

# Carolina Digital Repository Update 2019

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

## Contents

Introduction .....	1
Key Initiative: Launching a New Repository Platform .....	2
Goals .....	2
Community Involvement .....	2
Our Efforts .....	2
Growth .....	3
Usage .....	3
Collection Highlights .....	6
Faculty Work .....	7
Discussion and Recommendations .....	7

## Introduction

The Carolina Digital Repository (CDR) is a Samvera Hyrax-based institutional repository, which launched on June 5, 2019. Prior to June 5, 2019, the CDR platform was built on custom Fedora and housed both institutional repository and born-digital special collections materials. The former CDR platform now contains only born-digital special collections materials, as well as one highly customized institutional repository collection (Research Labs of Archaeology). Additionally, the UNC Libraries administers a CONTENTdm platform for digitized special collections content.

This report summarizes activity in the CDR from January 1, 2019 through November 30, 2019. It describes key initiatives, growth and usage of the repository as well as a description of future initiatives. We would like to thank Joy M. Perrin at Texas Tech University for the inspiration and format of this report.

## Key Initiative: Launching a New Repository Platform

In January 2016, the UNC-Chapel Hill faculty adopted an open access policy. In support of this policy, the UNC-Chapel Hill Libraries launched a new repository system to house institutional repository materials. The development process took approximately two years and involved over 16 people from five library departments. Although much of the development work took place prior to January 1, 2019, it is briefly described here for completeness and context.

### Goals

Our main goal for launching a new repository platform was to optimize CDR for scholarly content. Specifically, we wanted to:

- Improve the CDR's search engine optimization
- Optimize CDR metadata for content discovery, particularly for scholarly content indexing tools like Unpaywall and Google Scholar
- Take advantage of existing Hyrax features, especially user-facing work-level analytics
- Customize CDR's branding to emphasize UNC and the UNC Libraries' role

### Community Involvement

We chose Samvera Hyrax, an open source repository system with robust community involvement. We have contributed to the Samvera Community this year through participation in release testing efforts, code contributions and feature documentation. Our previous contributions include establishing a QA Lead, documentation and additional code contributions.

### Our Efforts

We utilized the core features of Hyrax but needed to customize the platform to support our existing use cases. Honors theses and several masters papers required custom workflows to accommodate departmental requirements. We also needed to support connecting people's names, departments and ORCIDs so that they appeared together on the work record page. Finally, we imported Google Analytics data so that users could see view and download statistics prior to the new platform launch.

Additionally, our graduate student conducted user experience testing on the new platform, which gave us valuable insights into the submission process. He also designed the look and feel of the site, which now emphasizes the CDR's connection to UNC and the UNC Libraries through branding.

In the first part of 2019, we focused on planning for the content migration from custom Fedora to Hyrax. Prior to 2019, our Metadata Librarian created Metadata Application Profiles for all our worktypes, which served as a template for the content migration. We identified collections to be migrated and the order in which they would migrate. After each collection was migrated, we completed quality assurance on a sample set of records.

Other features and initiatives completed this year include:

- Advanced search interface
- Automated DOI creation and registration with DataCite

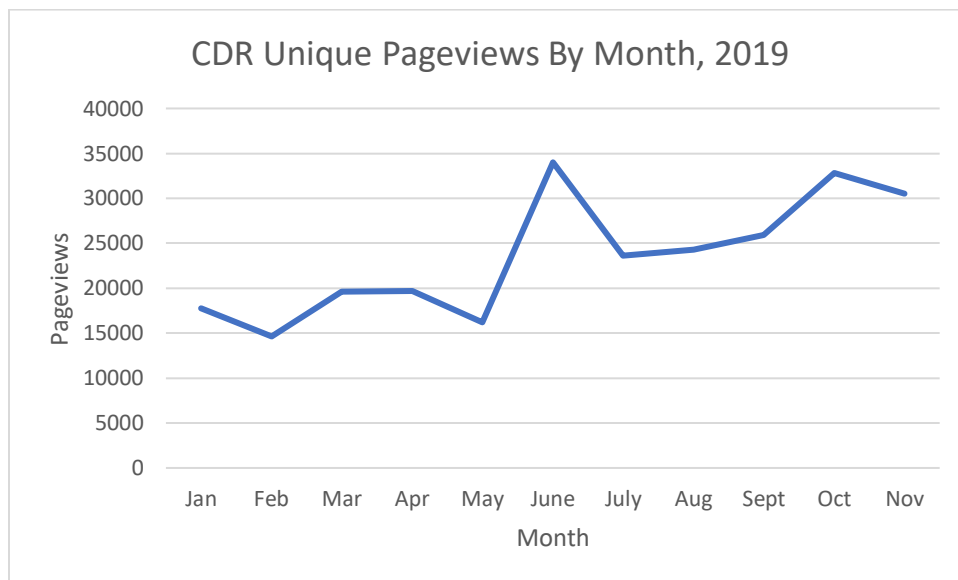
- Bug fixes
- Date slider
- Note field which is visible to CDR administrators only
- OAI-PMH feed
- Platform upgraded to 2.5.0 and 2.6.0
- Registration of CDR with Unpaywall, CORE and BASE
- User and CDR administrator-facing documentation

