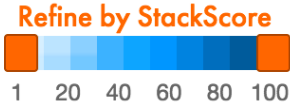
“As a subject librarian, I would like to see what resources in my subject area are heavily used at peer institutions but are not in my institution’s collection.”

UC5.2: Be guided in collection building by usage

- business analytics
- help libraries make best use of collection building activities and funds
- useful at both institutional or cross-institutional levels
- cross-institutional – need usage data that can be shared and merged

Usage data

brewing beer Go!
Advanced Search



Showing 0 to 15 of 685 results for "brewing beer"

- Format ▾
- Book (642)
 - Serial (29)
 - Video/Film (6)
 - Other (3)
 - Sound Recording (3)
 - Book Part (1)
 - Manuscript (1)

- Library ▾
- WID (379)
 - BAK (168)
 - NET (66)
 - HYL (42)
 - SCH (37)
 - CRL (31)
 - LAW (31)
 - FIG (23)
 - HOU (15)
 - ECB (14)
 - more

- Liquor and lab
- Goldwidows
- Beer
- A history of be
- The barbarian
- The U.S. brewi
- Beer in the Mic
- Ale & beer
- Beer
- War, wine, and

Title Author Year StackScore ▾

THE CHRONICLE OF HIGHER EDUCATION

February 20, 2015

Home News Global Opinion & Ideas Facts & Figures **Blogs** Advice Forums Jo



The Conversation
Opinion and ideas.

October 7, 2014 by [David Weinberger](#)



A Good, Dumb Way to Learn From Libraries

Too bad we can't put to work the delicious usage data gathered by libraries.

... may not know as much as click-obsessed Amazon does about how people interact with their books. What they do know, however, reflects the behavior of a community of scholars, and it's unpolluted by commercial imperatives. But privacy concerns have forestalled making library usage data available to application developers outside the library staff, and often even within. And the data are the definition of incompatible: Libraries collect them in different formats at different levels of granularity and at different time scales, making them hard to work with.

- Most work at Harvard
- Demonstration and discussion Tuesday morning (Paul Deschner) & brainstorming follows

CLUSTER 6. THREE SITE SERVICES

UC6.1: Cross-site search

As a scholar I don't want my discovery process to be constrained by the collection boundaries of my university yet I want to retain the detailed coverage of special collections that are important in my field.

*Bonus points: I want **results ranked by the scholarly value**, not simply popularity in the public eye.*

UC6.1: Cross-site search

- demonstrate the power of sharing library data as linked open data
 - opportunities for third parties
 - benefits for partners
- deal with dupes & near dupes

