Collaborating to Manage Research Data

Zheng (John) Wang Rick Johnson Hesburgh Libraries University of Notre Dame



Notre Dame Research Profile

- Research & Sponsored Programs Revenue
- FY09: \$120.9 million
- FY14: \$189.7 million
- \$68.8 million (56.9%) increase
- 9.4% annual growth rate



Data Growth Potential

- Investment in Research
- FY09: \$88.9 million
- FY14: \$161.9 million
- \$73 million increase (82.1%)
- 12.7% annual growth rate
- Advancing Our Vision
 - \$13 million recurring investment
 - 10 Disciplines (i.e. computational data, adult stem cell, nuclear physics)
 - 80 new faculty

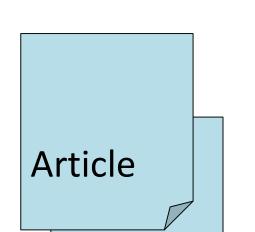


Nature of Research Data

- Often requires enormous storage (volume, accumulative)
- Often exists in diverse formats
- Increasingly owned by faculty across institutions
- Requires intensive resources and new expertise



Related Materials: Any Format



ETD





Statistical Data

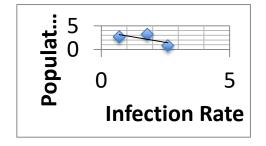




Image Data

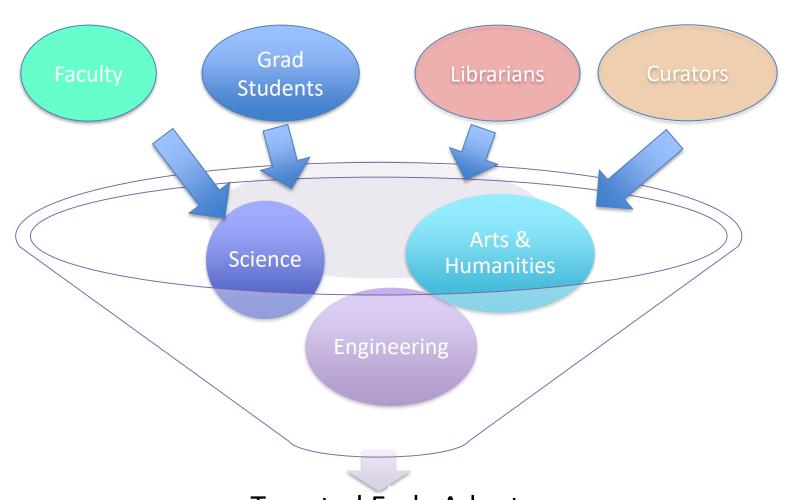


Library Context

- Institutional Digital Repository
- Part of Hydra community
- Vertical Successes
- High demand due to the Open Data Mandate and Public Access Policies
- Storage allocation internal to library needs
- Digital Initiative Program staffing level



User-Centric Personas



Targeted Early Adopters

Core Features



Design Strategy

- Design for everyone, but optimize for intermediate (Alan cooper)
 - Critical needs
 - What are the core? Other advanced services?