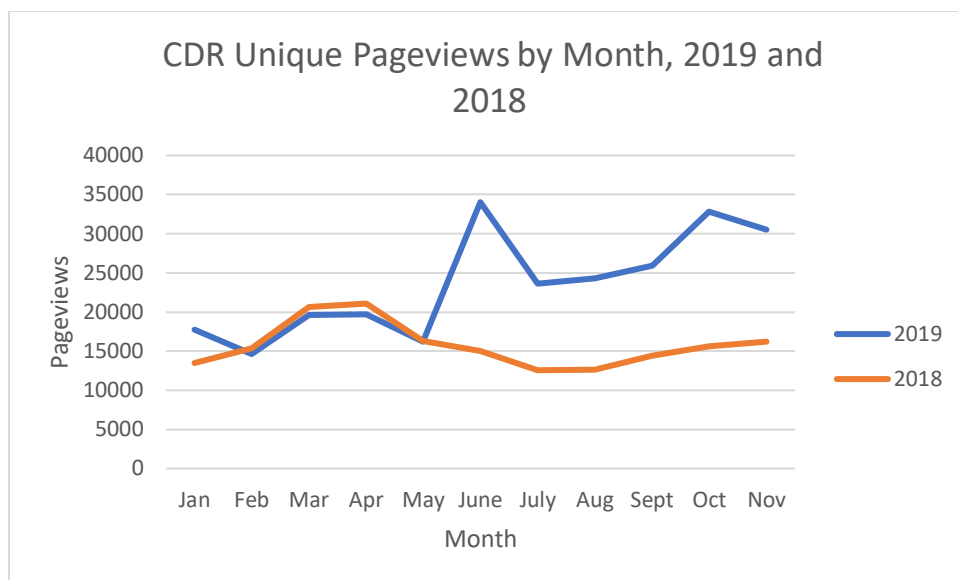
## Growth

CDR growth in 2019 is difficult to quantify due to our migration in the middle of the year. We migrated 27952 works from the custom Fedora platform to Hyrax. After the new platform launched, 826 works were added. Most of the new works were student papers, dissertations and masters theses, which is typical for CDR.

## Usage

The CDR recorded 259,024 unique pageviews between January 1, 2019 and November 30, 2019. Our busiest day was June 4, 2019 during which we received 5,680 pageviews. Most of these pageviews (4,467) corresponded to “Predicting the Difficulty of Trivia Questions Using Text Features”, a SILS masters paper whose author was a contestant on “Jeopardy” during this timeframe. A link to her masters paper was included in national media coverage of her winning streak, which translated into record traffic for CDR.



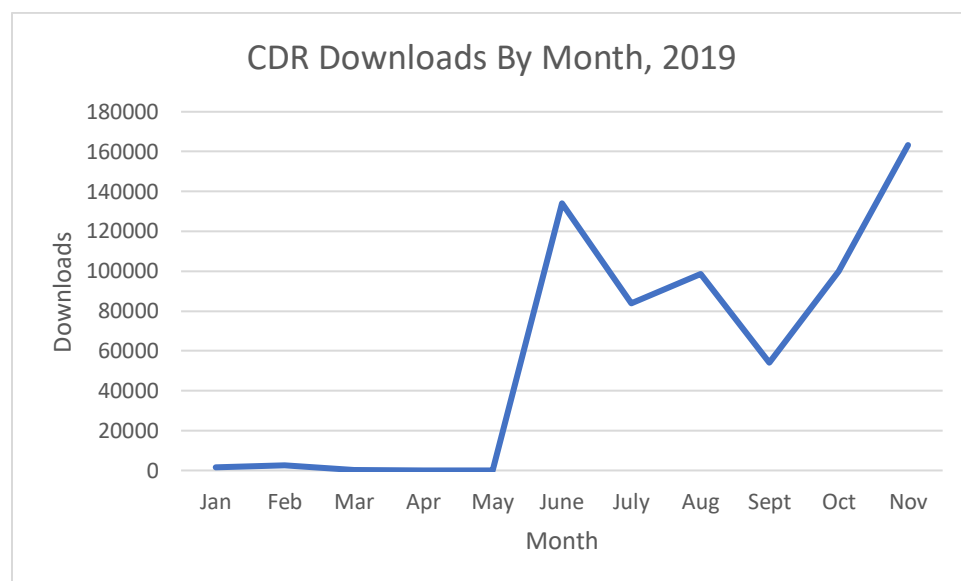


### Most Viewed Works, 2019

Work Title	Resource Type	Number of Unique Pageviews
Predicting the Difficulty of Trivia Questions Using Text Features	Masters Paper (SILS)	12191
Effects of heterogeneous and homogeneous grouping on student learning	Masters Thesis	1248
Filled-in vs. Outline Icons: The Impact of Icon Style on Usability	Masters Paper (SILS)	1078
Adolescent Adjustment in Rural Appalachian Youth	Dissertation	1008
On-Campus Parking: Problems and Solutions	Masters Paper (City and Regional Planning)	872
Cinema Archipelago: A Geography of Philippine Film and the Postnational Imaginary	Dissertation	678
The Complications of Liberty: Free People of Color in North Carolina	Dissertation	671

from the Colonial Period through Reconstruction		
The Kongo Rule: The Palo Monte Mayombe Wisdom Society	Dissertation	591
Choosing Between Multinomial Logit and Multinomial Probit Models for Analysis of Unordered Choice Data	Masters Thesis	575

