Policy Realities in Research Data Management

OCLC Research Library Partnership
Works in Progress Webinar
The Realities of Research Data Management

• Four-part series exploring how research universities are addressing the challenge of managing research data throughout the research lifecycle

• Examines RDM case studies at:
  – University of Edinburgh (UK)
  – University of Illinois at Urbana-Champaign (US)
  – Monash University (Australia)
  – Wageningen University & Research (the Netherlands)

• oc.lc/rdm
Research Data Management
Service Categories

EDUCATION
Raise awareness of RDM’s importance, encourage RDM skill-building, and disclose RDM tools and resources

EXPERTISE
Decision support for, and customized solutions to, specific research data management problems

CURATION
Technical infrastructure and related services that support data management throughout the research cycle

Figure 1. RDM Service Categories. From A Tour of the Research Data Management (RDM) Service Space. The Realities of Research Data Management, Part 1. (CC BY 4.0)
Heidi Imker
Director, Research Data Service
University of Illinois at Urbana-Champaign

Jacquelijn Ringersma
Head of Digital Production Centre
Wageningen University & Research
Realities of Research Data Repository Policies

OCLC Webinar
July 25, 2017

Heidi Imker, PhD
Director, Research Data Service
Associate Professor, University Library
University of Illinois at Urbana-Champaign
The Illinois Data Bank is a self-serve publishing platform that centralizes, preserves, and provides persistent and reliable access to Illinois research.

- Custom web application that interacts with the preservation system that supports all digital content at our library
- A sibling repository to our document repository, IDEALS
- Overall development is described in:
# Illinois Data Bank Policies

**Live Policies:** [https://databank.illinois.edu/policies](https://databank.illinois.edu/policies)

<table>
<thead>
<tr>
<th>Policy Title</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access and Use Policy</td>
<td><a href="http://hdl.handle.net/2142/91041">http://hdl.handle.net/2142/91041</a></td>
</tr>
<tr>
<td>Accession Policy</td>
<td><a href="http://hdl.handle.net/2142/91042">http://hdl.handle.net/2142/91042</a></td>
</tr>
<tr>
<td>Deposit Agreement</td>
<td><a href="http://hdl.handle.net/2142/91043">http://hdl.handle.net/2142/91043</a></td>
</tr>
<tr>
<td>Preservation Policy</td>
<td><a href="http://hdl.handle.net/2142/91044">http://hdl.handle.net/2142/91044</a></td>
</tr>
<tr>
<td>Preservation Review, Revision, Retention, Deaccession, and Withdrawal Procedure</td>
<td><a href="http://hdl.handle.net/2142/91045">http://hdl.handle.net/2142/91045</a></td>
</tr>
<tr>
<td>Withdrawal Guidelines</td>
<td><a href="http://hdl.handle.net/2142/91616">http://hdl.handle.net/2142/91616</a></td>
</tr>
<tr>
<td>Preservation Review Guidelines</td>
<td><a href="http://hdl.handle.net/2142/91617">http://hdl.handle.net/2142/91617</a></td>
</tr>
</tbody>
</table>
Me: Why can’t I just make all the decisions?
Me: Why can’t I just make all the decisions?

Elise: Because you’re not a dictator.
Policies Shmolicies

But also ...

- **Consistency as staff turns over**
- **De-personalization**
- **Clarity and transparency**
Proposed Timeline

**Drafted Milestones**

**August 15 2015**
- Minimal prototype

**September 15 2015**
- First round user testing (alpha 1)
- Policy pre-review w/ legal

**November 15 2015**
- Second round user testing begins (alpha 2)
- Policy review w/ campus group

**March 15 2016**
- Third round user testing begins (beta)
- Policy finalized with legal

**May 15 2016**
- Illinois Data Base ready for campus use
Proposed Timeline

Drafted Milestones

August 15 2015
- Minimal prototype

September 15 2015
- First round user testing (alpha 1)
  - Policy pre-review w/ legal

November 15 2015
- Second round user testing begins (alpha 2)
  - Policy review w/ campus group

March 15 2016
- Third round user testing begins (beta)
  - Policy finalized with legal

May 15 2016
- Illinois Data Base ready for campus use

Milestones for policy development literally coincided with milestones for technical development of the Illinois Data Bank.
Process

- Created draft policies largely based on existing policies
Process

- Created draft policies largely based on existing policies
- Identified campus representatives for a Illinois Data Bank Policy Review Group (list vetted by Library Administration and OVCR)
## Composition of Review Group

