Reinventing Invention
A view from OCLC

Richard Wallis
Technology Evangelist
@RJW
Reinventing Invention
A view from OCLC

A challenge from David Weinberger....

The net has won!
The Industrial Revolution
The Industrial Revolution

1779 (c.)

Image courtesy of: Shropshire County Council
Information Management: A Proposal

Abstract

This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext system.

Keywords: Hypertext, Computer conferencing, Document retrieval, Information management, Project control
Entity Based Data on the Web

Semantic Search

Google Semantic Search

Search Engine Optimization (SEO) Techniques That Get Your Company More Traffic, Increase Brand Impact, and Amplify Your Online Presence

Mount Everest

Facebook Amazon Apple
A Web of Data
A Web of Data

A Library Shaped Black Hole?
Sendak still rocks!
Where the wild things are / story and pictures by Maurice Sendak

Sendak, Maurice
[37] p. : col. ill. ; 23 x 26 cm
ISBN0060254920

$8.95
Caldecott Medal book - 1964

I loved this book as a child!!!!!
| 03465njm 22005891a 4500 001 ocm65198904 003 OCoLC 005 20060324022622.0 007 sd bsmenn----- 008 060324s197u ru opn rus d 028 12 $a C10-06739/3-1--C10-06740/3-1 $b Melodii"a-" 035 $a (Sirsi) o65198904 040 $a NGU $c NGU 041 0 $d rus $d ita 049 $a NGU 245 00 $a Poi"u"t solisty Bol'shogo teatra Soi"u"za SSR $h [sound recording] = $b Soloists of the Bolshoi Theatre. 246 31 $a Soloists of the Bolshoi Theatre 260 $a [Moscow] : $b Melodii"a", $c [197u?] 300 $a 1 sound disc : $b analog, 33 1/3 rpm, stereo. ; $c 12 in. 500 $a Opera excerpts; sung in the original languages. 505 0 $a Side one: Eugene Onegin. Onegin's area / P. Tchaikovsky (Yuri Masurok, baritone) -- Boris Godunov. Marina's Aria / M. Mussorgsky (Yelena Obraztsova, mezzo-soprano) -- Queen of Spades. Hermann's Arios (Vladimir Atlantov, tenor) -- The mermaid. The miller's aria / A. Dargomyzhsky (Alexander Vedernikov, bass) --The enchantress. Kuma's aria / P. Tchaikovsky (Tamara Milashkina, soprano) -- Ruslan and Lyudmila. Farlaff's Rondo / M. Glinka (Yevgeni Nesterenk, bass) -- 505 8 $a Side two: The maid of Orleans. Jeanne's aria / P. Tchaikovsky (Irina Arkhipova, mezzo-soprano) -- Tosca. Aria of Cavaradosi / G. Puccini (Zurab Soktkilava, tenor) -- Sadko. Volkova's lullaby / N. Rimski-Korsakov (Bela Rudenko, soprano) -- Demon. Demon's romance / A. Rubinstein (Alexander Ognivtsev, bass) --The Marriage of Figaro. Cherubino's aria / W. Mozart (Galina Borisova, mezzo-soprano) -- Iolantha. Roberto's arioso / P. Tchaikovsky (Yuri Gulyaev, baritone) 650 0 $a Operas $v Excerpts. 700 1 $a Mazurok, I'U'rili. $4 prf 700 1 $a Obrazt's'ova, Elena. $4 prf 700 1 $a Atlantov, Vladimir. $4 prf 700 1 $a Vedernikov, A. $4 prf 700 1 $a Milashkina, Tamara. $4 prf 700 1 $a Nesterenko, Evgenii. $4 prf 700 1 $a Arkhipova, Irina, $d 1925-2010 $4 prf 700 1 $a P't'a"vko, V. $4 prf 700 1 $a Soktkilava, Zurab. $4 prf 700 1 $a Rudenko, Béla, $d 1933- $4 prf 700 1 $a Ognivt's'ev, Aleksandr $4 prf 700 1 $a Borisova, Galina, $d 1941- $4 prf 700 1 $a Obrazt's'ova, Elena. $d 1942- $4 prf
record
/ˈrɛkɔːd/
noun
a thing constituting a piece of evidence about the past, especially an account kept in writing or some other permanent form.
entity
/ˈɛntɪti/
noun
a thing with distinct and independent existence.

relationship
/rɪˈleɪʃ(ə)ŋʃɪp/
noun
the way in which two or more people or things are connected
Record
Title: "War and Peace"
Author: "Leo Tolstoy 1828-1910"
ISBN: 0307266931

Entity (http://worldcat.org/entity/work/id/115206288)
Type: Work
Name: "War and Peace"
Author: http://worldcat.org/entity/person/id/1234

Entity (http://worldcat.org/entity/person/id/1234)
Type: Person
Name: "Leo Tolstoy"
Born: 1828
Died: 1910
Birthplace: http://worldcat.org/entity/place/id/8976

Entity (http://worldcat.org/entity/place/id/8976)
Type: Place
Name: "Yasnaya Polyana"
SameAs: http://geonames.org/468686
War and Peace

