

Cataloging, Assessment & The Future of Metadata Management

Ted Fons
Executive Director Data Services &
WorldCat Quality

ALCTS-CaMMs Heads of Cataloging Interest Group
ALA Annual, 2013

- ❖ The goal of the next generation of metadata management initiatives is to improve access to library collections where the user is: on the web
- ❖ Is it possible to align assessment to that goal?

Answer: Yes

How?
Measure different things

- Brief summary of Mugridge & Cheney presentations
 - Compare and contrast
- Brief summary of the next generation of metadata management initiatives
- What principles of assessment will be relevant in the new environment?
- Discussion

Summary of Presentations

Assessment in Practice

Rebecca L. Mugridge
American Library Association
CaMMS Heads of Cataloging Interest Group
January 28, 2013

Quantitative Assessment

Heads of Cataloging Interest Group – ALA Midwinter, 2013

Casey Cheney-Vice President, Bibliographic Services

BACKSTAGE LIBRARY WORKS

January 28, 2013

What is being assessed?

Assessment in Practice

- * Improve effectiveness
- * Identify areas for improvement
- * Communicate with customer
- * Communicate with administration
- * Lower costs
- * Help with decision making

Effectiveness

Customer Satisfaction

Costs

Quantitative Assessment

Heads of Cataloging Interest Group – ALA Midwinter, 2013

- Total project

Effectiveness

- Project efficiency

Costs

- Individual catalog

Efficiency

- Individual catalog

Efficiency

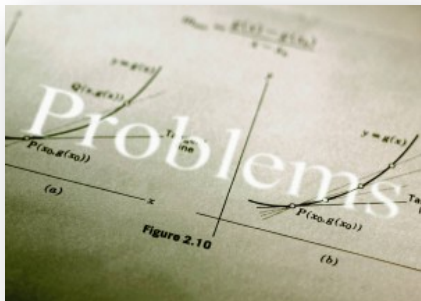
- Quality Control

Quality

Control

Issues & Problem Statements

New Bibliographic Initiatives



Problem 1: MARC21 does not allow for effective exposure of our data on the web so we can meet our users where they are.

Problem 2: We don't have Linked Data.

Problem 3: MARC is a legacy format that needs to be replaced.

Problem 1: MARC21 does not allow for effective exposure of our data on the web so we can meet our users where they are.



Library of Congress: Bibliographic Framework Initiative

Proposing a new model for library data exchange.

NISO: Bibliographic Roadmap Project

Coordinating stakeholders and monitoring overlaps and gaps between initiatives.

WC3 Community Group: Schema.org Bib Extend

Proposing extensions to Schema.org to capture critical bibliographic data elements.

The User Has Many Choices





How college students use information sources



Search engine

2005: 92%

2010: 83%



Wikipedia

2005: NA

2010: 7%



Library Web site

2005: less than 1%

2010: 0%

Library Web site use



2010: 57%



2005: 61%

The “problem” with access to library collections isn’t that the user isn’t using the library catalog. (That’s just a fact.)

The problem is that access to library collections is imperfect because we don’t expose our collections very well on the web.

