The Future of Metadata Management & Making Library Collections Discoverable on the Web

Ted Fons, OCLC

The National Library Descriptors Conference

Proposed Changes in the Development of Library Collections in the Era of the Semantic Web

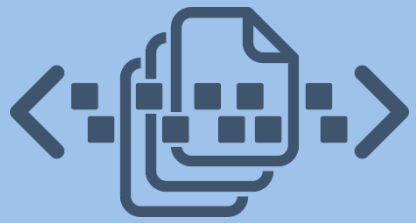
Warsaw - 21-22 April, 2015

OCLC The world's libraries. Connected.

BIBLIOTEKA NARODOWA

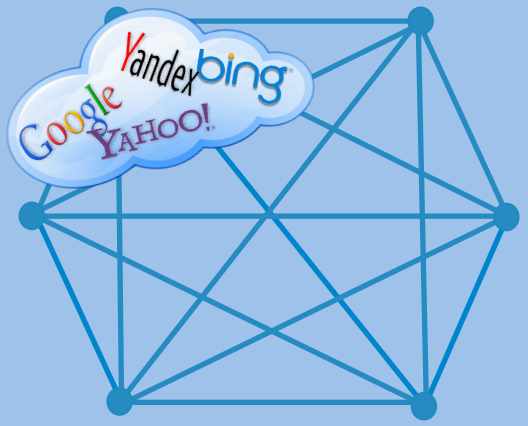
The Future

Cataloging



Harvesting

Datamining



Our Goal

The Goal

Connect the reader to content.

S.R. Ranganathan

1. Books are for use.
2. Every reader his [or her] book.
3. Every book its reader.
4. Save the time of the reader.
5. The library is a growing organism.

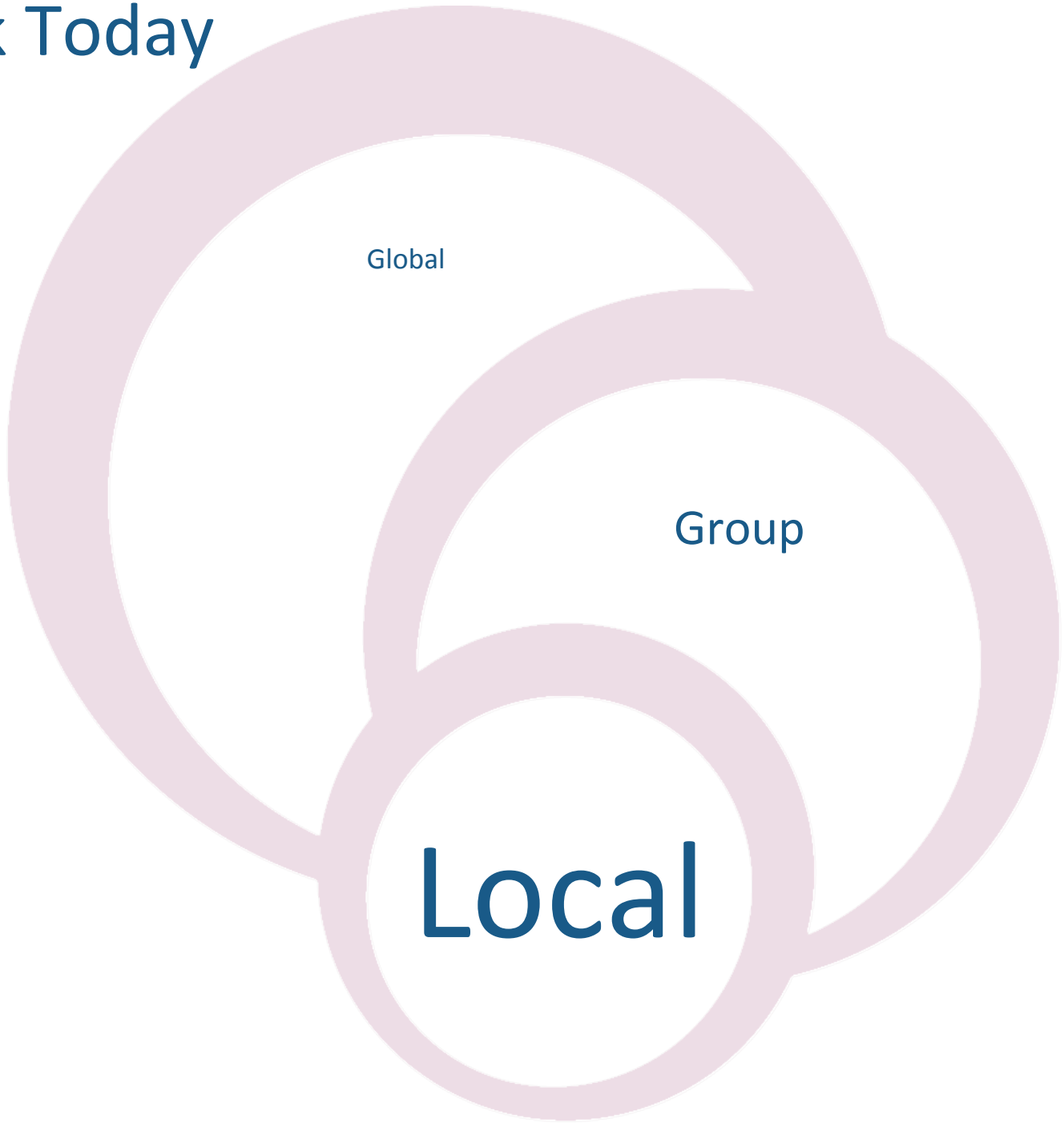
Image credit: http://static.guim.co.uk/sys-images/Guardian/Pix/pictures/2009/3/23/1237806064989/Young-man-reading-a-book-001.jpg

Cataloging

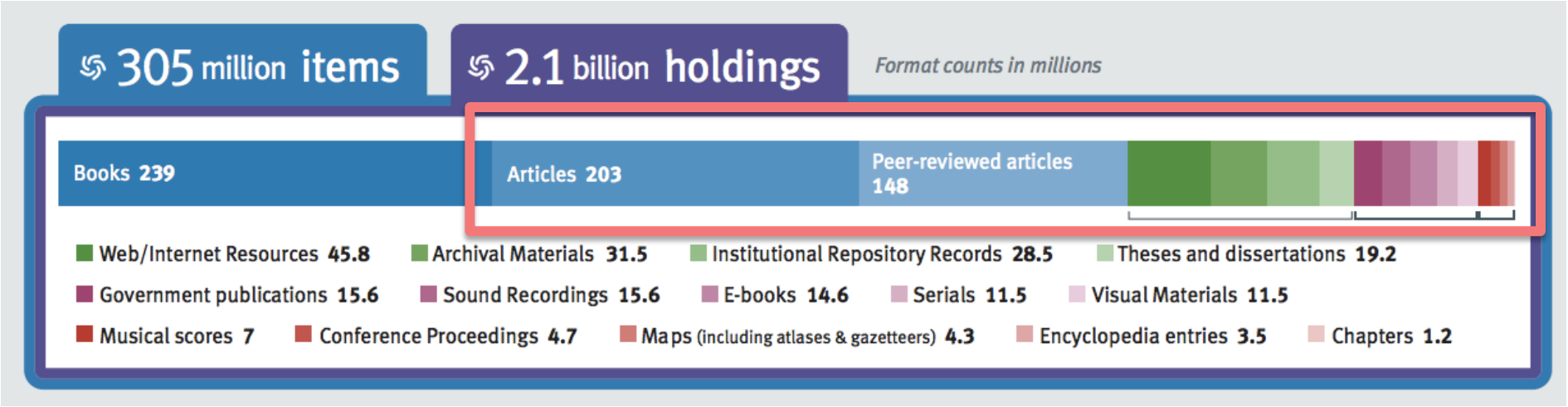
How We Work Today

We catalog:

- Books
- Music
- Journal titles
- Authorities



What is in a Global Discovery System



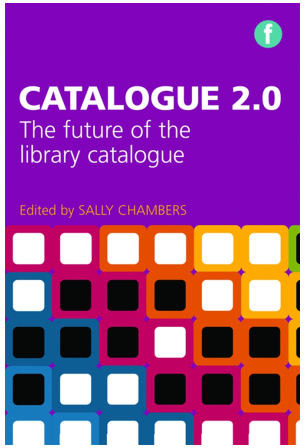
We catalog:

- Books
- Music
- Journal titles
- Authorities

Readers want:

- eBooks
- Articles
- Unique content

Catalogue 2.0 by Karen Calhoun



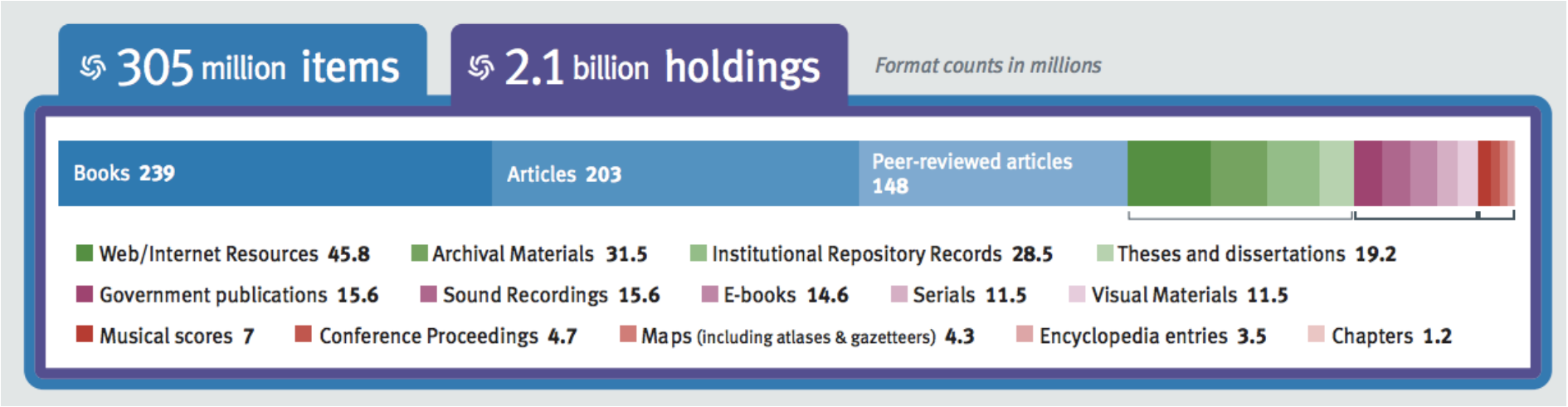
- Calhoun: “Metadata has changed as collections have changed. It remains important, but **it comes in many forms and from many sources**. The centrality of bibliographic control has been disrupted.” P. 15.
- And: “There is less need and place for traditional bibliographic control as a set of methods for providing [metadata] for discovery, access and management of the content of **mainstream books and serials**. “p. 24.

