Developing the Scholarly Communication Ecosystem: A CMU Perspective

David Scherer, Ole Villadsen, Keith Webster
@davidascherer @OleVilladsen @cmkeithw
#cni17f    #CMUecosystem
Strategic context
Creating a 21st century library at CMU
STRATEGIC RECOMMENDATION

Create a 21st century library that serves as a cornerstone of world-class research and scholarship.

In a world where digital is becoming the default format for information, the library will remain a vital presence on campus, sustaining serious scholarship and providing opportunities for interactive research and study environments. To support this important work for students, faculty, and staff, and to create 21st century library spaces for 21st century learners, the library will:

- Develop information specialists as partners in research, teaching, and learning.
- Collaborate with peer institutions to provide coordinated access to a global collection of information resources.
- Steward the evolving scholarly record, and champion new forms of scholarly communication.
- Be recognized globally as a leader in the development of the scholarly information ecosystem.
- Discovering & Planning
  - Researching scientific literature and patents
  - Designing Studies
  - Procuring funding
  - Managing reputations
  - Measuring impact

- Experimenting & Managing Data
  - Searching for collaborations
  - Experimenting
  - Managing and analysing data

- Driving & Establishing Impact

- Authoring, Publishing & Disseminating Findings
  - Writing up findings
  - Publishing article(s)
  - Disseminating outputs and outcomes
OPEN SCIENCE
Outcomes

Method
Evidence
Discussion

Discussion
Revision
Re-use

PROCESS

ANNNOUNCING THE PARTNERSHIP

DIGITAL SCIENCE
Strategic partnership announced

RESEARCH INFORMATION
Building the library of the future

LSE IMPACT BLOG
Reimagining the role of the library

DIGITAL SCIENCE PODCAST
The 21st century librarian
April 10 - 14: CELEBRATE National Library Week!

Join us for games, treats & buttons to learn how we can help you at every step of the research process.

**Discover**
MONDAY, 4/10
Hunt Library, 1-2:30pm

**Organize**
TUESDAY, 4/11
Sorrells Library, 1-3:00pm

**Create**
WEDNESDAY, 4/12
CUC, 11:30 - 1:30 pm

**Share**
THURSDAY 4/13
Hunt Library, 11 am - 1:00 pm

**Impact**
FRIDAY 4/14
Scott Hall, 11:30 - 1:30 pm

Carnegie Mellon University Libraries
Towards a Comprehensive Institutional Repository
“A number of participants noted that as they strategized about the direction of their repository should take, they had institutional goals in mind. For example, in an institution that is aiming to expand its research profile, the library wants to understand how it can assist with helping the university account for research, bring together outputs of research, and showcase those results.”
Rethinking Institutional Repository Strategies

Report of a CNI Executive Roundtable Held April 2 & 3, 2017

“There’s a clear path of evolution in architectural thinking, moving from the IR as platform to IR as a service to IR as a bundle of related services.”

“Although we are aware that at least one CNI member institution has begun using figshare as its repository platform, none of the participants in the roundtable reported the development at their institution. However, we know that many individual researchers in other institutions are choosing to deposit research outputs in figshare or other platforms rather than [sic] locally...”
Historical State of Open Science at CMU

- Traditional IR
- Archives
- Data Repository
KiltHub Repository

Discover research from Carnegie Mellon University –

WEAVING THE FABRIC OF YOUR RESEARCH

Carnegie Mellon University
KiltHub Repository

Provided by the University Libraries, KiltHub (kilthub.cmu.edu) is the comprehensive institutional repository and research collaboration platform for research data and scholarly outputs produced by members of Carnegie Mellon University and their collaborators. KiltHub collects, preserves, and provides stable, long-term global open access to a wide range of research data and scholarly outputs created by faculty, staff, and student members of Carnegie Mellon University in the course of their research and teaching.

About KiltHub

Contact Us

KiltHub Repository
kilthub@andrew.cmu.edu
412-268-2443
KiltHub Twitter

About

Why deposit to KiltHub?