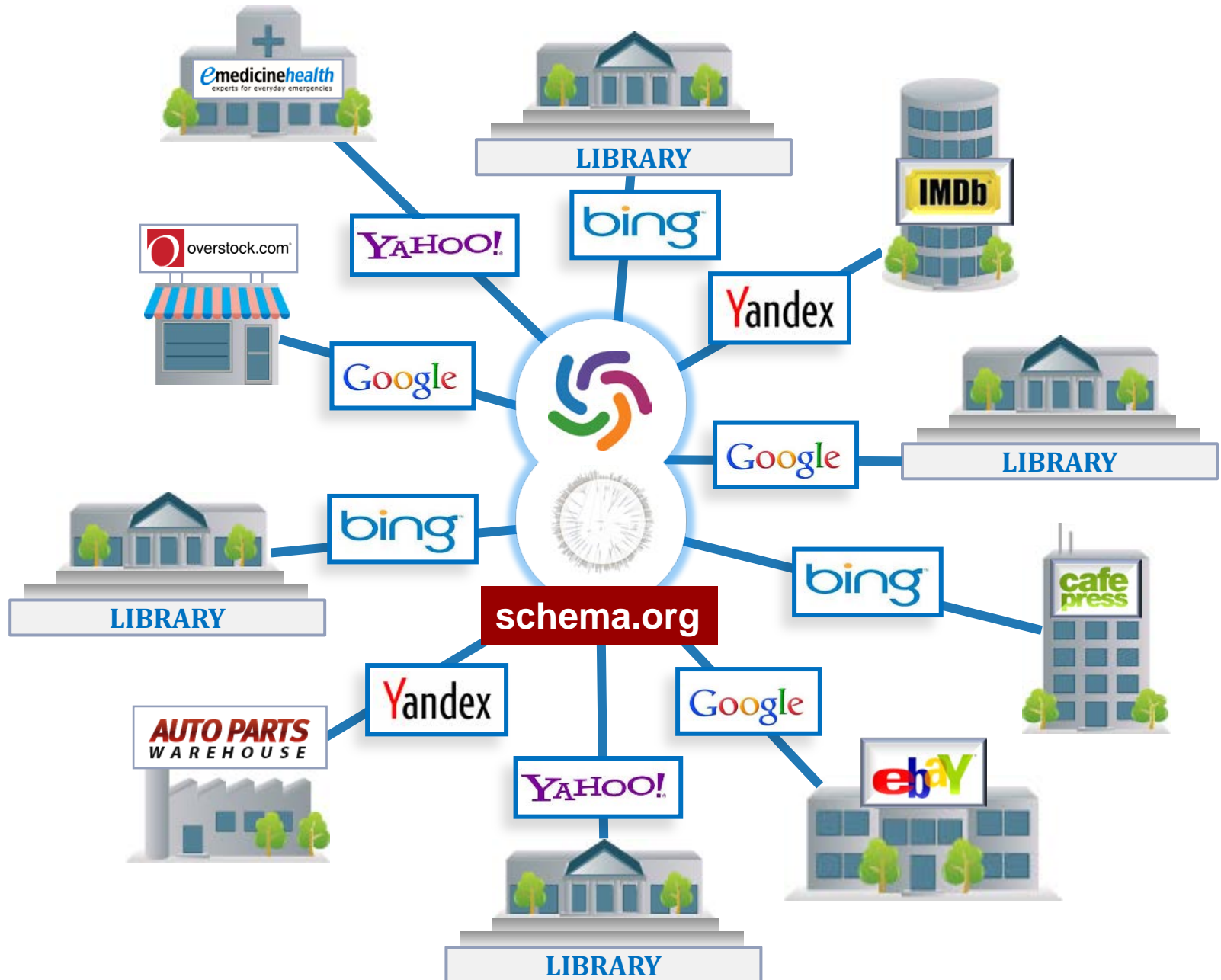
The fundamental question:

How to connect the user to library collections
via metadata on the web?

How do we improve access to our data on the web?