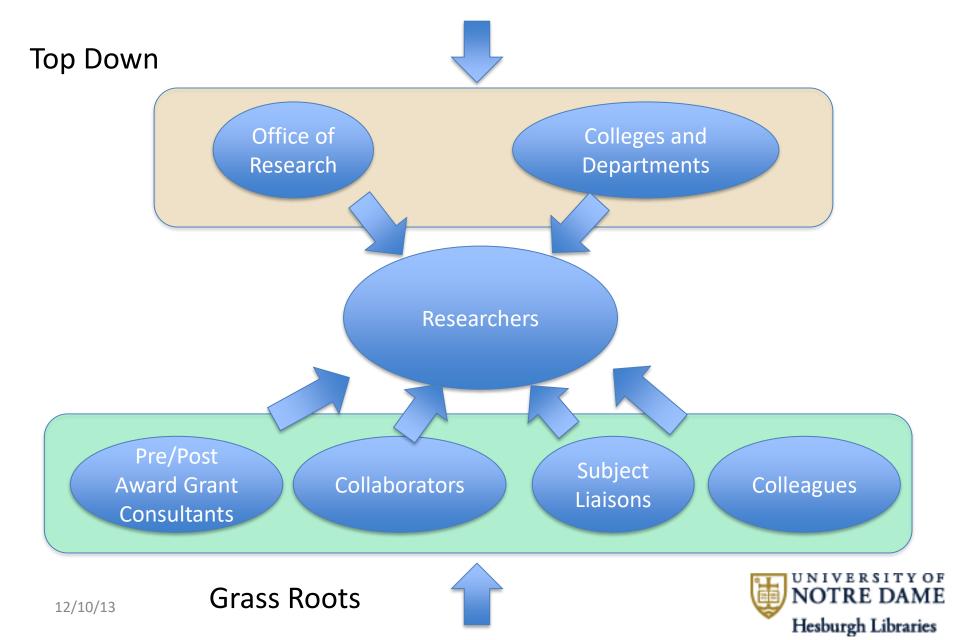


Partnerships are Key

- Increasing complexity demands support of experts
 - Funding Agency Requirements
 - Copyright and Intellectual Property
 - Metadata and Data Structuring
 - Data sharing and Preservation



How Do We Connect with Researchers?



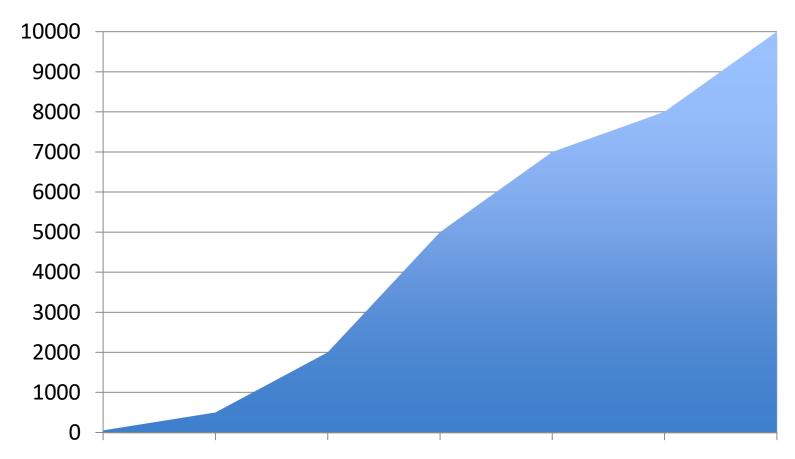
Priorities First Phase

- Wide format support, with mixed collections
- Focus on Preservation and Curation
- Get Users Engaged Early
- Release Early; Release Often



Grow Together

User Base





Features

Layered Continuous Improvement

Feedback determines Investment in most impactful advanced services



Continous Feedback from Growing User Base



Early User Feedback



Advanced Discovery, Harvesting, Curation, Analysis, Computation

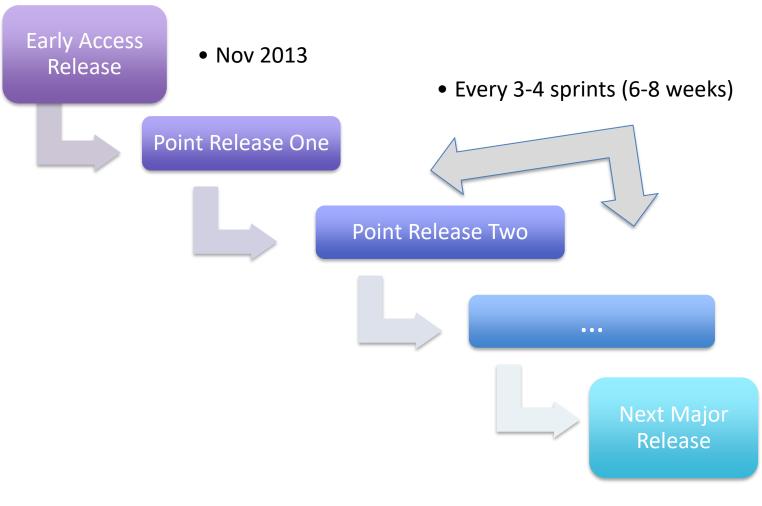
Enhanced Management, Presentation, Publishing, Collaboration

