

Metadata is a plural noun

Karen Coyle
NETSL, 2009

There is little question but that library catalog data is metadata. However, we tend to view our data in a rather narrow way, treating each record as a single unit on its own, and assuming that the record will live forever and only inside the library catalog. I want to challenge those concepts, because there is greater value in library metadata than is currently being served, and we are missing some very important opportunities.

Metadata

data about data

First I want to describe what metadata means today.

Metadata

constructed

constructive

actionable



metadata

real planet



longitude and latitude

Longitude and latitude are metadata about the planet. The planet doesn't really have lines about around it that define spaces.



“turn left at...”

navigation

You need longitude and latitude because there no sign posts, no markers, no point of reference.



red line, green line

subway navigation

A different kind of map, with a different purpose. Not finding your way on the planet, but in a subway system.



A notable aspect is that it does NOT represent the real area it covers: when compared to a map of the area, it doesn't match. But that's not its purpose. Its purpose is to get you around, mainly underground, following clues like colors and the ending point of the transit lines.

USATODAY.com

File Edit View Navigation Bookmarks Mail Window Help

http://www.usatoday.com/sports/baseball/playoffs/bos03pl.htm

baseball statistics 30%

Weather

Search

powered by Google

Inside Fantasy

[Fantasybaseball home](#)

[Playernotes](#)

[Playerprofiles](#)

[Playestatistics](#)

[Projections](#)

[Probablepitchers](#)

[Teamprofiles](#)

[Teamstatistics](#)

[Standings](#)

[Schedules](#)

[Adviceforum](#)

[Baseballnews](#)

Premium Products

[Fantasygames](#)

[LeagueManage](#)

[HotSheef](#)

Free Games

[MarchMania](#)

[Pigskin Pool](#)

More Tools

[Scores](#)

[ALInjuries](#)

[NLInjuries](#)

[Transactions](#)


[Gamematchups](#)

[Sheridan'sodds](#)

[Liveodds](#)

[Sagarinratings](#)

[Salarydatabase](#)



Boston Red Sox

[Statistics](#) • [Schedule/results](#) • [Roster](#) • [Depth chart](#)

More Stats

[Individual stats](#) • [Day/night](#) • [Grass/turf](#) • [Home/road](#) • [Runners on/bases empty](#)
[Left/righty](#) • [Late inning pressure](#)

BATTERS	BA	SLG	OBA	G	AB	R	H	TB	2B	3B	HR	RBI	BB	SO	SB	CS	E
Mirabelli	.364	.455	.364	5	11	2	4	5	1	0	0	0	0	4	0	0	0
Walker	.349	.767	.391	12	43	9	15	33	1	1	5	6	1	3	0	0	2
Nixon	.294	.676	.385	11	34	4	10	23	1	0	4	7	4	10	1	2	0
Varitek	.294	.706	.351	11	34	8	10	24	2	0	4	5	3	7	0	0	0
Garcia/parra	.265	.327	.333	12	49	4	13	16	1	1	0	1	5	10	1	0	2
Ramirez	.265	.469	.321	12	49	8	13	23	1	0	3	7	4	11	0	1	0
Damon	.256	.410	.356	10	39	3	10	16	3	0	1	4	5	4	3	0	0
Millar	.240	.300	.283	12	50	3	12	15	0	0	1	3	3	13	0	0	1
Ortiz	.191	.362	.283	12	47	4	9	17	2	0	2	8	5	15	0	0	0
Mueller	.174	.239	.255	12	46	1	8	11	3	0	0	0	5	11	0	0	0
Jackson	.125	.125	.125	9	8	0	1	1	0	0	0	1	0	3	0	1	1
Kapler	.059	.059	.059	7	17	0	1	1	0	0	0	0	0	6	0	1	0
Brown	.000	.000	.000	4	2	0	0	0	0	0	0	0	0	1	0	0	0
McCarty	.000	.000	.000	2	1	0	0	0	0	0	0	0	0	1	0	0	0

PITCHERS	W- L	ERA	G	GS	CG	GF	SH	SV	IP	H	R	ER	HR	BB	SO
Embree	1- 0	0.00	8	0	0	1	0	0	6.2	4	0	0	0	0	1
Jones	0- 0	0.00	1	0	0	0	0	0	0.1	1	0	0	0	1	1
Sauerbeck	0- 0	0.00	1	0	0	0	0	0	0.1	1	0	0	0	1	0
Timlin	0- 0	0.00	8	0	0	0	0	0	9.2	1	0	0	0	2	11
Williamson	2- 0	1.13	8	0	0	6	0	3	8.0	3	1	1	1	3	14
Arroyo	0- 0	2.70	3	0	0	2	0	0	3.1	2	1	1	1	2	5
Wakefield	2- 2	2.91	5	3	0	1	0	0	21.2	14	9	7	1	9	17
Lowe	0- 3	4.18	5	3	0	2	0	1	23.2	21	12	11	1	14	11
P. Martinez	1- 1	4.76	4	4	0	0	0	0	28.1	29	15	15	3	7	23
Burkett	0- 0	7.00	2	2	0	0	0	0	9.0	16	9	7	2	2	2
Kim	0- 0	13.50	1	0	0	0	0	0	0.2	0	1	1	0	1	1

This is metadata about baseball. It represents an entire season, something that doesn't really have an analog equivalent. It also represents the season as something very different from the actual game of baseball: using statistics. It allows you to compare players, and compare across teams and seasons. (For some people, this is more interesting than the game itself.)

metaDATA

Emphasis on DATA in metadata

Means:

rigorous

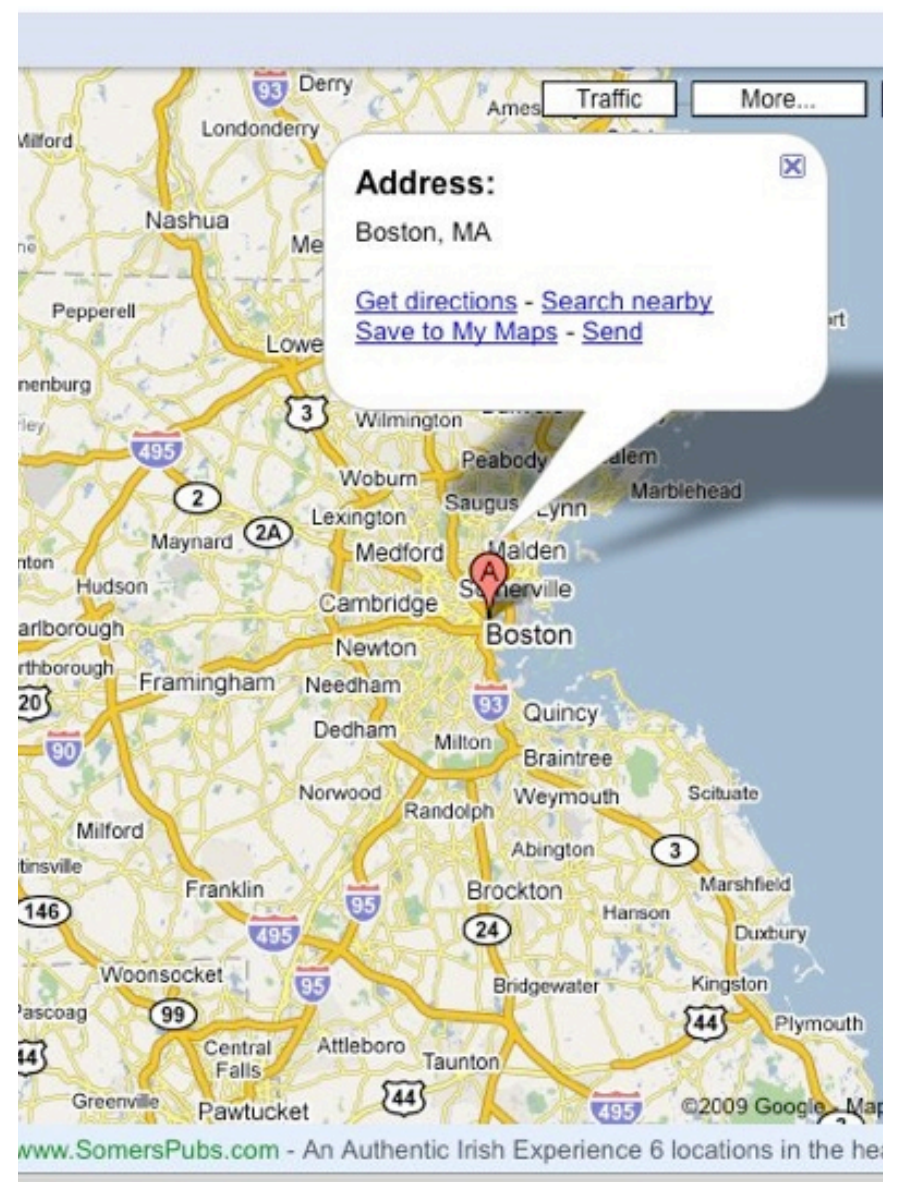
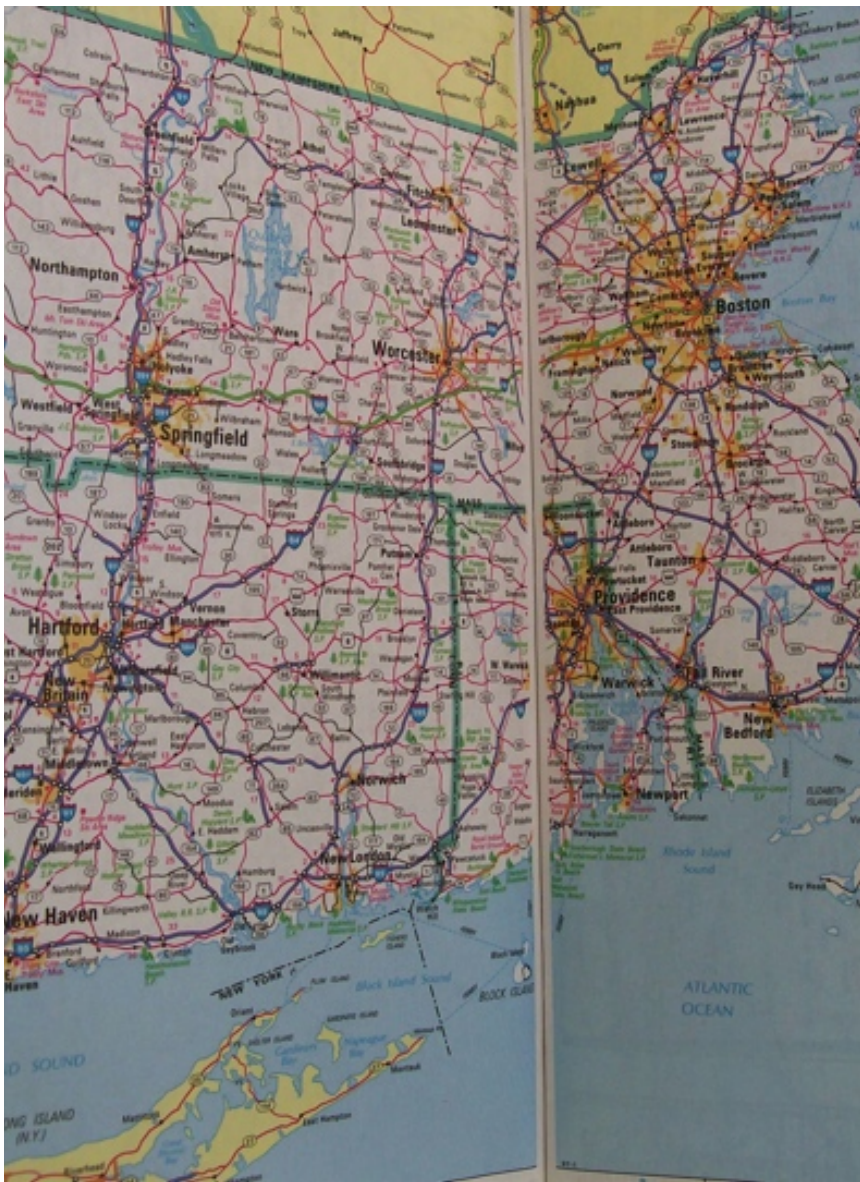
actionable

