Making Migration Less Mysterious:
Developing a Migration Plan for
ScholarWorks@UMassAmherst

Northeast Institutional Repository Day 2023
Erin Jerome, Library Publishing & Institutional Repository Librarian, University of Massachusetts Amherst
Background

ScholarWorks@UMassAmherst
Launched on bepress’ Digital Commons

Repository Assessment Task Force
Infrastructure assessment and stakeholder feedback

Repository Task Force 2.0
Identify platforms and make recommendations

July 2006
Elsevier acquires bepress

August 2017

January 2021
Interim Report presented to Library Leadership

August 2021

Fall 2022

January 2023
Final Report presented to Library Leadership
Migrate to Atmire-hosted DSpace 7.x and Janeway
January 2023: ScholarWorks overview

- ~ 58,000 works
- 1.2 TB in size
  - Audio files: 6GB
  - Video files: 97 GB
- 11 active journals/conference proceedings migrating to Janeway
Project roadmap

- Step one - interviews
- Step two - draft migration plan
- Step three - cleanup, cleanup, cleanup
Step one: Interviews

1. What kind of preparation did you do?
2. Did you leave content behind (i.e., choose not to migrate particular items)?
3. Did you migrate any journals?
4. What did your migration team look like?
5. Did you encounter anything unexpected during the process?
6. Are there things that you wish you’d handled differently?
Interview takeaways

- Keep your migration team small and agile
- Expect to find inconsistent metadata mappings to Dublin Core
- Bepress’ Amazon S3 backups are not truly real-time
- Digital Commons does not provide metadata to indicate hidden items
- Mapped items are not clearly represented in S3 exports
- Embargoes do not migrate well
- Collection level metadata does not migrate
Step two: Draft Migration Plan

- Schedule standing weekly 30-minute migration team check-ins
  - Four person team + consultants as needed
- No new collections or journals
- Maintain current organizational hierarchy in IR
- Focus on cleanup of duplicate items, link outs, and items with 50+ authors
- Identify metadata only items for possible deletion
- Track embargoed and mapped items for post-migration correction
Draft Migration Plan (continued)

- Save peer reviewer lists and manuscripts in process for journals
- Save all collection-related metadata and administrator information
- Develop content freeze workaround for May 2024 graduates
- Develop communication plan
- Migration plan as living document

Parking lot: what to do with the Expert Gallery Suite
Step Three: Cleanup

- Initial Content Inventory reports exceeded column limits in Excel
  - One metadata only article, uploaded in triplicate, listed 2279 individual authors
- Excel corrupted file warnings
  - Special characters in item level metadata were throwing errors and had to be removed
- 24,000 metadata only items
  - 3,000 of which were uploaded 2-4 times
- 700 items that listed 50 - 1,000 individual authors
- Hidden items
- Items that link out
- Author list
<table>
<thead>
<tr>
<th>Collection</th>
<th>Total number of items</th>
<th>Items w/ PDFs</th>
<th>Metadata only</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>pse_faculty_pubs</td>
<td>1002</td>
<td>21</td>
<td>981</td>
<td>polymer science &amp; engineering</td>
</tr>
<tr>
<td>ece_faculty_pubs</td>
<td>974</td>
<td>36</td>
<td>938</td>
<td></td>
</tr>
<tr>
<td>chem_faculty_pubs</td>
<td>1126</td>
<td>219</td>
<td>907</td>
<td></td>
</tr>
<tr>
<td>cs_faculty_pubs</td>
<td>1022</td>
<td>272</td>
<td>750</td>
<td>computer science</td>
</tr>
<tr>
<td>che_faculty_pubs</td>
<td>695</td>
<td>143</td>
<td>552</td>
<td></td>
</tr>
<tr>
<td>math_faculty_pubs</td>
<td>951</td>
<td>467</td>
<td>484</td>
<td></td>
</tr>
<tr>
<td>astro_faculty_pubs</td>
<td>941</td>
<td>461</td>
<td>480</td>
<td></td>
</tr>
</tbody>
</table>