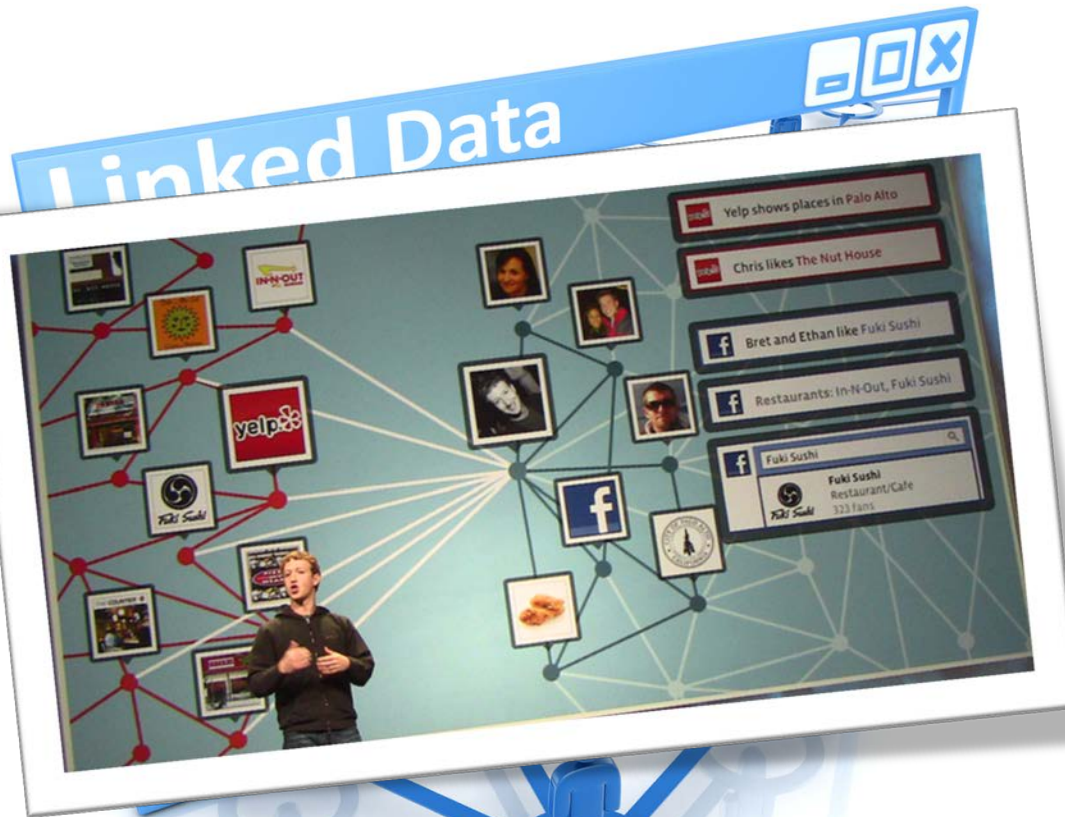
1. Aggregate: Gather all library collections together at the network level so the whole becomes more attractive to the web.
2. Syndicate: Make that data available to the web where the user is
3. Entify: Finally, we must change the way we describe our data to extract what is most interesting to consumers.
4. Manage: Manage all of the data (including articles and chapters), using authority files and controlled vocabularies.
 - Use canonical identifiers (linked data principles)

Syndication



Solutions

Next Generation Cataloging



Cataloging in a linked environment:

- Locating the resource within a network of useful links.

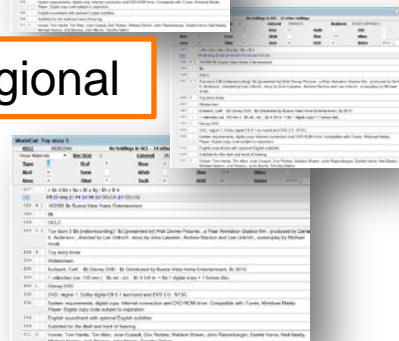


Managing Entities

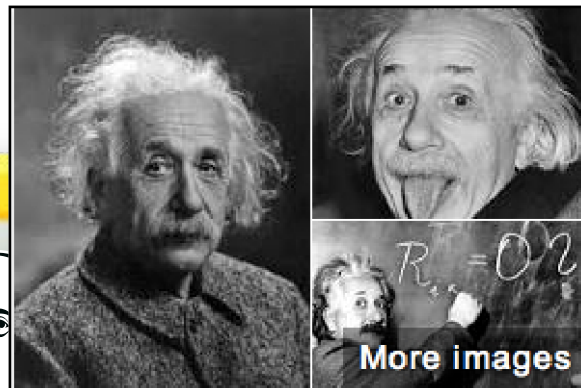
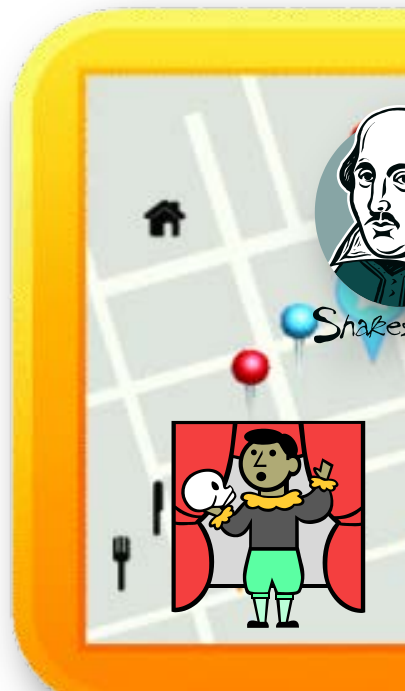