used on a computer

on a network (almost always)

Internet (if publicly available, as opposed to enterprise networks)

WWW is primary delivery/interaction application



Actionable

Here's what we mean by actionable. The map at the left is a paper map. I can use it to drive from Boston to Worcester. But there's a lot that I can't do with it. On the right we have an example from Google maps. Because it is represented as data, there are many things that you can do. You can get directions; you can move the map around, you can search the map area for restaurants, stores, airports. You can save it or email it.

Hotels in Boston

www.Expedia.com Save up to 50% on **Hotels** at Expedia And Now Earn ThankYou Points Too!

Sponsored Links

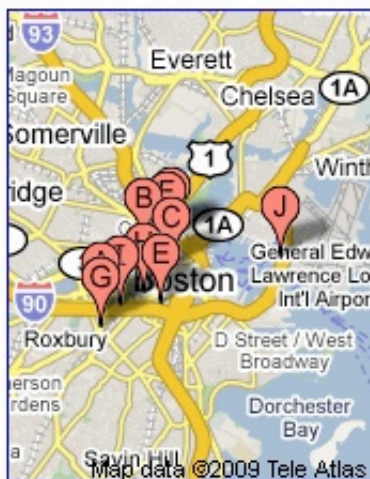
Boston Hotels

www.hotels.com/Boston Top **Boston Hotels** at Great Rates **Boston Hotels** Experts 800-449-4167

Boston Discount Hotels

LastMinuteTravel.com/Boston Hit the town for less. **Boston** From \$77. Book now!

Local business results for **hotels** near **Boston, MA**



- A. [Sheraton Boston Hotel](http://www.starwoodhotels.com) - www.starwoodhotels.com - (617) 236-2000 - [654 reviews](#)
- B. [The Liberty Hotel Boston](http://www.libertyhotel.com) - www.libertyhotel.com - (617) 224-4000 - [192 reviews](#)
- C. [Nine Zero Hotel in Downtown Boston](http://www.ninezero.com) - www.ninezero.com - (617) 772-5800 - [381 reviews](#)
- D. [Onyx Hotel](http://www.onyxhotel.com) - www.onyxhotel.com - (617) 557-9955 - [296 reviews](#)
- E. [Doubletree Hotel Boston Downtown](http://www.hiltonfamilyboston.com) - www.hiltonfamilyboston.com - (617) 956-7900 - [445 reviews](#)
- F. [Bulfinch Hotel](http://www.bulfinchhotel.com) - www.bulfinchhotel.com - (617) 624-0202 - [285 reviews](#)
- G. [Midtown Hotel](http://www.midtownhotel.com) - www.midtownhotel.com - (617) 262-1000 - [228 reviews](#)
- H. [The Boston Park Plaza Hotel & Towers](http://bostonparkplaza.com) - bostonparkplaza.com - (617) 426-2000 - [929 reviews](#)
- I. [Westin Hotel Copley Place Boston](http://www.starwoodhotels.com) - www.starwoodhotels.com - (617) 262-9600 - [447 reviews](#)
- J. [Hyatt Harborside at Boston's Logan International Airport](http://harborside.hyatt.com) - harborside.hyatt.com - (617) 568-1234 - [481 reviews](#)

[More results near Boston, MA »](#)

Interact(ionable)

And you can include it in other applications, in this case a list of hotels in Boston. Once you have the map as actionable software, there are many things you can do with it.



◀ almost any bike and a **Cardo scala-rider®** wireless headset for high quality voice instructions inside ▶

where you need it

This is where we are headed, IMO. While the web today is interactive and has many useful services, the next step is to take those services to the user wherever the user is; not requiring the user to be at a computer. This is the model for all of the services being developed (and that have been developed) for telephones. So you can walk down a street and find nearby restaurants, or locate a post box or a public restroom.

metadata are for use

with apologies to Ranganathan

and the more that metadata is used, the more value it creates, just as books create more value the more they are used.

from which a man, presumably the librarian, is selecting one. The ends of the rolls are furnished with tickets.

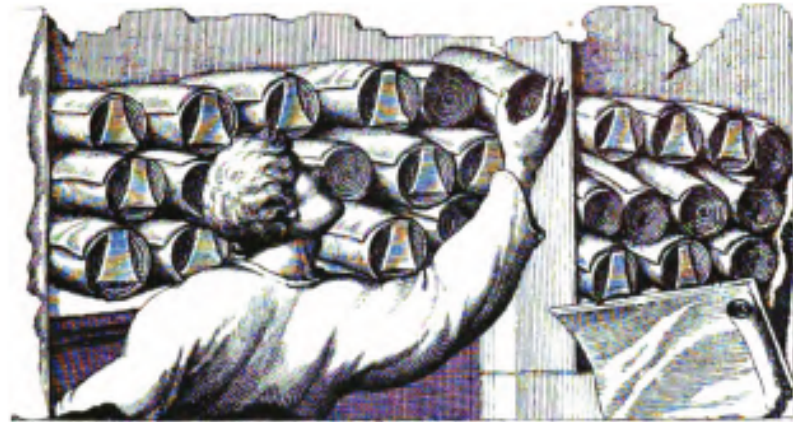


Fig. 11. A Roman taking down a roll from its place in a library.

The system of pigeon-holes terminated, in all probability, in a cornice. The explorers of Herculaneum depose to the discovery of such an ornament there.

John Willis Clare. The care of books. 1901

library metadata

Taking a look at library metadata. Has existed for hundreds, no, thousands of years. There has been library metadata since the days when the sun revolved around a flat earth. Tags on scrolls, because you couldn't see inside the scrolls to locate them.



card catalog was a great leap in technology. it made modern library catalog metadata possible.

Is it constructed? Yes! AACR/1&2

What is its purpose? Jewett & Cutter expressed this best: to help a user find a book owned by the library; by author, title or subject.

Folio Soderberg, Percy Measday, 1901-
636.6865 All about lovebirds / P. M.
Sod Soderberg. -- Hong Kong ; Neptune City,
N.J. : T.F.H. Publications, c1977.
96 p. : ill. (some col.) ; 21 cm.
Includes index.
ISBN 0-87666-957-7

1. Lovebirds. I. Title

OCoLC 19981124 #5370389 OCLDsl

typed, then printed

From: OCLC../cataloging/cards/

Other technology allowed the card to become a very efficient way of sharing our metadata creation duties.

card as the first ALA standard. LC card product beginning in 1902.

LC Control No.: 84100005

LCCN Permalink: <http://lccn.loc.gov/84100005>

000 00745nam a2200229 a 450
001 954379
005 19840215000000.0
008 840103s1983 njua b 001 0 eng
035 __ |9 (DLC) 84100005
906 __ |a 7 |b cbc |c orignew |d 2 |e ncip |f 19 |g y-gencatlg
010 __ |a 84100005
020 __ |a 0876669437
040 __ |a DLC |c DLC |d DLC
050 00 |a SF473.L6 |b R63 1983
082 00 |a 636.6/865 |2 19
100 1_ |a Roberts, Mervin F.
245 10 |a All about breeding lovebirds / |c Mervin F. Roberts.
260 __ |a Neptune, N.J. : |b T.F.H. Publications, |c c1983.
300 __ |a 96 p. : |b ill. (some col.) ; |c 21 cm.
504 __ |a Includes bibliographical references and index.
650 _0 |a Lovebirds |x Breeding.
991 __ |b c-GenColl |h SF473.L6 |i R63 1983 |t Copy 1 |w BOOKS

Machine-readable

catalog.loc.gov

MARC, mid-1960's. Designed for the printing of cards. machine-readable but still mainly textual. still mainly aimed at human readers of the information. machine-oriented information separate, and not easy to input, and therefore often not taken advantage of (fixed fields). but still very valuable.

