Data Management Programs

GETTING TO CONCRETE RESULTS

Paul J. Bracke
Associate Dean for Research and Assessment

ARL Membership Meeting
October 8, 2014
THE ROAD TO CONCRETE RESULTS

Response to 2003 Atkins Report on “Revolutionizing Science & Engineering Through Cyberinfrastructure

- nanoHUB funded in 2004 -> HUBzero
- Library/Researcher Collaborations
- Library Reorganization to Align with Campus (ADR) and Hiring
- 2008 - Data Curation Profiles (with UIUC GSLIS)

Campus Approach changes with May 2010 NSF Announcement of DMP Mandate

- Summer 2010 - Provost appoints faculty taskforce to address mandate, co-chaired by CIO and Dean of Libraries
- Fall 2010 – Commitment for joint development of PURR using HUBzero (Libraries, CIO, OVPR)
- 2011 – DMP Workshops begin
- 2011 – Data Information Literacy (with Minnesota, Cornell, Oregon)
- 2012 – OVPR Commits $1.2M over 2 years for development, Project Director from Libraries
- PURR soft launch in late 2012, full launch in 2013
- 2014 – NSF DIBBs
THE ROAD TO CONCRETE RESULTS

Response to 2003 Atkins Report on “Revolutionizing Science & Engineering Through Cyberinfrastructure

• nanoHUB funded in 2004 -> HUBzero
• Library/Researcher Collaborations
• Library Reorganization to Align with Campus (ADR) and Hiring
• 2008 - Data Curation Profiles (with UIUC GSLIS)

Campus Approach changes with May 2010 NSF Announcement of DMP Mandate

• Summer 2010 - Provost appoints faculty taskforce to address mandate, co-chaired by CIO and Dean of Libraries
• Fall 2010 – Commitment for joint development of PURR using HUBzero (Libraries, CIO, OVPR)
• 2011 – DMP Workshops begin
• 2011 – Data Information Literacy (with Minnesota, Cornell, Oregon)
• 2012 – OVPR Commits $1.2M over 2 years for development, Project Director from Libraries
• PURR soft launch in late 2012, full launch in 2013
• 2014 – NSF DIBBs
THE ROAD TO CONCRETE RESULTS

- nanoHUB funded in 2004 -> HUBzero
- Library/Researcher Collaborations
- Library Reorganization to Align with Campus (ADR) and Hiring
- 2008 - Data Curation Profiles (with UIUC GSLIS)

Campus Approach changes with May 2010 NSF Announcement of DMP Mandate
- Summer 2010 - Provost appoints faculty taskforce to address mandate, co-chaired by CIO and Dean of Libraries
- Fall 2010 – Commitment for joint development of PURR using HUBzero (Libraries, CIO, OVPR)
- 2011 – DMP Workshops begin
- 2011 – Data Information Literacy (with Minnesota, Cornell, Oregon)
- 2012 – OVPR Commits $1.2M over 2 years for development, Project Director from Libraries
- PURR soft launch in late 2012, full launch in 2013
- 2014 – NSF DIBBs
DATA SERVICES @ PURDUE LIBRARIES

- Purdue University Research Repository (PURR)
- Organizing and Describing Data
- Data Sharing, Publication, and Preservation
- Data Management Plan Consulting
- Data Education and Literacy
- Grant Collaboration
DATA ECOSYSTEMS

• Purdue Data Services
  • Research Computing Center
    • Research Data Depot (HPC Storage)
    • HPC Clusters
  • Data Management Solutions on Campus
    • HUBzero (25+ Subject Communities)
    • CRIS (Computational Research Infrastructure for Science)
• Disciplinary Repositories/Infrastructures
Partnerships
LEVELS OF ENGAGEMENT

External Engagement
DataCite, Research Data Alliance (RDA)
Partnerships on Disciplinary Approaches

Librarian/Faculty Collaborations
Data Specialists AND Liaison Librarians
Research Collaborations, Data Curation Profiles, Service and Instruction

Repository Development
Purdue University Research Repository (PURR) – http://purr.purdue.edu
Libraries/Information Technology/Vice President for Research/Faculty

Administrative/Executive Engagement
Dean of Libraries, Chief Information Officer, Vice President for Research, Provost
Campus Research Administration/ADs for Research
**PURR Governance:**

**Steering Committee**
- Chaired by Libraries ADR, Project Director, Representatives from IT and OVPR, Faculty Representatives
- High-level Operational Decisions, Policy Recommendations

**Executive Committee**
- Executive Vice President for Research and Partnerships, Dean of Libraries, Chief Information Officer
- Policy and Budgetary Decisions
Staffing:
- Research Data @ Purdue University Libraries– 7 Librarians, 1 Software Developer, 2.5 FTE GAs
- Rely on IT and Development Support from HUBzero team
- Liaisons - 14 working with data, including Bioinformatics & GIS Specialists
  * Computational Sciences Information Specialist, Bioinformatics partially funded through cluster hire program

