

CIC metadata aggregation: A collaborative initiative to enhance metadata sharing

Timothy W. Cole (t-cole3@uiuc.edu)
Muriel Foulonneau (mfoulonn@uiuc.edu)

Grainger Engineering Library
University of Illinois at Urbana-Champaign

*Charlottesville, VA
8 November 2005*





The project

- Sept. 2003 – June 2006
- All 13 CIC institutions are participating
 - Currently 19 data providers with 2 more repositories in test
 - Aggregation contains 520,000 resource descriptions
- Objectives
 - Support CIC mission of regional collaboration
 - Investigate issues surrounding metadata shareability
 - Test OAI – PMH as technical infrastructure for metadata sharing
 - Provide Web interface(s) to CIC aggregated material





Context of & motivations for collaboration

- Follow-up to Mellon-funded experiments with OAI
 - Had shown need for collaborations to build competencies & capacity and serve as catalyst to encourage institutions to implement OAI (M. Halbert et al., ECDL 2003)
 - 2 CIC institutions (Illinois & Michigan) had been involved
 - Homogeneity of CIC institutional missions provided opportunity to explore high level of interoperability
- Participants shared need to enhance discoverability and reusability of library-published digital & digitized content
- Attractiveness of shared funding model





Key Accomplishments

- Technical Validation of CIC OAI-PMH implementations
- Advances in metadata sharing
 - created incentive for data providers to modify the way they share metadata by showing the direct impact possible
 - increased understanding of DP & SP roles; demonstrated (again) dependence of each on the other
 - provided input to broader community initiatives – e.g., DLF-NSDL best practices for OAI and shareable metadata
 - demonstrated critical importance of retaining context and providing reliable, direct linkages to resources





Evolution in CIC metadata sharing

- Aggregation item count has grown by 50% so far during project
- Broader adoption of optional protocol features
 - More metadata formats
We reprocess UDC, QDC and MODS records for items;
MARC and DC-Collection records for collections
 - Increased use of OAI sets & <setDescription> element
Inclusion of collection-level descriptions within <setDescription>
Now harvesting set descriptions and transforming them as
collection level descriptions
- More consistent & robust links to resources being described
 - Enhanced user experience
 - Value-added services (e.g., automated capture of thumbnails)





CIC collaboration as a test bed for DPs

- Data providers are interested in impact and importance of what they do or don't do
 - Do I need to use resumption tokens?
 - How might including collection-level descriptions enhance the discoverability of my content?
 - What are the benefits of exposing metadata in multiple formats?
 - Which URL(s) for the resource should be shared in the metadata records?