Relevance: 🌟🌟🌟🌟

LC Control No.: 84100005

LCCN Permalink: <http://lccn.loc.gov/84100005>

Type of Material: Book (Print, Microform, Electronic, etc.)

Personal Name: [Roberts, Mervin F.](#)

Main Title: All about breeding lovebirds / Mervin F. Roberts.

Published/Created: Neptune, N.J. : T.F.H. Publications, c1983.

Description: 96 p. : ill. (some col.) ; 21 cm.

ISBN: 0876669437

Notes: Includes bibliographical references and index.

Subjects: [Lovebirds --Breeding.](#)

LC Classification: SF473.L6 R63 1983

Dewey Class No.: 636.6/865 19

catalog.loc.gov

A coded document, mainly designed for display to humans. Still the same document we've been displaying in catalogs for over 150 years. And only designed to be used inside a library catalog. No thought to how it might play with other day, such a Google maps, or Wikipedia.

LC Control No.: 84100005
LCCN Permalink: <http://lccn.loc.gov/84100005>
 000 00745nam a2200229 a 450
 001 954379
 005 19840215000000.0
 008 840103s1983 njua b 001 0 eng
 035 __ |9 (DLC) 84100005
 906 __ |a 7 |b cbc |c orignew |d 2 |e ncip |f 19 |g y-gencatlg
 010 __ |a 84100005
 020 __ |a 0876669437
 040 __ |a DLC |c DLC |d DLC
 050 00 |a SF473.L6 |b R63 1983
 082 00 |a 636.6/865 |2 19
 100 1_ |a Roberts, Mervin F.
 245 10 |a All about breeding lovebirds / |c Mervin F. Roberts.
 260 __ |a Neptune, N.J. : |b T.F.H. Publications, |c c1983.
 300 __ |a 96 p. : |b ill. (some col.) ; |c 21 cm.
 504 __ |a Includes bibliographical references and index.
 650 _0 |a Lovebirds |x Breeding.
 991 __ |b c-GenColl |h SF473.L6 |i R63 1983 |t Copy 1 |w BOOKS

records in a database/ catalog

- 1) locked away in a record; don't present our data as data; databases keep users out except for the user interface, so very little functionality available to programs
- 2) each library's information is separate (except in OCLC and some union catalogs); so the totality of library information isn't usable; still based on physically locating items, not on making use of the information that we have.

think outside the (catalog)
box

think about our data as data,
not as catalog entries.

think about our data interacting
with other data on the web.

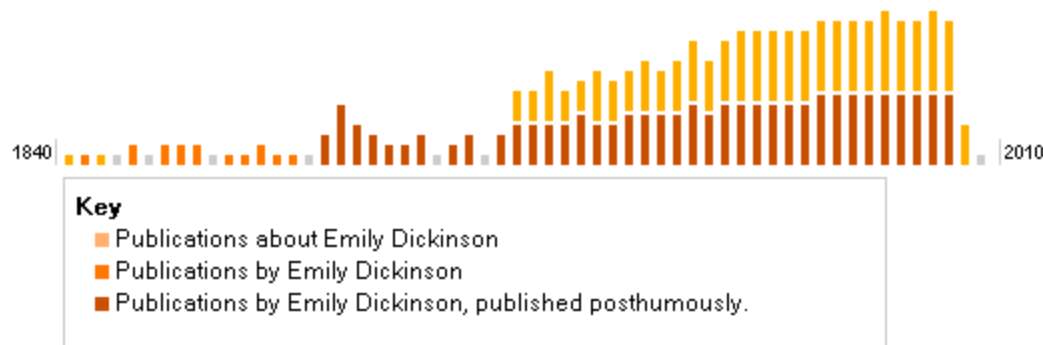
how could our data be used as data? think about how it could be used outside of the catalog – and not by putting marc records on the web, but allowing all of the individual data elements to be exploited. (yes, this will require using a data format other than marc.)

Dickinson, Emily 1830-1886

Overview

Works:	3,125 works in 4,975 publications in 39 languages and 291,643 library holdings
Genres:	Artists' books American poetry Motion pictures Filmstrips Biography Broadside Dissertations, Academic Miniature books Visual literature
Subject Headings:	Poets, American—19th century Women poets, American—19th century Agoraphobia—Patients—United States
Roles:	Lyricist , Librettist , Honoree , Other , Author in quotations or text abstracts , Originator , Author of dialog
Classifications:	ps1541,811.4

Publication Timeline

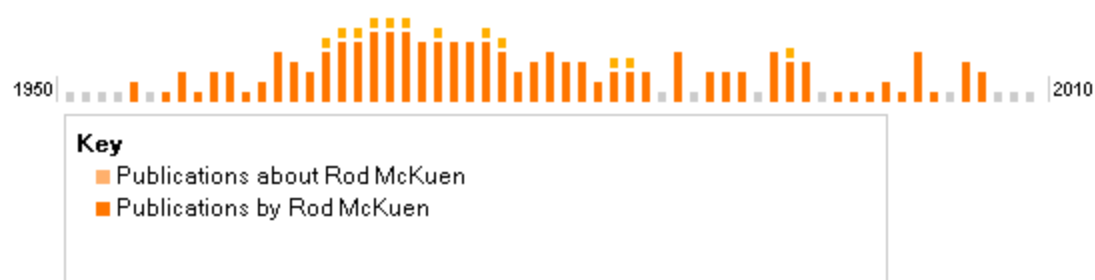


McKuen, Rod

Overview

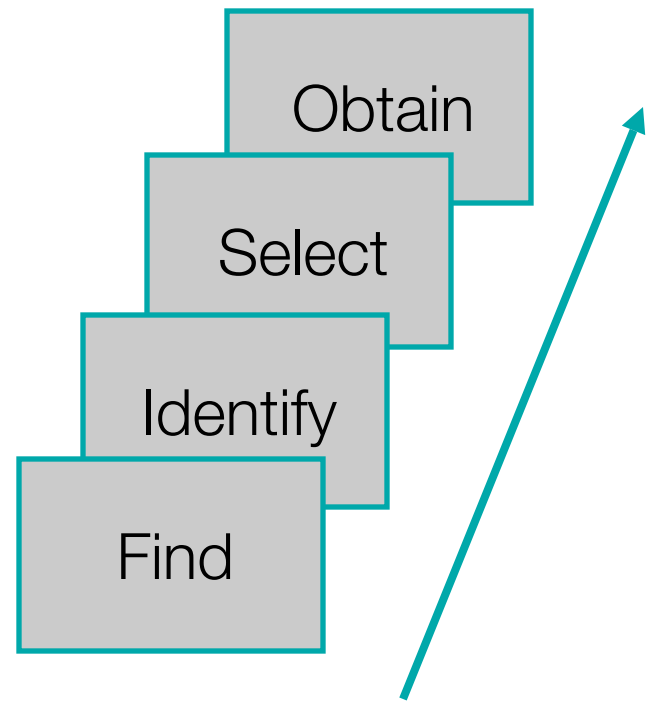
Works:	362 works in 548 publications in 7 languages and 28,628 library holdings
Subject Headings:	Poets, American—20th century
Roles:	Performer , Composer , Compiler , Editor , Lyricist , Illustrator , Narrator , Vocalist , Writer of accompanying material , Arranger , Conductor
Classifications:	ps3525.a264, 811.54

