

D4

Data Federation Of University Research

A Roadmap for Regional Collaboration CNI – December 10, 2013

Deborah Ludwig, University of Kansas
Ann Campion Riley, University of Missouri, Columbia

Partners & Funding



- IMLS National Leadership Planning Grant



- University of Kansas, Grant Leadership



- Great Plains Network



- Greater Western Library Alliance

Greater Western Library Alliance

Joni Blake, Ph.D. Executive Director

- Consortium of 33 research libraries.
- Focus on cost-effective and high-quality information services and resources.



GWLA provides national and international scholarly leadership for learning and research

The Great Plains Network

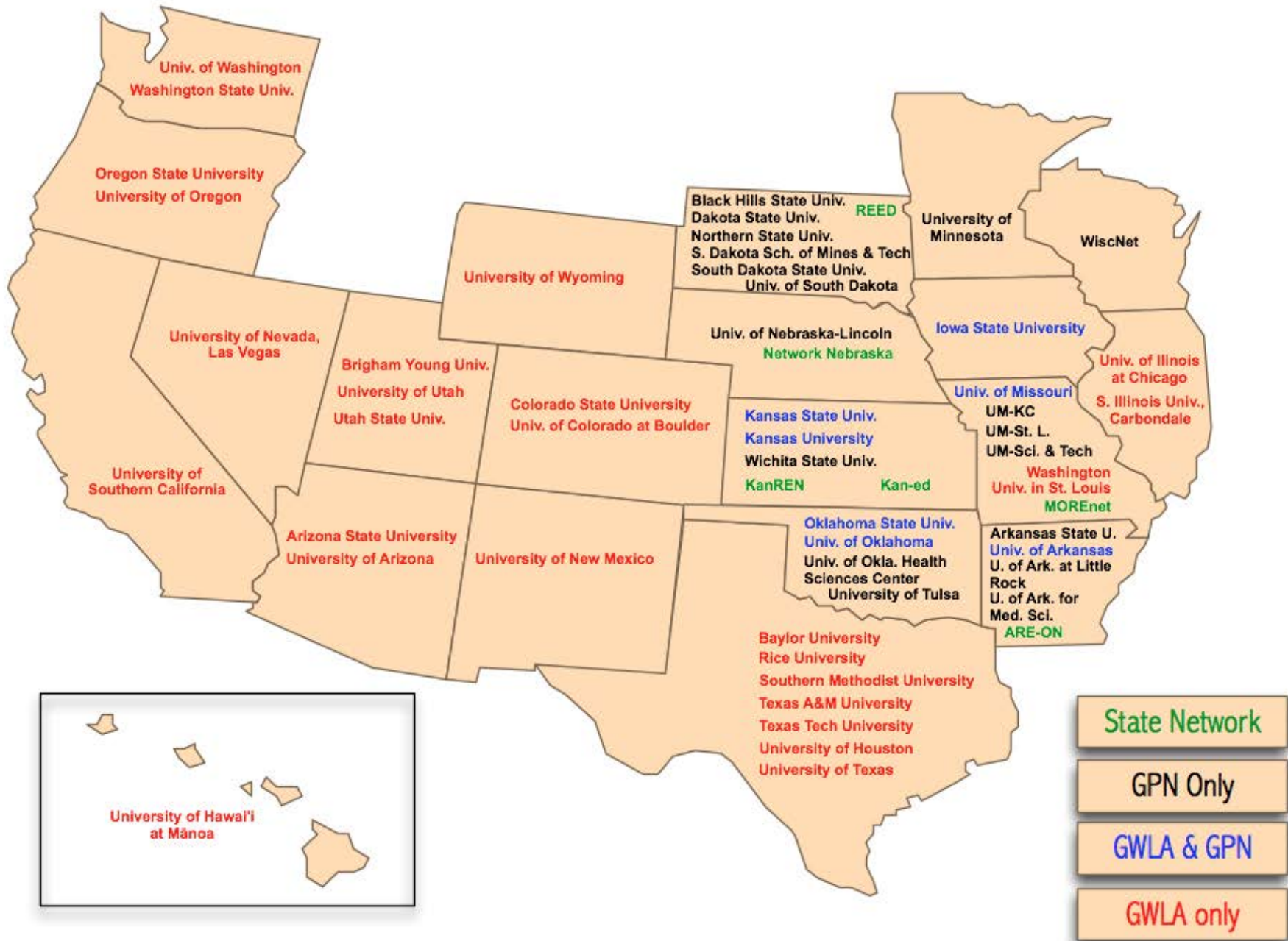
Greg Monaco, Ph.D.
Director for Research & CI



