

# DASPOS WORKSHOP

ACM/IEEE JOINT CONFERENCE ON DIGITAL LIBRARIES – JULY 24, 2013

## PURR: A University Data Repository Service



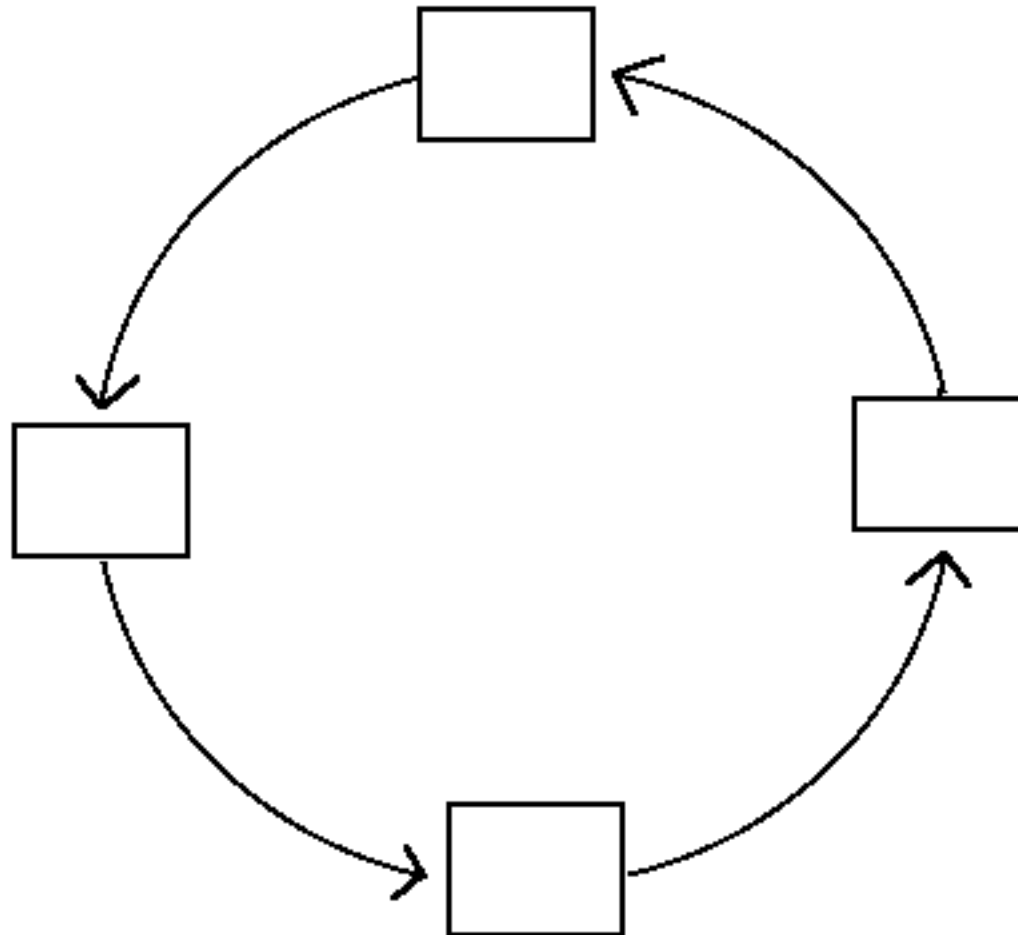
**PURDUE**  
UNIVERSITY  
LIBRARIES

Michael Witt  
Associate Professor of Library Science  
Web: <http://www.lib.purdue.edu/research/witt>  
E-mail: [mwitt@purdue.edu](mailto:mwitt@purdue.edu)