Publication Timeline

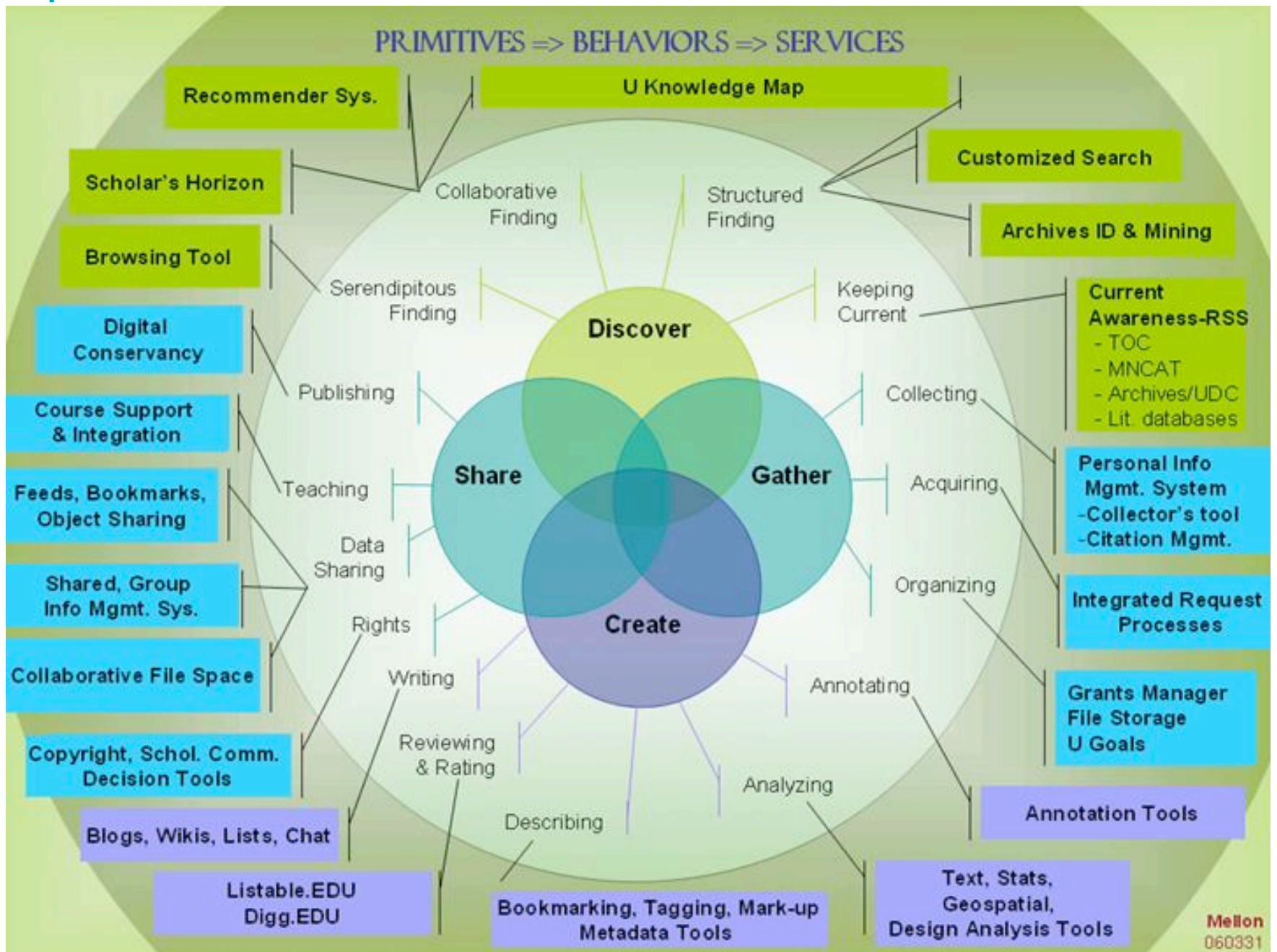


FRBR user tasks

A man walks up to a library catalog....



what is holding us back? In part, it's our vision of the function of our data. FRBR starts with someone approaching the catalog, and ends when they get something off the shelf. That's only a very small part of information seeking behavior, and less and less of it proportionally now that many resources are available online. This narrow view is a hindrance to development of services that users need.



Notice that it also has four primary activities, but they are very different from the four given by FRBR. As a matter of fact, the library catalog is directly implicated in only a small portion of these activities. The outer ring suggests services related to each activity. How many of these should library data interact with? Most, if not all. The services may not be directly managed by the library, e.g. courseware; EndNote and Zotero. And there are services beyond these that could make use of bibliographic data (e.g. Wikipedia; Google books).

How can we make the change we need to expand the usefulness of our data?

semantic web



One possible direction is offered to us by the semantic web work. (this cartoon will make sense in a moment)

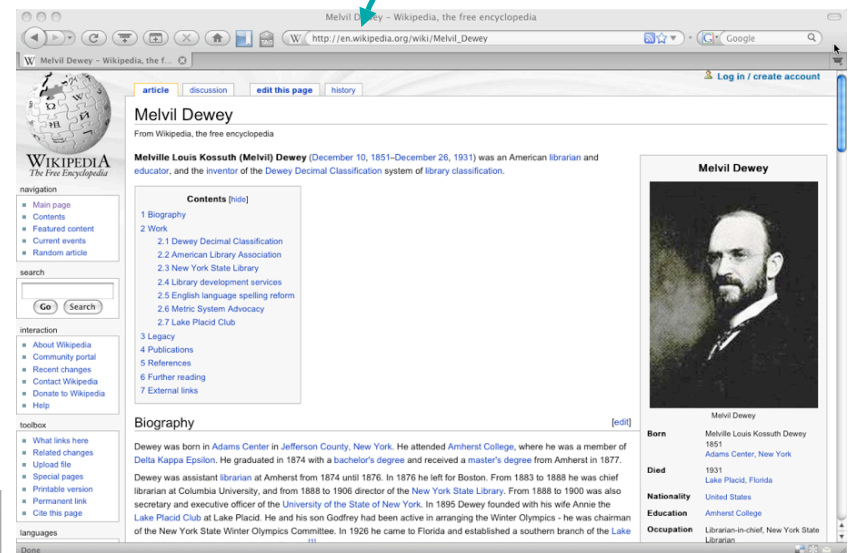


link

Web

links between documents

One meaning: "link"



The problem the semantic web is trying to solve: links between documents (just link); data within documents not actionable. The semantic web wants to get into the information that is contained within the documents.

blah blah Moby Dick blah blah
whales blah blah blah Herman
Melville blah blah blah New Bedford,
Massachusetts blah blah blah

web pages

are mostly text

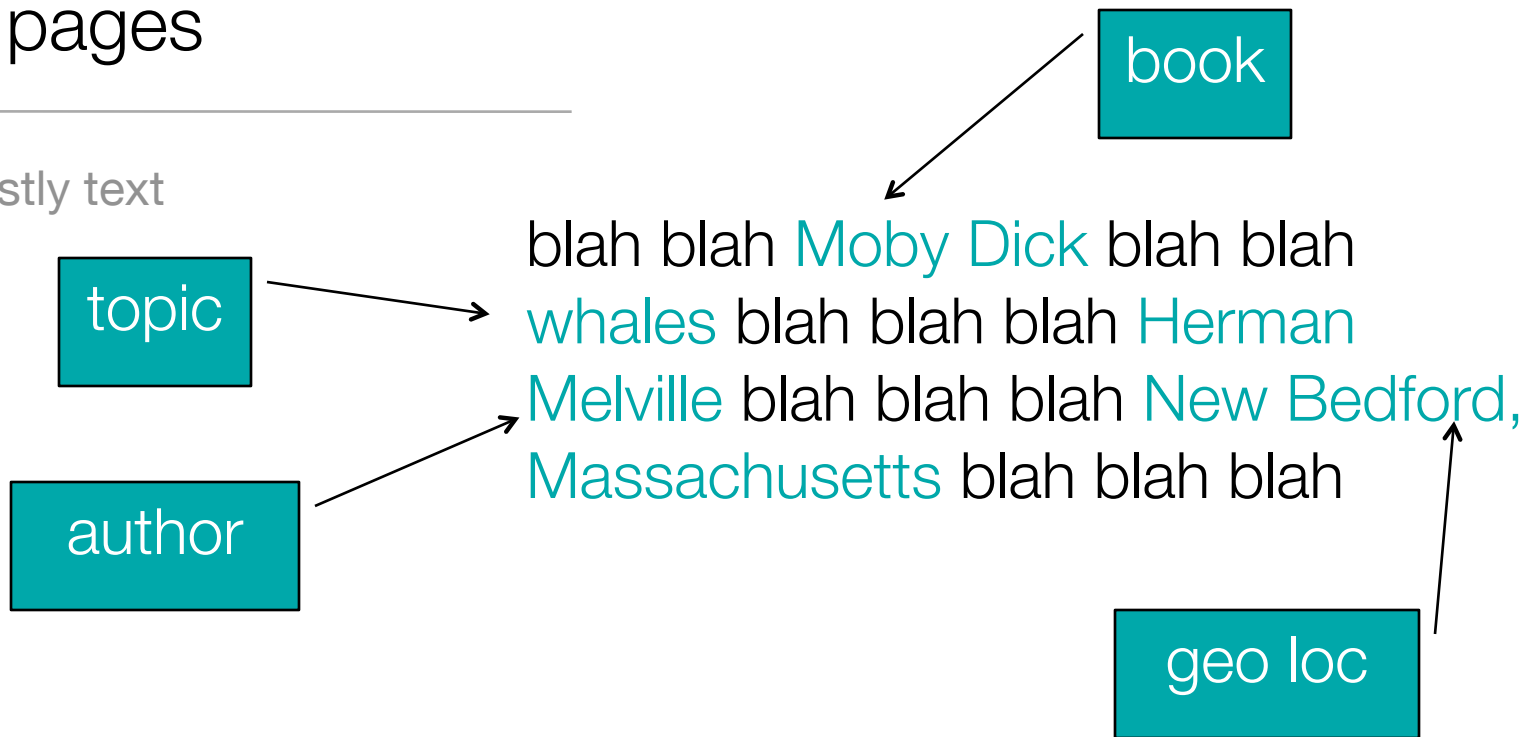
blah blah Moby Dick blah blah
whales blah blah blah Herman
Melville blah blah blah New Bedford,
Massachusetts blah blah blah

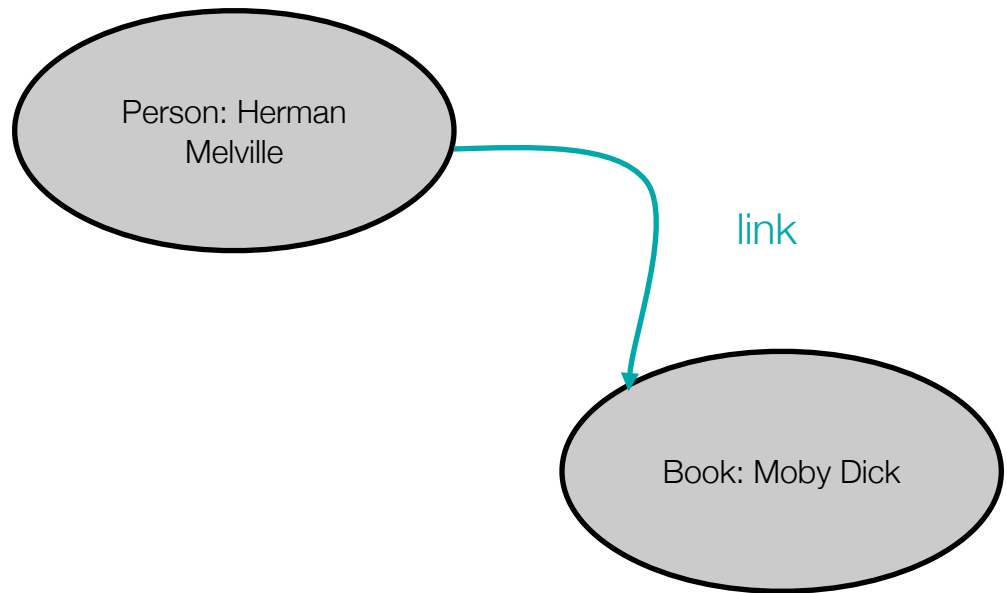
However, we can't use the information in the documents because it isn't identified as information – at least not in a way that a machine could recognize.