Core Features: Wide Format Support,
Preservation, Sharing

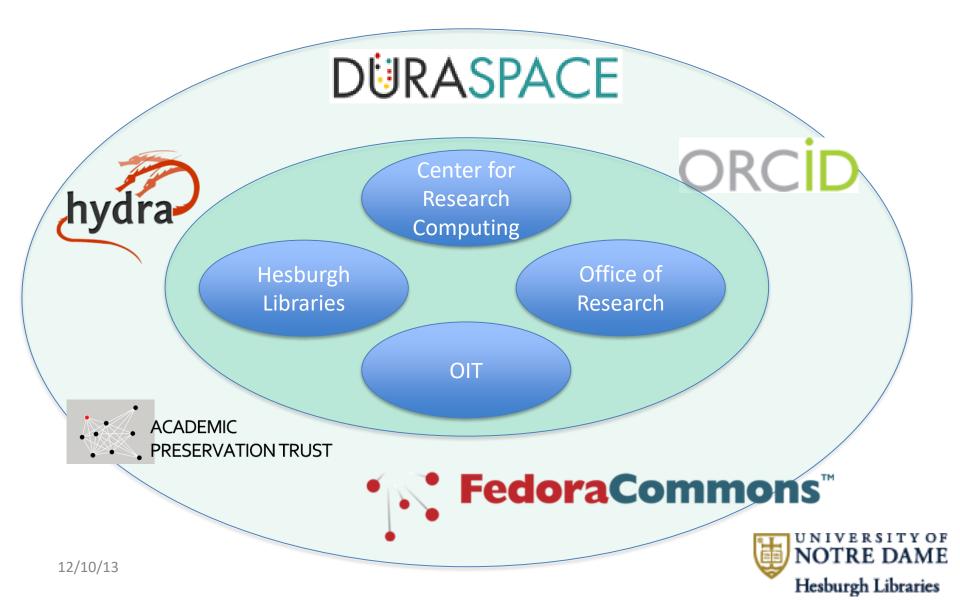
Identify Target Users and Most Common Needs



Release Early; Release Often



Sustain through Community



Sustain through Community

Shared IR Project

- Complete Multi-Institutional Collaboration from top to bottom within the Hydra Community
 - Shared Roadmapping and Governance
 - Community and Local Roles
 - Rotate Resources

DATA CURATION EXPERTS













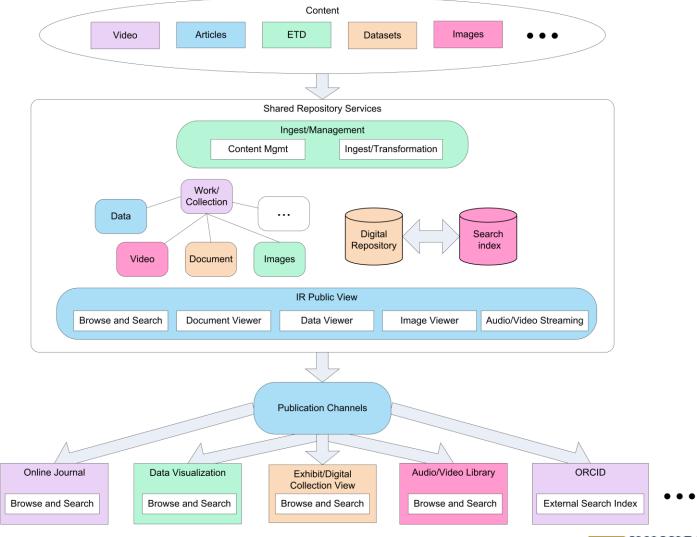


Community Development & Adoption **Notre Dame** CurateND Uva Libra2.0 Duke Cincinnati DCE Curate Va Tech LSE? Northwestern Penn State UNIVERSITY OF

