



# *What are Librarians in CIC doing?* Purdue's Interdisciplinary Research Initiative



Center for Library Initiatives Conference

**Librarians & e-Science: Focusing Towards 20/20**

May 12–13, 2008 Hosted by Purdue University

D. Scott Brandt  
Associate Dean for Research  
[techman@purdue.edu](mailto:techman@purdue.edu)

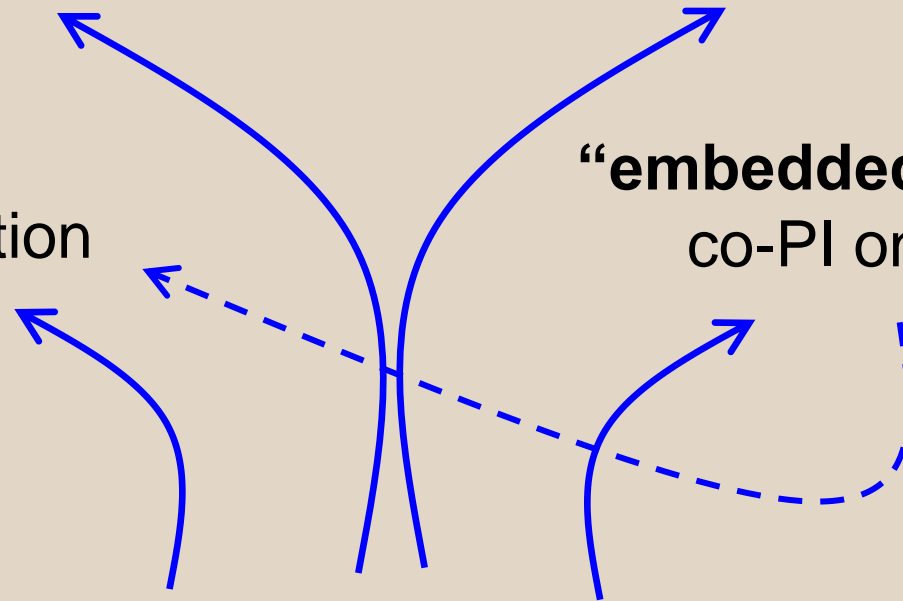


**Digital Initiatives** (products)  
digital tools/systems

**New Roles** (services)  
interdisc liaison

**D2C2**  
data curation

**“embedded librarian”**  
co-PI on grants



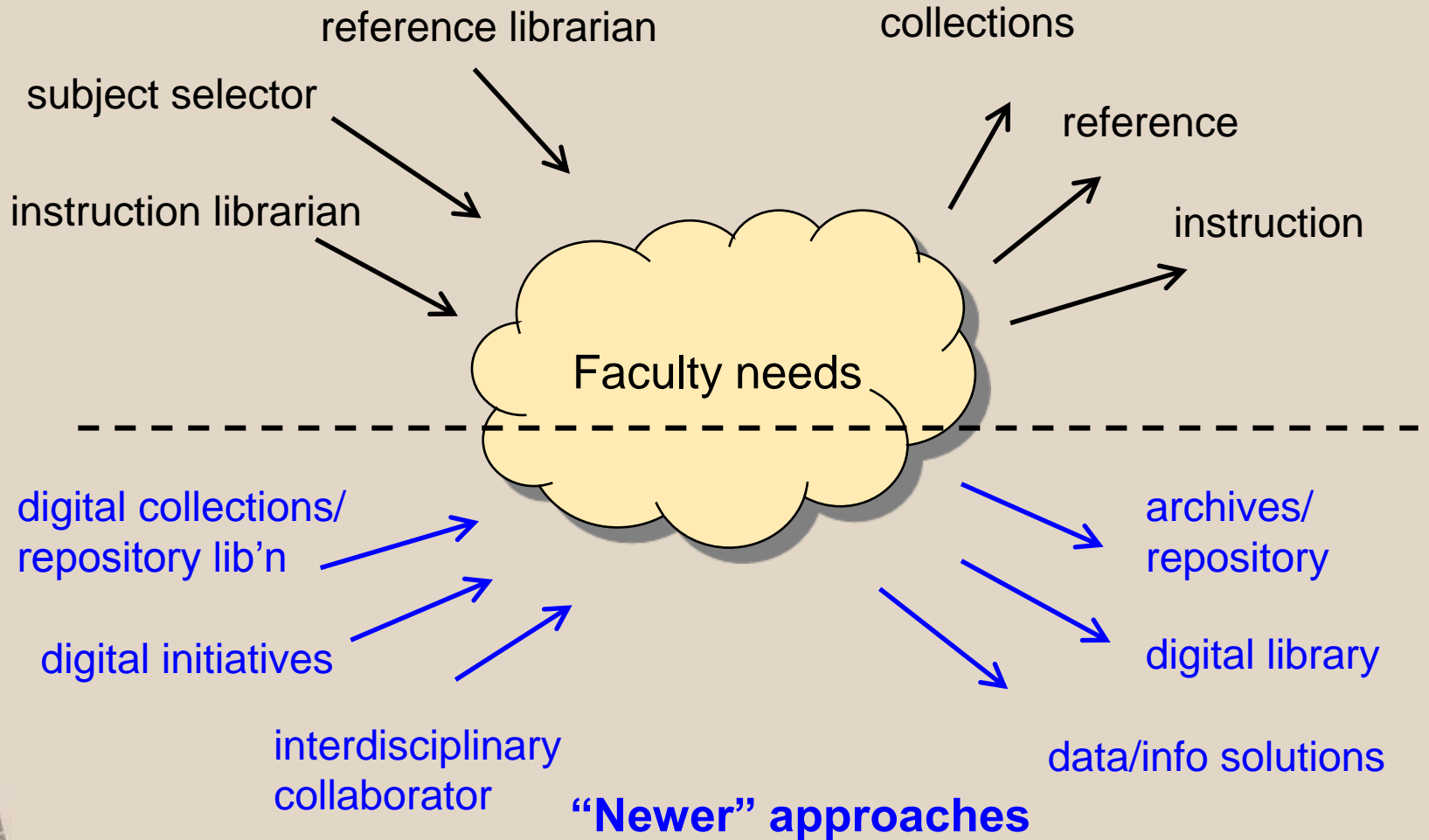
**interdisciplinary research initiative explores  
and partners to solve user data/info needs in  
cyber enabled environment**



**“Traditional” approaches**

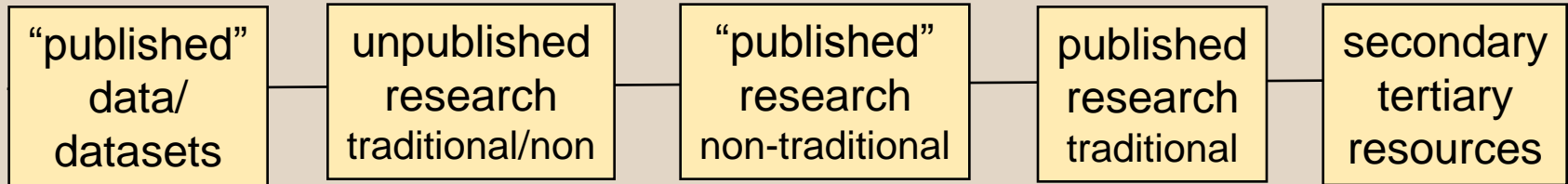
**Roles**

**Outcomes**





Research traditionally moves in this direction



analyzed data/datasets

processed data/datasets

“raw” data/datasets



Library service traditionally moves in this direction

**Given the changing nature of research and scholarly communication (i.e., e-research), how do librarians change to adapt to new ideas regarding what they collect to support research & education and how they provide access to it?**

*Modified from: Brandt, D.S. “Scholarly Communication” (in To Stand the Test of Time: Long-Term Stewardship of Digital Data Sets in Science and Engineering.: Final Report of Workshop New Collaborative Relationships: Academic Libraries in the Digital Data Universe. ARL, Washington, DC, September 2006.)*



# Background

- **2004**: Interdisciplinary Research Initiative revealed data needs on campus
- **2005**: Reorganization to parallel other academic units to facilitate interaction and collaborative research (ADR)
- **2006**: Created D2C2 as way to further explore domain, organize sponsored research and leverage collaborations