Duplicating records for local purposes

"Ken Chad examines the distinction betweeen **redundant cataloging** (re-editing records to suit local practices) and **redundant catalogs** [in the UK]. He enumerates the benefits of moving from … 160 standalone catalogues to a single shared catalogue at the network level for all of these libraries"

Karen Calhoun in Catalogue 2.0

What is in a Global Discovery System



We catalog:

- Books
- Music
- Journal titles
- Authorities

Readers want:

- eBooks
- Articles
- **Unique content**

The value of authorities

FRAD Tasks

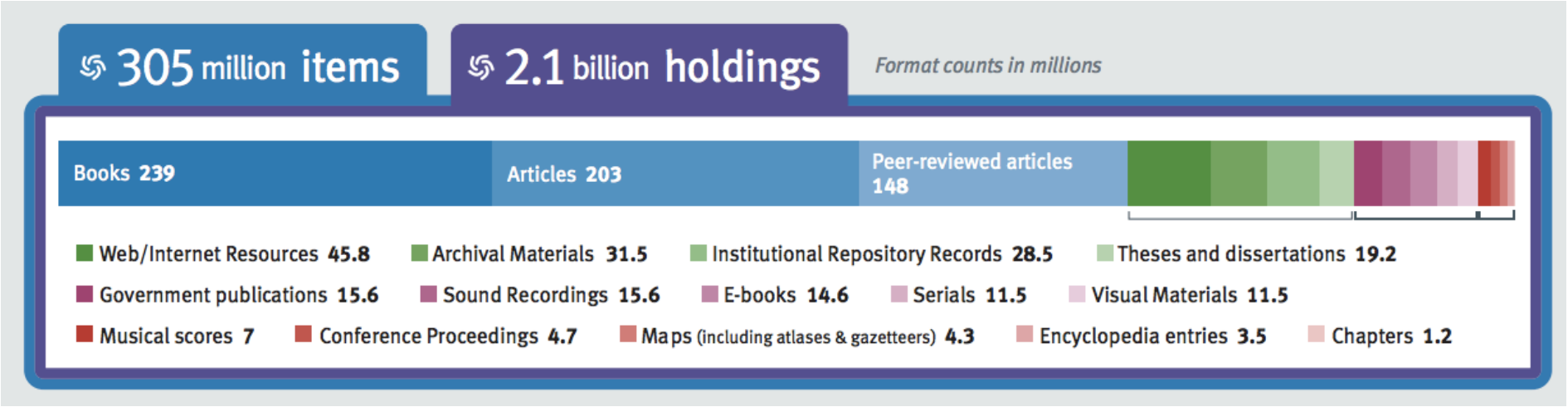
- Find
- Identify
- Clarify
- Contextualize

VIAF
Virtual International Authority File

isni

http://www.ifla.org/publications/functional-requirements-for-authority-data

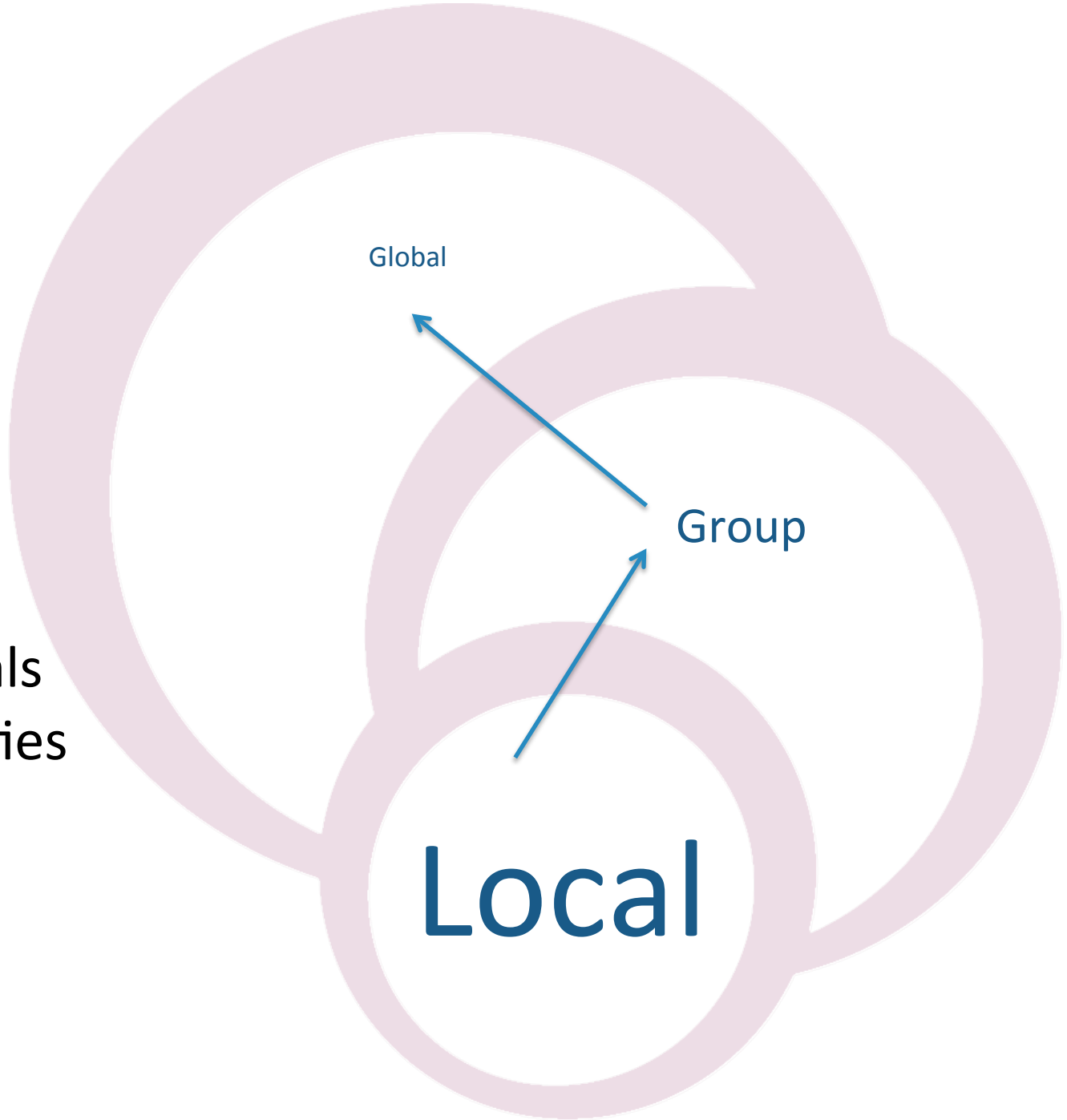
What is in a Global Discovery System



So, what should we do?

1. Catalog unique materials
2. Create authorities
3. Use harvesting and data mining for everything else

Where?

