INTERDISCIPLINARY RESEARCH

A Study of How Faculty and Graduate Students find and use Research Data

Contributors:

Grace Agnew (Rutgers)
Ron Jantz (Rutgers)
Ryan Womack (Rutgers)

Annie Johnson (Temple)
Robert K. Ohlendorf , (Penn State)

Virtual Data Collaboratory

Regional Cyberinfrastructure for Collaborative Data Intensive Science

- $4 Million grant, awarded by NSF, September 1, 2016-August 31, 2020.
- NSF DIBBS (Data Infrastructure Building Blocks) grants intended to build robust and shared data-centric cyberinfrastructure capabilities, to accelerate interdisciplinary and collaborative research in areas of inquiry stimulated by data.
- Rutgers University Libraries, Penn State University Libraries, and Temple University Libraries are designing and prototyping the VDC discovery service and data repository to facilitate access to data, researchers, tools, and research publications, focusing specifically on interdisciplinary research.
- Utilizes Samvera 2.X repository application, a linked data repository platform

Interdisciplinary Research - The Study

- Building on the work of the grant, this presentation discusses graduate student and faculty views of how they view their research and find and select interdisciplinary data.

- Results of the study are from in-depth interviews with 14 faculty who do data intensive and interdisciplinary research, two STEM graduate student focus groups (50+ students) at Rutgers and Temple and a follow-up survey to evaluate the DIBBS design resulting from the investigation

Interdisciplinary Research – A Quick View

Interdisciplinary research . . .

- Can be impeded by current academic structures.
- Is characterized by shifting boundaries, intersecting domains, and collaborating specialists.
- Involves exchanges of factual knowledge, new ideas, socialization, new technology, and a network of contacts.
- Is influenced by personal compatibility, work connections, incentives, and infrastructure
- Remains a vague concept where collaborations are difficult to achieve and rarely stable.

Views: Interviews Focus Groups and a Survey

- One on one interviews with 14 faculty with interdisciplinary research focus at Rutgers, Penn State and Temple
- Two focus groups, consisting of 15+ STEM postdoctoral students at Rutgers and 50+ STEM graduate students at Temple University. The following presents a brief synopsis of their response.

- We followed up with a Qualtrics survey to the students and to faculty at Rutgers, Penn State, and Temple.

FINDINGS - Faculty

- How do you engage with other disciplines?

  - **Manual and labor intensive**. *Attend conferences, call experts, Wikipedia*

  "I dug up my high school textbook to learn math"

- Not "what," but "who"

- *Preference for conversation over research. Professional networks are critical for finding someone authoritative to contact. Prefer local contact for face to face discussion, collaboration*

FINDINGS – Graduate Students

- What is your primary area of research?
  - *Some students (particularly at Temple) appeared to be confused about this question and did not think of themselves as belonging to a discipline.*
  - *Research is very multidisciplinary. Rutgers students answered more readily but stated they are primarily engaged in problem-based research across multiple disciplines.*

- Do you use research data that you didn't create?
  - *Students used data provided by their professor or discovered in the scholarly literature,* ***DOI is critical.***
  - *Or they used trusted third party databases, such as the Cambridge Crystallographic Data Centre database, and data they created.*

FINDINGS –Interdisciplinary Data Challenges

- Communication – Talking a different language
  - -- Methodology – theoretical vs. experimental
  - -- Terminology and approach – translating across disciplines
  - --- Actual language differences in international research

- Needs
  - -- Data hidden in disciplinary portals.
  - -- Cloud-based data and tools. Better interaction between cloud data and tools. HPC, in particular.
  - -- Better tools to abstract variables, portions of data
  - --- Better collaboration tools for sharing data. *GitHub often cited for ability to track contributions and version data and applications*
  - -- Need to trust that data is curated and archived for the long term. Need this support for their own data.

FINDINGS – Trusting Research Data

- Who created is most important

  Reputation of PI, group or lab

  "Thought leader" as identified by colleagues, publication in top journals

- Is the data from a grant more trustworthy?

  -- Faculty - Grant indicates peer evaluation of design

  -- Temple graduate students felt grant actually decreased trustworthiness; grantor might have bias, particularly a commercial funder.

  - Have other people used it? Is it cited?

  "If this data is wrong, at least I am in good company."

  - Trustworthy Creation

  -- Created with respected equipment/application

  -- Codebook, variables explained

Findings – Trusting Research Data

- Is knowing the hypothesis or research question important in evaluating data for reuse?
  - *"Yes, this is helpful." Why it was created is important*

Attribute Name	Values
Title	Incremental and Radical Innovations in Research Libraries. Data
Creator	Jantz, Ronald http://orcid.org/0000-0003-0552-1094 Rutgers, The State University of New Jersey rjantz@rutgers.edu Department: Rutgers University Libraries Position: Professor Discipline(s): Computer and Information Sciences
Identifier doi	http://dx.doi.org/doi:99.0000/396ea237-ba37-4d90-8025-3245e09e45cb
Genre	Data
Date created	2013
Abstract	In the analysis, it was found that five factors had a significant impact on the innovation performance of the library - behavioral integration, decision awareness, structural differentiation, organizational ambidexterity, and size of the organization. The data for the regression analysis is included here in SPSS format.
Research problem	Determine the variables that affect organizational innovation in research libraries.
Extent	2
Format	application/pdf application/octet-stream
Discipline	Computer and Information Sciences
Coverage spatial	50 Research Libraries in the United States
Coverage temporal	2013
Relation type	isReferencedBy
Relation uri	https://doi.org/doi:10.7282/T3ZP44PT
Rights	Open Data Commons Attribution License (ODC-By) v1.0
Use	Grace Agnew https://orcid.org/0000-0003-4189-4832
Use Purpose	Modify for use in original research

The Survey

Demographics of Survey Respondents (N=13)

Discipline/Area	Faculty	Graduate Student	Staff
Cancer	1		
Computer Science		1	1
Economics	1		
Environmental Eng.	1		
Epidemiology	1		
Information Science	1		
Life Sciences	1		
Media Studies	1		
Neuroscience		1	
Physical Chemistry		1	
Public Health	1		
Social Decision Making	1		

Survey Questions and Responses

- In finding trustworthy data, what is the most important for you?

	Least Important (5)	(4)	(3)	(2)	Most Important (1)
Granting Agency	5	2	4	1	1
Research Question or Hypothesis	0	3	2	4	4
Discipline of the Creator	2	6	4	1	0
Availability of an Abstract	0	0	3	5	5
Availability of a README	6	2	0	2	3

Survey Questions and Responses

How important is it to be able to contact the data creator?

-- *Very important (6)*
-- *Important (3)*
-- *Somewhat important (1)*
-- *Not important (2)*

- How important is it to know other researchers have used the data?
  - *Very important (2)*
  - *Important (4)*
  - *Somewhat Important (5)*
  - *Not important (0)*

- What are the primary reasons for data reuse?
  - *Cite data in scholarly publication: 10*
  - *Use for computational analysis: 8*
  - *Modify for use in original research: 5*
  - *Use for teaching a course: 5*
  - *Use unchanged: 5*
  - *Use for govt or institutional policy recommendations: 1*

Survey Questions and Responses

- How important is the information about the data creator?

	Not Important	Somewhat Important	Important	Very Important
Name of Creator	0	4	3	6
ORCID ID	3	2		
University Organization/ Affiliation	0	3		
Department within Organization	1	5		
Position	1	4	5	2
Discipline	0	3	4	4

The problem is the “meta-discipline.” The “formal” discipline or field of study is the lens (approach, knowledge, methodology) that you bring to the “meta-discipline.”