Research Data Management Infrastructure at Princeton: 
TigerData and Other Repositories

Curt Hillegas, PhD
Associate CIO
Research Computing & Princeton Institute for Computational Science and Engineering

Wind Cowles, PhD
Director of Data, Research, and Teaching, Princeton University Library
Head of the Princeton Research Data Service
RDM as a campus-wide service

2017 - Initial conversations between Dean for Research, Dean of Libraries, and CIO with department chairs and faculty

2018 - Provost-funded partnership: Princeton Research Data Service

4 New lines: 3 in the Library, 1 in Research Computing

2019 - PRDS launches
Princeton Research Data Service

Provide Princeton’s research community with the expert services and infrastructure needed to store, manage, retain, and curate digital research data, and to make it available to the broader research community and to the public.
<table>
<thead>
<tr>
<th>Consultations, resources, and tools</th>
<th>Training and Instruction</th>
<th>Research data repository services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consultations on data management, documentation, and stewardship, and open research practices</td>
<td>Live workshops/webinars on tools and best practices</td>
<td>Data curation guidance and assistance (regardless of repository)</td>
</tr>
<tr>
<td>Feedback and advice on Data Management Plans</td>
<td>Research Data Stewardship Program</td>
<td>Host and manage Princeton’s Research Data Repository</td>
</tr>
<tr>
<td>Tools for open data and code practices and data management, including ORCID and Code Ocean</td>
<td>Custom workshops available for labs, departments, and centers</td>
<td>Re-use support: Data curation services</td>
</tr>
<tr>
<td></td>
<td>Recorded webinars and tutorials</td>
<td>Discovery and Citation support: DOIs for datasets</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Large data support</td>
</tr>
</tbody>
</table>
RDM as a campus-wide service

2017 - Initial conversations between Dean for Research, Dean of Libraries, and CIO with department chairs and faculty

2018 - Provost-funded partnership: Princeton Research Data Service

4 New lines: 3 in the Library, 1 in Research Computing

2019 - PRDS launches

2019 - Princeton Working Research Data Service conversations begin
Why this new “working research data” service?

- Massive and growing amounts of digital data
- Where to store masses of data, multitude of options
- Storage hardware and software constantly need to be updated
- Backing up data appropriately
- Need to store data for analysis and preservation
- Uptake of computational methods
- Added requirements for security and audit options
Why this new “working research data” service?

Growing need for multidisciplinary collaboration

Next generation planning

People leave the university

Facilitation of retrieval and examination by future researchers
capacity, reliability, functionality, and performance describe, tag, seamlessly move, easily find, re-use & share with collaborators
What does TigerData look like?

Advanced Data and Storage Mgmt (RC, OIT, PUL IT)
- New data management system (MediaFlux)
- New heterogenous storage pool

Custom User Interface (PUL IT)
- New web front-end for data management (integrated with MediaFlux)
- “Presentation Layer” includes metadata functionality

Data Management Training (PUL PRDS)
- New series of live workshops, trainings and on-demand material to support data management
- Build a community of practice across campus
Governance

TigerData Steering Committee
Co-chairs: Curt Hillegas, Wind Cowles

Operations Working Group
Chair: Beth Holtz*
Co-chairs: Dan Chin, Steph Ayers

Metadata Working Group
Chair: Matt Chandler
Co-chair: Esmé Cowles

Departmental IT
Faculty
Researchers
OIT Management
Security
Risk Management
General Counsel
A Community of Practice

Plan
Acquire
Process
Analyze
Preserve
Share Results
Discover & Reuse
Where are we?

- Centrally funded
- Initial pools of storage have been purchased
- MediaFlux is installed in Staging, staff have been trained, and testing is being done
- Initial architecture design of TigerData
Where’s our focus now?

☐ Continuing to understand and define functional requirements and use cases served by TigerData

☐ Working with early adopters to understand value

☐ TigerData service management - ServiceNow/RT

☐ Establishing how we will collaborate to administer and support TigerData

☐ Acquire the next round of storage hardware
2023 Proposed Milestones

May
TigerData Pilot
- Production Environment Configured
- Plan for early adopters

June
Early Adopters
- Start migrating data onto TigerData storage for early adopters
- Define scope for UI
- Understand storage need, open RFP for storage purchase

Nov
TigerData UI Testing
- Iteratively test User Interface developed by the PUL Software Engineers
- Early Adopter UX testing, feedback

Dec
TigerData Go Live
- TigerData UI goes live
- New storage hardware installed
- Launch of formal training program for data managers
Thanks!

Curt Hillegas, PhD
Associate CIO
Research Computing & Princeton Institute for Computational Science and Engineering
curt@princeton.edu

Wind Cowles, PhD
Director of Data, Research, and Teaching
Princeton University Library
Head, Princeton Research Data Service
windcowles@princeton.edu