# interdisciplinary research

- Much research has become cross, multi and interdisciplinary related: nanomedicine, healthcare engineering, social computing
- This is especially true because computing now underlies most research and data is becoming not only the outcome of research, sometimes it is the genesis or driver of it
- There are many possibilities for Library Science applied to domain problems



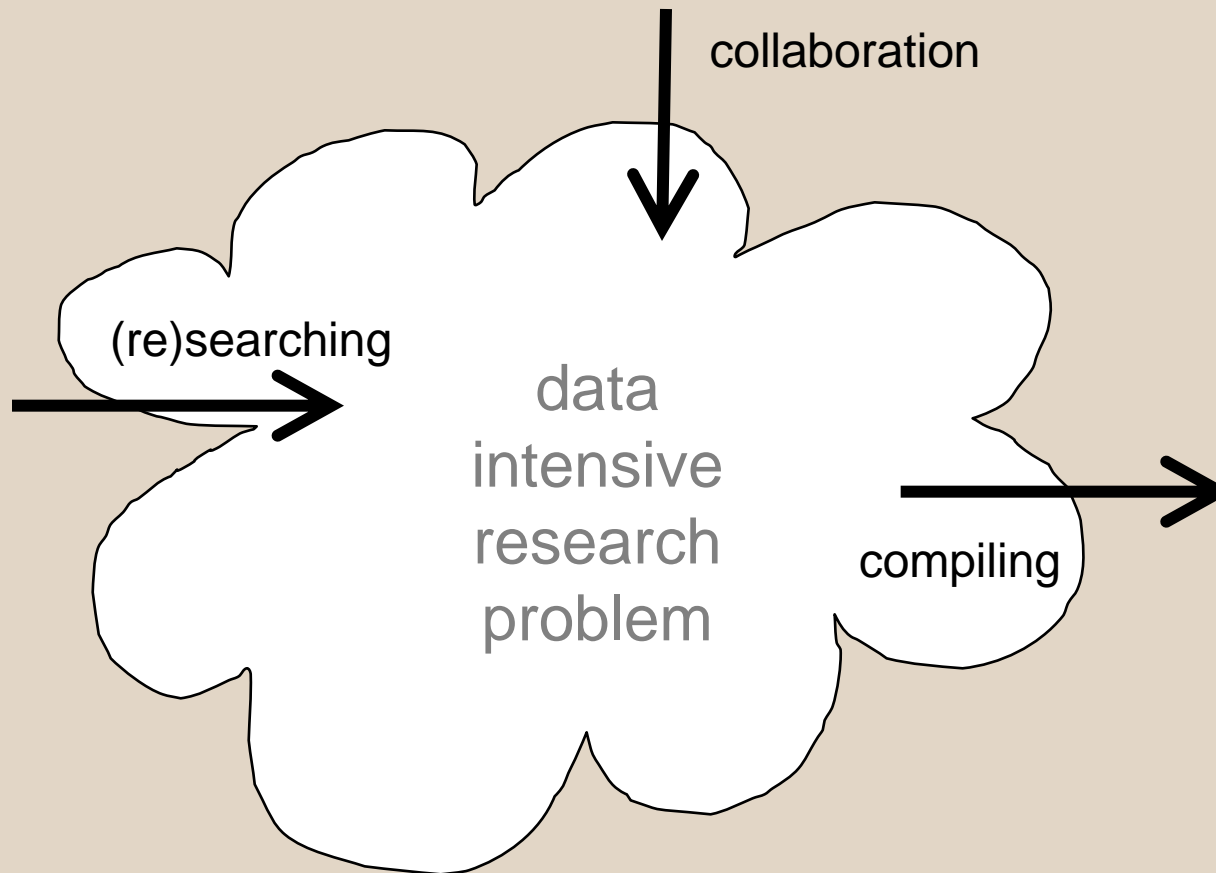
# 2004: “interdisciplinary” research

Initiative for Libraries faculty to collaborate with other faculty across campus—apply library science knowledge and expertise to research questions and problems:

**collect, organize, describe, curate,  
archive, and disseminate:  
data and information**



# How a scientist sees it...







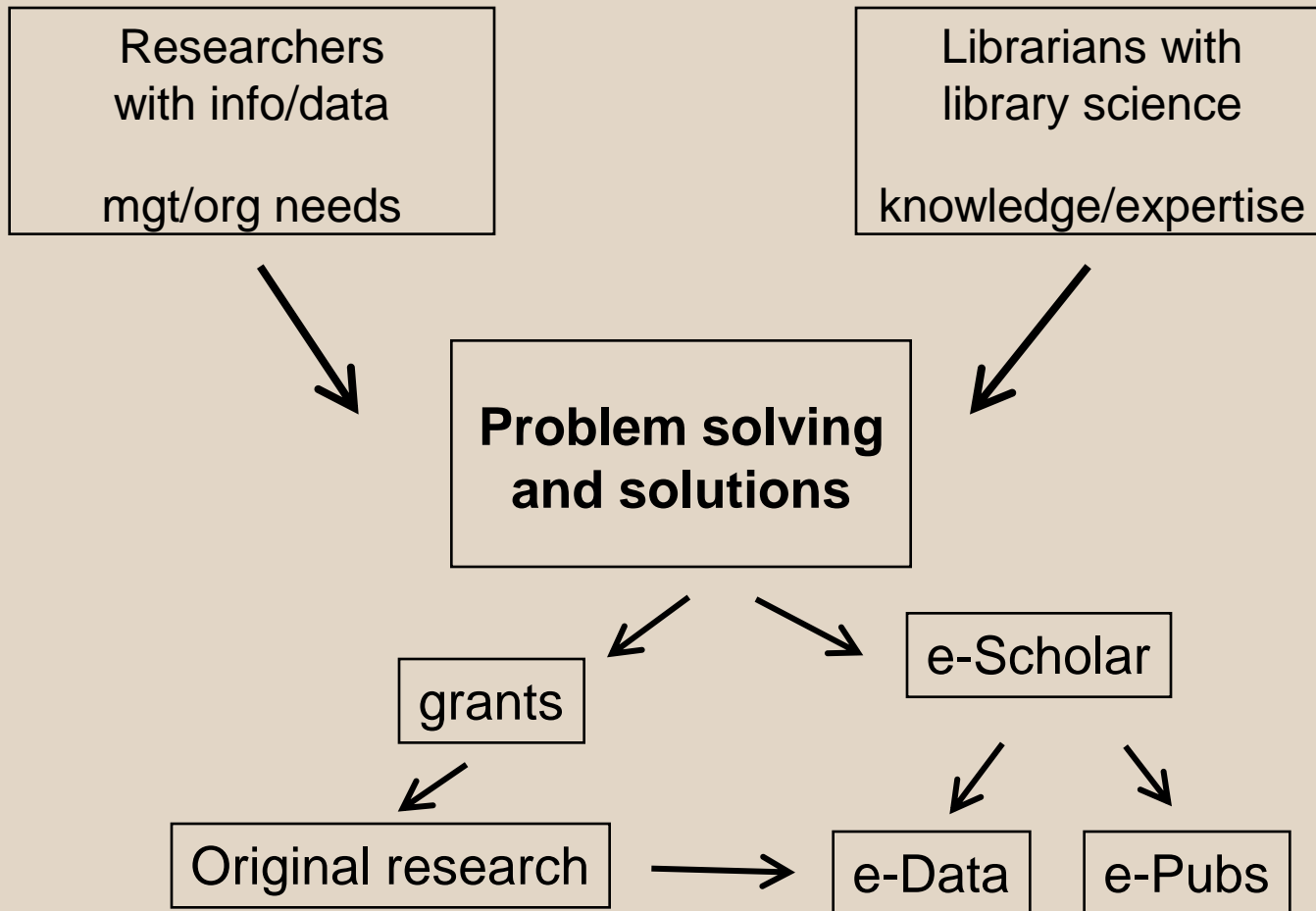
# Building relationships

Agronomy, Biology,  
Center for Environment,  
Chemical Engineering,  
Chemistry, Cyber Center,  
Discovery Learning  
Center, Earth  
& Atmospheric Sciences,  
IT at Purdue (IT@P),  
Oncological Sciences

- Not sure how or whether to share data
- Lack of time to organize data sets
- Need help describing data for discovery
- Want to find new ways to manage data
- Need help archiving data sets/collections



# interdisciplinary research





# Distributed Data Curation Center



The Distributed Data Curation Center (D2C2) investigates and pursues innovative solutions for curation issues of organizing, managing, facilitating access to, archiving for and preserving research data and data sets in complex environments.