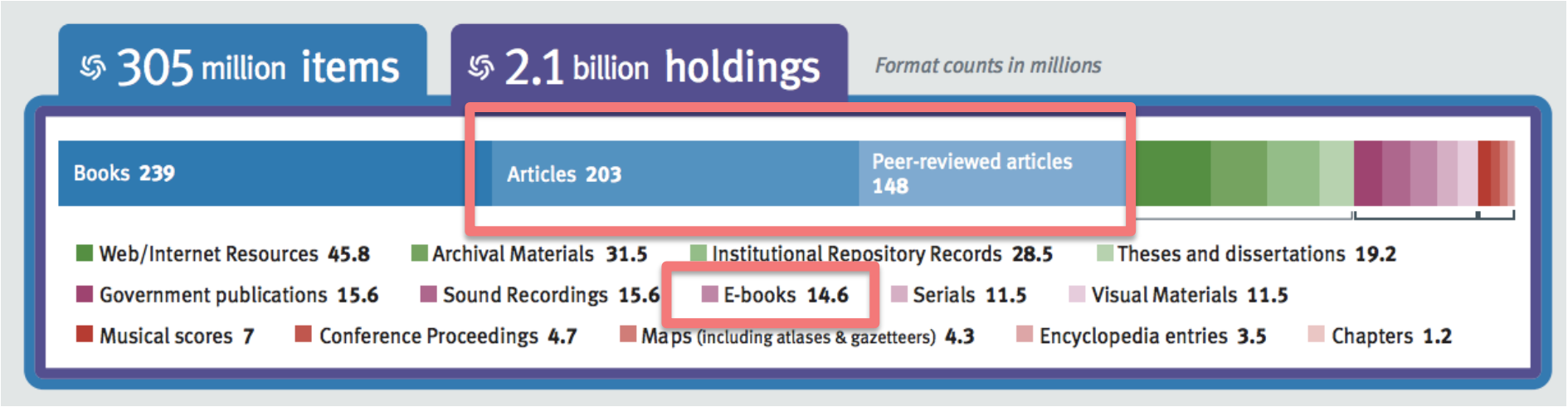


Catalog

- Unique Materials
- Create Authorities

Data Harvesting

What is in a Global Discovery System

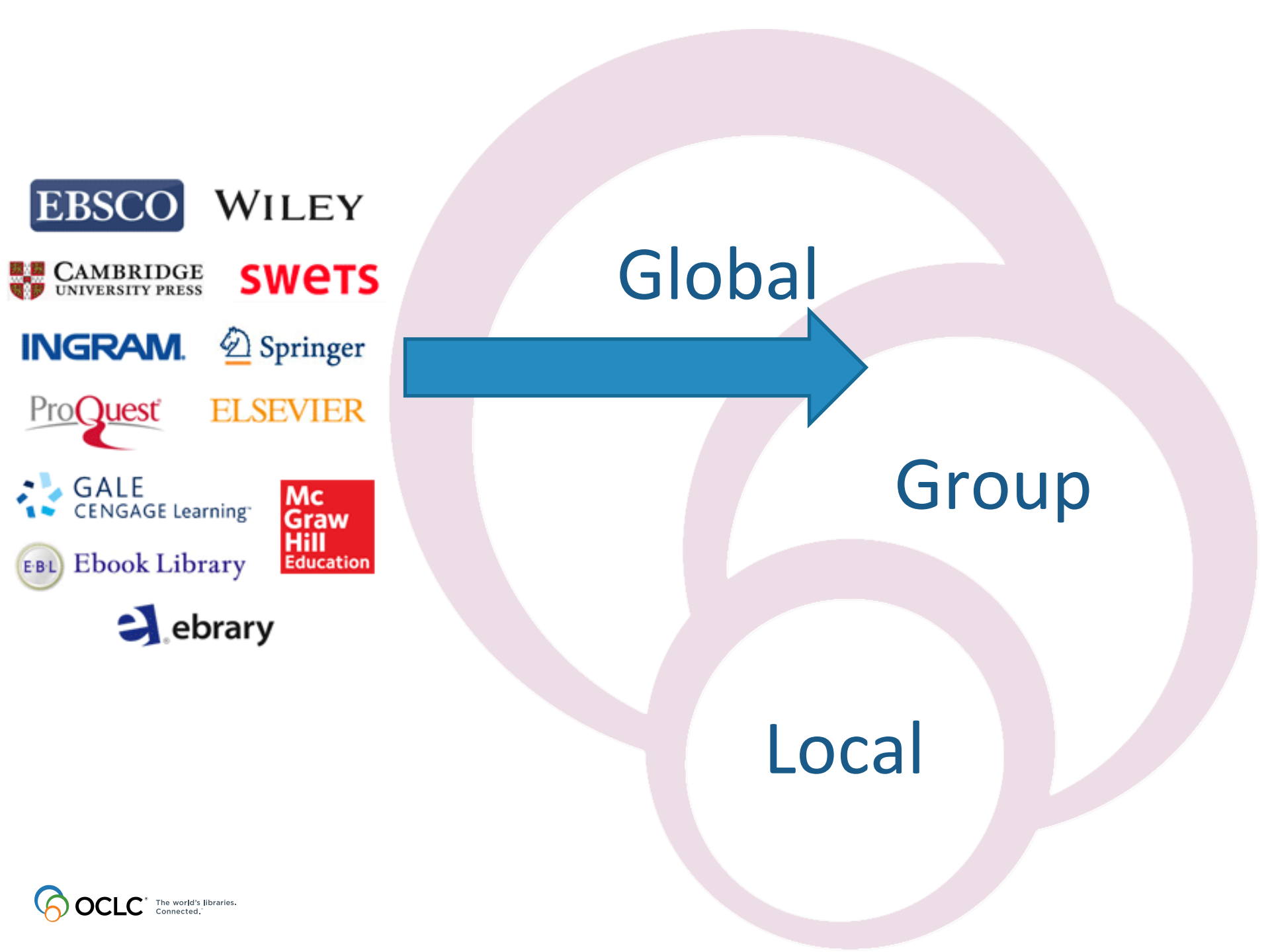


We catalog:

- Books
- Music
- Journal titles
- Authorities

Readers want:

- eBooks
- Articles
- Unique content


EBSCO
WILEY
CAMBRIDGE UNIVERSITY PRESS
swets
INGRAM
Springer
ProQuest
ELSEVIER
GALE CENGAGE Learning
Mc Graw Hill Education
EBL Ebook Library
ebrary
Global
Group
Local