blah blah Moby Dick blah blah
whales blah blah blah Herman
Melville blah blah blah New Bedford,
Massachusetts blah blah blah

web pages

are mostly text

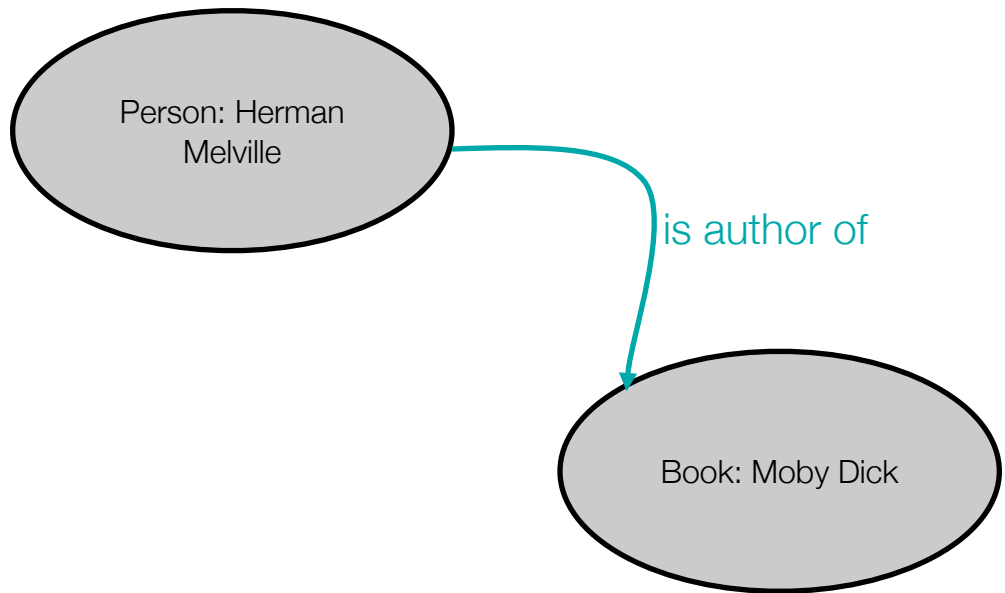




identify your entities

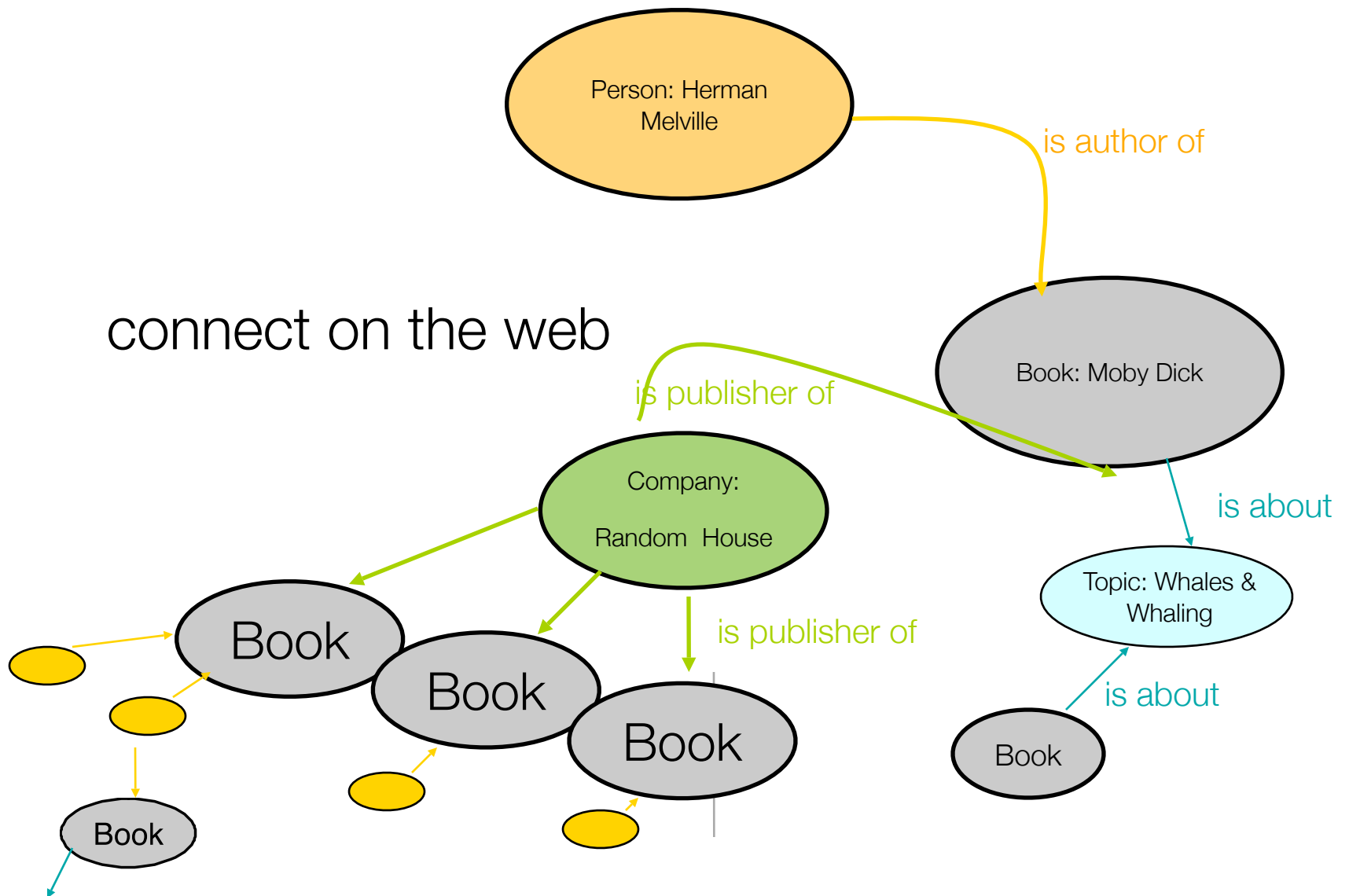
give them names, identifiers

For the semantic web you need to identify your entities in a standard way that is machine actionable.



make links meaning

Links need to have meaning; provide relationships between entities. Links will have to be standardized, and this is a very complex operation.

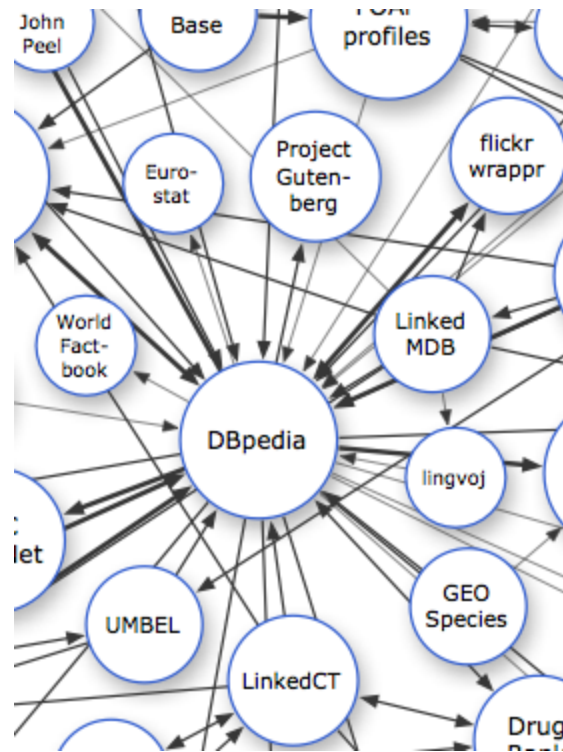


Connect outside of any individual web page. You end up with a web of information, rather than a web of documents. In the end, it looks very much like a huge, online database. And the SemWeb community has developed a language that can be used to query the cloud of information, although it's just in its early stages of development.