# Curation

- **Curation** is the activity of managing and promoting *the use of data*, starting from the point of creation, to ensure its fitness for contemporary purposes and availability for discovery and re-use.
- **Archiving** is a curation activity which ensures that data is properly selected and stored, can be easily *accessed* and that its logical and physical integrity is maintained over time.
- **Preservation** is an archiving activity in which specific items of data are *maintained* over time so that they can still be accessed and understood through succession and obsolescence of technologies.

Lord, P. Macdonald, Lyon & Giaretta (2004) "From data deluge to data curation." Proceedings of the UK e-Science All Hands Meeting 2004, 31st August - 3rd September, Nottingham UK.



# Research Goals

## Investigate landscape

- Researcher needs
- Data mgmt problems
- Distributed nature...

## Enhance Discoverability

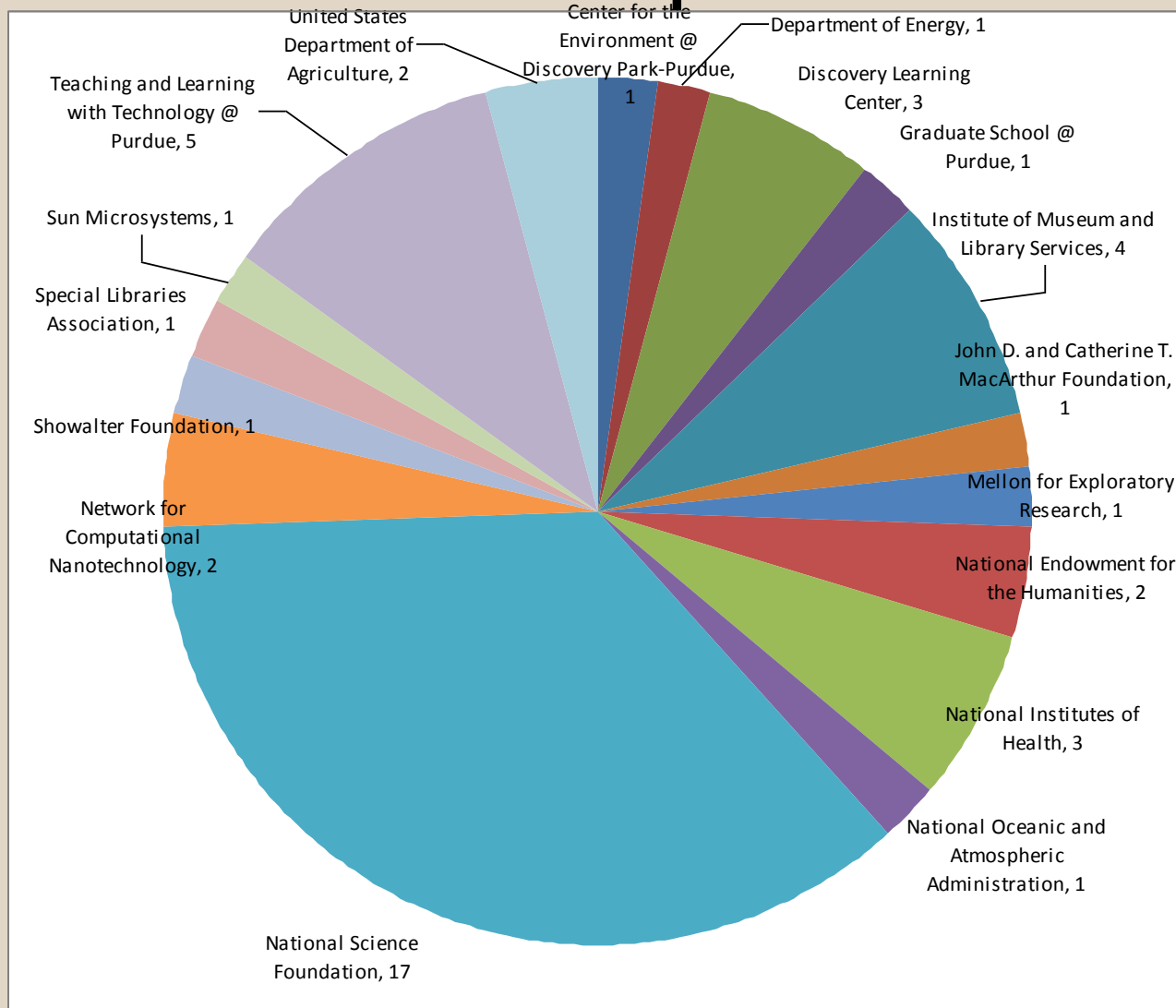
- Descriptive navigation
- Mapping data schemas
- Exposing metadata

