Managing the Collective Collection: library print collections in the networked environment

Constance Malpas

Program Officer

RLG Programs

ARL RLLF @ OCLC

Dublin, OH

1 November 2007

Moving 'into the cloud'

...the next wave of corporate productivity gains should be paced by Web 2.0 driven collaboration tools that use *the network as the platform* to enable users to connect 'any device to any content over any combination of networks'

John Chambers, CEO, Cisco Systems, (May 2007)

As quoted by Mary Meeker, Morgan Stanley technology analyst, at the Web 2.0 Summit in San Francisco, October 2007

Moving '*into the cloud*': modes of transport

For some, the ascent is easy.

"...*he passed with the cloud from the sight of men*"

The Ascension of Christ
Hans Suess von Kulmbach (German, born about 1480, died 1521/22)
Rogers Fund, 1921 (21.84)
The Metropolitan Museum of Art

Moving '*into the cloud*': modes of transport

In most cases, special assistance is needed.

The Glory of Saint Marziale by Sebastiano Ricci (1659-1734); Saint Marziale, Venice, Italy; Cameraphoto/Art Resource, New York
Ceiling fresco in the Church of San Marziale. Venice, Italy

OCLC Programs & Research Themes

Managing the Collective Collection

Renovating Descriptive & Organizing Practices

Modeling New Service Infrastructure

Supporting New Modes of Research, Teaching & Learning

Architecture & Standards

Measurements & Behaviors

The RLG Programs approach:

- *Assess the landscape*: capture community priorities and trends through face-to-face encounters, working groups, interviews and data analysis
  - Group meetings, conference presentations, white papers

- *Strategize:* determine where community investment is most likely to contribute near-term value
  - "Calls to action", internal and external advisory groups, participation in internal product planning

- *Collaborate:* work with partners to implement change at the local, group and global level
  - Formulate policies, prototype services, partner with external agencies to leverage collective assets
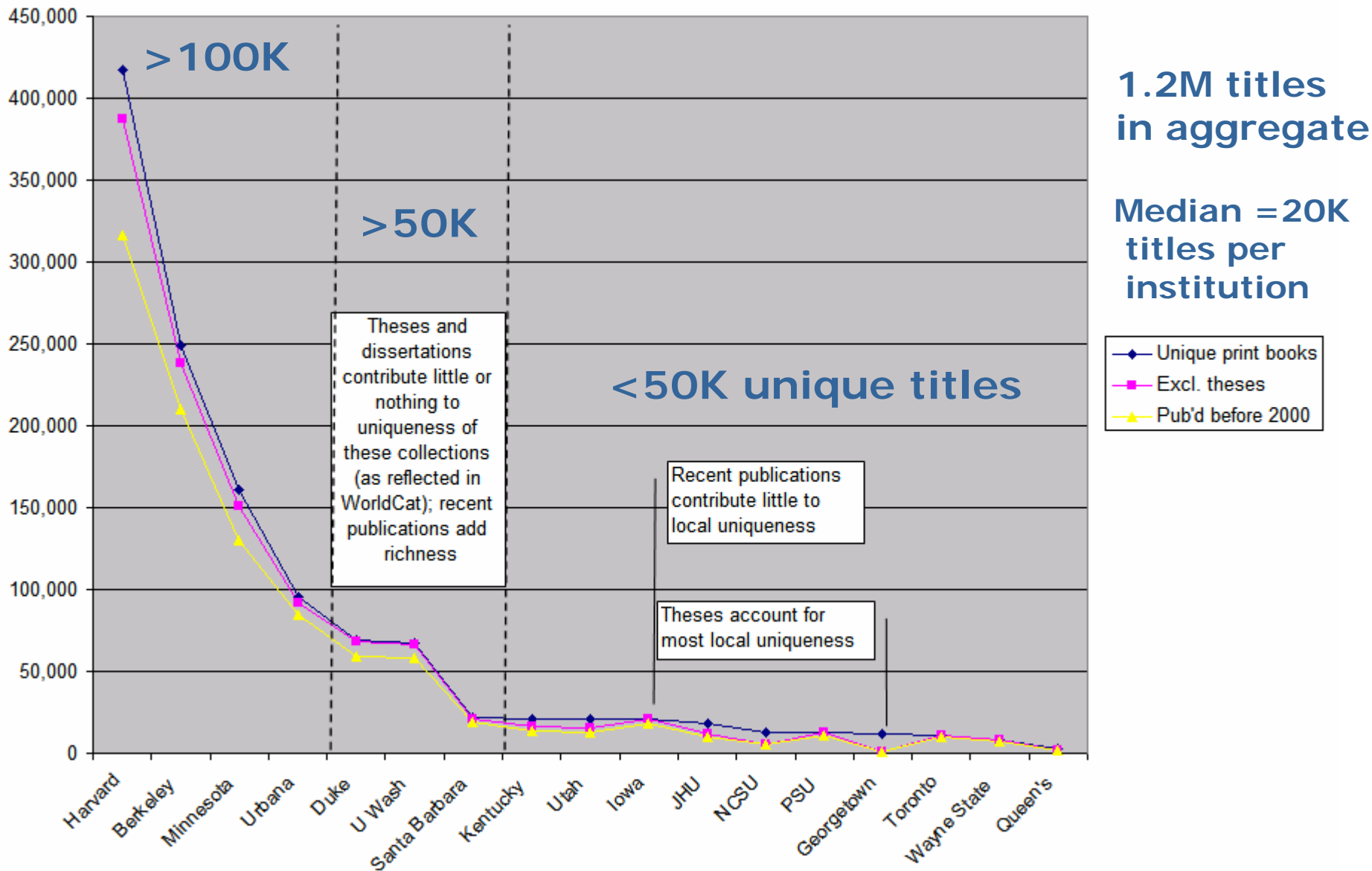
Last Copies - Managing Risk