<table>
<thead>
<tr>
<th>Role</th>
<th>Campus Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Senior IT Security Engineer</td>
<td>Technology Services</td>
</tr>
<tr>
<td>Export Compliance Officer</td>
<td>Office of the Vice Chancellor for Research</td>
</tr>
<tr>
<td>Research Ethics Officer</td>
<td>Office of the Vice Chancellor for Research</td>
</tr>
<tr>
<td>Deputy CIO of Research IT and research leader at National Center for Supercomputing Applications</td>
<td>Technology Services; National Center for Supercomputing Applications</td>
</tr>
<tr>
<td>Director, Institutional Review Board</td>
<td>Office for the Protection of Research Subjects</td>
</tr>
<tr>
<td>Associate Vice Chancellor for Research and Director of Sponsored Programs</td>
<td>Office of the Vice Chancellor for Research; Office of Sponsored Programs</td>
</tr>
<tr>
<td>Director, Records and Information Management Services</td>
<td>Records and Information Management Services</td>
</tr>
<tr>
<td>Senior Associate University Counsel</td>
<td>University Counsel</td>
</tr>
<tr>
<td>Senior Technology Manager (1)</td>
<td>Office of Technology Management</td>
</tr>
<tr>
<td>Senior Technology Manager (2)</td>
<td>Office of Technology Management</td>
</tr>
<tr>
<td>Assistant Dean for the Graduate College</td>
<td>Graduate College</td>
</tr>
<tr>
<td>Associate University Librarian for Research</td>
<td>University Library</td>
</tr>
<tr>
<td>Manager, Scholarly Communication and Repository Services</td>
<td>University Library</td>
</tr>
<tr>
<td>Engineering, Physics and Astronomy Librarian</td>
<td>University Library</td>
</tr>
</tbody>
</table>
Process

- Created draft policies largely based on existing policies
- Identified campus representatives for a Illinois Data Bank Policy Review Group (list vetted by Library Administration and OVCR)
- Gave contact in campus legal a preview of the draft policies for pre-feedback
Process

• Created draft policies largely based on existing policies
• Identified campus representatives for a Illinois Data Bank Policy Review Group (list vetted by Library Administration and OVCR)
• Gave contact in campus legal a preview of the draft policies for pre-feedback
• Assembled the Policy Review Group for an “orientation meeting”
Process

• Created draft policies largely based on existing policies
• Identified campus representatives for a Illinois Data Bank Policy Review Group (list vetted by Library Administration and OVCR)
• Gave contact in campus legal a preview of the draft policies for pre-feedback
• Assembled the Policy Review Group for an “orientation meeting”
• Gave members 1 month to review policies and provide feedback prior to subsequent meeting
Process

- Created draft policies largely based on existing policies
- Identified campus representatives for a Illinois Data Bank Policy Review Group (list vetted by Library Administration and OVCR)
- Gave contact in campus legal a preview of the draft policies for pre-feedback
- Assembled the Policy Review Group for an “orientation meeting”
- Gave members 1 month to review policies and provide feedback prior to subsequent meeting
- Incorporated any minor feedback into revised drafts and selected only major feedback items for group discussion at meeting
Process

- Created draft policies largely based on existing policies
- Identified campus representatives for a Illinois Data Bank Policy Review Group (list vetted by Library Administration and OVCR)
- Gave contact in campus legal a preview of the draft policies for pre-feedback
- Assembled the Policy Review Group for an “orientation meeting”
- Gave members 1 month to review policies and provide feedback prior to subsequent meeting
- Incorporated any minor feedback into revised drafts and selected only major feedback items for group discussion at meeting
- Repeated review/meet process 3 times total
Process

• Created draft policies largely based on existing policies
• Identified campus representatives for a Illinois Data Bank Policy Review Group (list vetted by Library Administration and OVCR)
• Gave contact in campus legal a preview of the draft policies for pre-feedback
• Assembled the Policy Review Group for an “orientation meeting”
• Gave members 1 month to review policies and provide feedback prior to subsequent meeting
• Incorporated any minor feedback into revised drafts and selected only major feedback items for group discussion at meeting
• Repeated review/meet process 3 times total
• Presented finalized polices to Library Administration and OVCR
Actual Timeline

**Drafted Milestones**

- **August 15 2015**
  - Minimal prototype

- **September 15 2015**
  - First round user testing (alpha 1)
  - Policy pre-review w/ legal

- **November 15 2015**
  - Second round user testing begins (alpha 2)
  - Policy review w/ campus group

- **March 15 2016**
  - Third round user testing begins (beta)
  - Policy finalized with legal

- **May 15 2016**
  - Illinois Data Base ready for campus use

**Actual Milestones**

- **August 15 2015**
  - Minimal prototype

- **October 09 2015**
  - Policy pre-review w/ legal

- **January 13 2016**
  - 1st Policy Review w/ campus group

- **February 19 2016**
  - 2nd Policy Review w/ campus group

- **April 01 2016**
  - 3rd Policy Review w/ campus group

- **May 03 2016**
  - Policies vetted by OVCR and Library Administration

- **May 15 2016**
  - Illinois Data Base ready for campus use
Why did this process work?