- Consortium of over 20 universities in the Midwest, covering 8 states
- Partners to facilitate the use of advanced cyberinfrastructure (Gigabit point of presence – gigaPOP, for Internet2, network, storage, computation) for research computing

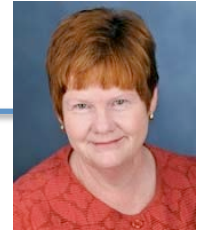
GPN mission to provide outreach to the research & education community

Geographic Membership, Great Plains Network/Greater Western Library Alliance



The University of Kansas

- **Deborah Ludwig, Assistant Dean, KU Libraries** (principle investigator)
- **Bob Lim, Chief Information Officer**
- **Paul Farran, IT Chief of Staff**
- **Scott McEathron, Head of the Center for Graduate Initiatives & Engagement**
- **Nikki Potter, Grant Project Coordinator**



Grant Steering Committee

Remaining Members of the Steering Committee:

Adrian Alexander, Dean of Libraries, University of Tulsa



Judy Ganson, Director of Collection Management Services & Systems, University of Arkansas Libraries


Rick McMullen, Director, Arkansas High Performance Computing Center



Ann Champion Riley, Associate Director of the Access, Collections, and Technical Services Division, University of Missouri Libraries



The IMLS Grant - Teams



Advisory Council	<ul style="list-style-type: none">• Deans• CIOs
Steering Committee	<ul style="list-style-type: none">• GPN• GWLA
Working Groups	<ul style="list-style-type: none">• Research Teams• Strategy Groups

The IMLS Grant

Purpose

To leverage collective strengths and create a proposal for a scalable and federated approach to the lifecycle management of research data based on the needs of GPN and GWLA member institutions.

Grant: 3 Goals

Goal #1: in-depth environmental scan [Oct 2-2012 – April 2013]

Goal #2: GPN and GWLA member forum and two-day workshop focused on understanding challenges and solutions in managing, sharing, and preserving research data. [May 2013]

Goal #3: Plan for a scalable multi-institutional approach to research data management to support. [June 2013 -]

Environmental Scan (Goal 1)

1. Literature review
2. Technologies & Projects review
3. Survey of member institutions

The IMLS Grant – Lit. Review

1. General works on data lifecycle issues
2. Assessments of Researcher behavior, attitudes and needs
3. Services, roles and responsibilities
4. Sharing, reuse, publication and citation
5. Data management planning
6. Policies and standards
7. Institutional repositories, approaches, and issues
8. Disciplinary or subject repositories, approaches, and issues
9. Federated approaches
 - National and regional collaborations
 - International collaborations
10. Economics of data curation and conceptual and theoretical frameworks and approaches
11. Archiving and preservation
12. Metadata and description

(available from the **D4** web site, <http://imls.gwla.org>)

Technical Reports

- Major Research Data Initiatives and Approaches to Curation
- Working Repository Examples (Institutional or Disciplinary)
- Repository Technologies
- MetaData Services & Schema
- Data Discovery / Access / Sharing
- Storage
- Preservation Services / Systems

(available from the **D4** web site, <http://imls.gwla.org>)



IMLS - KU GWLA GPN Survey

Lifecycle Management of Research Data

Some universities provide services to support the use and management of research data. Please indicate by checking any or all boxes that apply what research data services are available at your institution and who is providing that service.

General Support and Services for Research Data

	University Library	Central IT	University Office of Research	Department or Research Center	Other	Not Available	Don't Know
2.01) Provides assistance to researchers in developing data management plans for research proposals	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.02) Offers formal training (e.g. workshops or courses) for researchers on data management practices	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Survey Highlights

- Spring, 2013 to all GWLA / GPN institutions
 - multiple contacts (Dean, CIO, VPR)
 - 23 responses, 20 institutions
- Most common roles for Libraries
 - Data management plans & consultation on practices
 - Training & Workshops
 - Dedication of resources (staff, funding)