Local
Catalogs

Regional



Global



More images

Albert Einstein

Theoretical Physicist

Albert Einstein was a German-born theoretical physicist who developed the general theory of relativity, one of the two pillars of modern physics. [Wikipedia](#)

Born: March 14, 1879, [Ulm](#)

Died: April 18, 1955, [Princeton](#)

Education: [University of Zurich](#) (1905), [ETH Zurich](#) (1901), Aargau Cantonal School (1895–1896), Luitpold Gymnasium

Children: [Hans Albert Einstein](#), [Eduard Einstein](#), [Lieserl Einstein](#), [Ilse Einstein](#), [Margot Einstein](#)

Spouse: [Elsa Einstein](#) (m. 1919–1936), [Mileva Marić](#) (m. 1903–1919)

Awards: Nobel Prize in Physics, Copley Medal, Max Planck Medal, [More](#)



Metadata Management in the Future

Cataloging

A digital form for cataloging. It has a blue header bar. Below it, there are two input fields: one labeled 'Title:' and another labeled 'Subject:'. At the bottom right of the form, there are two circular buttons, one grey and one green.

Algorithmic Data
Loading & Mining

Progress

Loading Data...



The Bibliographic Framework Initiative

Definition & Goals

News and Announcements



A Bibliographic Framework for the Digital Age (October 31, 2011)


« [Back to News and Announcements](#) | [Download and Print](#) [PDF, 58 KB]

The [Working Group of the Future of Bibliographic Control](#), as it examined technology for the future, wrote that the Library community's data carrier, MARC, is "based on forty-year-old techniques for data management and is out of step with programming styles of today." [1] The Working Group called for a format that will "accommodate and distinguish expert-, automated-, and self-generated metadata, including annotations (reviews, comments) and usage data." [2] The Working Group agreed that MARC has served the library community well in the pre-Web environment, but something new is now needed to implement the recommendations made in the Working Group's seminal report. In its recommendations, the Working Group called upon the Library of Congress to take action. In recommendation 3.1.1, the members wrote:

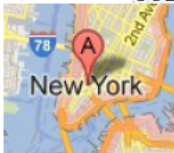


"Recognizing that Z39.2/MARC are no longer fit for the purpose, work with the library and other interested communities to specify and implement a carrier for bibliographic information that is capable of representing the full range of data of interest to libraries, and of facilitating the exchange of such data both within the library community and with related communities." [3]

This same theme emerged from the recent test of the Resource Description and Access (RDA) conducted by the National Agricultural Library, the National Library of Medicine, and the National Library of Health.

Deconstructing MARC – Identifying MARC Resources (MARCR)



001 26132811
 008 920608s1992 nyu 0 eng
 010 \$a92021087
 020 \$a1555701205
 040 \$aDLC\$cDLC\$dAGL
 049 \$aHNKA
 050 00 \$aZ678.9\$b.C58 1992
 082 00 \$a025/.00285\$220
 1 \$aCohn, John M
 10 \$aPlanning for automation :\$ba how-to-do it manual for librarians /\$cJohn M. Cohn, Ann L. Kelsey, Keith Michael Fiels
 260 \$aNew York :\$bNeal Schuman Publishers,\$cc1992
 300 \$avi, 116 p. :\$c28 cm
 440 0 \$aHow-to-do it manuals for libraries ;\$uno. 25
 504 \$aIncludes bibliographical references and index
 650 0 \$aLibraries\$xAutomation\$xManagement\$xHandbooks, manuals, etc
 650 0 \$aLibrary planning\$xHandbooks, manuals, etc
 700 10 \$aKelsey, Ann L
 700 10 \$aFiels, Keith Michael

Presentation Excerpt:

- *Bibliographic Framework Initiative Approach for MARC Data as Linked Data*
- Sally McCallum
- Sept 13, 2012
- IGeLU Conference, Zurich Switzerland