Data Mining & The Web

Global

Group

Local

The Web of …

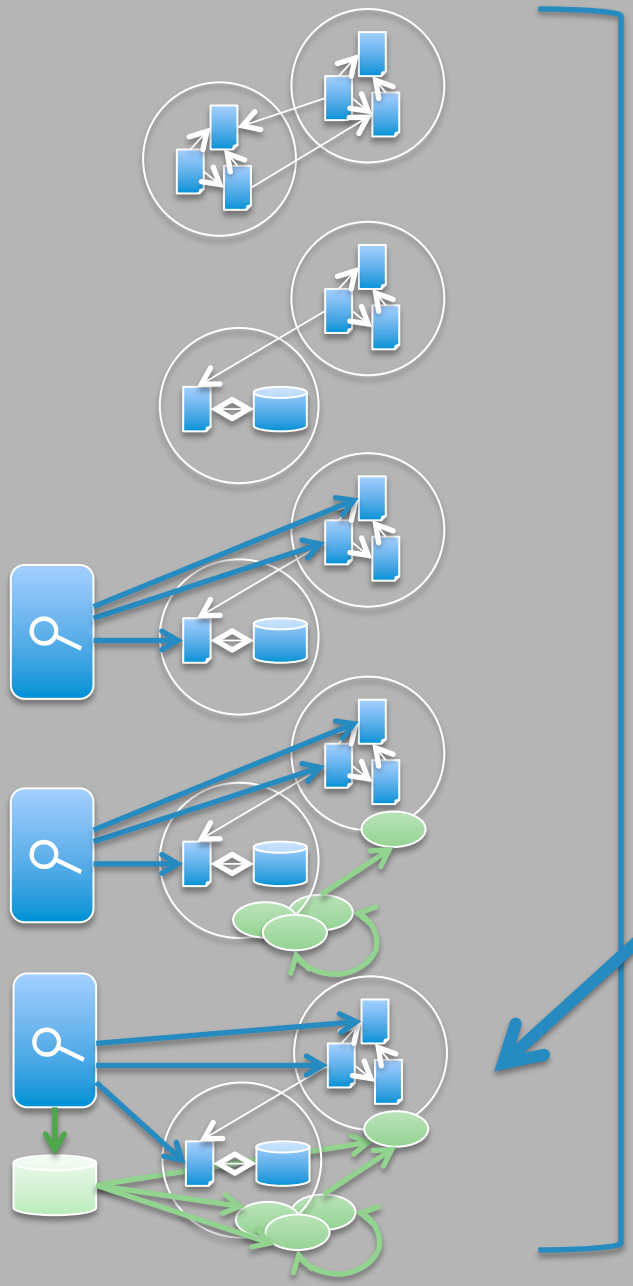
Documents

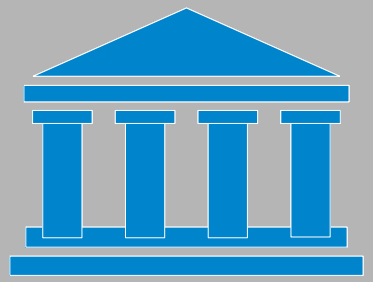
Active Documents

Discovery

Data

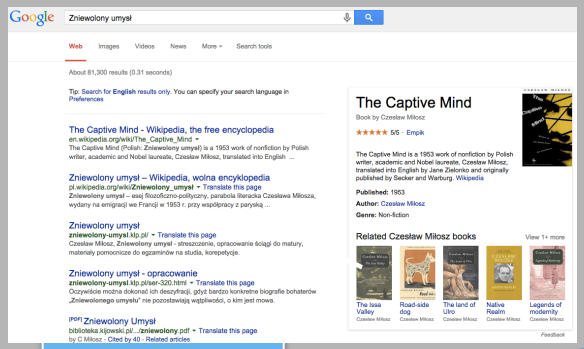
Knowledge

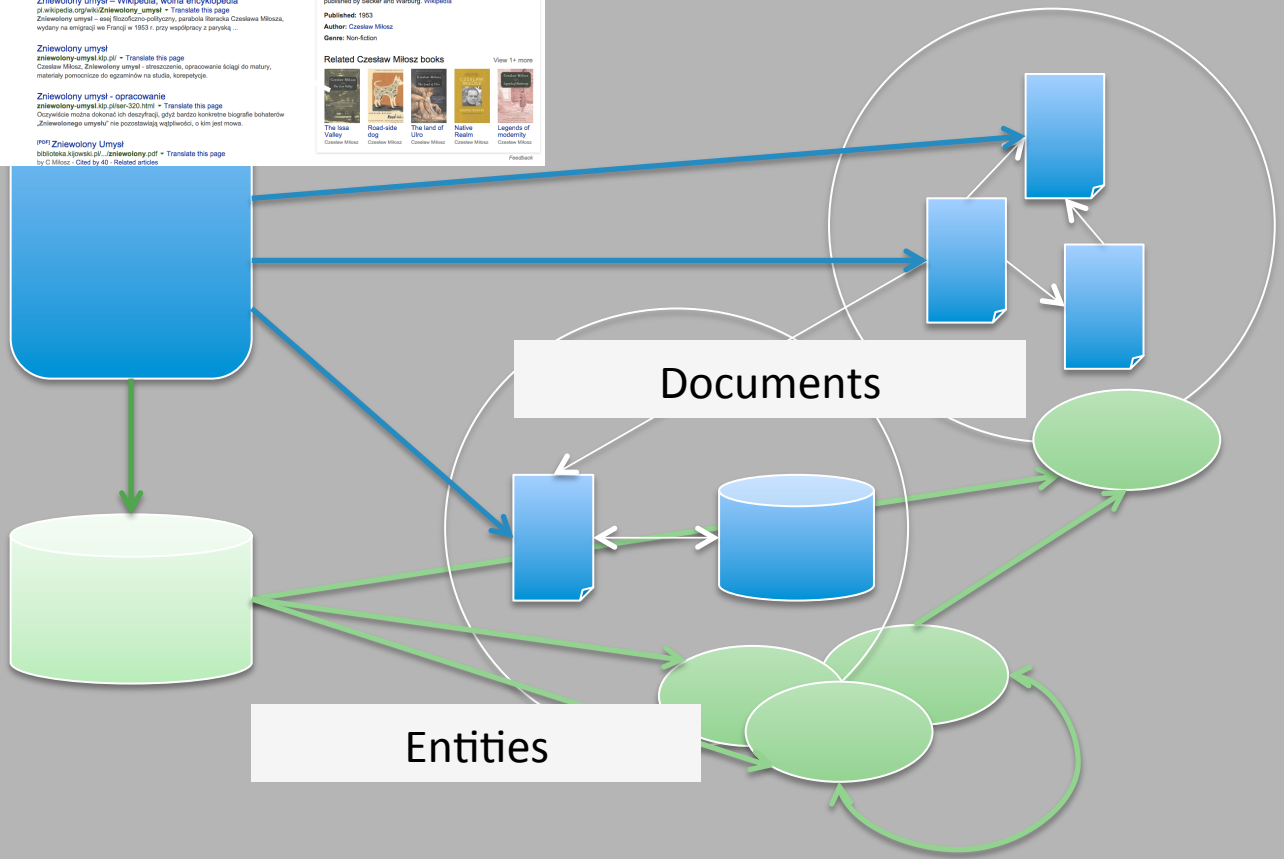


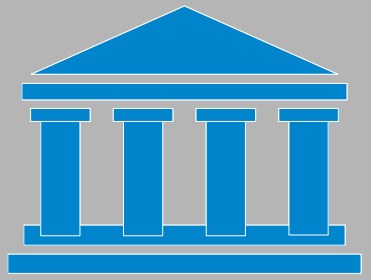


Libraries can connect to the web of knowledge

The Knowledge Graph







Libraries can connect to the web of knowledge

Libraries can create a knowledge graph

Google

Zniewolony umysł

Web Images Videos News More Search tools

About 81,300 results (0.31 seconds)

Tip: Search for **English** results only. You can specify your search language in Preferences

The Captive Mind - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/The_Captive_Mind
The Captive Mind (Polish: **Zniewolony umysł**) is a 1953 work of nonfiction by Polish writer, academic and Nobel laureate, Czesław Miłosz, translated into English ...

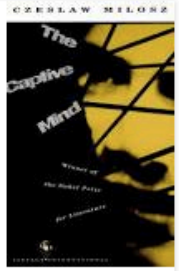
Zniewolony umysł – Wikipedia, wolna encyklopedia
pl.wikipedia.org/wiki/**Zniewolony_umysł** Translate this page
Zniewolony umysł – esej filozoficzno-polityczny, parabola literacka Czesława Miłosza, wydany na emigracji we Francji w 1953 r. przy współpracy z paryską ...

Zniewolony umysł
zniewolony-umysl.klp.pl/ Translate this page
Czesław Miłosz, **Zniewolony umysł** - streszczenie, opracowanie ściągi do matury, materiały pomocnicze do egzaminów na studia, korepetycje.

Zniewolony umysł - opracowanie
zniewolony-umysl.klp.pl/ser-320.html Translate this page
Oczywiście można dokonać ich deszyfracji, gdyż bardzo konkretne biografie bohaterów „**Zniewolonego umysłu**" nie pozostawiają wątpliwości, o kim jest mowa.

[PDF] Zniewolony Umysł
biblioteka.kijowski.pl/.../**zniewolony**.pdf Translate this page
by C Miłosz - Cited by 40 - Related articles

The Captive Mind



Book by Czesław Miłosz

★★★★★ 5/5 · Empik

The Captive Mind is a 1953 work of nonfiction by Polish writer, academic and Nobel laureate, Czesław Miłosz, translated into English by Jane Zielonko and originally published by Secker and Warburg. Wikipedia

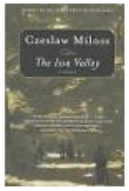
Published: 1953

Author: Czesław Miłosz

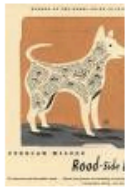
Genre: Non-fiction

Related Czesław Miłosz books

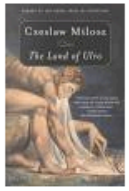
View 1+ more



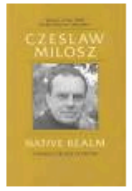
The Issa Valley
Czesław Miłosz



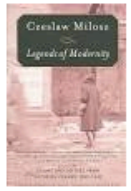
Road-side dog
Czesław Miłosz



The land of Ulro
Czesław Miłosz



Native Realm
Czesław Miłosz



Legends of modernity
Czesław Miłosz

Feedback

Google lazienki palace

Web Images Maps Videos News More Search tools

About 339,000 results (0.46 seconds)

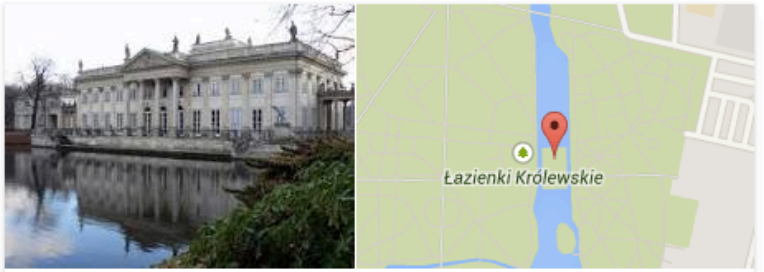
Łazienki Królewskie
www.lazienki-krolewskie.pl/ Translate this page
w Muzeum Łazienki Królewskie w Warszawie ... Muzeum Łazienki Królewskie Profil Facebook dla Ośrodek Edukacja Muzealnej - Muzeum Łazienki Królewskie ...
4.7 ★★★★½ 25 Google reviews · Write a review

Agrykoli 1, 00-460 Warszawa
22 506 00 28
Godziny otwarcia - Kalendarz wydarzeń - Kontakt

Łazienki Palace - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Łazienki_Palace
The Łazienki Palace is a Baroque palace in Warsaw's Royal Baths Park. The Łazienki Palace is located in the largest park in Warsaw that occupies over 76 ...
History - Architecture - Interiors - See also

Images for lazienki palace Report images

More images for lazienki palace



Łazienki Palace

Palace in Warsaw, Poland

Directions Write a review

The Łazienki Palace is a Baroque palace in Warsaw's Royal Baths Park. The Łazienki Palace is located in the largest park in Warsaw that occupies over 76 hectares of the city center. Wikipedia

Address: Agrykoli 1, 00-460 Warszawa

Opened: 1689

Hours: Closing soon · 9:00 am – 6:00 pm

Phone: 22 506 00 28

Architectural style: Neoclassical architecture

Function: Palace

Architect: Domenico Merlini

OCLC The world's libraries. Connected.

Establishing Semantic Identity

For Accurate Representation on the Web

Kenning Arlitsch, Dean of the Library
Patrick OBrien, Semantic Web Research Director

Kenning Arlitsch
Dean of the Library

The Point

Libraries are poorly defined and represented on the Semantic Web…

…but we know how to fix that problem…

…mostly

Google's Perception of MSU Lib - 2012

Google montana state university library SIGN IN

Web Images Maps Shopping More Search tools

About 13,200,000 results (0.16 seconds)

Montana State University Library
www.lib.montana.edu/
Montana State University in Bozeman. Ask A **Librarian** ask a **librarian** . Renne **Library**, 1st floor. Friday, May 17: 10:00-5:00. Saturday, May 18: 10:00-1:00 ...

Articles & Research Databases :: **Montana State University Library**
www.lib.montana.edu/resources/
Articles & Research Databases. Home » Resources. By title. A B C | D E F ...

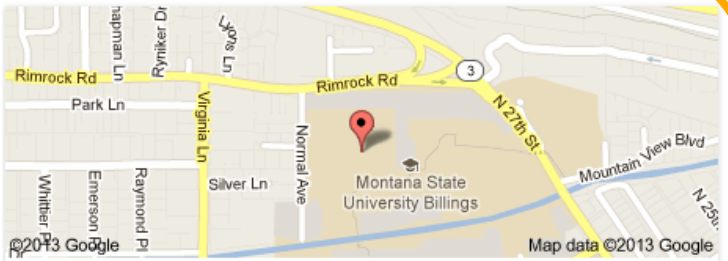
Digital Collections - **Montana State University Library**
www.lib.montana.edu/digital/
MSU Library Digital Initiative collections can be full digital object retrieval or ...

Special Collections & Archives :: **Montana State University Library**
www.lib.montana.edu/archives/
Special Collections & Archives. The Merrill G. Burlingame Special Collections ...

Contact Us :: **Montana State University Library**
www.lib.montana.edu/contact.php
Contact Us. Comments and Feedback. General Questions and Comments for ...

A-Z Site Index :: **Montana State University Library**
www.lib.montana.edu/sitemap.php
A-Z Site Index. Browse by page title: A B C | D E F | G H I | J K L | M N O | P Q R ...

Montana State University (MSU) Library, Mobile
www.lib.montana.edu/m/
MSU Library (Mobile). Search; Databases; Hours; Ask a Librarian; About ...