## Facilitate Curation

- Archiving and preserving data
- Community driven specs for sharing
- Tools and services (“middleware”) to connect systems



# Collaborative Proposals 2005-08





Purdue University - West Lafayette, IN

Award Amount: **\$421,068**

Grant Category: Research and Demonstration

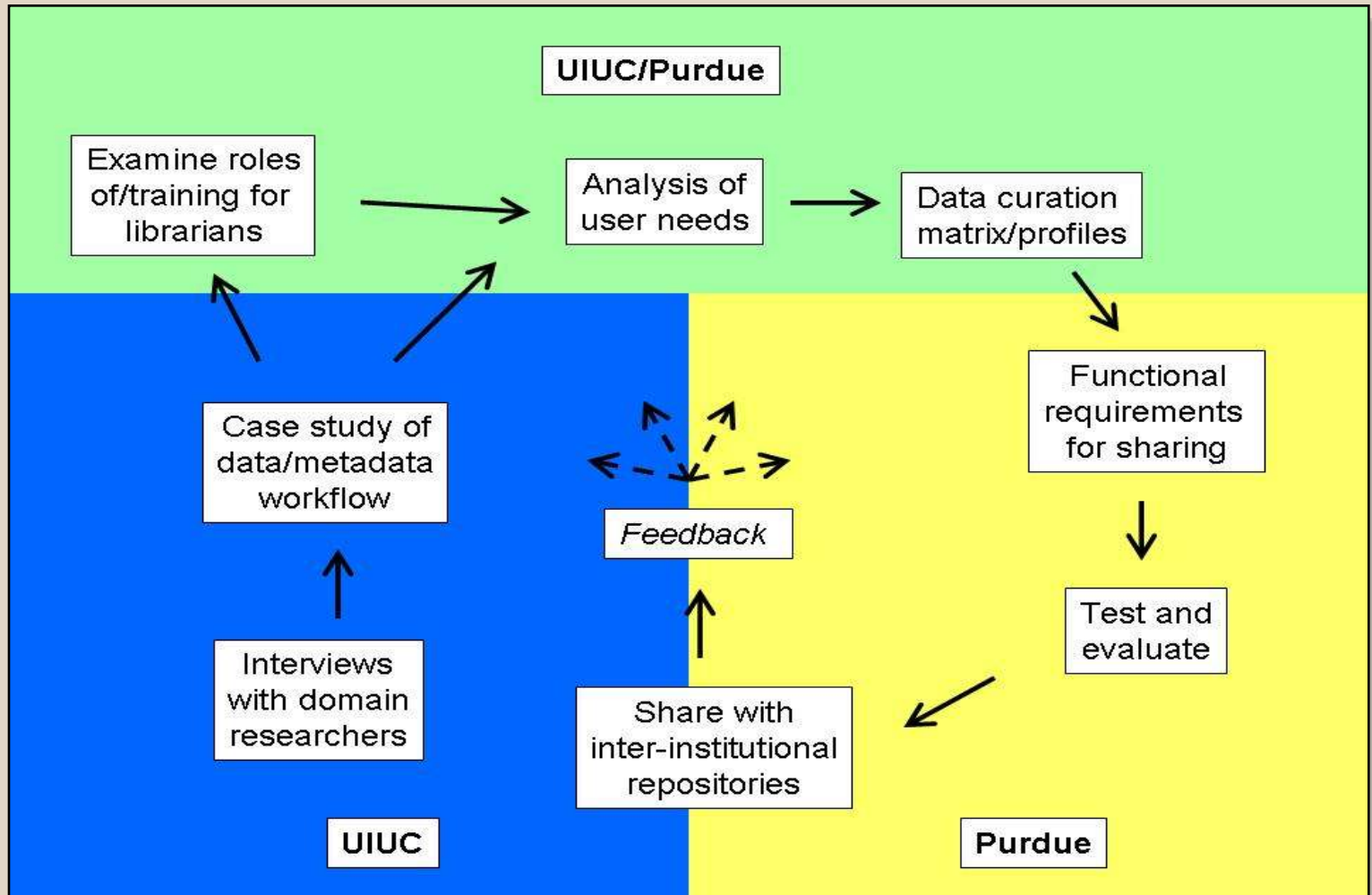
**Project Title: "Investigating Data Curation Profiles Across Multiple Research Disciplines."** Investigators in the Distributed Data Curation Center in the Libraries at Purdue University, and the University of Illinois, Urbana-Champaign will address the question “which researchers are willing to share data, when, with whom, and under what conditions?” The team will produce case studies of researcher data/metadata workflow, curation profiles describing policies for archiving and making available research data, a matrix to compare parameters across disciplines, system requirements for managing data in a repository, and recommendations for implementing results under diverse systems. The project will describe the roles of librarians and identify the skill sets they need to facilitate scholarly communication and data sharing. With UIUC GSLIS & Libs.