What Does It Look Like?

```
@prefix schema: <http://bibframe.org/> .
```

```
<http://worldcat.org/works/12345>
```

```
  a bf:work;  
  bf:title "war and Peace";  
  bf:subject <http://id.loc.gov/authorities/names/n79054933>;  
  bfp:creator <http://viaf.org/viaf/96987389>;  
  bf:languageofwork "en";  
  bfp:genre "History";  
  bfp:genre "Fiction";  
  bf:description "An epic novel featuring the Russian role in the Napoleonic wars and p  
  bfp:hasInstance <http://www.worldcat.org/oclc/38264520>;  
  bfp:hasInstance <http://www.worldcat.org/oclc/620344>.
```

```
<http://www.worldcat.org/oclc/38264520>
```

```
  a bf:Instance;  
  bf:isbn "9780553479430";  
  bf:isbn 0553479431;  
  bf:title "war and Peace";  
  bf:instanceof <http://exampleworks.org/work/12345>.
```

```
<http://www.worldcat.org/oclc/620344>
```

```
  a bf:Instance;  
  bf:title "war and Peace";  
  bf:subject <http://id.loc.gov/authorities/names/n79054933>;  
  bfp:creator <http://viaf.org/viaf/96987389>;  
  bfp:genre "History";  
  bfp:genre "Fiction";  
  bf:description "An epic novel featuring the Russian role in the Napoleonic wars and pr  
  bf:instanceof <http://exampleworks.org/work/12345>
```

Back to Assessment...

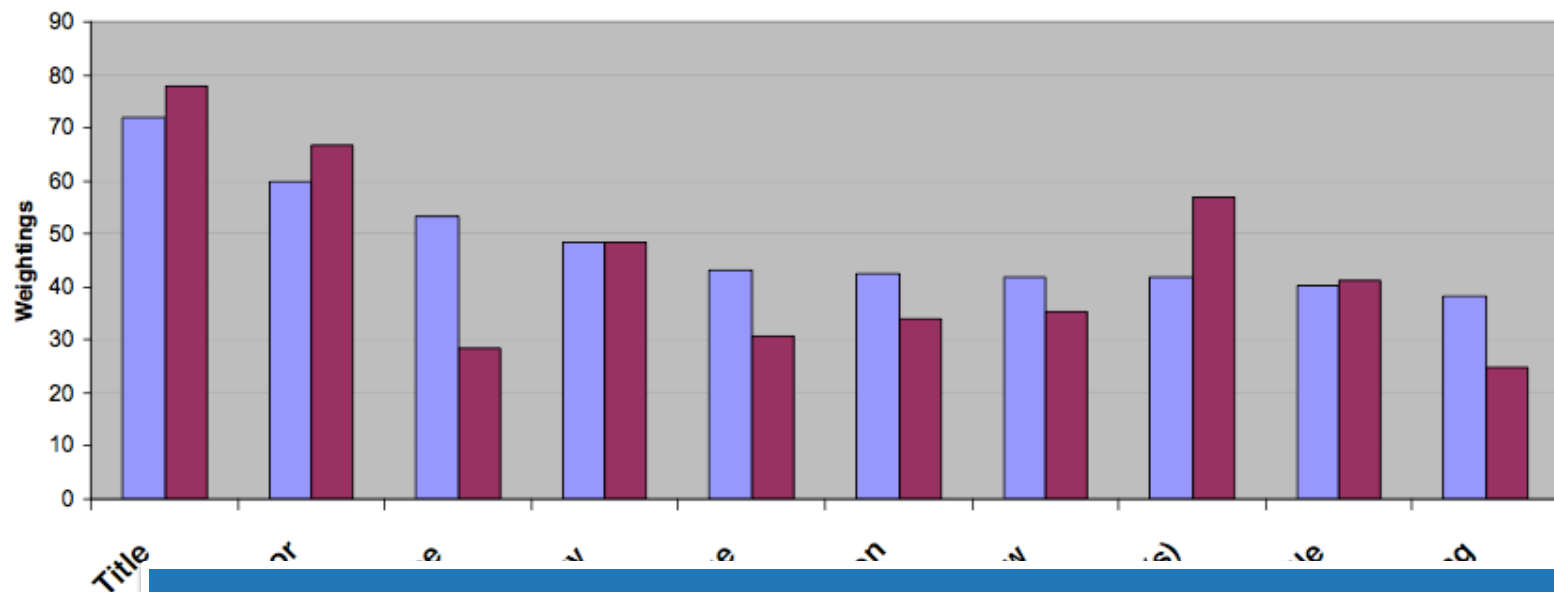
“Evidence-Based Cataloging”