War and Peace is a novel by the Russian author Leo Tolstoy, first published in 1869. The work is epic in scale and is regarded as one of the most important works of world literature. It is considered as Tolstoy's finest literary achievement, along with his other major prose work, Anna Karenina. War and Peace delves into the lives of five Russian aristocratic families, thus seen through the eyes of five Russian aristocratic families. Portions of an earlier version of the novel, then known as The Year 1805, were serialized in the magazine The Russian Messenger between 1865 and 1867. The novel was first published in its entirety in 1869. Newsweek in 2009 ranked it first in its list of the Top 100 Books. In 2003, the novel was listed at number 20 on the BBC's survey The Big Read. Tolstoy himself, somewhat enigmatically, said of War and Peace that it was "not a novel, even less is it a poem, and still less a historical chronicle". Large sections of the work, especially in the later chapters, are philosophical discussions rather than narrative. [−]
<table>
<thead>
<tr>
<th>Also known as:</th>
<th>Bengali</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pierre Besouchoff</td>
<td>Andrei Bolkonksi</td>
</tr>
<tr>
<td>Andrey Nikolaevich Bolkonksi</td>
<td>War and peace</td>
</tr>
<tr>
<td>Sonya Rostov</td>
<td>War and Society</td>
</tr>
<tr>
<td>War &amp; Peace</td>
<td></td>
</tr>
<tr>
<td>Princess Marya</td>
<td></td>
</tr>
<tr>
<td>War and Peace</td>
<td></td>
</tr>
<tr>
<td>Count Nikolay Rostov</td>
<td></td>
</tr>
<tr>
<td>Petya Ilyitch Rostov</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Statements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>author</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Commons category</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>GND identifier</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Yasnaya Polyana

Yasnaya Polyana - Wikipedia, the free encyclopedia
Yasnaya Polyana (Russian: Ясная Поляна, literally: "Bright Glade") was the home of the writer Leo Tolstoy, where he was born, wrote War and Peace and Anna Karenina, and is buried. Tolstoy called Yasnaya Polyana his "inaccessible literary stronghold". Wikipedia

Address: Yasnaya Polyana, Tula Oblast, Russia, 301214
Phone: +7 487 517-60-73
Burial: Leo Tolstoy

Reviews
4.6 ★★★★★ 26 Google reviews

More reviews: trivago.de

People also search for

Yasnaya Polyana
Tula
Anna Karenina
Optina Monastery
Yasnaya Polyana
Tula
Kremlin
Optina Monastery
Moscow
Vysotsky Monastery
Open Linked Data - Silos
Behind A Vocabulary Barrier
A general purpose vocabulary for describing things on the web

- Linked Data
- Embedded in HTML
- RDFa, Microdata, JSON-LD
- Descriptive data
- Active links

"15% of the Web"
Why Catalog?
So we can find things

Why Share on the Web?
So today’s users can find our things
A Web of Data
Trend 1: Shift from Collections to Services

**Explain:** Libraries shift their focus from collections on the shelf to the services they offer. Pressure on collection and staff budgets.

Trend 2: Search for Distinctiveness

**Explain:** As licensed content is available in commodity collections, academic libraries strive to declare their distinctiveness.

Trend 3: Be Found on the Web

**Explain:** As users search everywhere, libraries want their collections to be found on the web.
So, what are we going to do?
The solution starts here.

The library knowledge graph
A graph of relationships
The library knowledge graph

A graph of relationships
What will be better when we have that graph thingy?

I’m gonna make it better
What will be better?

Improved coherence of the data

*Explanation:* Clustering related things (manifestations, items, formats, etc.)

Improved library applications

*Explanation:* Improved discovery. Make it more web-like.
Reduce redundancy of data by managing entities in large aggregations.

Be Found on the Web

*Explanation:* Using the data models and structures familiar to the web.
The library knowledge graph
Lots of things....if we do it right.
The library knowledge graph

Discovery works...
Cataloging will be different...

- Managing the quality of Works
- Improving clusters
- Managing the quality of Persons
- Links to works, Other IDs

Consistent with RDA
So what progress have we made?
• 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
• 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
Data from one converted record does not make an entity.

Transformation into Linked Data is just a beginning ...

- Mine and analyse the aggregate
- Identify, map, merge - evidence based
- Relate to external sources
- Establish the entities

Getting from here to there
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.
Can we measure impact?
Monthly Unique Visitors

- May 2014
- Jun 2014
- Jul 2014
- Aug 2014
- Sept 2014
- Oct 2014
OCLC’s Entity-Based Data Strategy

- VIAF, ISNI, FAST Publish Linked Data
- WorldCat.org Linked Data Release – using Schema.org
- Internal agreement on data strategy
- Evangelism
- Research & Design with Data Architecture Group
- Data mining of WorldCat resources
- BIBFRAME Early Experimenters

2012

2013

2014

WorldCat Works Released

- Application Integration
  - WorldCat Discovery
  - Analytics
  - Discovery API
  - Cataloging

- More Entities Released
  - Person
  - Manifestation
  - Organization
  - Concept

- New Products
- Continuing Evangelism
- New Services
- Continuing Innovation

...
Add your holdings & metadata to WorldCat

Schema – Bib Extend
http://www.w3.org/community/schemabibex/

OCLC – Data
http://www.oclc.org/data
Reinventing Invention
A view from OCLC

Richard Wallis
Technology Evangelist
richard.wallis@oclc.org
@rjw