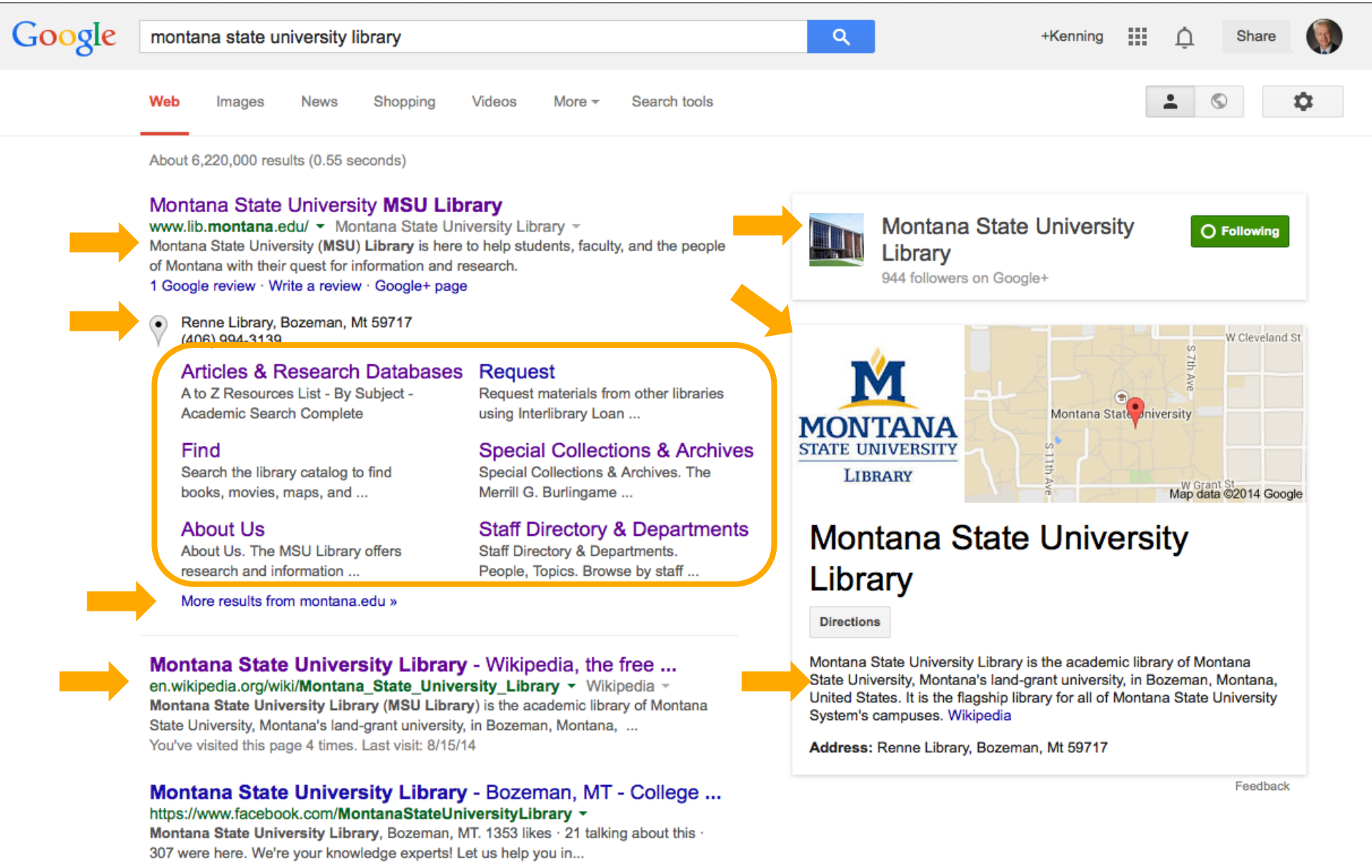
Montana State University Library

Directions Be the first to review

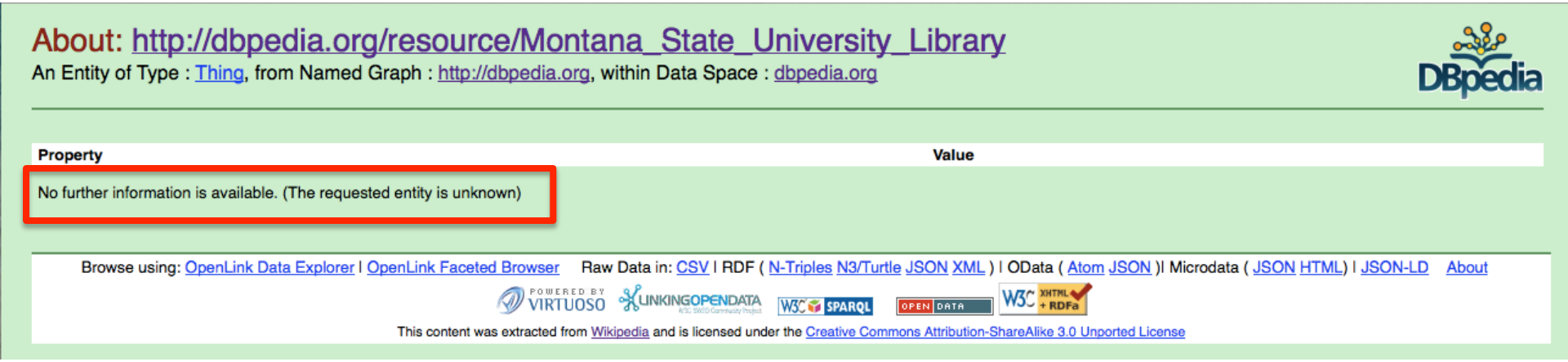
Address: 1500 University Dr, Billings, Mt 59101
Phone: (406) 657-1662
Hours: Wednesday hours 7:30 am–10:00 pm - See all

Feedback

MSU Library - 2014



DBPedia entry - 2012



2014 DBpedia entry

About: Montana State University Library

An Entity of Type : architectural structure, from Named Graph : http://dbpedia.org, within Data Space : dbpedia.org

Montana State University Library (MSU Library) is the academic library of Montana State University, Montana's land-grant university, in Bozeman, Montana, United States. It is the flagship library for all of Montana State University System's campuses. In 1978, the library was named the Roland R. Renne Library to honor the sixth president of the university. The library supports the research and information needs of Montana's students, faculty, and the Montana Extension Service.

Property	Value
dbpedia-owl:abstract	Montana State University Library (MSU Library) is the academic library of Montana State University, Montana's land-grant university, in Bozeman, Montana, United States. It is the flagship library for all of Montana State University System's campuses. In 1978, the library was named the Roland R. Renne Library to honor the sixth president of the university. The library supports the research and information needs of Montana's students, faculty, and the Montana Extension Service.
dbpedia-owl:thumbnail	http://commons.wikimedia.org/wiki/Special:FilePath/Looking_SE_at_Renne_Library_-_Montana_State_University_-_2013-07-09.jpg?width=300
dbpedia-owl:wikiPageExternalLink	http://www.lib.montana.edu/ http://www.lib.montana.edu/digital/ http://arc.lib.montana.edu/brook-0771/ http://arc.lib.montana.edu/msu-extension/
dbpedia-owl:wikiPageID	33974954 (xsd:integer)
dbpedia-owl:wikiPageRevisionID	604080365 (xsd:integer)
dbpprop:budget	7735592.0
dbpprop:director	Kenning Arlitsch
dbpprop:established	January 1894
dbpprop:itemsCollected	Books 523,937; E-Books 230,104; Gov Docs 114,000; Microforms-Audio-Video 2,210,828; Serial Subscription 15,615; Archives linear
dbpprop:libraryName	Montana State University Library
dbpprop:location	Bozeman, Montana, US
dbpprop:numEmployees	52 (xsd:integer)
dbpprop:phoneNum	994 (xsd:integer)
dbpprop:type	dbpedia:Academic_library
dcterms:subject	category:Academic_libraries category:Montana_State_University_–_Bozeman category:University_and_college_academic_libraries_in_the_United_States
grs:point	45.66670283333333 -111.04849322222222
rdf:type	geo:SpatialThing owl:Thing http://schema.org/EducationalOrganization http://schema.org/Library http://schema.org/Organization http://schema.org/Place d0:Location dul:Agent dul:SocialPerson dbpedia-owl:Agent dbpedia-owl:ArchitecturalStructure dbpedia-owl:Building

DBPedia entry - 2012

About: http://dbpedia.org/resource/Montana_State_University_Library

An Entity of Type : Thing, from Named Graph : http://dbpedia.org, within Data Space : dbpedia.org

Property	Value

No further information is available. (The requested entity is unknown)

Browse using: OpenLink Data Explorer | OpenLink Faceted Browser Raw Data in: CSV | RDF (N-Triples N3/Turtle JSON XML) | OData (Atom JSON)| Microdata (JSON HTML) | JSON-LD About

This content was extracted from Wikipedia and is licensed under the Creative Commons Attribution-ShareAlike 3.0 Unported License

About: Montana State University Library

An Entity of Type : architectural structure, from Named Graph : http://dbpedia.org, within Data Space : dbpedia.org

Montana State University Library (MSU Library) is the academic library of Montana State University, Montana's land-grant university, in Bozeman, Montana, United States. It is the flagship library for all of Montana State University System's campuses. In 1978, the library was named the Roland R. Renne Library to honor the sixth president of the university. The library supports the research and information needs of Montana's students, faculty, and the Montana Extension Service.