Assessing and <u>characterizing 'uniqueness'</u> in the aggregate book collection, enabling institutions to make better-informed decisions about the <u>disposition of local holdings</u>

- Content characterization: WorldCat sample – English cataloging only
  - Material and format/genre types: theses, gov't docs, pamphlets, conference proceedings, literature etc.
  - Unique works vs. unique manifestations; extending <u>previous work</u>
  - Ghosts: local holdings no longer exist
  - Watch for findings in November/December
- Geographic and institutional distribution: ARL community
  - 6.9 M unique print book titles held by 123 ARL institutions in North America
  - 14% of unique holdings were published after 1999 – uniqueness of 'first copies' subject to erosion
  - 7% of unique holdings represent theses and dissertations – intrinsically unique content
  - Distribution patterns vary by tier, suggesting shared collection priorities and (perhaps) opportunities for collaboration

Research lead: Ed O'Neill
Programs lead: Constance Malpas
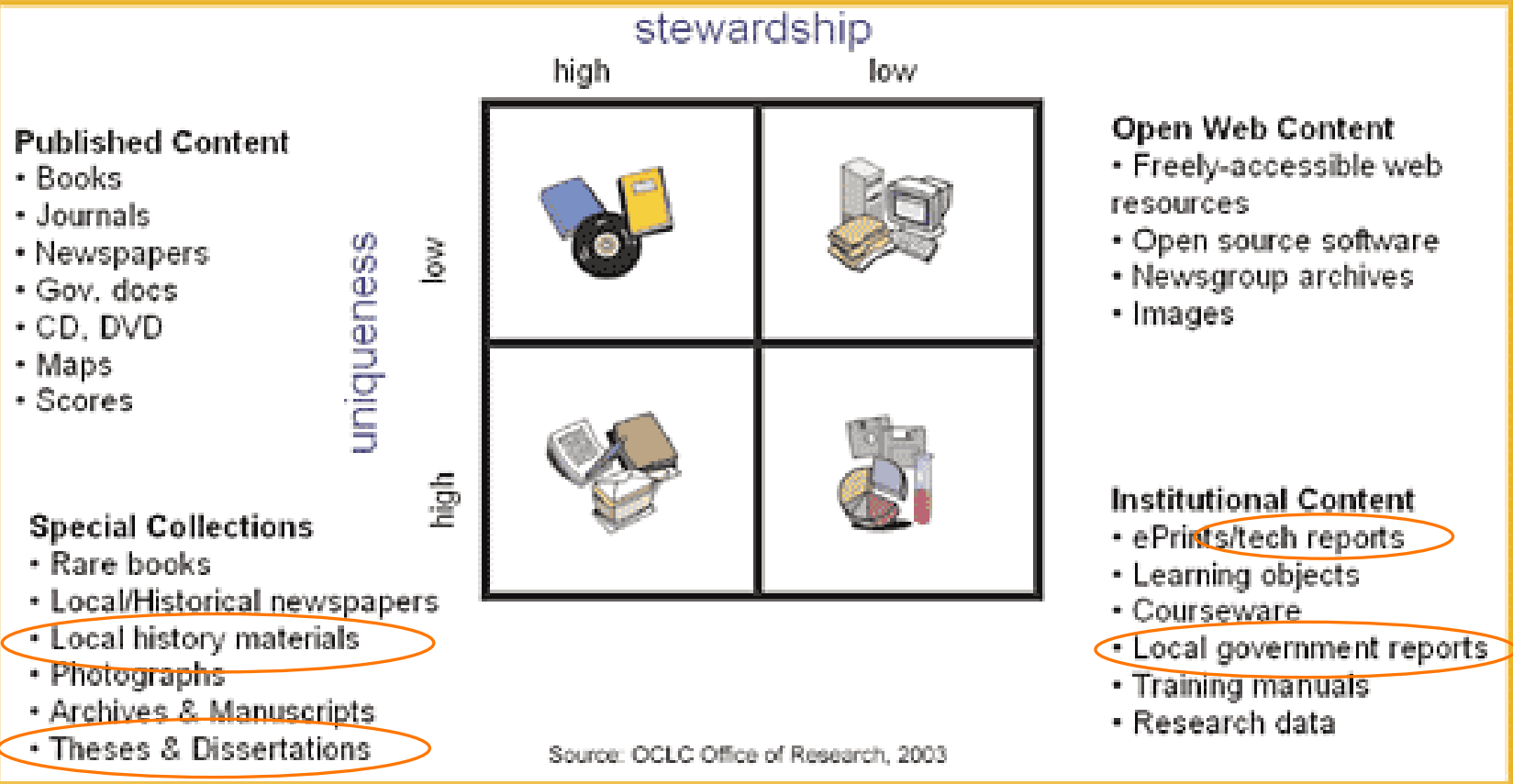
Distribution of Unique Titles: ARL RLLF sites

1.2M titles in aggregate

Median =20K titles per institution

Legend:
- Unique print books
- Excl. theses
- Pub'd before 2000

Chart labels:
- >100K
- >50K
- <50K unique titles

Text boxes:
- Theses and dissertations contribute little or nothing to uniqueness of these collections (as reflected in WorldCat); recent publications add richness
- Recent publications contribute little to local uniqueness
- Theses account for most local uniqueness

X-axis institutions: Harvard, Berkeley, Minnesota, Urbana, Duke, U Wash, Santa Barbara, Kentucky, Utah, Iowa, JHU, NCSU, PSU, Georgetown, Toronto, Wayne State, Queen's

Last Copies – Content Characterization

Collections Grid

A framework for representing content

stewardship

high   low

Published Content
- Books
- Journals
- Newspapers
- Gov. docs
- CD, DVD
- Maps
- Scores

uniqueness

low

high

Special Collections
- Rare books
- Local/Historical newspapers
- Local history materials
- Photographs
- Archives & Manuscripts
- Theses & Dissertations

Open Web Content
- Freely-accessible web resources
- Open source software
- Newsgroup archives
- Images

Institutional Content
- ePrints/tech reports
- Learning objects
- Courseware
- Local government reports
- Training manuals
- Research data