The CDR received 638,464 downloads during the same time period. Note that downloads for part of March and all of April and May 2019 were not recorded due to technical issues. Unsurprisingly, given the extensive media coverage, “Predicting the Difficulty of Trivia Questions Using Text Features” was the most downloaded item, with 12,118 downloads.



### Most Downloaded Works, 2019

Work Title	Resource Type	Number of Downloads
Predicting the Difficulty of Trivia Questions Using Text Features	Masters Paper (SILS)	12118
Fire and Stone	Book	3845

Excavations in the Archaic Civic Buildings at Azoria in 2005-2006	Article	3081
The Archaeology of Urbanization	Book Chapter	3079
Some Thoughts on the Conservation of an Archaic Greek City on Crete	Article	3072
The Archaic Houses at Azoria	Book Chapter	3018
Stratigraphic Excavations at Azoria in 2014	Article	2956
Excavations at Azoria and Stratigraphic Evidence for the Restructuring of Cretan Landscapes ca. 600 BCE	Book Chapter	2950
Excavations at Azoria, 2003-2004, Part 2: The Early Iron Age, Late Prepalatial and Final Neolithic Occupation	Article	2948
Urbanization and the Emergence of the Greek Polis	Book Chapter	2946

## Collection Highlights

Although the first half of 2019 was devoted to migrating existing works to the new repository, we deposited additional works in the second half of the year. We continue to support student paper deposit, as well as self-deposit of faculty work. Our notable new partnerships in 2019 included:

### [Academy of Educators](#)

The UNC School of Medicine Academy of Educators started depositing articles into the CDR in the latter part of Fall Semester 2019. This group of practitioners consists of faculty and educators in the School of Medicine and aims to promote teaching and faculty development.

### [IDST 190-007 Student Projects](#)

Students enrolled in IDST 190-007: Experiencing Latin America: Bodies, Belonging, Nature deposited portions of their final portfolios into the CDR at the end of Fall Semester 2019. As part of their final project students shot a movie as a group, and each student deposited a still from the movie and a short self-reflection into the CDR. The professor, Gabriela Valdivia, wanted to teach her students the importance of archiving their work in a trusted repository.

### [RTI Press](#)

The Research Triangle Institute Press (RTI) started to deposit their open access books into the CDR, in order to preserve them and help them reach a wider audience. Additionally, RTI hopes to strengthen their connections to UNC-Chapel Hill.

### [Summer Undergraduate Research Fellowship](#)

Students participating in the Center for Undergraduate Research's Summer Undergraduate Fellowship program deposited their research project slides into the CDR.

### [Faculty Work](#)

After the migration, CDR only saw 16 self-deposits of articles from faculty and staff. This was largely due to our new partnership with the Academy of Educators and their staff member's advocacy to his faculty.

To address the low rates of self-deposit, we are continuing initiatives which we piloted in 2019. The Highly Cited Authors project aims to deposit high-impact research from UNC authors into the CDR. The Author Citations project (now "GreenOA") aims to identify recent, paywalled articles authored by UNC researchers and deposit them into the CDR. Both approaches utilize mediated deposit rather than self-deposit, which we hope will increase uptake.

Additionally, we will load recent articles from a 1foldr report into the CDR. We had hoped to have completed this task already but encountered serious metadata issues during the assessment and quality assurance process. Therefore, we have had to adjust our timeframe to address these problems.

## [Discussion and Future Steps](#)

The number of views and downloads of CDR content has increased since our platform migration. Some of this traffic is no doubt a side effect of the increased media coverage surrounding "Predicting the Difficulty of Trivia Questions Using Text Features" but the sustained increase suggests that SEO and discovery in the new platform is working well. Most of the views and download activity occurred after the platform transition. We do not have Google Scholar fully integrated yet and expect to see more traffic once that occurs.

Rates of faculty self-deposit are low, which is consistent with previous years. To address this, we will continue to pursue mediated deposit strategies (see "Faculty Work" for details). We hope that increased

article deposits will shift the focus of CDR from a student paper focused repository to a faculty research focused repository. Nevertheless, we anticipate that student papers and other student work will continue to be a main focus of the repository and one of our primary areas of service work.

In the coming year, we will focus on feature development and enhancements for CDR, particularly to optimize the repository for Google Scholar and other indexers. Also, we will investigate strategies for batch upload and editing, which will aid us in our content uploads.