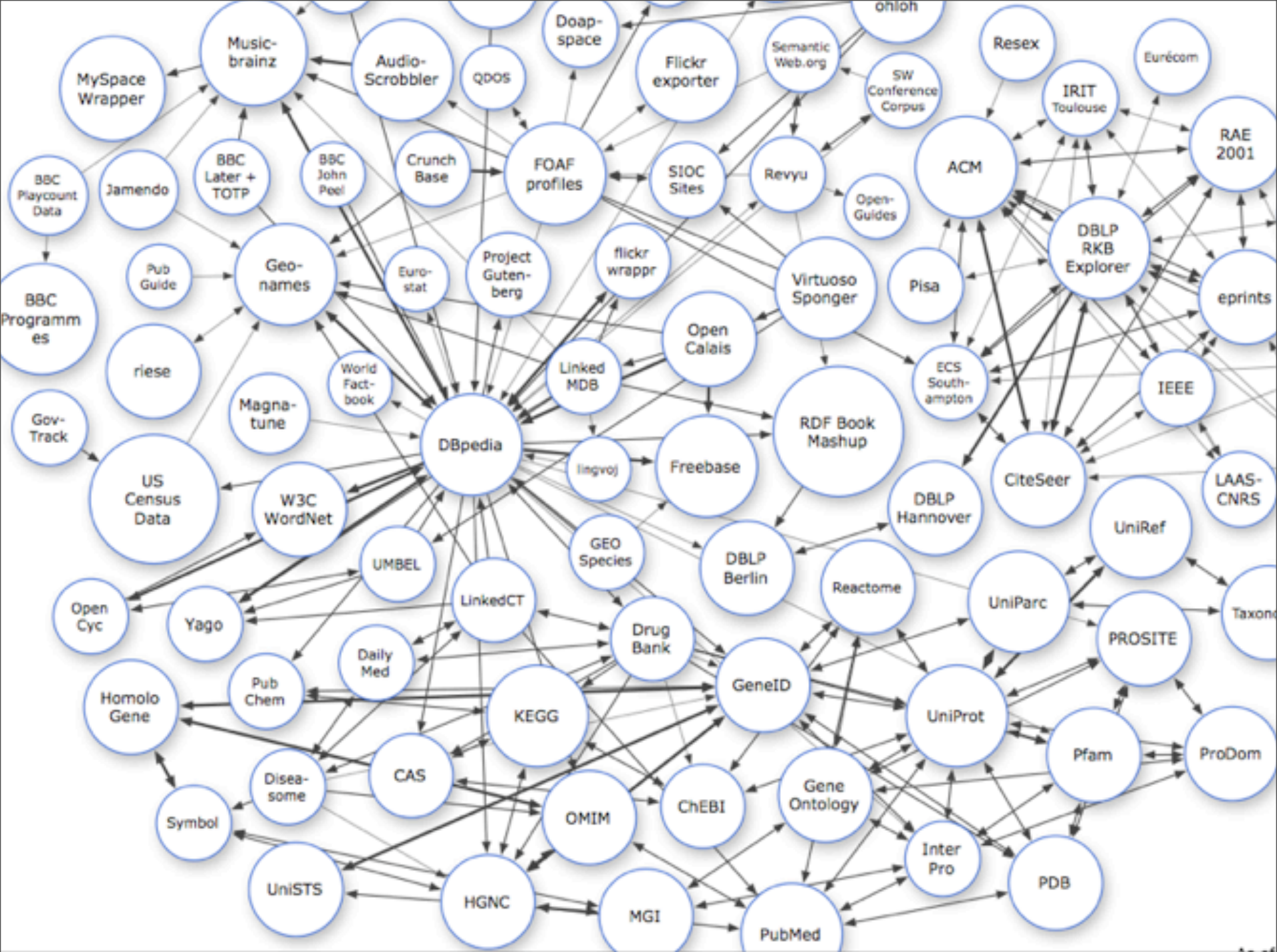
meanwhile

linked open data

<http://linkeddata.org>



Actually taking documents on the web and making them semantic is huge. But there are lots of folks who have data – already identified data. And there are standards for how to make this data visible and usable on the web.



What used to be separate data hidden away in databases now becomes a huge teeming mass of interactive data. And guess what – libraries have some of the best data ever created, and definitely have THE data about published resources.

But libraries aren't represented here – why not?

LC Control No.: 84100005

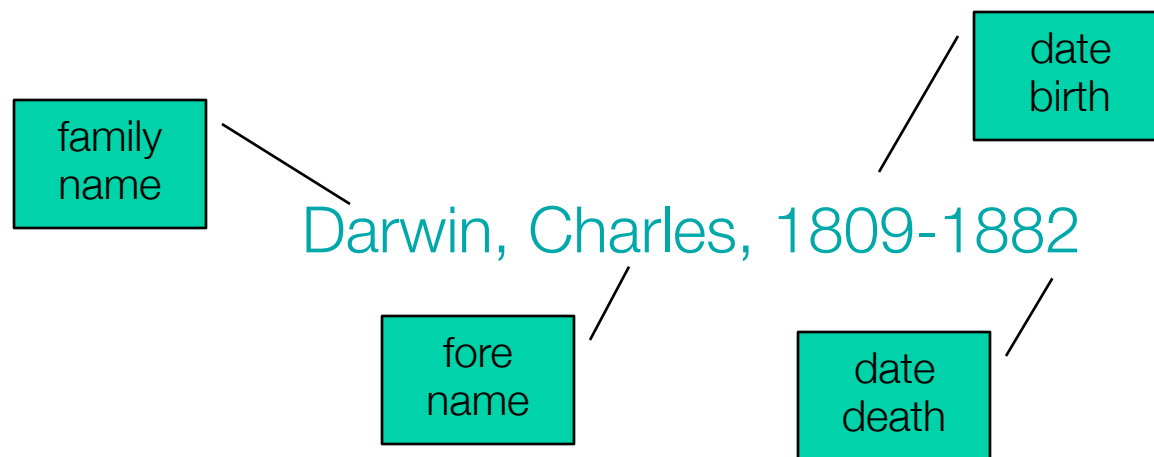
LCCN Permalink: <http://lccn.loc.gov/84100005>

000 00745nam a2200229 a 450
001 954379
005 19840215000000.0
008 840103s1983 njua b 001 0 eng
035 __ |9 (DLC) 84100005
906 __ |a 7 |b cbc |c orignew |d 2 |e ncip |f 19 |g y-gencatlg
010 __ |a 84100005
020 __ |a 0876669437
040 __ |a DLC |c DLC |d DLC
050 00 |a SF473.L6 |b R63 1983
082 00 |a 636.6/865 |2 19
100 1_ |a Roberts, Mervin F.
245 10 |a All about breeding lovebirds / |c Mervin F. Roberts.
260 __ |a Neptune, N.J. : |b T.F.H. Publications, |c c1983.
300 __ |a 96 p. : |b ill. (some col.) ; |c 21 cm.
504 __ |a Includes bibliographical references and index.
650 _0 |a Lovebirds |x Breeding.
991 __ |b c-GenColl |h SF473.L6 |i R63 1983 |t Copy 1 |w BOOKS

Machine-readable

catalog.loc.gov

Why not libraries? We have the information, but like documents on the web it isn't coded in a way that it can join the cloud. And the MARC record doesn't use properties and relationships, as required by linked data.



Library data as LOD

what important world events happened during his lifetime?

who was writing on science at the same time?

show me a map of important places in his life

What could you do with it?

With some adjustment (ok a fair amount of adjustment), the data in library catalogs could connect with data in the cloud and you could ask questions and get answers:

What makes good sense to humans isn't necessarily sensible to machines and algorithms. If I want to say: give me all of the authors born between 1800 and 1830, I can't do it with this data because birth data isn't a data element, even though the information is there.

Beta Site



Open Library

One web page for every book. (More...)

GO



Scanned books only ▶ [Advanced Search](#)

featuring **22,845,290** books
(including **1,064,822** with full-text)

Open Library Upgrade

Open Library has upgraded to new software. Improvements include readable URLs, better default cover images, an improved type system and more APIs for developers, including the **RESTful API**. More improvements are coming soon - check our **blog** for updates.

Find Books

- ▶ Request scan
- ▶ Recent changes
- ▶ Example search
- ▶ Add a book
- ▶ Index of books

Build The Library

- ▶ About the project
- ▶ About the technology
- ▶ About the librarianship
- ▶ A guided tour

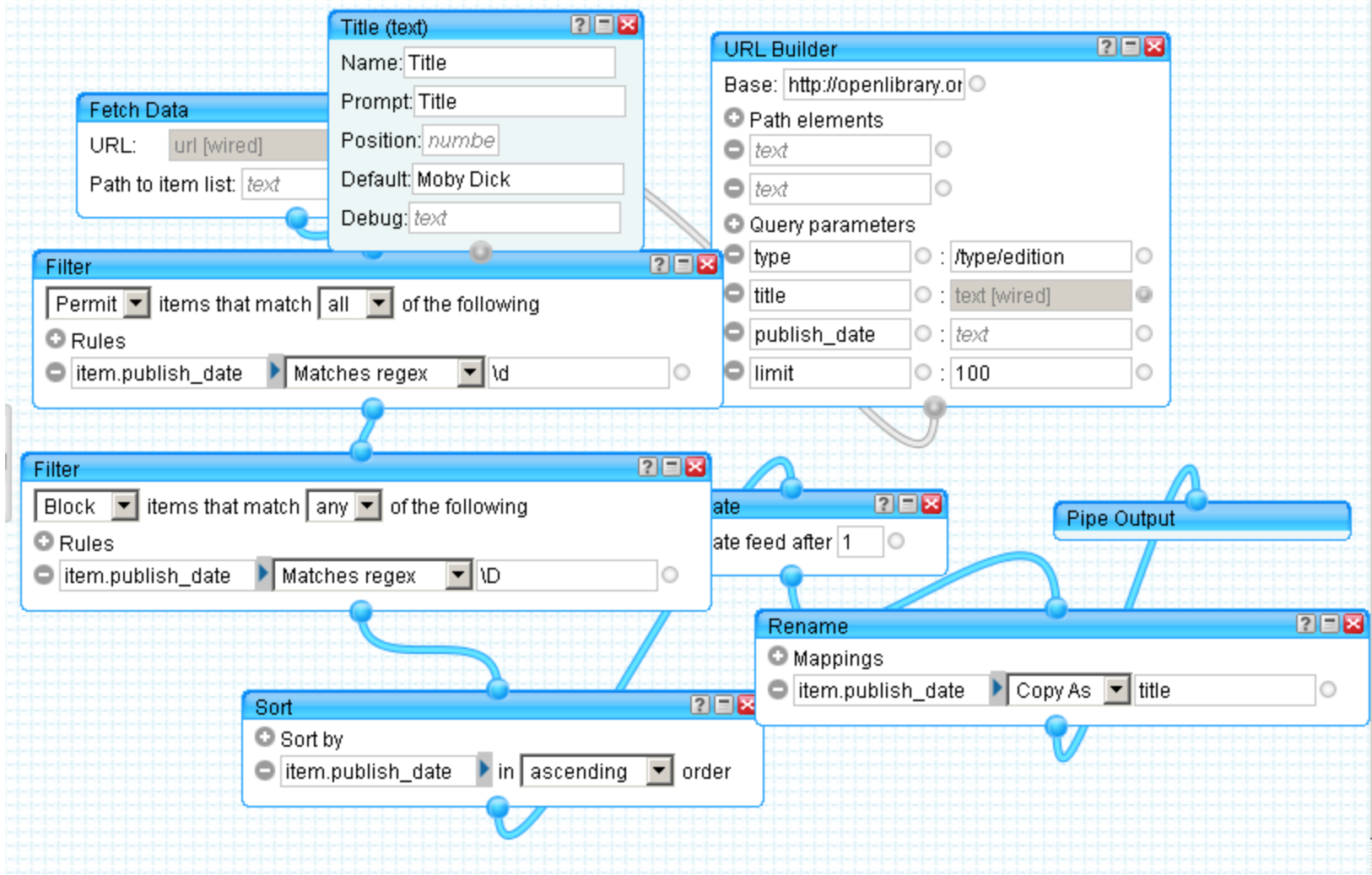
Develop The Site

- ▶ Bug tracking
- ▶ Source code repository (infogami)
- ▶ Developer documentation
- ▶ Translate the site