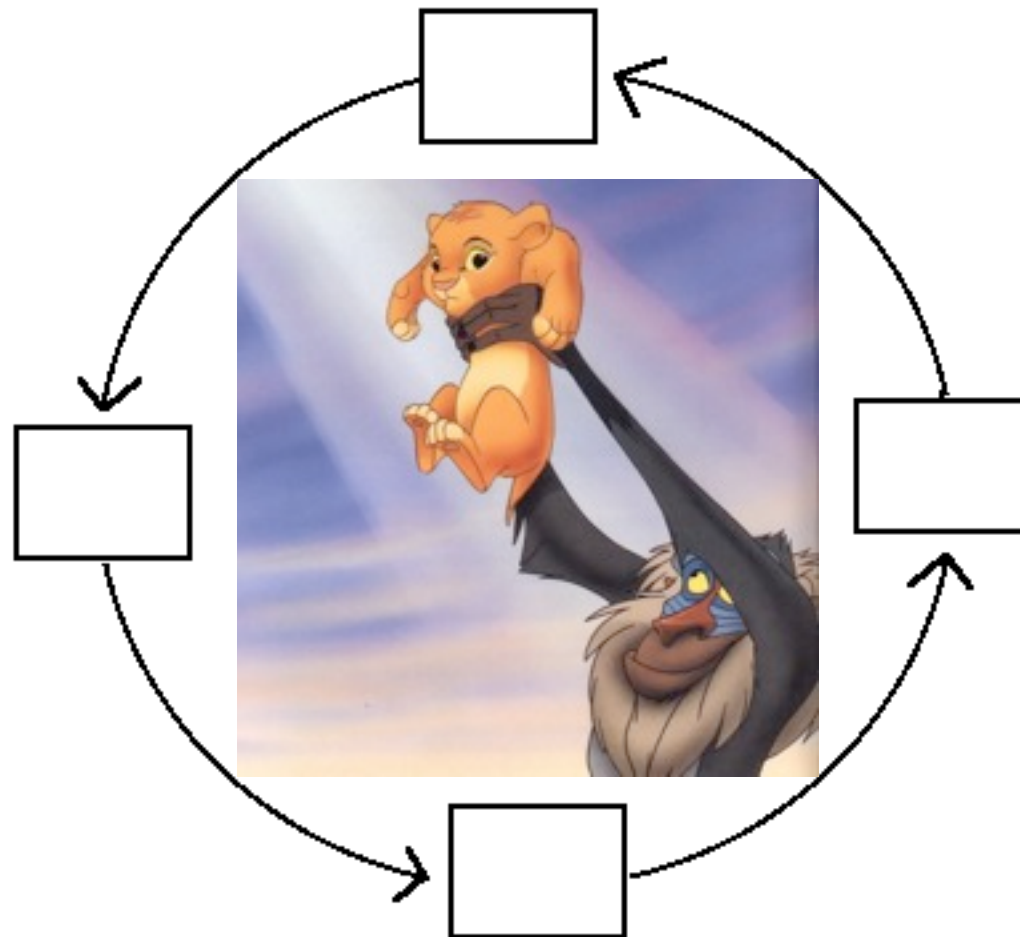
# SOME TYPES OF DATA REPOSITORIES

- Publisher, e.g., [Dryad](#)
- Sub/Disciplinary, e.g., [RKMP](#)
- Consortium, e.g., [ICPSR](#)
- Country, e.g., [Research Data Australia](#)
- Government, e.g., [Data Portal India](#)
- Research center, e.g., [NASA GES DISC](#)
- Instrument, e.g., [CHANDRA](#)
- General-purpose, e.g., [FigShare](#)
- Roll-your-own, e.g., [DataVerse](#)
- University, e.g., [PURR](#)

# INSERT LIFECYCLE DIAGRAM HERE



# INSERT LIFECYCLE DIAGRAM HERE





## Publish Datasets with DOIs

Use PURR to publish datasets with Digital Object Identifiers (DOI) that make it easier for people to cite your data and give you credit. Purdue is a founding member of DataCite, the international agency that registers DOIs for data.

Learn More

header  
<identifier> <http://dx.doi.org/10.4231/D39P2W550>  
<datestamp>2012-10-24T12:28:53-04:00</datestamp>  
<getSpec>Dataset</setSpec>  
header  
metadata  
dc:identifier http://dx.doi.org/10.4231/D39P2W550  
dc:creator Khan, Ishita K.; Chitale, Catherine; Rayon, Catherine; Kihara, Daisuke  
dc:subject Tansinectin  
dc:subject moonlighting proteins

### Start Your Research Project



**Create a Data Management Plan**  
Learn about the detailed requirements for your data management plan (DMP). Funding agency requirements are very specific and our DMP resources can help you to clear up any confusion. [Get Started](#)



**Upload Research Data to Your Project**  
Create a project to upload and share your data with collaborators using our step-by-step form to guide you through the process. Invite collaborators from other institutions to join your project. [Create a Project](#)



**Publish your Dataset**  
Package, describe, and publish your dataset to receive a DOI. Publishing will ensure your dataset is citable, searchable, and archived for the long-term. [See Publication Process](#)

### Featured Dataset

**Evaluation of Function Predictions for Moonlighting Proteins**

By Ishita K. Khan, Meghana Chitale, Catherine Rayon, Daisuke Kihara  
*Purdue University, Université de Picardie Jules Verne*

Supplemental datasets used for evaluation of function predictions for moonlighting proteins.

### Do you have a question?

**Ask a Librarian**

Send us your email question.

Enter your email address

Your Question/Message

Send

<http://purr.purdue.edu>

# PURR TEAM

- **Executive Committee:** Dean of Libraries, Vice President for Research, Chief Information Officer
- **Steering Committee:** 2 from libraries, 2 from IT, 2 from research office and sponsored programs, 3 domain faculty researchers
- **Personnel:** Project Director (.50), Technologists (3.85), HUBzero Liaison (.35), Metadata Specialist (.20), Digital Archivist (.25), Digital Data Repository Specialist (1.0)

# PURR BY THE NUMBERS

- Included in 720 data management plans (DMPs) with grant proposals
- 49 grants awarded
- 761 registered researchers
- 217 research projects

DMP analysis (n=111 most recent NSF proposals from our university)

- 49% PURR
- 29% Local computer or server
- 14% Disciplinary repository (e.g., ICPSR, Protein Data Bank, nanoHUB, NEES)
- 8% No data or not applicable

# PRESERVATION DESIGN

Initial commitment of 10 years

- data producer or dept can fund for longer
- otherwise remanded to library collection

Design guided by ISO 16363 / TRAC

- Organization infrastructure
- Digital object management
- Technical infrastructure & Security Risk Management



# ARCHIVAL INFORMATION PACKAGE

Bagit “bag” contains:

- bag declaration file, manifest file, data files

Metadata file (XML):

- METS wrapper
- Dublin Core and MODS (descriptive metadata)
- PREMIS (preservation metadata)

MetaArchive: LOCKSS replication network (7 copies)

# IN THE NEXT YEAR...



## Online software development and publication:

- git integration
- shared online development environment with [Rappture toolkit](#)
- software publication workflow and archive
- best example: [nanoHUB.org](#)