Survey Highlights

Most common roles for Libraries

- Data management plans & consultation on practices
- Training & Workshops
- Dedication of resources (staff, funding)
- Consulting on options for data licensing
- Consulting on data archiving (metadata, deposit)
- Identifying [remote] data repositories for archiving

Survey Highlights

Most common roles for Office of Research

- Ensure university compliance for research data in accordance with commercial licenses, government regulations, and funding agency mandates

Survey Highlights

Most common roles for IT

- Short & long term data storage
- High performance computing environments

Survey Highlights

Most commonly NOT offered by university:

- Repository on site to store metadata and data together

Survey Institutional Challenges

FUNDING FOR SUSTAINABILITY OF EFFORTS

- *Business model for costs*
- *Funding for both the service provider and the researcher.*

COORDINATED DECISIONS, POLICIES & ROLES

- *“everyone thinks its a good idea, but no one wants to take ownership”*

AWARENESS OF ISSUES IN PRESERVATION & ARCHIVING

LACK OF TECHNOLOGY & INFRASTRUCTURE

RESEARCHER ATTITUDES

STAFF EXPERTISE

- *Not enough personnel to handle all of it.*
- *Lack of knowledge and subject expertise with the actual data*

Workshop & Forum (Goal 2)



Dr. Marilu Goodyear

- Director, KU School of Public Affairs & Admin.
- Former KU Chief Information Officer
- Former KU Associate Dean of Libraries



Dr. Noel Rasor

- Assistant Director, KU Public Management Center

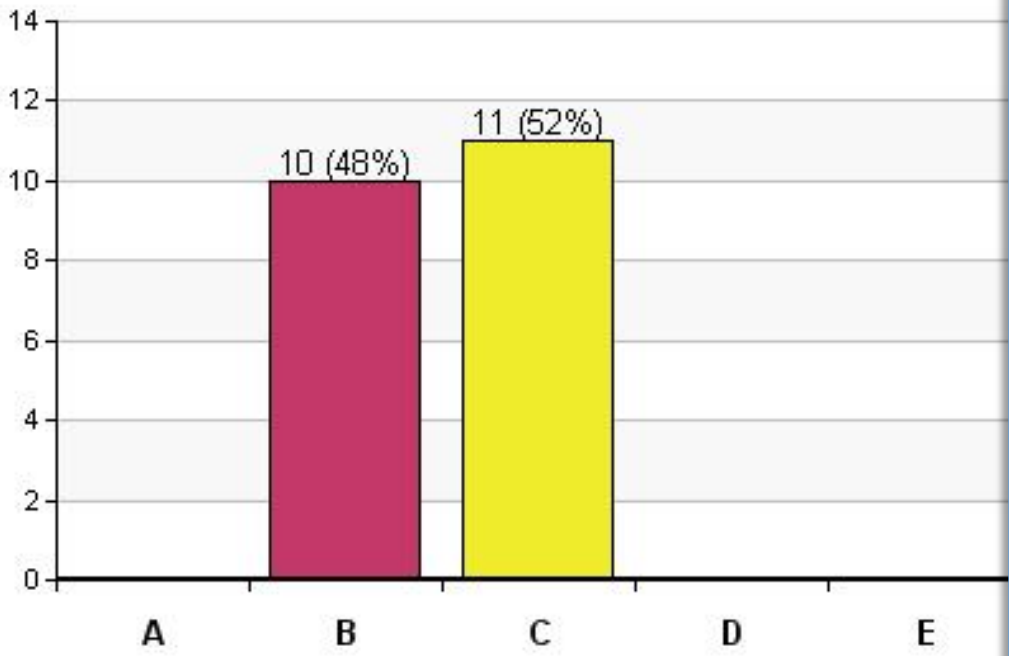
Advisory Council Meetings, Kansas City

May 29-30, 2013

- Arizona State University
- Iowa State University
- Kansas State University
- Oklahoma State University
- Texas A&M
- University of Arkansas
- University of Colorado
- University of Kansas
- & Medical Center
- University of Missouri
- University of Missouri – KC
- University of New Mexico
- University of Oregon
- University of Tulsa
- University of Washington
- Wichita State University

*Deans of Libraries, CIO's,
Chief Research Officers*

Advisory Council Meeting Findings



Advisory Council Member Responses

- I believe that research data issues should be addressed by:**
- A. Each individual campus alone.
 - B. Each **individual campus with assistance** from national and regional groups.
 - C. **Regional groups should provide services** to collaborating campuses, with each campus implementing the services they see as important.
 - D. National groups should take full responsibility.