# Curation research

- “Who will share what with whom, and when (and how)?”
- Are there differences between scientists who want to share (and those who don't) and levels/types of data?
- Needs/requirements → management policies → curation profiles → rules → systems implementation







# Other projects...

## Awarded or Completed Projects [Export to Microsoft Excel](#)

▲ Status	End Date	Project Title	Librarian	Funding Org
Awarded	04.01.2008	Ingest, Preservation and Access for Water Quality Datasets in an Institutional Repository	Witt, Michael; Bracke, Marianne	Center for the Environment @ Discovery Park-Purdue
Awarded	00.00.0000	Investigating Data Curation Profiles Across Multiple Research Disciplines	Carlson, Jake; Brandt, Scott; Witt, Michael	Institute of Museum and Library Services
Awarded	00.00.0000	Metasearch Technologies for the NSDL Distributed Community	Bracke, Paul	National Science Foundation
Awarded	06.01.2008	Thought Ark: A Community-Oriented Idea Space	Bracke, Paul	Teaching and Learning with Technology @ Purdue
Awarded	05.07.2008	Investigate and Implement Persistence for HUB Resources	Witt, Michael	Network for Computational Nanotechnology
Awarded	12.30.2008	Purdue Center for Serious Games and Learning in Virtual Environments	Witt, Michael	Discovery Learning Center
Completed	05.31.2007	An Expert System Multimedia Tutorial for Locating Technical Information	Fosmire, Michael; Sapp, Megan; Van Epps, Amy	Teaching and Learning with Technology @ Purdue
Completed	05.31.2009	E-coli Resource	Brandt, Scott	National Institutes of Health
Completed	01.15.2008	Identifying Factors of Success in CIC Institutional Repository Development	Brandt, Scott	Mellon for Exploratory Research
Completed	12.31.2006	Deploying Electronic Thesis and Dissertation (ETD) Linking	Witt, Michael	Graduate School @ Purdue
Completed	09.01.2008	Lewis and Clark and the Indian Country: A Traveling Exhibition for Libraries	Hovde, David; Alcorta, Marissa; Riehle, Catherine; Rodriguez, Elisabeth; Yazza, Valerie	National Endowment for the Humanities
Completed	12.31.2007	Development of an OAI-PMH Interface for nanoHUB.org	Witt, Michael	Network for Computational Nanotechnology
Completed	05.31.2008	Sun STK 5800 Storage System	Witt, Michael	Sun Microsystems

Number of Projects: 13



# Metadata standards

- Users of cyber-enabled instruments need to manage, disseminate and preserve their data while providing undergrad students a rich research experience.
- Design tools to provide standardized metadata to enable interoperability and long term preservation of data sets.



# Automate metadata

- A small science research group does not catalog or preserve or do archiving, but wants to increase accessibility of datasets collected to share/reuse.
- Create a functional proof-of-concept system for automating the generation of descriptive metadata and the ingestion of datasets into an accessible repository.



# Search-ability

- A researcher decides to develop a specialized portal to leverage political, environmental and technology information to spur economic development in a specialized area.
- Develop a collection of non-traditional resources and provide a layer of metadata to improve organization & search results.