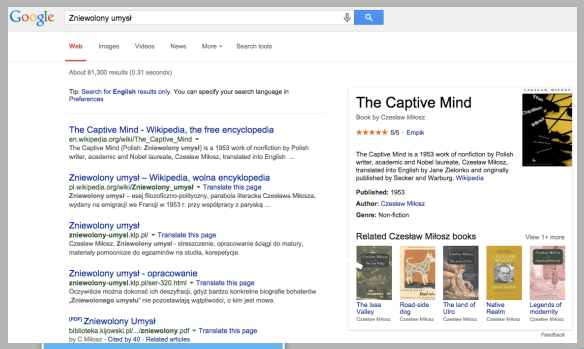
Property	Value
dbpedia-owl:abstract	Montana State University Library (MSU Library) is the academic library of Montana State University, Montana's land-grant university, in Bozeman, Montana, United States. It is the flagship library for all of Montana State University System's campuses. In 1978, the library was named the Roland R. Renne Library to honor the sixth president of the university. The library supports the research and information needs of Montana's students, faculty, and the Montana Extension Service.
dbpedia-owl:thumbnail	http://commons.wikimedia.org/wiki/Special:FilePath/Looking_SE_at_Renne_Library_-_Montana_State_University_-_2013-07-09.jpg?width=300
dbpedia-owl:wikiPageExternalLink	http://www.lib.montana.edu/ http://www.lib.montana.edu/digital/ http://arc.lib.montana.edu/brook-0771/ http://arc.lib.montana.edu/msu-extension/
dbpedia-owl:wikiPageID	33974954 (xsd:integer)
dbpedia-owl:wikiPageRevisionID	604080365 (xsd:integer)
dbpprop:budget	7735592.0
dbpprop:director	Kenning Arlitsch
dbpprop:established	January 1894
dbpprop:itemsCollected	Books 523,937; E-Books 230,104; Gov Docs 114,000; Microforms-Audio-Video 2,210,828; Serial Subscription 15,615; Archives linear
dbpprop:libraryName	Montana State University Library
dbpprop:location	Bozeman, Montana, US
dbpprop:numEmployees	52 (xsd:integer)
dbpprop:phoneNum	994 (xsd:integer)
dbpprop:type	dbpedia:Academic_library
dcterms:subject	category:Academic_libraries category:Montana_State_University_–_Bozeman category:University_and_college_academic_libraries_in_the_United_States
grs:point	45.66670283333333 -111.04849322222222
rdf:type	geo:SpatialThing owl:Thing http://schema.org/EducationalOrganization http://schema.org/Library http://schema.org/Organization http://schema.org/Place d0:Location dul:Agent dul:SocialPerson dbpedia-owl:Agent dbpedia-owl:ArchitecturalStructure dbpedia-owl:Building

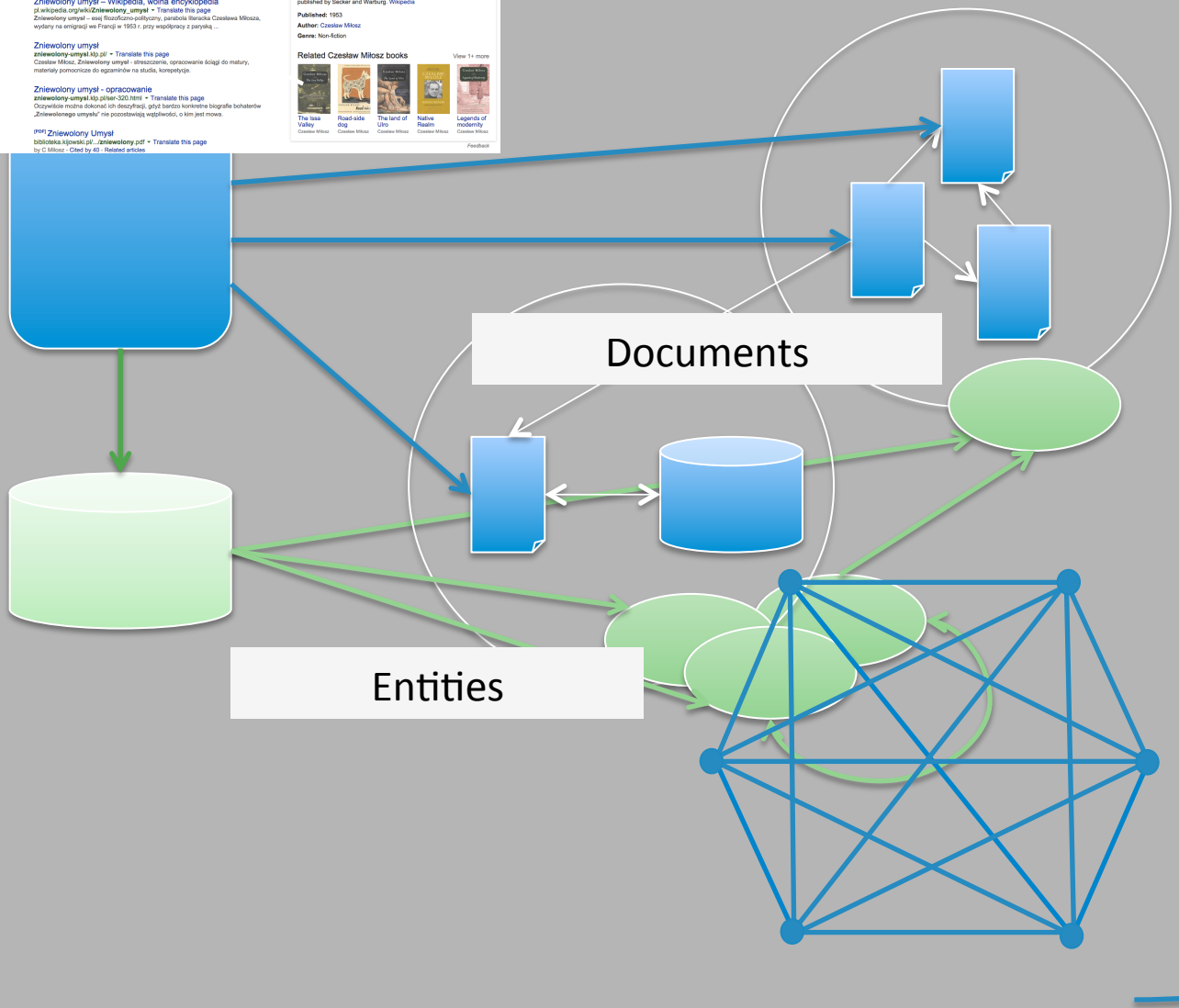
2014 Dbpedia entry

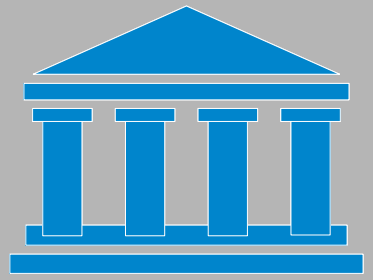
Summary

- Define library organization in Wikipedia
  - Beware of *pedia culture and process
- Engage with other trusted data sources
  - FreeBase
  - Google Places/Google My Business
  - Google+
- Mark-up metadata with Schema.org