Potential Paths Forward

	Critical Red	Feasible Blue
The Argument	11	12
Funding Models	9	2
Business Requirements/ Policy	8	5
Faculty Incentives	8	0
Data Management Plans	3	9
Discovery and Access	2	0
Storage	1	1
Education and Training	0	7
Preservation	0	4

Dot-based voting on key issues by Advisory Council, May 29-30, 2013.

Red = critical
Blue - feasible

Data Summit Participants

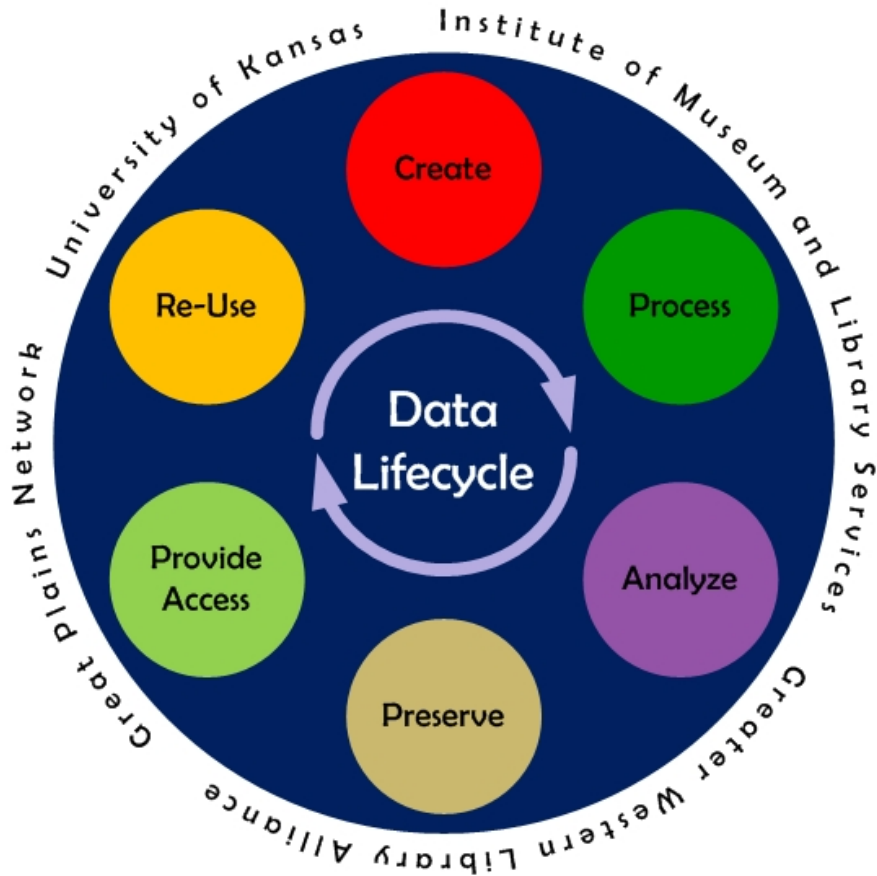
- Name some things you would envision as elements of a good solution to issues of research data management at your institution.
- Name some stakeholders (people, positions or roles) that must be involved for efforts to succeed.
- Please share any ideas you have on how institutions could cooperate on data management and curation.

The IMLS Grant – Goal 3

Goal #3: Create and disseminate a plan for a scalable multi-institutional approach to research data management to support the university members of GPN and GWLA and advance this plan for funding. *

* “Endorse and support has been request to date.”

Grant Goal 3 – The Plan



3 Strategies

Strategy 1: Outreach, Education, & Training

Strategy 2: Strengthen Infrastructure

Strategy #3: Communication & Advocacy

Strategy #1 – Outreach, Education, Training

*GWLA & GPN will offer a collaborative program of **outreach, education, and training** for member universities, addressing common institutional needs related to data management, data **policy development**, and **best practices** for data lifecycle curation.*

Potential Paths Forward

	Critical Red	Feasible Blue
The Argument	11	12
Funding Models	9	2
Business Requirements/ Policy	8	5
Faculty Incentives	8	0
Data Management Plans	3	9
Discovery and Access	2	0
Storage	1	1
Education and Training	0	7
Preservation	0	4