NOTRE DAME

Hesburgh Libraries

Mixed Architecture





Future Priorities

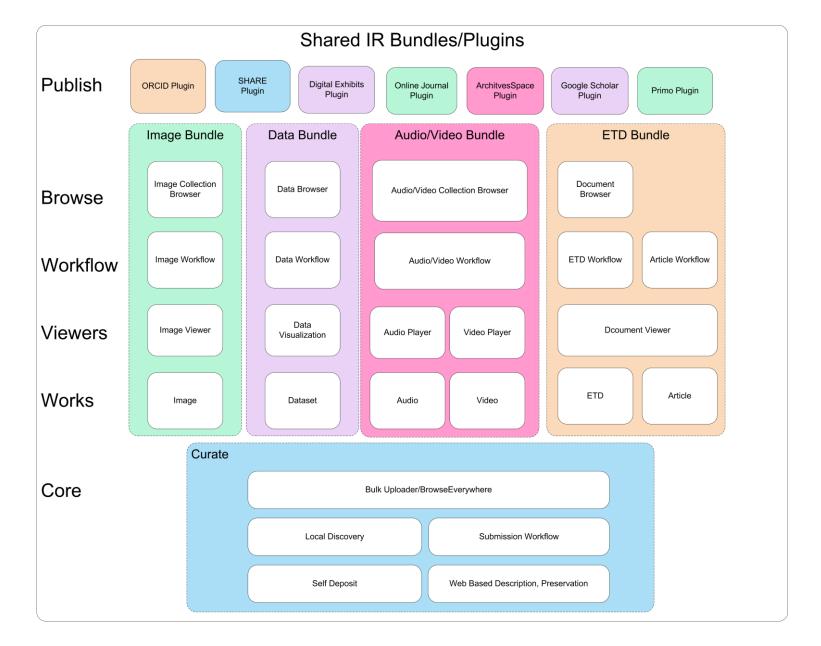
- By April
 - Improved Organization and Managed Collection
 Support
 - Submission, Review, and Publish Workflows
 - Integration with ORCiD
 - Bulk Ingest Support



Future Priorities

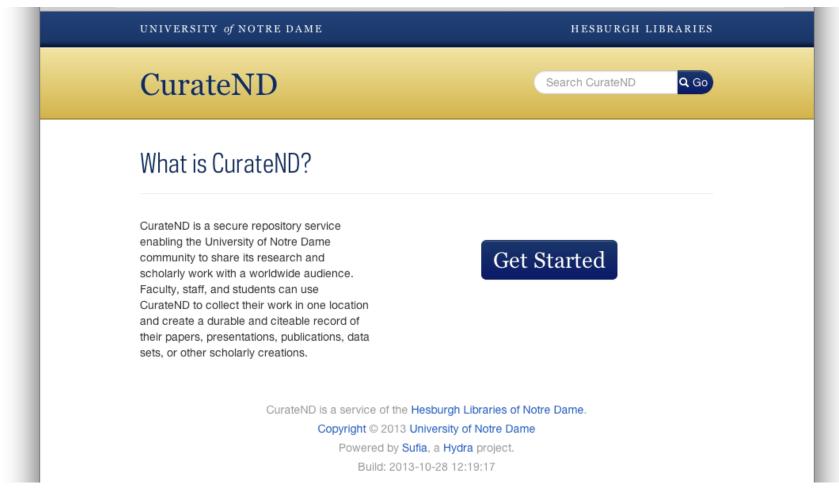
- April and beyond
 - Enhanced Publishing Layer to Digital Exhibits,
 Online Journals
 - Tuned and Optimized for large datasets
 - Integration with SHARE
 - Pluggable support for solution bundles







Quick Demo





FAQ

- Zheng (John) Wang, <u>zheng.wang@nd.edu</u>
- Rick Johnson, <u>rick.johnson@nd.edu</u>
- Hydra Shared IR Project Wiki:
 - https://wiki.duraspace.org/display/hydra/Shared+ IR+project
- http://curate.nd.edu