The Knowledge Graph

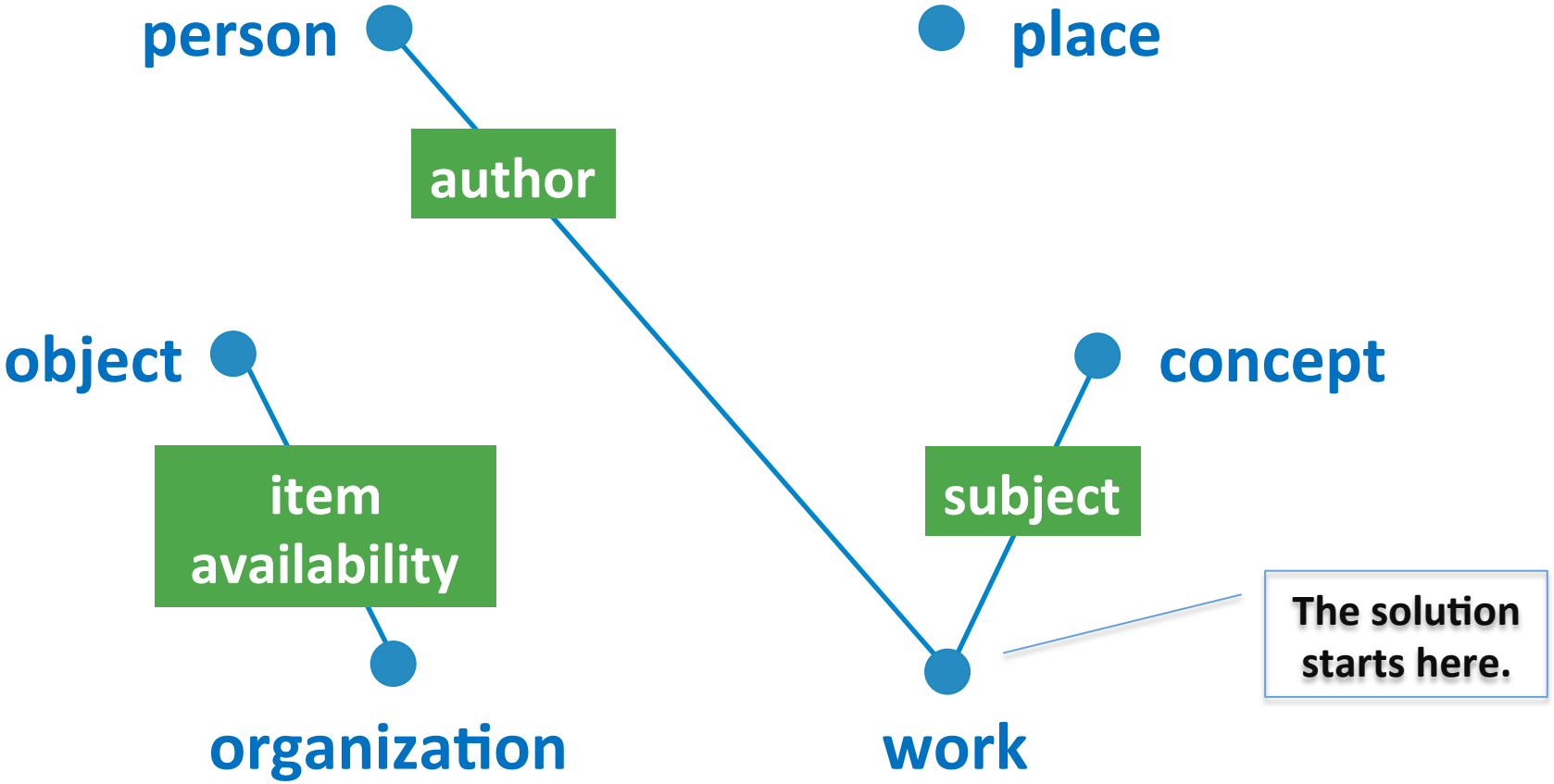




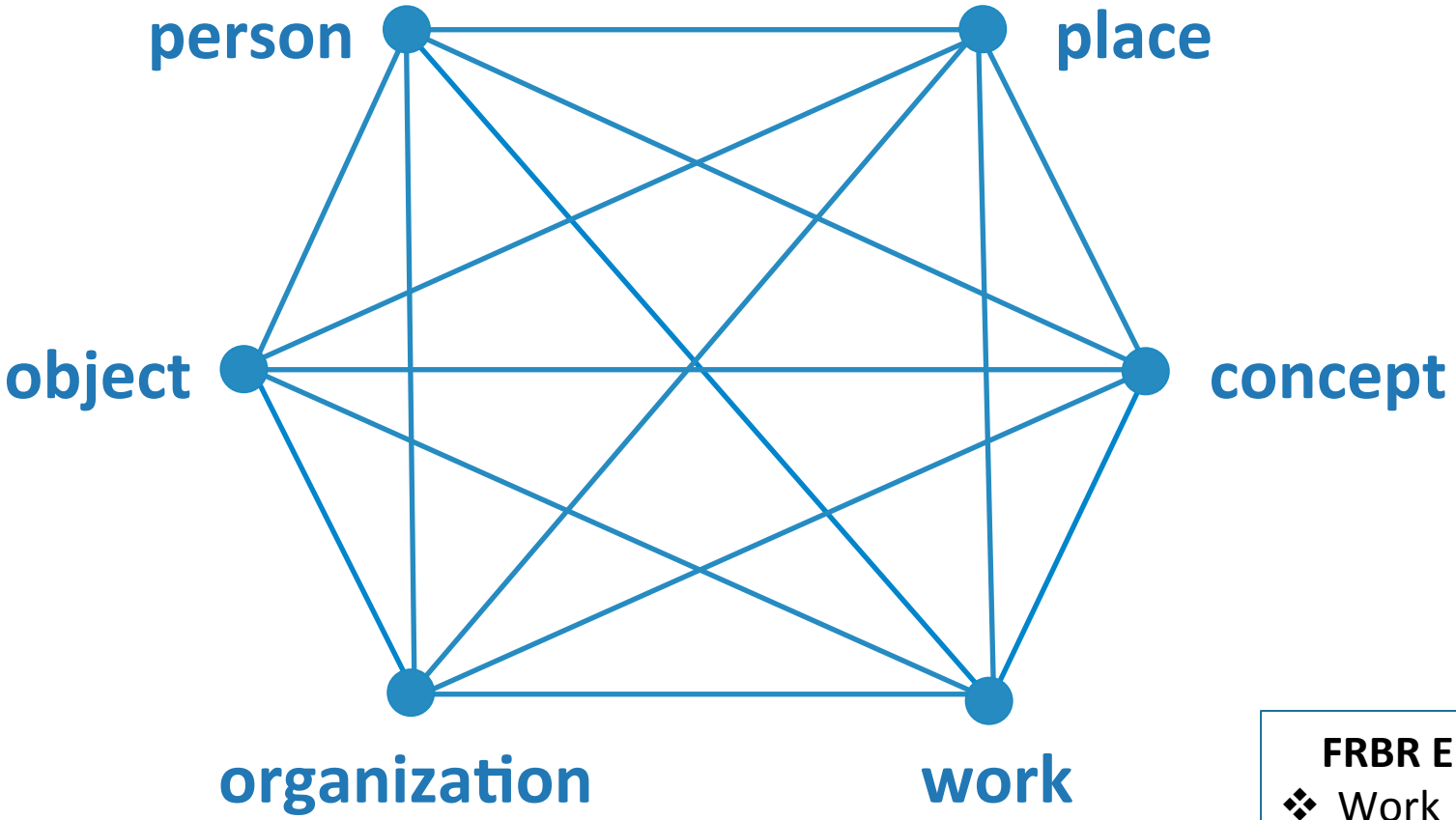


Libraries can connect to the web of knowledge

Libraries can create a knowledge graph



The library knowledge graph



FRBR Entities

- Work
- Expression
- Manifestation
- Item

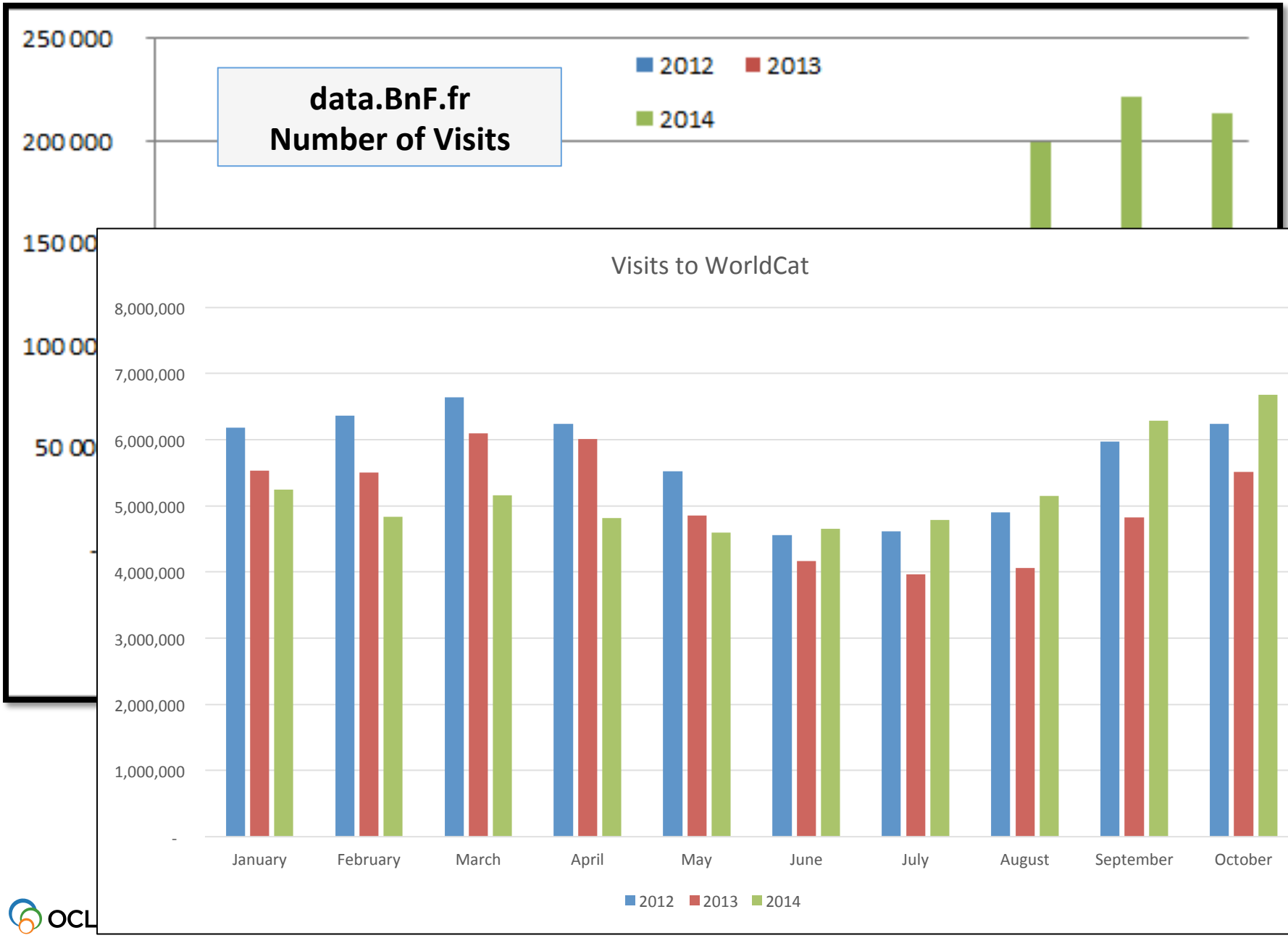
The library knowledge graph

http://www.ifla.org/publications/functional-requirements-for-bibliographic-records

Example of benefits…

Discovery

Example of Benefits: Web Exposure



How does data mining work?

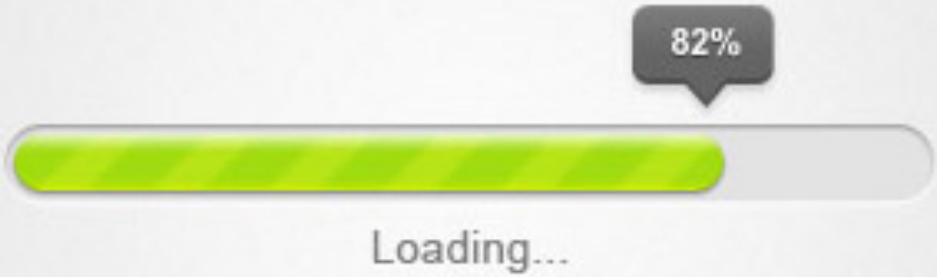
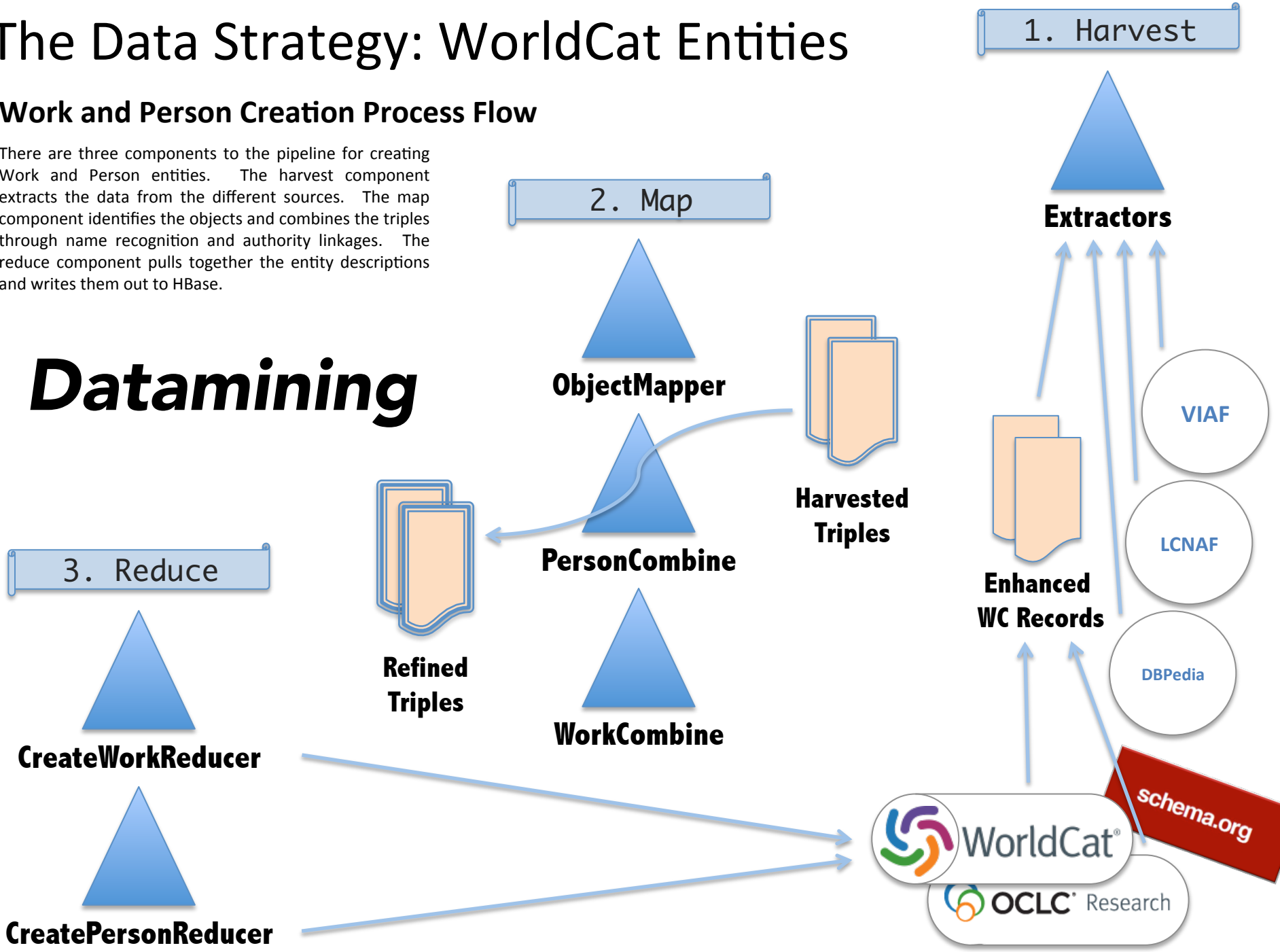