# Learning tools

- Educators want to integrate GIS experiences in environmental sciences from the field to the class.
- Help build lightweight, intuitive online GIS applications to access and use large amounts of data in a secure archive.



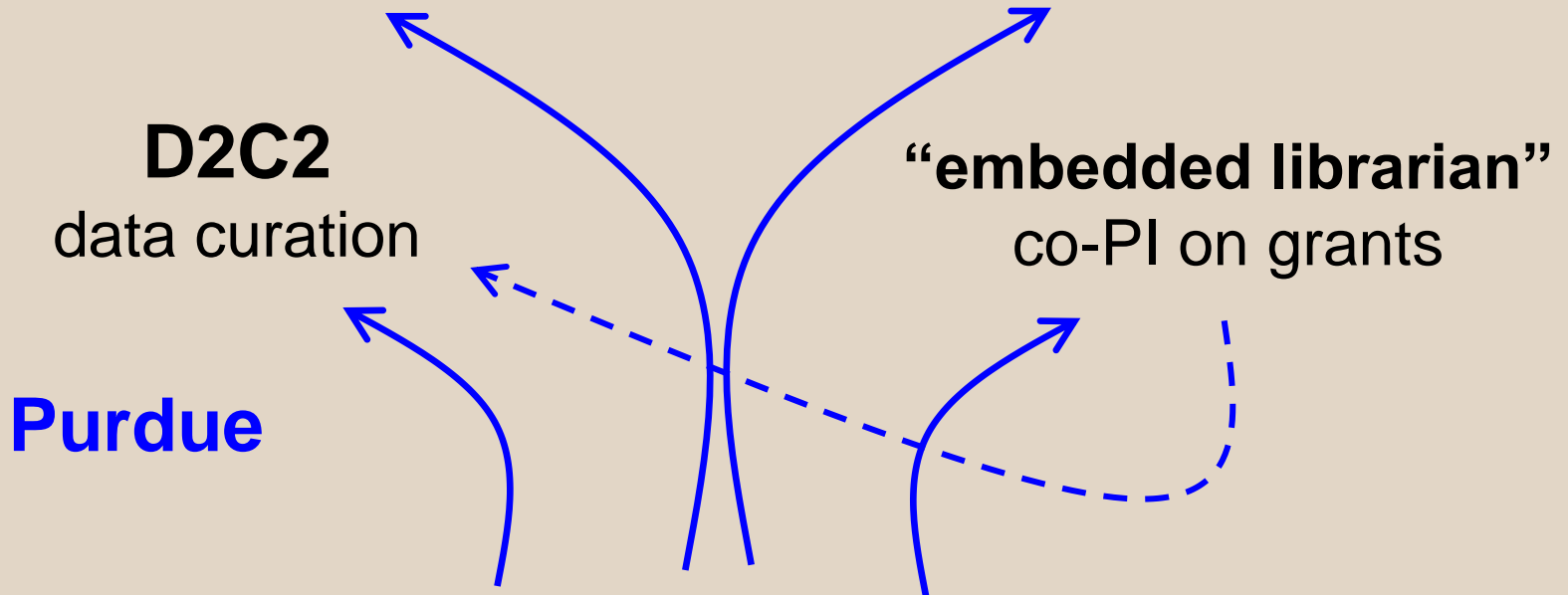
# Approaches to do it?

- Attend research seminars, callouts, and discussions to identify collaboration and funding opportunities. Also create or host them.
- Build relationships. Find researchers who understood that collecting, organizing and providing access to data and information are not only important, but critical, and thus, librarians need to be involved.
- Find problems to solve, then collaborate on solutions.
- Talk about what we knew—organizing data and information (different meanings to different groups).
- Bring something to the table. Have to be prepared to demonstrate something tangible (initially a proof-of-concept or a prototype).



**Digital Initiatives** (products)  
digital tools/systems

**New Roles** (services)  
interdisc liaison



**interdisciplinary research initiative explores  
and partners to solve user data/info needs in  
cyber enabled environment**





# 2008 CIC Center for Library Initiatives Conference Librarians & e-Science: Focusing Towards 20/20

## Thank you

D. Scott Brandt  
Associate Dean for Research  
[techman@purdue.edu](mailto:techman@purdue.edu)