PURR:
- $1.2M in start-up funds (Included HUBzero hosting, staff time from Libraries and IT, term hires)
- Recurring Operating Budget of ~$380,000 from OVPR with acknowledgement that budget will need to be adjusted in the future to cover storage growth
- Designated as a “Research Core” – Eligibility for supplementary funding
- Grants (e.g., NSF DIBBs) are used to research and develop new functionality
- Quota system for data with fee structure for extra space
## PURR Project Space Allocation and Pricing

### Storage space for PURR projects currently covered by central resources

<table>
<thead>
<tr>
<th>Option</th>
<th>For what</th>
<th>Space Available</th>
<th>For how long:</th>
<th>For who:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default/trial projects</td>
<td>Just trying things out, or don’t need much space</td>
<td>10 GB</td>
<td>3 years</td>
<td>Any Purdue faculty, staff, and graduate student</td>
</tr>
<tr>
<td>Default data publications</td>
<td>Great for small publications</td>
<td>1 GB</td>
<td>10 years</td>
<td>Any Purdue faculty, staff, and graduate student</td>
</tr>
<tr>
<td>Supported projects</td>
<td>Funded projects with PIs from Purdue</td>
<td>100 GB</td>
<td>10 years or life of grant</td>
<td>Purdue faculty with a verifiable grant or account number*</td>
</tr>
<tr>
<td>Supported data publications</td>
<td>Publishing work done on a funded project</td>
<td>10 GB</td>
<td>10 years</td>
<td>Purdue faculty with a verifiable grant or account number*</td>
</tr>
</tbody>
</table>

### Estimated cost of additional space **

<table>
<thead>
<tr>
<th>Option</th>
<th>Space Available</th>
<th>For how long:</th>
<th>For who:</th>
<th>Cost:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extra project space</td>
<td>As needed</td>
<td>Per year</td>
<td>Any Purdue faculty, staff, and graduate student</td>
<td>$2.10 per GB per year ***</td>
</tr>
<tr>
<td>Extra publication space</td>
<td>As needed</td>
<td>10 years</td>
<td>Any Purdue faculty, staff, and graduate student</td>
<td>$14.30 per GB ***</td>
</tr>
</tbody>
</table>

### Frequently Asked Questions

- What’s the difference between a default project and a supported project?
- How do I budget for data storage?
- How do I upgrade my project to a supported project?
- How long will my data stay online?
Reception
PURR BY THE NUMBERS

Soft launch in 2012, 2013 was first full year

- 1166 proposals that include PURR in DMP
- 108 grants awarded
- 431 total research projects
- 1,466 registered researchers
- 68% of projects have more than one member
- Average project team size: 4
- Average number of files per project: 110
- 559 reference questions/tickets
- 100 datasets published
THE PROPOSAL PROCESS AT PURDUE

START HERE

1. Do you need help identifying funding opportunities? NO
   YES
      2. Do you want to do research for industry? NO
         YES
            The Discovery Learning Research Center offers expertise in learning research and transformative project management, evaluation, and assessment. Tania Burgess, tburgess@purdue.edu
      4. Do you want to partner on a teaching and learning research project for a multi-investigator proposal? NO
         YES
            VPR’s Office of Research Communications provides proposal graphics. Please work with Research Development: Sally Bond, sdbond@purdue.edu
      2. Do you need help with proposal graphics? NO
         YES
            If you have been awarded funding, you may contact Research Development Services at rds@purdue.edu
      3. Do you need to assemble a team from multiple units or disciplines? NO
         YES
            Research Development Services can assist with team building around national initiatives in thematic areas. Perry Kirkham, pkirkham@purdue.edu
      7. Do you need help developing a data management plan, creating online collaborative spaces, and/or creating data-sharing platforms? NO
         YES
            Purdue University Research Repository (PURR) will help you create a data management plan, upload research data and publish your dataset. http://purr.purdue.edu
      8. Do you need help developing a data management plan, creating online collaborative spaces, and/or creating data-sharing platforms? NO
         YES
            Research CORES provides access to state-of-the-art cost-effective technology and data analysis. https://engineering.purdue.edu/VPR/CON
      13. Has your proposal been accepted? NO
         YES
            Congratulations! You’re ready to proceed with your award.
      14. Are you creating a new center or institute? NO
         YES
            OWPR provides startup support to approved centers and institutes, as well as hiring of key staff and securing initial goals. Dave Rotterman, drotter@purdue.edu
      12. Does your project involve human subjects, vertebrate animals, export controls, biohazards, or any other non-financial regulatory requirements? NO
         YES
            Research Regulatory Compliance will walk you through compliance issues. vpregulatory@purdue.edu and/or exportcontrols@purdue.edu

Good luck with your next proposal!
DATA MANAGEMENT AND CREDIBILITY

Paul J. Bracke
Associate Dean for Research and Assessment
Purdue University Libraries
pbracke@purdue.edu