

Content Liberation Update, August 2019

Rebekah Kati
Institutional Repository Librarian
University of North Carolina at Chapel Hill

Introduction

In 2018, the Open Access Implementation Team (the Team) conducted three Content Liberation pilot projects to ingest UNC-authored content into the Carolina Digital Repository (CDR). A summary of each project is available in [Content Liberation Overview](#).

This report summarizes our current progress on these projects and suggests recommendations for progress in the next academic year. A follow-up report will be available in August 2020.

- [August 2018 Content Liberation Update](#)

Author Citations

As reported last year, this pilot project has been on hold due to the purchase of 1science's 1foldr report. After the new CDR launched in June 2019, the Team began the process to upload 1foldr content from 2016 to 2018. During our assessment, we found several issues that required remediation before content could be loaded into the CDR, specifically:

- Author names appeared in several places and were not consistent
- The author departmental affiliation metadata, when present, did not conform to our internal metadata schema
- Embargo information was not provided by 1foldr
- A bug in the 1foldr download script overwrote article file names, which caused PDFs to be matched to incorrect metadata records. Our developers needed to rewrite the script to fix this issue and we needed to re-download the batch

We anticipate that the loading process for 2016 to 2018 1foldr content will be complete by early September 2019. Additional 1foldr content will be loaded as staff time permits. As pre-2016 content is not covered by UNC's Open Access policy, we will need to do additional assessments to determine the version that the publisher's policy allows us to deposit. Despite these challenges, we feel that the work was worth it, as the 1foldr content enabled us to quickly switch the collection focus of CDR from student papers to faculty scholarship.

We still feel that the Author Citations project holds great value to identify post-2018 articles, as well as articles which are behind paywalls. We are continuing to collect metadata for UNC-authored articles. In the next year, we will continue to load articles identified by 1foldr and evaluate the Author Citations project workflow with the new repository platform in mind.

CV Review

The Team developed and documented the workflow for locating publications, comparing the publication list with the SHERPA/RoMEO API to determine which publications were eligible for deposit in the CDR, and batch uploading deposits in the CDR.

The team attended several faculty meetings during the Fall 2018 semester. Due to the departure of our Open Access Librarian, we did not visit any faculty meetings during Spring 2019.

We still feel that CV Review is a useful service, but are concerned about the amount of time and labor that is required to review and deposit the contents of a CV. We will continue to offer it upon request but will not promote it widely.

Highly Cited Authors

The Team continued to review articles authored by UNC researchers named on Clarivate's Highly Cited Authors list. Ten additional researchers appeared on the 2019 list and they authored 167 highly cited articles. The Team defined "highly cited article" as an article which received over 300 citations in Scopus. We determined that 51 out of the 167 highly cited articles were eligible for deposit in the CDR. While we were analyzing the list, we were preparing to move content to the new CDR platform. Therefore, we determined that it would be best to wait until after the new CDR platform was launched to deposit the articles. We will work on this in the coming months.

The Team performed an assessment to determine whether any of the 167 articles were already available openly prior to deposit in the CDR. Using Unpaywall's Simple Query Tool, we found that:

- 84 articles were not available openly in a repository or on the publisher's website (50.3% liberated)
- 83 were already available openly in a repository or on the publisher's website (49.7%)

Additionally, we also determined whether the 166 articles were available on an academic social network or author's institutionally affiliated website:

- 80 articles were not available on an academic social network or an author's institutionally affiliated website (47.9%)
- 87 were available on academic social networks or an author's institutionally affiliated website (52.01%)

We will continue to contact the remaining researchers and deposit their work. Although much of the scholarship is open elsewhere, we feel that it is important enough to preserve in the CDR. Additionally, the project will raise awareness of the CDR to highly influential faculty, which could lead to increased deposits in the future.

We are happy with the progress that we have made this year, although it was slow as we concentrated on launching the new CDR platform. We feel that we are concentrating on content that is most important to add and look forward to making additional progress in the coming year.