openlibrary.org

Find earliest publication date; feed to copyright analysis program

We are beginning to have some examples of bibliographic data as linked open data. OL is not quite LOD, but



what is the earliest
 publication date for a book
 with the title Moby Dick?

http://pipes.yahoo.com/pipes/pipe.info?_id=216efa8c3b04764ca77ad181b1cc66e4

tag: openlibrary

Earliest publication date

Searches the Open Library for a title, returns the first 100 items, and selects the earliest publication date from that group. Can be used to help determine copyright status based on earliest date of publication.

Pipe Web Address: <http://pipes.yahoo.com/kcoyle/216efa8c3b04764ca77ad181b1cc66e4> ([edit](#))

☆ [Edit Source](#) [Delete](#) [Re-publish](#) [Unpublish](#) [Clone](#)

[Configure this Pipe](#)

Title [Run Pipe](#)

[Use this Pipe](#)

[Get as a Badge](#) [MY YAHOO!](#) [Google](#) [Get as RSS](#) [Get as JSON](#) [More options](#)

List

1 item

1851

what is the earliest
publication date for a book
with the title Moby Dick?

pipes.yahoo.com

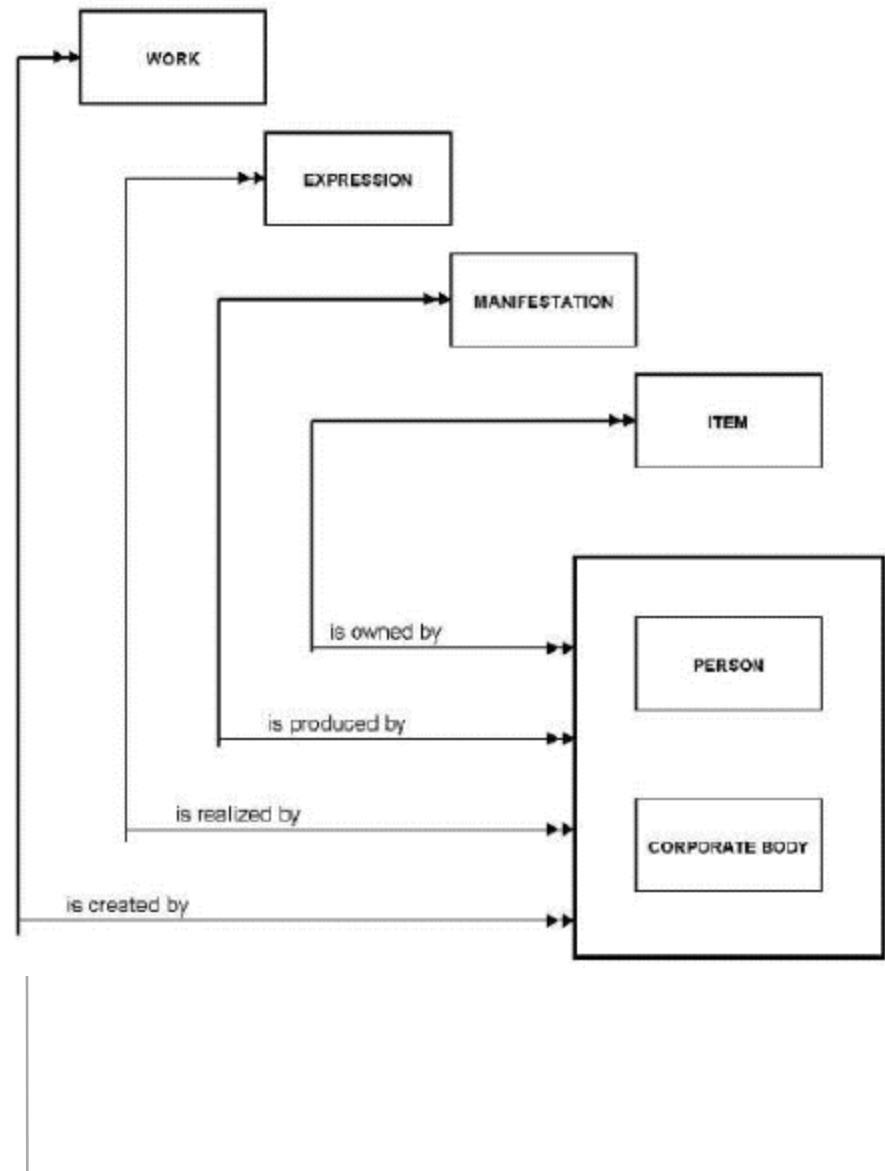
v · d · e		Works by Vladimir Nabokov	[hide]
Novels	Russian	<i>Mary</i> · <i>King, Queen, Knave</i> · <i>The Defense</i> · <i>The Eye</i> · <i>Glory</i> · <i>Laughter in the Dark</i> · <i>Despair</i> · <i>Invitation to a Beheading</i> · <i>The Gift</i> · <i>The Enchanter</i>	
	English	<i>The Real Life of Sebastian Knight</i> · <i>Bend Sinister</i> · <i>Lolita</i> · <i>Invitation of a Beheading</i> · <i>Priglasenie na kazn</i> · <i>Pale Fire</i> · <i>Ada or Ardor: A Family Chronicle</i> · <i>Transparent Things</i> · <i>Look at the Harlequins!</i> · <i>The Original of Laura</i> (unpublished fragment)	
Short stories	Russian	"The Wood-Sprite" · "Russian Spoken Here" · "Sounds" · "Wingstroke" · "Gods" · "A Matter of Chance" · "The Seaport" · "Revenge" · "Beneficence" · "Details of a Sunset" · "The Thunderstorm" · "La Veneziana" · "Bachmann" · "The Dragon" · "Christmas" · "A Letter That Never Reached Russia" · "The Fight" · "The Return of Chorb" · "A Guide to Berlin" · "A Nursery Tale" · "Terror" · "Razor" · "The Passenger" · "The Doorbell" · "An Affair of Honor" · "The Christmas Story" · "The Potato Elf" · "The Aurelian" · "A Dashing Fellow" · "A Bad Day" · "The Visit to the Museum" · "A Busy Man" · "Terra Incognita" · "The Reunion" · "Lips to Lips" · "Orache" · "Music" · "Perfection" · "The Admiralty Spire" · "The Leonardo" · "In Memory of L. I. Shigaev" · "The Circle" · "A Russian Beauty" · "Breaking the News" · "Torpido Smoke" · "Recruiting" · "A Slice of Life" · "Spring in Fialta" · "Cloud, Castle, Lake" · "Tyrants Destroyed" · "Lik" · "Vasily Shishkov" · "Ultima Thule" · "Solus Rex" · "The Word" · "Natasha"	
	French	"Mademoiselle O"	
	English	"The Assistant Producer" · "That in Aleppo Once..." · "A Forgotten Poet" · "Time and Ebb" · "Conversation Piece, 1945" · "Signs and Symbols" · "First Love" · "Scenes from the Life of a Double Monster" · "The Vane Sisters" · "Lance" · "Easter Rain"	
Plays	Russian	<i>Death</i> · <i>The Grandfather</i> · <i>The North Pole</i> · <i>The Tragedy of Mr. Morn</i> · <i>The Man from the USSR</i> · <i>The Event</i> · <i>The Waltz Invention</i>	
Non-fiction	Memoirs & Letters	<i>Speak, Memory</i> · <i>The Nabokov-Wilson letters</i> · <i>Selected Letters, 1940-1977</i>	
	Criticism	<i>Strong Opinions</i> · <i>Nikolai Gogol</i> · <i>Lectures on Literature</i> · <i>Lectures on Russian Literature</i> · <i>Lectures on Don Quixote</i> · <i>Notes on Prosody</i>	
Miscellanea		<i>Poems and Problems</i> · <i>Lolita: A Screenplay</i> · <i>The Annotated Lolita</i> · <i>Carrousel</i> · <i>Eugene Onegin</i> (translation) · <i>Father's Butterflies</i>	

Wikipedia

Vladimir Nabokov

From library catalog to LOD

FRBR leads the way



Not because it changes the data we have, but it gives us a view of the data as data, not as records. And it talks about the data in terms of entities and relationships, which is the same view that lod takes.

What can we do?

