Texas ScholarWorks

- Created in 2008
- Consists mostly of library managed collections
- Content from all areas of the University
- Not a lot of faculty content
Texas ScholarWorks

- Conference proceedings/journals: 159
- Student Work (not ETDs): 1,525
- UT Communities: 11,949
- ETDs: 17,983
- Faculty/Researcher Works: 2,609
- Total: 34,225
Faculty publications

Hard to demonstrate our value without use

https://hatponicsblog.wordpress.com/tag/children/
http://www.mackinacbridge.org/photo-gallery-10/
We Decided to Proactively Add Content

- Faculty/researcher published articles
- Where allowed by license (ex: CC-BY) or by publisher policy
- Only PDFs for now
- Created an FAQ to handle any questions about content
Where do we find this content & licensing info?

SHERPA/RoMEO is a service run by SHERPA to show the copyright and open access self-archiving policies of academic journals.
Using Google Script to Automate Sherpa/Romeo Copyright Check

- Find affiliation on [Web of Science](https://webofknowledge.com)
- Problem: WoS only allows a download of 500 records at a time.
- Import Web of Science .csv list to Google Spreadsheet
Using Google Script to Automate Sherpa/Romeo Copyright Check

- Insert Google Script created by Stephen X. Flynn (The College of Wooster) into Google script editor
- Register for an API key then insert the key below in the "Sherpa API key" section of the code. Without an API key, you will be limited to 500 requests per day, per IP address.
- Add and check the four functions produced (pubpdf, finaldraft, embargo, checkOAmandate)
- Create or sort list according to pubpdf for lowest hanging fruit
Processing

- Double check copyright & permissions
- Downloading metadata from Web of Science and open with EndNote
- Export EndNote data to Excel
- Find and download articles freely available from WoS or Google Scholar
- Input missing metadata into and ‘clean-up’ spreadsheet prior to batch upload
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Journal/Conference</th>
<th>Year</th>
<th>Volume/Issue</th>
<th>Pages</th>
<th>Abstract/Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finished Batch Ready for Upload</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abstract:**

Finished Batch Ready for Upload

**Keywords:**

batch processing, upload, readiness, preparation, completion.
Issues and Things to Consider

- Web of Science changes frequently
  - Finicky, the kind of metadata it provides can vary
- Time spent manually inputting some metadata
  - All caps, accent marks, ||
  - Finding and adding citations
  - Sources of funding in WoS
- Journal and publisher copyrights change often
  - Good faith - Check SHERPA/RoMEO often
Getting content into DSpace

- Until recently - used a program created in-house to generate simple archive format (SAF)
- Repository staff run program and hand off SAF to Texas Digital Library (TDL) for ingest into repository
- TDL still had to test batches and write scripts to get content into DSpace
Problems with Process

- Batch process running slow
  - Better with SAF Creator - much faster
- Upgrade halted process of batches
- Still potential for negative feedback from faculty
Results

• Currently 2,609 items in Faculty/Research collection
  • Of these, 1,695 have been added in the last 18 months
• 718 waiting to be ingested
• 262,647 downloads since April 2014
• 0 negative comments
Next Steps

- Continue processing batches
- Go beyond just the PDFs
- Use usage statistics to recruit more content
- Hire more students
Questions?

http://www.kuder.com/2013/01/raise-your-hand-adopt-a-classroom/