Photo credit: http://media02.hongkiat.com/freebies-for-web-designers-2011/progress-bar.jpg

The Data Strategy: WorldCat Entities

Work and Person Creation Process Flow

There are three components to the pipeline for creating Work and Person entities. The harvest component extracts the data from the different sources. The map component identifies the objects and combines the triples through name recognition and authority linkages. The reduce component pulls together the entity descriptions and writes them out to HBase.



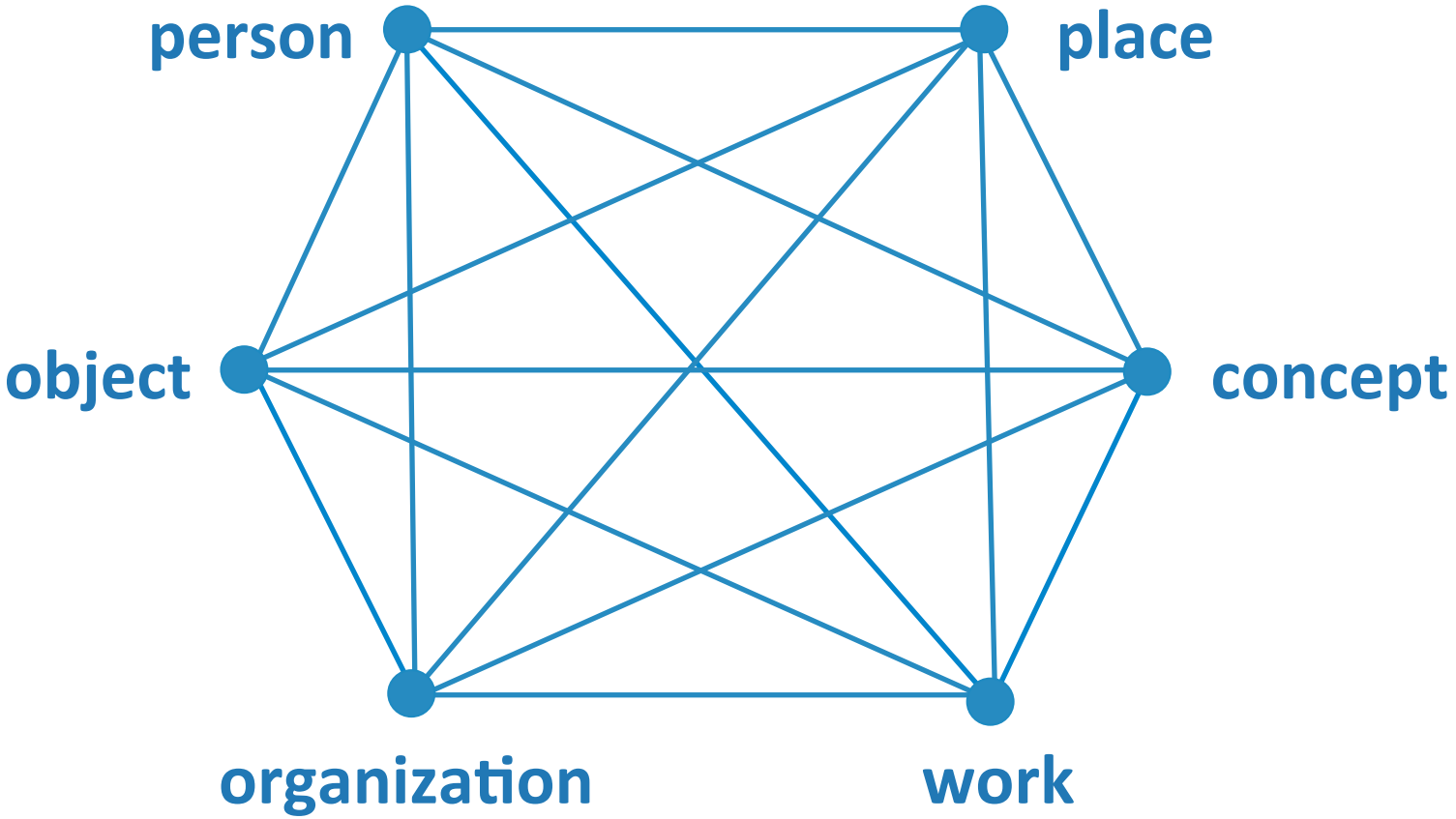
Datamining

The Work Entity

- **197+ million Work descriptions and URIs**
- Schema.org + BiblioGraph.net
- RDF Data formats
  - RDF/XML, Turtle, Triples, JSON-LD
- Links to WorldCat manifestations
- Links to Dewey, LCSH, LCNAF, VIAF, FAST
- Open Data license via Linked Data Explorer
  - 2015: Discovery API, Metadata API
- Released April 2014

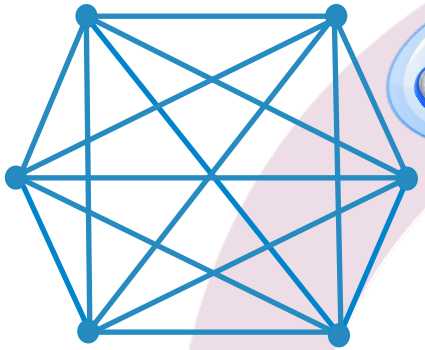
The Person Entity

- **98+ million Person descriptions and URIs**
  - **Person entities with authority: 20.2 million**
  - **Person entities without authority: 78.3 million**
- Schema.org + BiblioGraph.net
- Harvested from WorldCat data and enriched from other hubs RDF Data formats
  - RDF/XML, Turtle, Triples, JSON-LD
- Links to WorldCat Works. Added links from WC Works.
- Open Data license via Linked Data Explorer
  - 2015: Linked Data Explorer, Discovery API



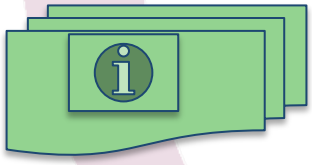
The library knowledge graph

Datamining



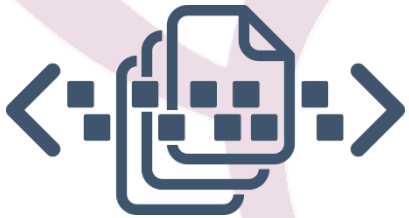
Global

Harvesting



Group

Cataloging

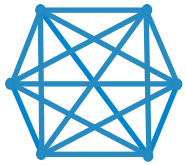


Local

The Future

So, what should we do?

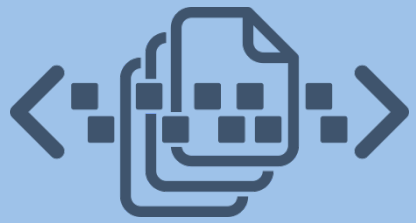
1. Catalog unique materials
2. Create authorities
3. Use harvesting and data mining for



everything else

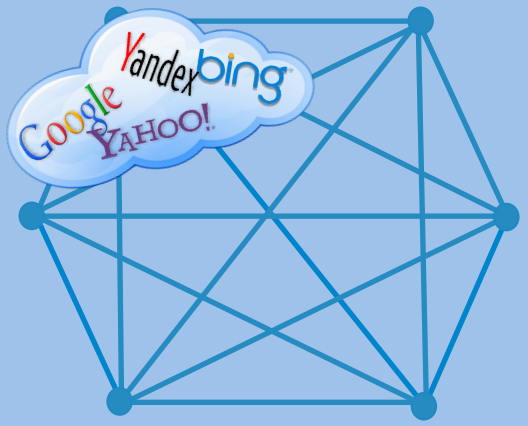
The Future

Cataloging



Harvesting

Datamining



Ted Fons
Executive Director, Data Services & WorldCat Quality
fonst@oclc.org

Discussion