Source: OCLC Office of Research, 2003

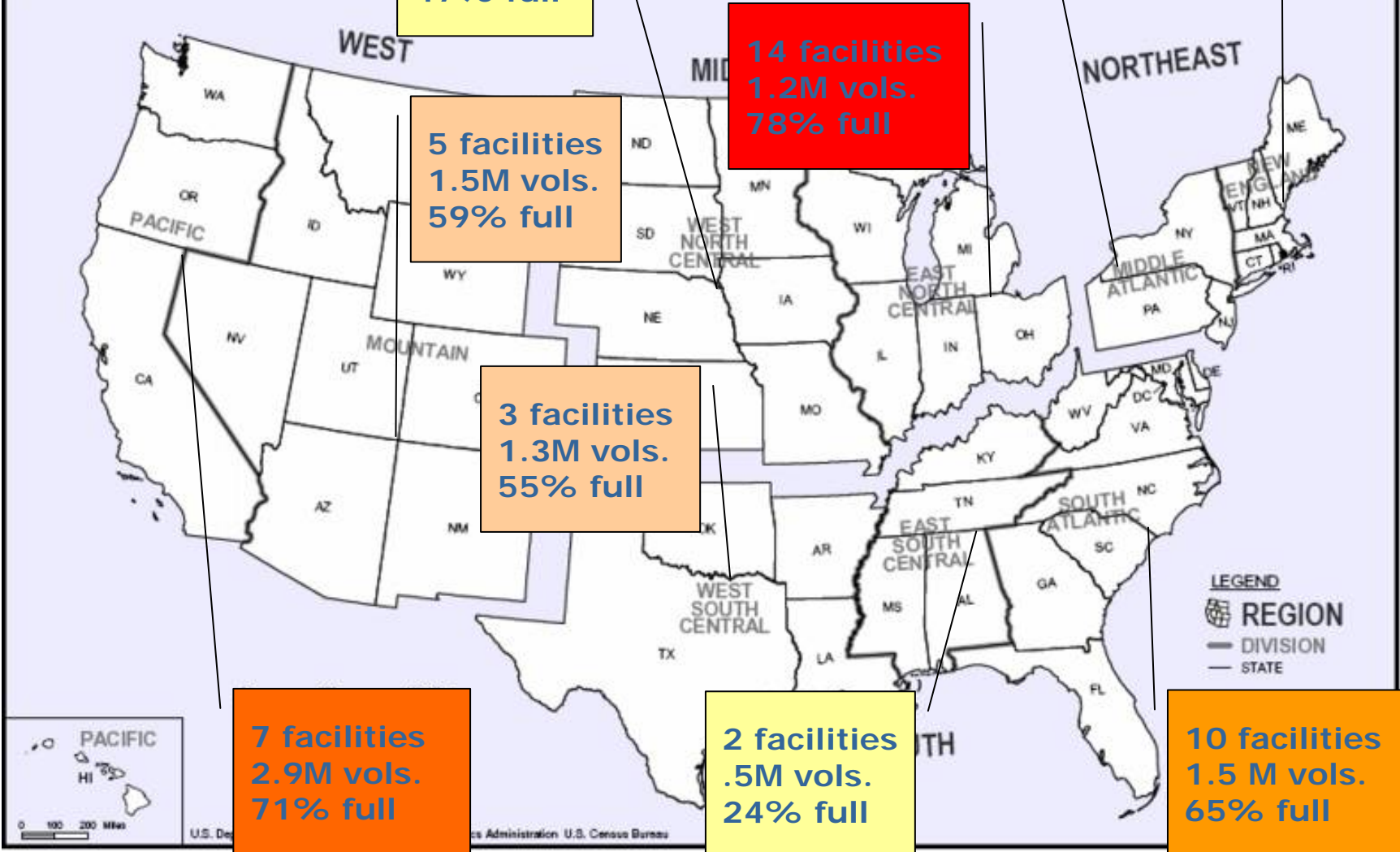# Shared Print Collections

Promoting <u>effective disclosure and management</u> of aggregate holdings, moving institutional holdings and services 'into the flow', <u>building library Web scale</u> on a global scale

- <u>North American Storage Trust</u> – community engagement
  - Leverage existing network of off-site storage collections to manage system-wide redundancy while reducing local costs
- Collaborative Collection Management Trust – pilot project
  - Explore use of aggregated holdings and circulation data to support local and group decision-making about growth and management of library print collections
- "<u>Library Offsite Storage and the Future of Library Print Collections</u>" - commissioned white paper by Lizanne Payne
  - Evidence-based assessment of library storage practices in North America with recommendations for future action

Programs lead: Constance Malpas
Product management lead (CCMT): David Millikin

System-wide library storage capacity

Collections and Divisions in the United States

PACIFIC (inset map: AK)