“[Catalogers] need to practice *evidence-based cataloging*. They need to catalog based on the evidence that they can find for the effectiveness of particular practices, and they need to judge their output according to this evidence.”

Hilder, Philip and Tan, Kah-Ching. 2008. Constructing Record Quality Measures Based on Catalog Use. *Cataloging & Classification Quarterly* 46:4.

Top Ten Elements for Identification and Selection at NLB, Compared to SLV Ratings for Those Elements



Evidence-based cataloging argument:
Focus on the elements of the record
that are highly relevant to retrieval.

What is being assessed?

Assessment in Practice

- * Improve effectiveness
- * Identify areas for improvement
- * Communicate with customers
- * Communicate with administration
- * Lower costs
- * Help with decision making

Effectiveness

Customer Satisfaction

Quantitative Assessment

Heads of Cataloging Interest Group – ALA Midwinter, 2013

- Total project output

- Project efficiency

Costs

- Individual cataloger

Efficiency

- Individual cataloger efficiency

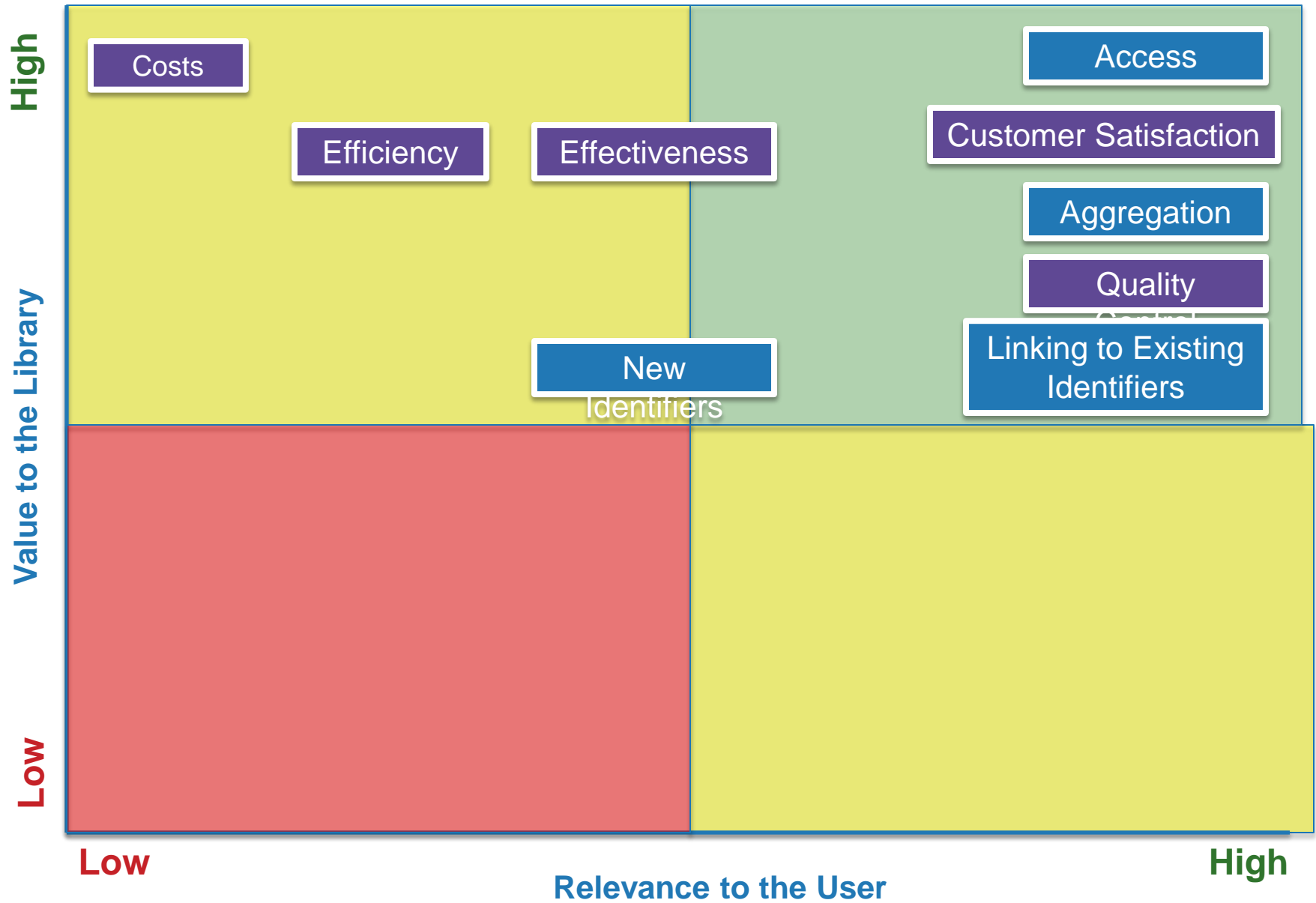