1. Make our data harvestable so that others can use it
2. Identify our data structures and vocabularies for use on the web

1. this isn't something we necessarily have to do ourselves – we can just let it happen. it would be better if we did get involved – mainly to provide some quality services that others may not
2. one of the big barriers is that folks don't know what data elements we have and what they mean. there are standard ways to identify data for use as linked data and we can provide that. this isn't a change to our data, just a way of making it more understandable and more usable to folks outside of the library world. it will also help library systems developers, because the data will be more available than it is today, where it is buried in documents and web sites.

languages

roles

places

resource types

names

subjects

vocabularies

easy to 'webify'

the easiest thing that we can do is to make our vocabularies available on the open web. what's a vocabulary? all of those controlled lists that we use in the creation of our metadata.

this is great stuff, and the most import of all is our name authorities. everyone is struggling with how to clearly identify persons with different names, and we've done it – not for everyone, but for millions of people who are either creators or subjects of creative works.

you may have heard about 'lsh.info' – which was an experiment in making lsh available in semantic web format. it was online for about 6-8 months, and although it wasn't highly publicized it got used in a number of interesting ways before lc took it down, because it wasn't official. but it did prove that if you make the data available, people will use it.



Resources for...

- [Kids, Families](#)
- [Librarians](#)
- [Publishers](#)
- [Researchers](#)
- [Teachers](#)
- [Visitors](#)

General Information

- [About the Library](#)
- [Concerts & Events](#)
- [Jobs/Fellowships](#)
- [Support the Library](#)
- [Shop the Library](#)
- [Inspector General](#)

More Library Resources

- [American Folklife Center](#)
- [Braille, Audio Materials](#)
- [Copyright Office](#)
- [Copyright Royalty Board](#)

This site serves as a placeholder for forthcoming web services that will enable both humans and machines to programmatically access authority data at the Library of Congress. The initial services offered are influenced by -- and therefore implement -- the [Linked Data](#) movement's approach of exposing and inter-connecting data on the Web via dereferenceable URIs. We aim to make resources available on this site within 6-8 weeks. Check this site regularly for more updates as we continue to develop this service!

Scope

The scope of this web service is to provide access to commonly found standards and vocabularies promulgated by the Library of Congress. This includes data values and the controlled vocabularies that house them. The main application will provide resolvability to values and vocabularies by assigning URIs. Each vocabulary will possess a resolvable URI, as will each data value within it.

URIs accessible at id.loc.gov will only link to authority data -- controlled vocabularies and the values within them. Therefore, users will not find identifiers for electronic bibliographic resources. The Library of Congress uses other identifier schemes such as [Handles](#) for this purpose.

Goals

- The Library of Congress will strive to communicate a clear usage policy for all vocabularies accessible through id.loc.gov. Particular emphasis will be made to include vocabularies which Library of Congress already exposes without restriction on use.
- By minting a URI for either a vocabulary or an individual data value, the Library of Congress promotes these resources to first-class, dereferenceable web resources.

Which vocabularies will be included?

Initially, within 6 to 8 weeks, the Library of Congress will release its first offering: the ~~Library of Congress Subject Headings~~. This will be an almost verbatim release

MARC vocabularies

id.loc.gov

lc has promised to make lcsch and other vocabularies, mainly those used in marc records, available online in the appropriate format. they announced this at ala midwinter in denver and said that lcsch would be available shortly (since it had already been done and the code was available). unfortunately, as of this date they haven't yet done so, and it's important to let them know that this really matters for libraries.

NSDL REGISTRY

Supporting Metadata Interoperability

Vocabulary: Show detail for RDA Groups of Instruments

Detail

Concepts

History

Versions

Maintainers

Preferred Label	URI	Status	Updated	Actions
Woodwinds	.../termList/groupsInstr/1001	New-Proposed	2009-03-25 13:28	
Brasses	.../termList/groupsInstr/1002	New-Proposed	2009-03-25 13:28	
Winds	.../termList/groupsInstr/1003	New-Proposed	2009-03-25 13:28	
Plucked instruments	.../termList/groupsInstr/1004	New-Proposed	2009-03-25 13:29	
Keyboard instruments	.../termList/groupsInstr/1005	New-Proposed	2009-03-25 13:26	
Strings	.../termList/groupsInstr/1006	New-Proposed	2009-03-25 13:26	
Instrumental ensemble	.../termList/groupsInstr/1007	New-Proposed	2009-03-25 13:31	

7 results

RDA vocabularies

metadataregistry.org

Because RDA is new and isn't already codified in a record format, it offers some interesting opportunities. In the development of the RDA online service, the vocabularies and data elements of RDA are being added to a metadata registry that will have a lot of necessary capabilities. For example, each controlled list will be registered with the current RDA terms, but it will also be possible to provide display forms of those terms in other languages.

Vocabulary: RDA multilingual vocabulary

Concepts: Werk

Detail

Properties

History

Detail

Preferred Label: Werk

Language: German

URI: http://metadataregistry.org/uri/rda_multilingual/1002

Top Concept?:

Status: New-Proposed

Properties

related to	Manifestation		New-Proposed
related to	Work		New-Proposed
preferred label	work	English	New-Proposed
preferred label	obra	Spanish	New-Proposed
preferred label	oeuvre	French	New-Proposed
has broader	Entität		New-Proposed

List | Get RDF

RDA vocabularies

metadataregistry.org

The German national library is already experimenting with multilingual entries for FRBR and RDA elements.

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:skos="http://www.w3.org/
2004/02/skos/core#">
  <rdf:Description rdf:about="http://RDVocab.info/termLst/ groupsInstr">
<dc:title>RDA Groups of Instruments</dc:title>
<rdf:type rdf:resource="http://www.w3.org/2004/02/skos/
core#ConceptScheme"/>
</rdf:Description>
<rdf:Description rdf:about="http://RDVocab.info/termLst/groupsInstr/1004">
<rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
<skos:inScheme rdf:resource="http://RDVocab.info/termLst/ groupsInstr"/>
<skos:prefLabel>Plucked instruments</skos:prefLabel>
</rdf:Description> </rdf:RDF>
```

RDA vocabularies

metadataregistry.org

NSDL REGISTRY

Supporting Metadata Interoperability

Schema: Show detail for RDA Elements

- Detail
- Properties
- History
- Maintainers

Label	URI	Status	Updated	Last Updated by	Actions
Abbreviated title	.../Elements/abbreviatedTitle	New-Proposed	2008-11-07 10:29	kcoylenet	
Accessibility content	.../Elements/accessibilityContent	New-Proposed	2008-11-07 10:31	kcoylenet	
Additional scale information	.../additionalScaleInformation	New-Proposed	2008-11-07 10:33	kcoylenet	
Address of the corporate body	.../addressOfTheCorporateBody	New-Proposed	2009-02-25 15:55	kcoylenet	
Address of the person	.../Elements/addressOfThePerson	New-Proposed	2009-02-25 15:57	kcoylenet	
Affiliation	.../Elements/affiliation	New-Proposed	2009-02-25 15:57	kcoylenet	
Applied material	.../Elements/appliedMaterial	New-Proposed	2009-02-24 14:31	DianeH	
Artistic and/or technical credits	.../artisticAndOrTechnicalCredits	New-Proposed	2008-11-07 18:09	kcoylenet	
Aspect ratio	.../Elements/aspectRatio	New-Proposed	2008-11-07 18:12	kcoylenet	
Associated institution	.../associatedInstitution	New-Proposed	2009-02-25 15:58	kcoylenet	
Awards	.../Elements/awards	New-Proposed	2008-11-07 18:12	kcoylenet	
Base material	.../Elements/baseMaterial	New-Proposed	2009-02-24 14:31	DianeH	
Base material for microfilm, microfiche and motion picture film	.../MicroficheAndMotionPictureFilm	New-Proposed	2009-02-24 14:19	DianeH	
Biographical information	.../biographicalInformation	New-Proposed	2009-02-25 15:58	kcoylenet	
Book format	.../Elements/bookFormat	New-Proposed	2008-11-07 18:14	kcoylenet	
Broadcast standard	.../Elements/broadcastStandard	New-Proposed	2008-11-07 18:14	kcoylenet	
Carrier type	.../Elements/carrierType	New-Proposed	2009-02-24 14:32	DianeH	
Cataloguer's note	.../Elements/cataloguersNote	New-Proposed	2008-11-07 18:18	kcoylenet	
Chronological designation of first issue or part	.../DesignationOfFirstIssueOrPart	New-Proposed	2008-11-07 18:18	kcoylenet	
Chronological designation of first issue or part of new sequence	.../FirstIssueOrPartOfNewSequence	New-Proposed	2008-11-07 18:19	kcoylenet	

334 results 1 2 3 4 5

Done

registry of RDA data elements

rdaonline.org

What can we do?

1. Make our data harvestable so that others can use it
2. Identify our data structures and vocabularies for use on the web
3. Make our data more 'data-
4. ^{like'} Make our systems more 'cloud-like'

Eventually we need to make our data more data-like. There are too many fields that are primarily textual, such as 'includes bibliographic references' or 'translation of ...' Where there are fixed field codes for such statements they require a separate input, and many libraries just don't bother. In fact, if we view our data as data, then 'includes bib references' is a bit of data that can be displayed however we want (and in whatever language) for our human users, but can be usable by machines if coded appropriately.

There is much in the MARC record that we haven't made good use of. it's really a shame to be spending the time entering data that isn't used. Perhaps our data practices were ahead of their time. We may just now be getting to a point where there is a context where our data can be fully used. There is no question in my mind that to make this work we have to develop systems that provide much more support for the metadata creator.

Eventually, our systems need themselves to become more cloud like. This cloud will grow, and if it develops in the way I'm hoping it will it will provide a great environment for information exploration. We should be in it.

metadata are a growing organism

with *MORE* apologies to Ranganathan

and the more that metadata is used, the more value it
creates

Thank you

kcoyle@kcoyle.net
<http://www.kcoyle.net>

