How automated workflows helped us ingest 600 faculty publications in three months in LMU’s institutional repository!

Shilpa Rele
*Loyola Marymount University*, rele@rowan.edu

Jessea Young
*Loyola Marymount University*, jessea.young@lmu.edu

Follow this and additional works at: https://digitalcommons.lmu.edu/librarian_pubs

Part of the Scholarly Communication Commons

Repository Citation

This Conference Presentation is brought to you for free and open access by the William H. Hannon Library at Digital Commons @ Loyola Marymount University and Loyola Law School. It has been accepted for inclusion in LMU Librarian Publications & Presentations by an authorized administrator of Digital Commons@Loyola Marymount University and Loyola Law School. For more information, please contact digitalcommons@lmu.edu.
How Automated Workflows Helped Us Ingest 600 Faculty Publications in Three Months in LMU’s Institutional Repository!

Shilpa Rele + Jessea Young
Digital Library Program
William H. Hannon Library

Loyola Marymount University
LAUNCHED IN 2010

- Bepress Digital Commons
- Kickstart project
BUILD + OUTREACH

- Building personal relationships with faculty
- Promote Digital Commons at New faculty orientation
- Word of mouth, updates
- Inform about access to usage statistics, monthly reports - impact of their scholarship
OLD WORKFLOW

Step 1: Get CV from faculty.

Step 2: Search one citation at a time by using SHERPA/RoMEO or emailing publishers for permission.

Step 3: Ingest the legal version.
CHALLENGES

- Adding content to IR: Slow, Labor Intensive
- No dedicated position within Digital Library Program for scholcomm
- Other work responsibilities
• Hired Library Assistant

**Exploration:**

• Google Script that searches Sherpa/Romeo
• Web of Science datasets
• OA_DOI
• 1Science datasets
• UnPaywall and OA Button
Step 1: Bypass the CV.

Download citations from a bibliographic database retrieved by an organization-enhanced search.
By inserting the ISSN number into the script, it reveals four functions:

1. Does the journal permit uploading the publisher PDF to an IR?
2. Does the publisher permit uploading the author’s final draft after peer review, and if so, with embargo restrictions?
3. Does the word “embargo” appear anywhere in the SHERPA/RoMEO record for this publisher?
4. Does this journal exclude authors from institutions which have passed an open access mandate?

---

**Step 2: SHERPA/RoMEO Google Spreadsheet Script**

Created by the College of Wooster

(Code: [https://pastebin.com/sXknBHDq](https://pastebin.com/sXknBHDq))
Step 3: Research ISSN or Publisher’s Emails

**Issue:** The script requires an ISSN. Sometimes there is no information about a journal in SHERPA/RoMEO.

**Solution:** A LMU “Journals Database” for ISSNs or Publisher’s Emails

\[ vlookup \]
<table>
<thead>
<tr>
<th>Publication</th>
<th>Journal Title</th>
<th>ISSN or Email</th>
<th>pubpdf</th>
<th>finaldraft</th>
<th>embargo</th>
<th>check OA Amandate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication</td>
<td>Journal Title</td>
<td>ISSN or Email</td>
<td>pubpdf</td>
<td>finaldraft</td>
<td>embargo</td>
<td>check OAmandate</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>----------------------------------------------------</td>
<td>---------------</td>
<td>--------</td>
<td>------------</td>
<td>---------</td>
<td>-----------------</td>
</tr>
<tr>
<td>“Intrinsic Linking and Knotting in Virtual Spatial Graphs,” with T. Fleming, Algebraic and Geometric Topology, vol. 7, 2007, pp. 583-601</td>
<td>Algebraic and Geometric Topology</td>
<td></td>
<td>Loading...</td>
<td>Loading...</td>
<td>Loading...</td>
<td></td>
</tr>
<tr>
<td>“Finite Type Link Homotopy Invariants,” Journal of Knot Theory and Its Ramifications, vol. 8, no. 6, 1999, pp. 773-787</td>
<td>Journal of Knot Theory and Its Ramifications</td>
<td></td>
<td>Loading...</td>
<td>Loading...</td>
<td>Loading...</td>
<td></td>
</tr>
<tr>
<td>Publication</td>
<td>Journal Title</td>
<td>ISSN or Email</td>
<td>pubpdf</td>
<td>finaldraft</td>
<td>embargo</td>
<td>embargo date</td>
</tr>
<tr>
<td>----------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------</td>
<td>----------------------</td>
<td>------------</td>
<td>----------</td>
<td>--------------</td>
</tr>
<tr>
<td>“Intrinsic Linking and Knotting in Virtual Spatial Graphs,” with T. Fleming, Algebraic and Geometric Topology, vol. 7, 2007, pp. 583-601</td>
<td>Algebraic and Geometric Topology</td>
<td>1472-2747</td>
<td>Publisher's version/PDF may be used after an embargo period</td>
<td>Final draft allowed</td>
<td>no</td>
<td>all good!</td>
</tr>
</tbody>
</table>
Step 5: Discover.

UnPaywall or OA Button to discover open access versions of articles
Old workflow: 3+ days to clear copyright for each CV
New workflow: About 1 day to clear copyright for each CV and contact publishers
CHALLENGES

1. Still Need Faculty’s Publication List
2. Backlog of 3,800 citations
3. Time consuming metadata cleanup
4. Batch Upload Structure of Digital Commons
5. Policy: Only full-time professors currently employed at LMU.
Freely Available vs. Legally Available
ResearchGate + Academic.edu
Lack of Policies
Ambiguity about how the datasets are pulled for LMU
OAdoi and R to find pre-prints or post-prints

OpenRefine to handle metadata cleanup

Zotero to integrate Web of Science and Google Scholar LMU affiliation alerts into workflow to ingest new publications faster

Maintain master inventory of LMU Scholarship
Growth

Between Oct. 2016 - Jan. 2017, we added 722 full-text materials in Faculty Work Series.

This was a 74% increase from last quarter.
FUTURE

- Share information from LMU’s Journal database
- Experiment with new ways to receive CVs
- Outreach Campaigns
Slides & Documentation:

goo.gl/Xnh71B

Shilpa Rele

@ShilpaRele

shilpa.rele@gmail.com

Questions?

Jessea Young

@itsjessea

jessea.young@lmu.edu