- Quality Control

Quality

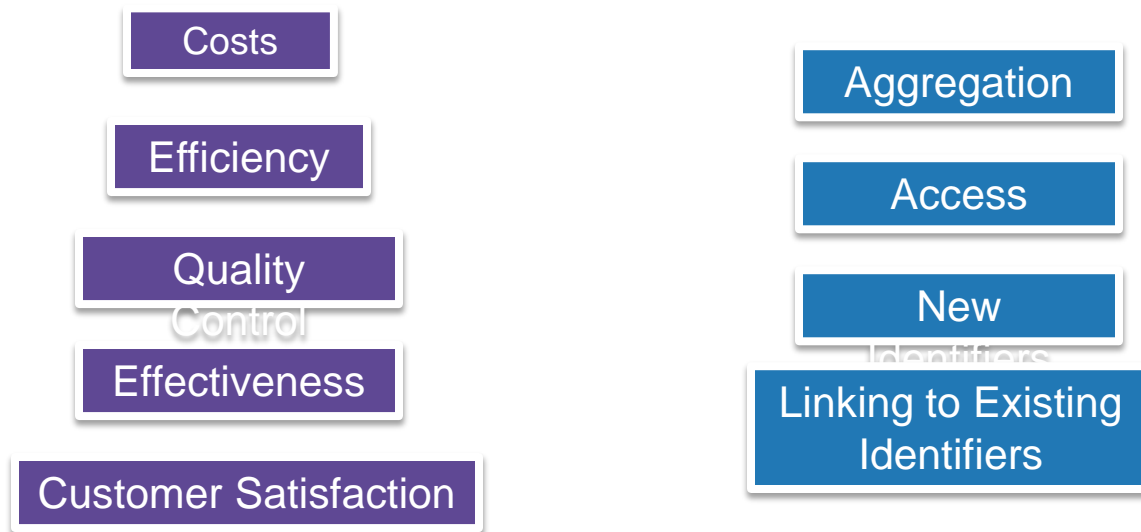
Control

How do we improve access to our data on the web?

1. Aggregate: Gather all library collections together at the network level so the whole becomes more attractive to the web. Aggregation
2. Syndicate: Make that data available to the web where the user is Customer Satisfaction Access
3. Entify: Finally, we must change the way we describe our data to extract what is most interesting to consumers. New Identifiers
4. Manage: Manage all of the data (including articles and chapters), using authority files and controlled vocabularies.
 - Use canonical identifiers (linked data principles) Quality Control Linking to Existing Identifiers



What to Assess?



- ❖ The goal of the next generation of metadata management initiatives is to improve access to library collections where the user is: on the web
- ❖ Is it possible to align cataloging assessment to that goal?

Answer: Yes

Aggregation

New

Identifiers

Access

Linking to Existing
Identifiers

How?
Measure different things

Ted Fons
fonst@oclc.org