- We gave ourselves (and others) a fairly reasonable amount of time
- We started with solid drafts
- It was supported by the OVCR and Library Administration
- We made it clear that we didn’t want to waste peoples’ time
- We made it clear that their input was necessary and important
- We backed that up by communicating well and being very organized
Have the policies been useful?

- Unexpectedly huge benefit to developing policies in concert with development of the repository itself
- Clarified positions and processes *for ourselves*
- Engagement opportunity with members of the Review Group
- Uses so far:
  - Avoiding emotional decision-making
  - Showing our dedication, professionalism, and thoughtfulness
  - Setting an example of transparency
What’s to come?

• Policy review process
• Tweaks
• Testing
To-Be Tested Review Strategy

Calculate a “Review Indicator” (RI) to help determine which data needs review.

\[ RI = \frac{\text{Downloads} \times \text{Relationships} \times \text{Format}}{\text{Bytes}} \]

where format is represented with a scale that increases with a format’s preservation “difficulty.”

Should We Keep Everything Forever? Determining Long-Term Value of Research Data
B Anderson, S Braxton, E Dunham, H Imker, K Rimkus
iPres 2016 Poster  http://hdl.handle.net/2142/91659
Thank you for the opportunity and interest!

Heidi Imker
@imkerinfo
imker@illinois.edu
researchdataservice.illinois.edu
databank.illinois.edu

• Elise Dunham - RDS Curator
• Steve Anderson - University Counsel
• Beth Namachchivaya - AUL for Research
• Sarah Shreeves - IDEALS Coordinator

We’re hiring!  go.illinois.edu/joinRDS
Wageningen University & Research
Wageningen University & Research
Wageningen University & Research

University (food, environment, plant, animal, social)
- ~ 10,000 students
- ~ 2,500 employees
- Yearly ~ 250 PhD student graduations

Research (food, environment, plant, animal, social)
- ~ 2,500 employees
- National research institutes
National context: RDM
RDM policy history
RDM policy history: Engage with researchers
RDM policy history

- 2010: first discussion on data between researchers, graduate schools and library
- 2014: first RDM policy: all PhD students and University chairs must have a Data Management Plan
RDM policy history

- 2010: first discussion on data between researchers, graduate schools and library
- 2014: first RDM policy: all PhD students and University chairs must have a Data Management Plan
- 2015: RDM services by Library, IT, legal and archive
Developing RDM services

Before	During	After

Research project time line
Developing RDM services

Data archiving at DANS:

- Trusted archive, PID, sustainable formats, access right help services

DMP training:

- We developed our Wageningen DMP training with the Graduate Schools. The training was copied by other Universities in the Netherlands.

Software archiving:

- On specific researcher request we created a software archiving environment on GitHub, which is now also used by non-Wageningen researchers.
Registration of datasets

Why:
One view for all Wageningen datasets
No institutional data archive; data are archived in multiple locations
Ready for Evaluation protocols

What:
Dataset has a PID
Publication has a PID
We combine the two in one new record: the enriched publication type
Recent developments
RDM policy history

- 2010: first discussion on data between researchers, graduate schools and library
- 2014: first RDM policy: all PhD students and University chairs must have a Data Management Plan
- 2015: RDM services by Library, IT, legal and archive
- 2016: Evaluation of RDM behaviour
- 2017: Additional RDM policy
Additional RDM policy

Dean of Science asked RDM support services

Make a proposal for a new, more advanced policy for Data Storage, Data Archiving and Data Registration

- National frameworks
- Diversity in the organisation
- Best practices
Use cases

Matching use cases to criteria

Defining policy

Easy adoption

Criteria from frameworks
Brief on the new RDM policy

All Wageningen University & Research researchers should:

*During research:*
Store their data on WUR servers, or on servers approved by IT (RDM services)

*After research:*
Archive all data underlying a publication in a trusted archive, which meets FAIR criteria
Register their data in the institutional repository
Advantages of engaging researchers

It created RDM awareness!

Build on what is already done
  It avoids building or buying new infrastructure
  It meets researchers workflows, and practices
  It facilitates adoption

Less rules, better services, happy clients/researchers
I share my tree research data online at WUR Library

Wageningen University & Research - Library
Accelerate your research
www.wur.eu/library