Strategy 1 Education, Training & Outreach

Areas of interest:

- library consultants
- data management skills
- common language
- Compliance
- researcher support before, during, and after
- basic records management
- links to tools for researchers



Strategy 2: Strengthen Infrastructure

*Strengthen networked **repository** and **storage** infrastructure for data derived from research conducted by member institutions. Improve **interoperability** and **opportunity** between institutions, disciplinary organizations, and other research data hosts for working collaboratively with data and for **access**, **archiving**, and long-term **preservation**.*

Potential Paths Forward

	Critical Red	Feasible Blue
The Argument	11	12
Funding Models	9	2
Business Requirements/ Policy	8	5
Faculty Incentives	8	0
Data Management Plans	3	9
Discovery and Access	2	0
Storage	1	1
Education and Training	0	7
Preservation	0	4

Strategy 2 -- Strengthen Infrastructure

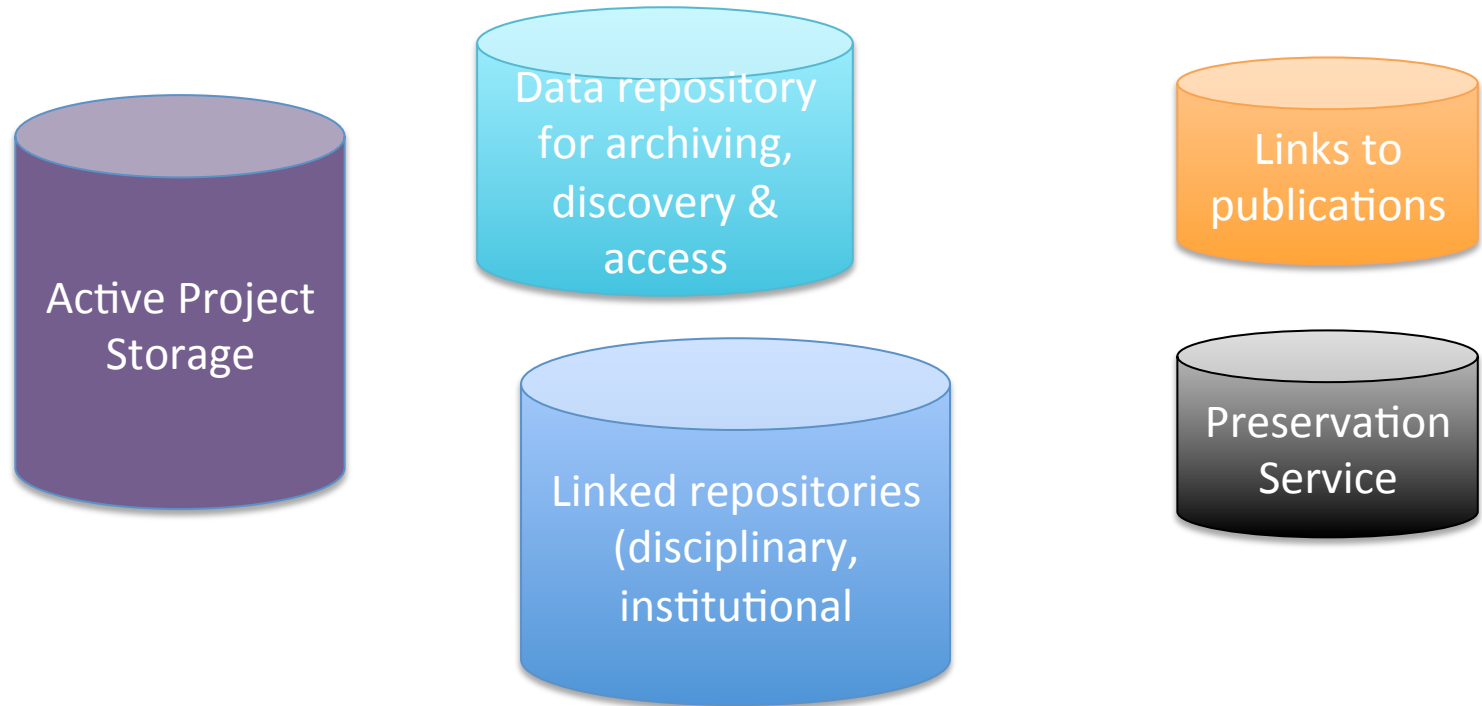
Initial dot voting did not seem to indicate that infrastructure was a pressing concern, but subsequent discussion and feedback from the advisory council advanced this as more important.