What Kinds of Materials does KiltHub Accept?

What Are The Requirements For Depositing Materials in KiltHub?
Why Deposit to KiltHub?

**Make it Open** - Share your data and other scholarly outputs, making your research and scholarship more visible to the world.

**Simplify Your Research Workflow** - It’s easy to upload, manage, share, and publish your data and scholarly outputs with a DOI.

**Get Credit for Your Work** – Track the usage of the DOI, Downloads, and Views of your work.

**Comply with Publishers and Funders** - Deposit your data, publications, and other grant outputs to show compliance with funder open access requirements.

**We Can Help** - Liaison Librarians can meet with you to guide you through the repository and deposit process. Contact your Liaison Librarian to setup a consultation.
KiltHub team

**Ole Villadsen**
Research Liaison in the fields of Cybersecurity and Information Systems

**David Scherer**
Scholarly Communications and Research Curation Specialist

**Lisa Zilinski**
Research Data Consultant

**Katie Behrmann**
Repository Specialist

**Ann Marie Mesco**
Data Coordinator

**Eric Kaltman**
CLIR Fellow for Data Curation in the Sciences

**Emma Slayton**
CLIR Fellow for Data Visualization and Curation
What Kind of Materials Does KiltHub Accept?

**Research Data (Data Repository)**
A dataset is typically a file or collection of files and its corresponding metadata. This metadata may include preservation and presentation information, as well as annotations and ancillary content that is associated with the data. KiltHub is limited in scope to the raw data, or data that supports the claims made in publications/presentations/figures.

**Scholarly Outputs (Traditional IR)**
When evaluating scholarly outputs for inclusion in KiltHub, scholarly outputs should be broadly understood to mean the published or otherwise publicly presented outputs, created by individuals or units within Carnegie Mellon University in the course of their scholarly or professional work that represents the intellectual life and academic endeavor of the university, and may have enduring value.
Towards an “Enterprise Repository”

Collaboration Space and Projects

Integrations with Researcher Tools
Research Information Management (RIM) Systems and Repositories
RIMs and IRs
RIMs and IRs
RIMs and IRs
IR to RIM

Metadata Harvesting and Data Match
RIMs to IRs
RIMs and IRs

Monitoring OA Status and Engagement
Interacting with the Faculty

Challenges and Lessons Learned
KiltHub Benefits to the Researcher

- **Compliance** - Deposit data, publications, and other grants to show compliance with funder and/or publisher open access requirements

- **Discoverability** - figshare’s large footprint on the Web increases discoverability through search engines (e.g. Google)

- **Make it Open** - Share your data and other scholarly outputs, making your research and scholarship more visible to the world.

- **Simplify Research Workflow** - Easy to upload, manage, share, and publish data and scholarly outputs with a DOI.
Funder and Publisher Compliance

2. **Data Accessibility**: The NOAA Program recommends (or requires) that public access to grant/contract-produced data be enabled as follows (one or more of the following, or alternate text as appropriate, can be provided as guidance; Data Management Plans submitted with Proposals should reflect the option(s) provided by the Program):

   - Funding recipients are expected to submit data to NOAA National Centers for Environmental Information (NCEI), which will provide public access and permanent archiving.
   - The NOAA Program has held preliminary consultation with NCEI regarding these pending data.
   - Data are to be submitted to the following International Council for Science (ICSU) World Data System facility: ____________________________.
   - (See list at [https://www.icsu-wds.org/community/membership/regular-members](https://www.icsu-wds.org/community/membership/regular-members))
   - The following NOAA facility (other than NCEI) will operate a publicly accessible online data server for these data: ____________________________.
   - An existing publicly accessible online data server at the funded institution is to be used to host these data (describe in proposal).
   - Funding recipients will establish their own data hosting capability (describe in proposal).

