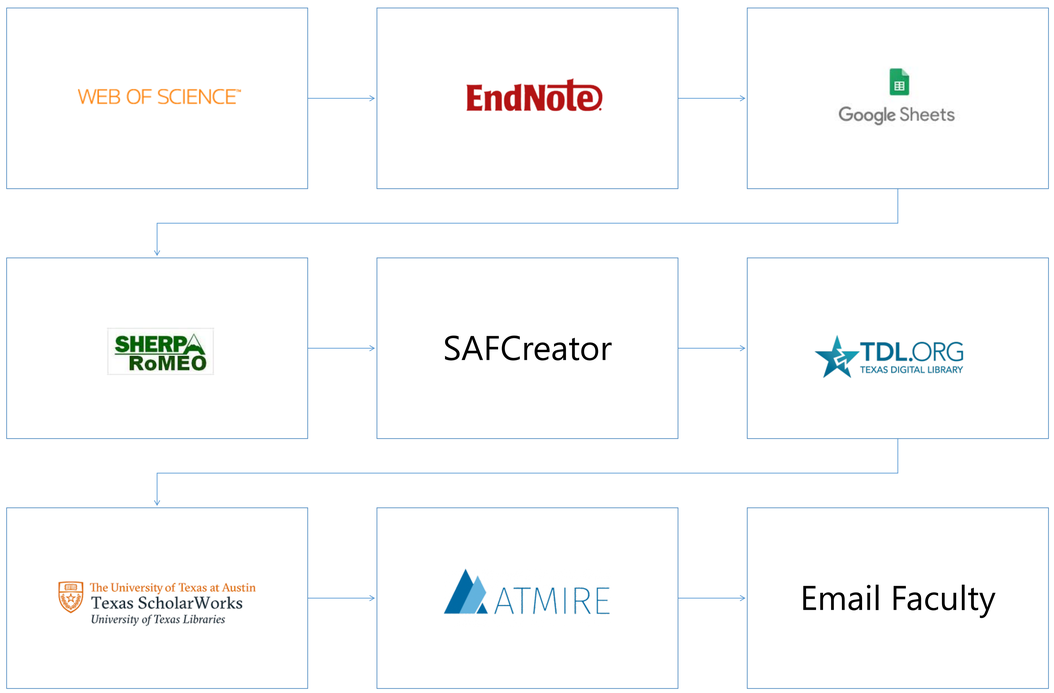
Outreach Outcomes and Batch Processing Tools for IR Deposited Faculty Work

Colleen Lyon, University of Texas at Austin

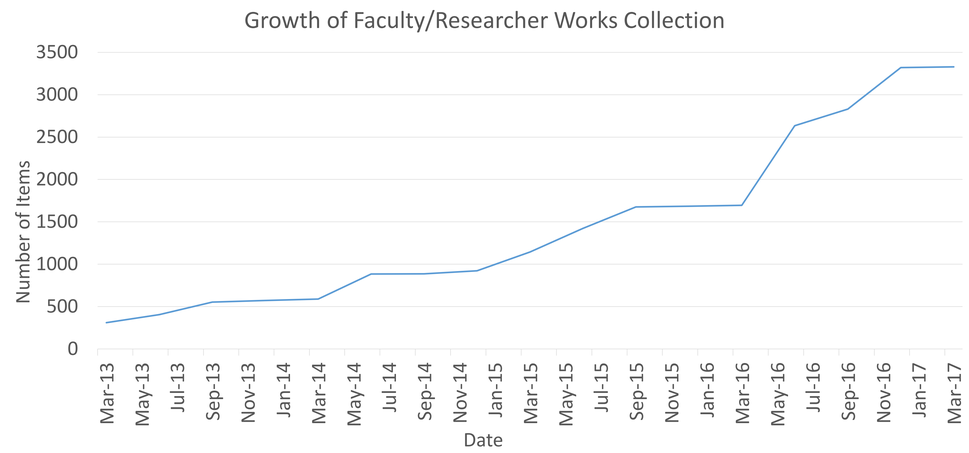
Problem:
Getting the word out to campus about our institutional repository, Texas ScholarWorks (TSW)[1], can be difficult, and even interested faculty may not have time to follow through with uploading their work. We wanted to find a way to use our existing repository services to increase the amount of UT Austin research being openly shared and increase awareness among faculty of our services.

Solution:
We developed a workflow that takes advantage of CC licenses and publisher policies, to allow us to share faculty articles via TSW, without needing to ask for permission from the authors.

Workflow:
1. Export UT Austin content from Web of Science into EndNote
2. Use an EndNote template to create a formatted csv
3. Import csv into Google Sheets
4. Insert Google script[2] into Google script editor
5. Register for Sherpa Romeo API key that allows for more than 500 requests per day
6. Run the script against the spreadsheet and sort by low-hanging fruit (e.g. Publisher PDF ok)
7. Double check copyright and permissions (for batches done months after the initial Sherpa Romeo check)
8. Find and download articles
9. Input missing metadata and clean up spreadsheet for batch
10. Use SAFCreator[3] to make the simple archive format needed for a DSpace batch ingest
11. Pass the batch off to Texas Digital Library for the ingest into Texas ScholarWorks
12. Use the Atmire Content & Usage Analysis (CUA) module to create download reports that can then be emailed to faculty



What worked:
We wanted to increase the number of articles being shared with the general public via our repository. This outcome was successful. We've added 2,406 items since we started the project. We did our initial export from Web of Science in Nov. 2014. We spent the early part of 2015 smoothing out the workflow and getting our students trained. We upgraded our repository in fall of 2015 which resulted in a hold on all batch ingest for 4-5 months.



What didn't work:
We hoped that in addition to adding more items to TSW, we might increase faculty awareness of and engagement with our repository. This wasn't as successful.
- We emailed 100 faculty with reports detailing downloads of their repository content and have received 0 responses.
- On the bright side, we've also had 0 negative responses to our article sharing workflow.

We used the CUA module from Atmire to help us simplify the process of running statistics reports for deposited content. We used the statistics reports as an outreach tool to increase faculty awareness of the repository, and hopefully help us solicit more content (including non-article content). The module wasn't as automated as we would have liked – at least for our purposes – so the process of running reports, while much improved from default Dspace options, still requires manual intervention.

Next steps:
- Work with our subject librarians to offer them usage statistics assistance as a way of communicating with their faculty. The message may be more meaningful coming from a known contact.
- Continue our outreach strategy of personalized emails regarding specific content.

References:
1. Texas ScholarWorks: https://repositories.lib.utexas.edu/
2. Google Script: Flynn, S.X., Oyler, C., & Miles, M. (Jan. 2013). Using XSLT and Google Scripts to Streamline Populating an Institutional Repository. *Code4Lib*, Issue 19. http://journal.code4lib.org/articles/7825
3. Creel, J. SAFCreator. https://github.com/jcreel/SAFCreator

Special thanks to Gilbert Borrego, Texas Digital Library, and all of our student workers