Survey results also indicated this area represents a gap for member institutions.



Strategy 2: Strengthen Infrastructure

Working group formed. Interest in connecting efforts with SHARE. Initial focus on storage, and discovery / access.



Strategy #3 – Communication & Advocacy

*Develop and execute a strategic **communications** plan and program of **advocacy** to advance member initiatives in data management and to advocate for **data sharing** as part of a greater ecology of scholarship.*

Strategy #3 – Communication & Advocacy

*Develop and execute a strategic **communications** plan and program of **advocacy** to advance member initiatives in data management and to advocate for **data sharing** as part of a greater ecology of scholarship.*

Working group not yet formalized:

- Advocacy for SHARE proposal
- Communications via Twitter, Facebook, Web presence, Email lists
- Annual event, the *Big Data Summit* in Kansas City
- GWLA Exec. Director on SHARE communications working group

Potential Paths Forward

	Critical Red	Feasible Blue
The Argument	11	12
Funding Models	9	2
Business Requirements/ Policy	8	5
Faculty Incentives	8	0
Data Management Plans	3	9
Discovery and Access	2	0
Storage	1	1
Education and Training	0	7
Preservation	0	4

Strategy 3
Communication & Advocacy
 Administrative awareness; ability to address “big issues;” faculty issues; costs; prestige; leadership & roles; managing university assets

Vision for Services

- ...a **community** of research universities engaged together in discovering and mutually supporting best practices, training, education, and tools for the effective management of research data across their intended lifecycle.
- ...rooted in the **larger ecosystem** of existing and emerging national and international data initiatives
- ...**aligned** with the research sharing efforts of ... other collaborative research data and publication hosts
- ...seeking to **leverage the best ideas** for research data management at scale

Connect with D4

Website: <http://imls.gwla.org>

Twitter: [@DataFOUR](https://twitter.com/DataFOUR)

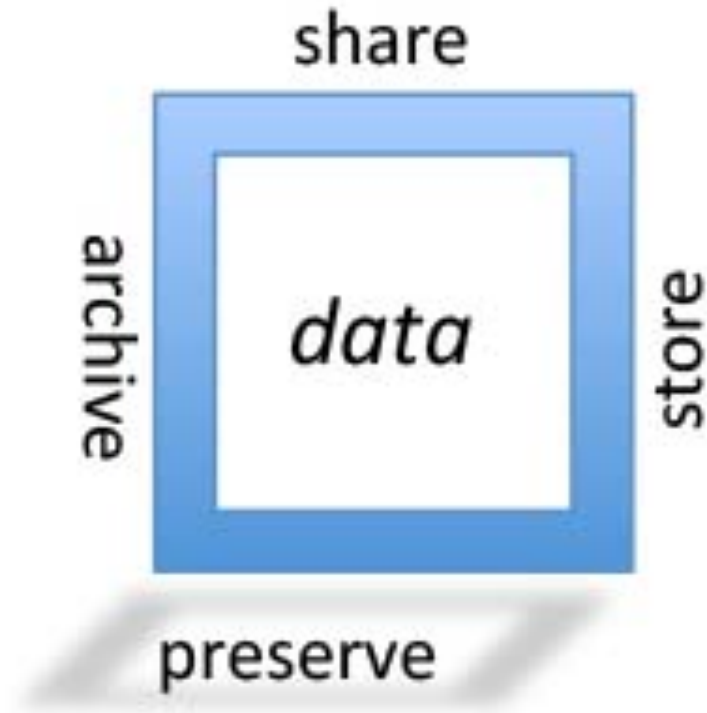
Facebook: <https://www.facebook.com/D4DataFOUR>

Email List: data@gwla.org.

Send note to npotter@kgs.ku.edu to be added to email

Upcoming Event: Big Data Summit 2
May 28, 2014 in Kansas City, MO

Questions?



Contact US

Deborah Ludwig (dludwig@ku.edu)

Assistant Dean for Innovation & Strategy

University of Kansas Libraries

Ann Champion Riley (rileyac@missouri.edu)

Associate Director of the Access, Collections, &

Technical Services Division

University of Missouri Libraries