1 facility
.5M vols.
47% full

9 facilities
2M vols.
69% full

5 facilities
1.7M vols.
41% full

14 facilities
1.2M vols.
78% full

5 facilities
1.5M vols.
59% full

3 facilities
1.3M vols.
55% full

7 facilities
2.9M vols.
71% full

2 facilities
.5M vols.
24% full

10 facilities
1.5 M vols.
65% full

LEGEND
REGION
DIVISION
STATE

Mass Digitization

Exploring the challenges of <u>bringing library content online</u> in an environment where <u>discovery is mediated by large information hubs</u>; understanding the <u>impact of large-scale text aggregations</u> on scholarship and learning

- Model agreements for public/private partnerships
  - Invitational meeting held in New York, June 2007
  - Explored the costs and benefits of partnering with commercial entities to bring library collections online; pooled expertise on negotiating tactics
  - Final report to appear in D-Lib magazine, November/December 2007
- Large-scale conversion of special collections
  - Standing-room-only pre-conference at SAA (August 2007), hosted by Newberry Library, focused on how to move more special collections online, more quickly
  - Culminated in "Shifting Gears" – a recommendation to adopt a digitization mandate driven by access, unleashed from traditional selection and cataloging workflows, and focused on moving content to the Web

Programs lead: Ricky Erway

Shifting Gears, Getting into 'the Flow'
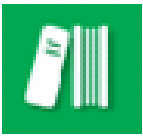
Shifting Gears: Gearing Up to Get Into the Flow

The essay, like the forum, focuses on digitization and related processes, but intentionally does not encompass technical specifications for various formats, born digital materials, nor rights issues (which warrant similar essays for each topic). It intends to be provocative. Not all of the ideas presented here will apply to a particular situation, but hopefully they will stimulate consideration of appropriate ways to move forward.

Special collections are stuck in an eddy, while the mass of digitized books drift by in the current of the mainstream. We need to jump into the flow or risk being left high and dry.

1. Access vs. preservation—Access wins!

Many of our digital initiatives have stressed the importance of preservation, leaving access as an afterthought (the idea being if you capture preservation-quality; you can always derive an access copy). In reality, due to the very special nature of these often unique materials, we will always preserve the originals to the best of our ability. In light of recent programs for the mass digitization of books, if special collections and their funding continue to be marginalized, our administrations may not keep us around to attend to the originals.

In the past, we've soothed our doubts by repeating the mantra, "we'll only get one chance to do it, so it's got to be done right." Experience has shown that that is not in fact the case. Often we do go back when the technology improves or when we better understand our users' needs. We need to put on our helmets now and go for the biggest bang for the buck in terms of access.

Collaborative Collection Management

Leveraging the <u>networks of trust</u> that we've built over 15 years of collaboration in the resource-sharing community, to move collection-sharing practices '<u>beyond borrowing</u>'

- <u>RLG Collections Summit</u> (November, 2007)
  Host: University of Pennsylvania
- Invitational meeting will explore cultural, technological and organizational obstacles to the cooperative solution of 4 key objectives
  - Retention of last copies
  - Ensuring back-up to online access
  - Expanding coverage, reducing duplication
  - Shared access to low-use collections
- Collections and Access managers from 34 research institutions in North America, the UK and Ireland
- Outcomes: recommended strategies at the local, group, and network levels; concrete next steps and "assignments"

Programs lead: Dennis Massie

"Shared access to low-use collections"

Approaches	Obstacles	Measures of Success	
▪ physical consolidation ▪ 'virtual' consolidation	▪ loss of ownership ▪ loss of browsing	Collections ▪ increased coverage, reduced cost	Access ▪ direct, unmediated delivery
Strategies ▪ fully disclose local and group storage collections and services ▪ implement single-copy policy to de-duplicate at ingest ▪ monitor system-wide usage to establish optimal overlap			
Local ▪ policies	Group ▪ coordination	Global / Network ▪ registries	

Solutions

RLLF Discussion Topics

- North American Storage Trust and the Cooperative Collection Management Trust pilot

- Moving SHARES 'beyond borrowing'

- OCLC/WorldCat's role in shared collection building, acquisitions and management

- Collection sharing beyond the library