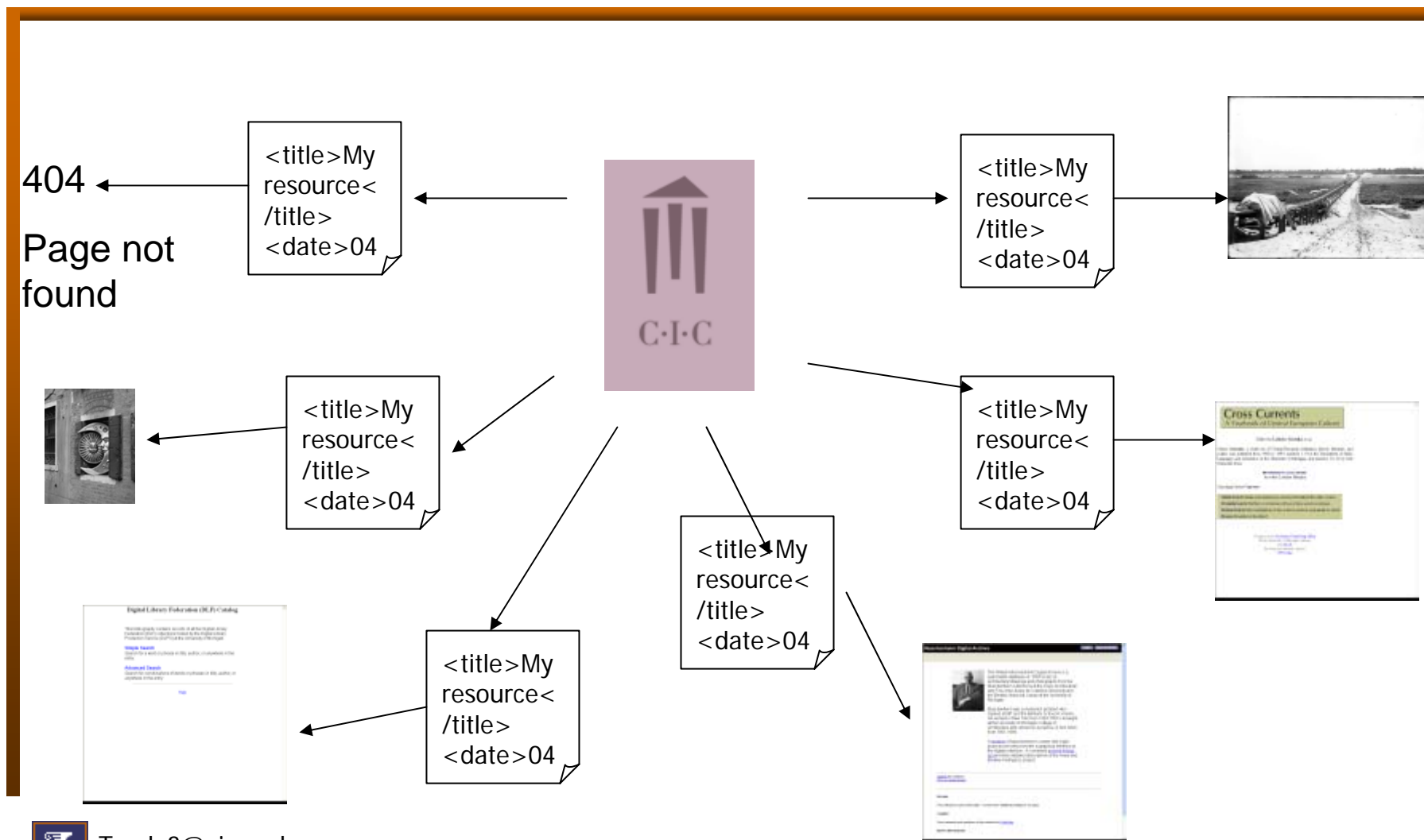
How much normalization can SP do?

Concept	% of records in the repository
Type	99%
Format	66%
Language	60%
Collection	100%
Resource URL	72%
Temporal Coverage	71%
Spatial Coverage	21%

These results may or may not be generalizable for other SPs



Examining resource URLs embedded in metadata



Generating or including thumbnails

CIC-OAI Metadata Search Portal

Home
Search
View Institutions
Help

Your search was in **all fields** for **"russian"**. The search was limited to the **"University of Illinois Library DIMTI"** archive.

You found **83 records**.

- * [Revise your search](#) to retrieve fewer records.
- * View your results, starting with **records 1 to 25 of 83**.

Results by Institution

Indiana University:
Indiana University Digital Library Program
[90 records](#)

Michigan State University:
Michigan State University Digital and Multimedia Center
[52 records](#)

Ohio State University:
Ohio Link Electronic thesis and dissertations
[10 records](#)

Pennsylvania State University:
Pennstate Libraries Image Collections
[3 records](#)

Pennstate thesis and dissertations
[2 records](#)

University of Chicago:
University of Chicago Library
[4 records](#)

University of Illinois at Urbana-Champaign:
Engineering Documents Center Collection
[1 record](#)


University of Illinois Finding Aids
[49 records](#)

University of Illinois Sheet Music
[24 records](#)


University of Illinois Library DIMTI
[83 records](#)

[Next 10 Records](#)
[Jump to Records: 1 | 26 | 51 | 76](#)

Record 1 of 83

Title	 <p>Dievochka pokupaet yarenye grushi. Une petite fille a chetant des poires. Ein kleines mÄdchen kauft Birnen.</p>
Author/Creator	(unknown)
Description	(unknown)
URL	http://images.library.uiuc.edu:8081/RussianPublics/image/97752201452003_fonar60-5.jpg
Collection	Early 19th Century Russian Readership & Culture

Record 2 of 83

Title	 <p>Dvornik i Trubochist. Le portier et le ramoneur. Der Hausknecht und der Schornsteinfeger.</p>
Author/Creator	(unknown)
Description	(unknown)