Submissions from 2010

- Polymerization of Mersomer-Based Ferrofluids, P Bian and TJ McCarthy
- An Age-Old Printing Process Goes Nano, KR Carter
- Enhancement of anhydrous proton transport by supramolecular nanochannels in comb polymers, YB Chen, M Thorn, S Christensen, C Versek, A Poe, RC Hayward, MT Tuominen, and S Thayumanavan
- Precise placements of metal nanoparticles from reversible block copolymer nanostructures, H Che, H Park, TP Russell, and S Park
- Aspects of Network Formation in Glassy Thermosets, AT Dotwiler and AJ Lesser
- Adsorption Energy of Nano- and Microparticles at Liquid-Liquid Interfaces, X Du, E Glogowski, T Emrick, TP Russell, and AD Dinsmore
- Manipulating Protein Adsorption using a Patchy Protein-Resistant Brush, S Gon, M Bendersky, JL Ross, and MM Santore
- Topological Defects in Twisted Bundles of Two-Dimensionally Ordered Filaments, GM Grason
- Thermoreversible Gelation of an Ionic Liquid by Crystallization of a Dissolved Polymer, TM Harner and DA Hoagland
Multi-author items: after cleanup

 Civil and Environmental Engineering
 Civil and Environmental Engineering Faculty Publication Series

Measurement of the $B \to J/\psi K^*(892)$ decay amplitudes

BABAR Collaboration

Publication Date
2001

Journal or Book Title
Physical Review Letters

DOI
https://doi.org/10.1103/PhysRevLett.87.241801

Volume
87

Issue
24

Recommended Citation
https://doi.org/10.1103/PhysRevLett.87.241801

Submissions from 2001

PDF Measurement of branching fractions and search for CP-violating charge asymmetries in charmless two-body $B$ decays into pions and kaons, BABAR Collaboration

PDF Measurement of CP-violating asymmetries in $B$-$\pi$ decays to CP eigenstates, BABAR Collaboration

PDF Measurement of $J/\psi$ production in continuum $e^+e^-$ annihilations near root s=10.6 GeV, BABAR Collaboration

PDF Measurement of the $B$-$0$ and $B^+$ meson lifetimes with fully reconstructed hadronic final states, BABAR Collaboration

PDF Measurement of the $B \to J/\psi K^*(892)$ decay amplitudes, BABAR Collaboration

Fracture modeling with a microstructural mechanics approach, CS Chang, TK Wang, LJ Sluys, and JGM van Mier

Slab effects in SMRF retrofit connection tests, SA Civjan, MD Engelhardt, and JL Gross
Adventures in metadata mapping

ISSN

Enter ISSN:

ISBN

Enter ISBN:

978-1-945764-02-8

Author Biography

This article was co-authored by Deborah Kelisch and Tim Scott. Deborah Kelisch is a cultural anthropologist who has worked in the field of education for over two decades as a practitioner, researcher and activist. She was a founding producer of Education Radio, a documentary-style radio program that shared the stories of students, teachers, scholars and activists experiencing the impacts of corporate education reform policy. She is also a founder and organizer with the Public Schools Action Coalition, a group of parents and teachers fighting for equity in their local public schools. Deborah has a Masters in Education from Teachers College, Columbia University and a Ph.D. in Anthropology from UMass Amherst. Tim Scott is a psychotherapist at the University of Massachusetts Amherst. He worked for a number of years in NYC, first as a school counselor in a public junior high school, then as a youth harm reduction clinician with the Special Health Outreach to Urban Teens program. Tim also worked as a union organizer for seven years and was a community organizer for more than a decade. He was a founding producer of Education Radio, a documentary-style radio program that shared the stories of students, teachers, scholars and activists experiencing the impacts of corporate education reform policy. Originally from Ogden, Utah, Tim holds an MSW degree and his current work is in social justice education.

Acknowledgements

Our heartfelt thanks goes to the activists and scholars engaged in the struggle for a just and equitable public education system.

Creative Commons License

- empty -

DOI

- empty -

DOI Link

- empty -
Lessons Learned

- Avoid custom metadata fields
- Add collections that you can shepherd all the way through the process
- Adhere to best practices for your platform (and for IRs in general)
- If the thing you want to do requires a lot of customization, maybe it’s not the right fit
- Keep future you in mind when approaching big changes/big projects
- Timelines are aspirational
Next steps

- Continue with metadata mapping
- Outreach to departments with large metadata only collections
- Clean up the “Browse by Author” page (https://scholarworks.umass.edu/authors.html)
- Prepare communication plan for release in January 2024
- And...figure out what we’re doing with all of our EGS profiles
Thank you!

Email me: ewjerome@umass.edu