**Recommended Repositories**

PLOS requires that authors comply with field-specific standards for preparation and recording of data and select repositories appropriate to their field, for example deposition of microarray data in ArrayExpress or GEO; deposition of gene sequences in GenBank, EMBL or DDBJ; and deposition of ecological data in Dryad. Authors are encouraged to select repositories that meet accepted criteria as trustworthy digital repositories.

PLOS has identified a set of established repositories below, which are recognized and trusted within their respective communities. For further information on environmental and biomedical science repositories and field standards, we suggest utilizing FAIRsharing; we have also created a FAIRsharing page of PLOS recommended data repositories. Additionally, the Registry of Research Data Repositories (RoRData) is a full-scale resource of registered repositories across subject areas. Both FAIRsharing and RoRData provide information on an array of criteria to help researchers identify the repositories most suitable for their needs (licensing, certificates and standards, policy, etc.).

Authors are encouraged to select the repository most appropriate for their research. PLOS does not dictate repository selection for the data access policy. If authors use repositories with stated licensing policies, the policies should not be more restrictive than the Creative Commons Attribution (CC BY) license. More information about the content license can be found in our licenses and copyright policy.

If no specialized community-endorsed open repository exists, institutional repositories that use open licenses permitting free and unrestricted use or public domain, and that adhere to best practices pertaining to responsible data sharing, sustainable digital preservation, proper citation, and openness are also suitable for data deposition.

**Cross-disciplinary repositories**

- Dryad Digital Repository
- Figshare
- Harvard Dataverse Network
- Open Science Framework
- Zenodo
How are researchers finding datasets?


Culture and Breaking Change: A Survey of Values and Practices in 18 Open Source Software Ecosystems

Software ecosystems have become one of the most important ways to organize software development, and to maintain and reuse code packages. But coordination can be a major challenge in software ecosystems when packages change, since packages tend to be highly interdependent yet independently maintained. The culture of an ecosystem includes those values and practices associated with managing change. We conducted a survey of thousands of developers in more than a dozen ecosystems, asking them about the values and practices that make up their communities' distinctive cultures; as well as the perceived power of different stakeholders and perceived health of the community.

This dataset release shares anonymized data from the survey.

Culture and Breaking Change: A Survey of Values and Practices in 18 Open Source Software Ecosystems

Two Critical Components Of Your Innovation Ecosystem - Forbes

Aug 14, 2013 - Finally, this framework will ensure that culture -- the most critical piece of innovation -- is on everyone's agenda. Again, it's easy to get distracted from doing hard work, through the allure of the less complex. Driving innovation is a matter of staying focused on culture changes (software). Yes, you upgrade...

Is culture-change the same as software-change? -- Tom Graves...
RIM Implementation Challenges

Curating profiles

Decentralized campus

Technology Integration


CC BY-ND 2.0, https://www.flickr.com/photos/epublicist/3546059144
Lessons Learned

- More limited discovery on Google Scholar than expected (under investigation)
- Formalizing data submission and deposit requirements & boilerplate descriptions of figshare for DMPs
- Balancing “carrots and sticks” when marketing RIM and Repository
  - Position the Libraries as “H&R Block, not the IRS”: We really are here to help
Next Steps and Future Expansions of the Ecosystem

Next Steps
Continued Rollout and Engagement of KiltHub

Deployment of Elements and Use Cases:
- Faculty Profiles/Documents
- Annual Reporting
- Support for Promotion Review and Tenure

Future Expansions of the Ecosystem
Completing the Research Lifecycle Support Loop
- E.g. Electronic Lab Notebooks, Protocol Repository, Collaborative Writing Tools, etc.
Thank you
Questions / Comments

David Scherer
daschere@andrew.cmu.edu
@davidascherer

Ole Villadsen
ovillads@andrew.cmu.edu
@OleVilladsen

Keith Webster
kwebster@andrew.cmu.edu
@CMKeithW

Pages
KiltHub:
https://kilthub.cmu.edu

About Page: http://library.cmu.edu/kilthub/about

Guides
http://library.cmu.edu/kilthub/deposit
Deposit Guide
Projects Guide
Collections
Data Deposit Guide