DIGITAL RESEARCH WORKFLOWS:
STRATEGIC ISSUES FOR HIGHER EDUCATION

CNI
San Diego
April 12-13, 2018
Roger Schonfeld and Donald Waters
Scholarly Communications

Stuff (the scholarly and cultural record)

- Preservation
- Description Organization Curation
- Discovery and Access
- Use in research and teaching
- New insights, new data
- Scholarly Publishing
Stuff (the scholarly and cultural record)

- Preservation
- Description Organization Curation
- Discovery and Access
- Use in research and teaching
- New insights, new data
- Scholarly Publishing
Publishing

Stuff (the scholarly and cultural record)

Scholarly Publishing

New insights, new data

Use in research and teaching

Preservation

Description Organization Curation

Discovery and Access
Preservation

Scholarly Publishing

New insights, new data

Use in research and teaching

Preservation

Description Organization Curation

Discovery and Access

Stuff (the scholarly and cultural record)
Scholarly Communications Pre-1990s

- Preservation
- Description Organization Curation
- Discovery and Access
- Use in research and teaching
- New insights, new data
- Scholarly Publishing

Stuff (the scholarly and cultural record)

ICPSR

OCLC
SCHOLARLY COMMUNICATIONS
1990s

Preservation

Description Organization Curation

Discovery and Access

Use in research and teaching

New insights, new data

Scholarly Publishing

Stuff (the scholarly and cultural record)

arXiv.org

PROJECT MUSE
JSTOR
INTERNET ARCHIVE
OCLC
SCHOLARLY COMMUNICATIONS
EARLY 2000S

- Scholarly Publishing
- Preservation
- Description Organization Curation
- Discovery and Access
- New insights, new data
- Use in research and teaching
- Stuff (the scholarly and cultural record)

arXiv.org

Logos: Project MUSE, JSTOR, LOCKSS, Internet Archive, PORTICO, OCLC
Research

Stuff (the scholarly and cultural record)

Use in research and teaching

Description Organization Curation

Discovery and Access

ARTstor

The Open Library Environment

Open, Flexible, Extensible, Sustainable

Wikipedia

YouTube

flickr

PROJECT MUSE

STOR

collectionspace

ArchivesSpace

Google

ExLibris

LibraryThing

collection space

OCLC

DPLA

borrowdirect

ScienceDirect

Uborrow

Scopus

Fedora

VIVO

ORCID

Khan Academy

InCommon

Mukurtu

github
Publishing

Scholarly Publishing

Stuff (the scholarly and cultural record)

New insights, new data

Use in research and teaching

PubPeer
Open Library of Humanities
slideshare
Oxford University Press
online resource centres
The Walt Whitman Archive
Michigan Publishing
University of Michigan Library
RePEc
ResearchGate
arXiv.org
hivebench
ACADEMIA
Apache Taverna
figshare
Mendeley
Thomson Reuters
EndNote
Hypothes.is
myexperiment
GitHub
Zotero
MethodsX
Dryad
RefWorks
Research workflow tools are of growing importance to scientific research and the universities that support it.

See:
www.sr.ithaka.org/blog/what-is-researcher-workflow/
LABORATORY SCIENCES RESEARCH WORKFLOW

Current Awareness → Research Design → Funding → Research Collaboration → Lab Safety & Management

Experiment Design → Data Collection → Analysis → Writing → Sharing

Submission → Review & Selection → Publication → Showcasing → Assessment
Emerging Verticals

<table>
<thead>
<tr>
<th></th>
<th>Clarivate</th>
<th>Elsevier</th>
<th>Digital Science</th>
<th>Open stack</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery</td>
<td>Web of Science</td>
<td>Scopus</td>
<td>(ReadCube)</td>
<td>Open Science Framework</td>
</tr>
<tr>
<td>Reference Management</td>
<td>Mendeley</td>
<td></td>
<td>(FigShare)</td>
<td>VIVO</td>
</tr>
<tr>
<td>Social Networking</td>
<td>WoS Profiles</td>
<td>Expert Finder</td>
<td>Symplectic Elements</td>
<td>SHARE</td>
</tr>
<tr>
<td>Profiles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publications / Citations</td>
<td>InCites</td>
<td>SciVal</td>
<td>Dimensions</td>
<td></td>
</tr>
<tr>
<td>Evaluation / Analysis</td>
<td>Converis</td>
<td>Pure</td>
<td></td>
<td>Zenodo</td>
</tr>
<tr>
<td>CRIS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funding Opportunities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digital Repository</td>
<td>Mendeley Data / SSRN</td>
<td>FigShare</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Point Solutions of Note:
- Academia.edu
- Research Gate
- ORCID
- CrossRef
- Google Scholar

HUMANITIES RESEARCH WORKFLOW

- Gather sources
- Catalog
- Transcribe / Translate
- Identify people, etc.
- Annotate and interpret
**Humanities Research Workflow**

1. **Gather sources**
2. **Catalog**
3. **Transcribe / Translate**
4. **Identify people, etc.**
5. **Annotate and interpret**

- **Digitization**
- **Encoded Archival Description**
- **TEI / Optical Character Recognition**
- **LOD / Encoded Archival Context**
- **W3C Web Annotation Model**
HUMANITIES RESEARCH WORKFLOW

Gather sources
Catalog
Transcribe / Translate
Identify people, etc.
Annotate and interpret

Digitization
Encoded Archival Description
TEI / Optical Character Recognition
LOD / Encoded Archival Context
W3C Web Annotation Model

Hidden Collections / Tropy
Archivist’s Toolkit
eMOP: Early Modern OCR Project
SNAC Cultures of Knowledge
Hypothes.is Mirador Scalar
**SO WHAT′S THE PROBLEM**

How uneven is the tooling of the workflow across the sciences, social sciences, and humanities?

What are the institutional interests in the tooling? The interests of libraries and IT?

Does it matter that commercial companies, like Digital Science, are stacking the tools into a workflow-based product or set of products? Does it matter that publishers, like Elsevier, are doing so?

Do the returns to scale of the web make it inevitable that the capital required to build and harden web-based tools can only come from commercial firms like Elsevier?