Using CLD to enhance discoverability

Menu

[Home](#)

Search the metadata portal

Show collections only

- ▶ [CIC keyword search](#)
- ▶ [CIC geographic browse](#)
- ▶ [CIC collection browse](#)

Project details

- ▶ [Project presentations & reports](#)
- ▶ [CIC collection development](#)
- ▶ [CIC Metadata guidelines](#)
- ▶ [Schemas and crosswalks](#)
- ▶ [Metadata aggregation](#)
- ▶ [OAI Content Providers](#)
- ▶ [Institutions participation](#)
- ▶ [The DLXS software](#)
- ▶ [Resources](#)

Contacts

[Partner section](#)

[Show all results](#) ✓

page 1 / 7 [Go to page 1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [Next page >>](#)

title [Representative American poetry, edited by William Stanley Braithwaite \[1\] & Henry Thomas Schnittkind.](#)

creator Braithwaite, William Stanley, 1878-1962. ed.

type text

Collection Making of America (MOA) Books - *University of Michigan*

[More info](#)

title [Index to American poetry and plays in the collection of C. Fiske Harris.](#)

creator Harris, Caleb Fiske, 1818-1881.

type text

Collection Making of America (MOA) Books - *University of Michigan*

[More info](#)

title [Poets and poetry of Vermont.](#)

creator Hemenway, Abby Maria, 1828-1890.

type text

Collection Making of America (MOA) Books - *University of Michigan*

[More info](#)

[9541 matches like this ...View](#)

title [The Automobile and American Poetry](#)

creator Goldstein, Laurence
Laurence Goldstein

type serial

Collection Michigan Quarterly Review Archive - *University of Michigan*

[More info](#)

title [Arguing \(What Else?\) about Jewish-American Poetry](#)

creator Laurence Goldstein

type Book Review

See relevant collections

▶ **[Database of American Poetry](#)**
University of Michigan
The Database of American Poetry features the works of more than 200 American poets, along with six landmark anthologies of American poetry. The database gathers the works of the most influential American poets, from the Colonial period to the early and [1288 items matching](#)

▶ **[Database of Twentieth Century American Poetry](#)**
University of Michigan
The Database of Twentieth Century American Poetry contains modern and contemporary American poetry from the early twentieth century to the present. It includes the works of most major poets of the twentieth century, beginning with the and [475 items matching](#)

▶ **[Database of Twentieth Century African-American Poetry](#)**
University of Michigan
The Database of Twentieth Century African-American Poetry includes collected poems and individual volumes from all the major movements and schools of twentieth century African American poetry from 1902 to the present day. and [173 items matching](#)

▶ **[Database of African-American Poetry](#)**
University of Michigan
The Database of African-American Poetry covers the works of 54 African-American poets writing in the eighteenth and



What we've not been able to do (yet)

- Adequate involvement of end-users – will require
 - Better integration with traditional & commercial resources
 - Development of focused portals --> manual classification or automatic clustering across institutional boundaries
 - Normalization / cross-walking across vocabularies; i.e., repurposing metadata for new use

- Bring collaborators to the same level of sophistication
 - Wide range of technical solutions / dependence on vendors
 - Wide range of descriptive practice & richness (considerable static content not being further developed)





References

- Halbert, M., Kaczmarek, J., Hagedorn, K., Findings from the Mellon Metadata Harvesting Initiative, ECDL 2003
<http://www.springerlink.com/app/home/contribution.asp?wasp=6590eb7adfc74a07acf9a25e6c518fcd&referrer=parent&backto=issue,7,47;journal,780,2099;linkingpublicationresults,1:105633,1>
- Foulonneau, Muriel and Timothy W. Cole. Strategies for reprocessing aggregated metadata. *ECDL 2005*
- Foulonneau, Muriel, Timothy W. Cole, Thomas G. Habing, Sarah L. Shreeves (2005). Using Collection Descriptions to Enhance an Aggregation of Harvested Item-Level Metadata. In JCDL 2005
- CIC metadata portal <http://cicharvest.grainger.uiuc.edu>
- DLF-NSDL best practices for OAI and shareable metadata wiki

