

Purdue University
Purdue e-Pubs

Charleston Library Conference

Open Access/Closed Coffers: Repositioning an Institutional Repository to Reflect Reality

Anna R. Craft

University of North Carolina at Greensboro, arcraft@uncg.edu

Follow this and additional works at: <http://docs.lib.purdue.edu/charleston>



Part of the [Library and Information Science Commons](#)

An indexed, print copy of the Proceedings is also available for purchase at: <http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

Anna R. Craft, "Open Access/Closed Coffers: Repositioning an Institutional Repository to Reflect Reality" (2012). *Proceedings of the Charleston Library Conference*.

<http://dx.doi.org/10.5703/1288284315125>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Open Access/Closed Coasters: Repositioning an Institutional Repository to Reflect Reality

Anna R. Craft, Metadata Cataloger, University of North Carolina at Greensboro

Abstract

During 2011, a combination of budgetary and staffing factors led the NC DOCKS institutional repository system at The University of North Carolina at Greensboro (UNCG) to restructure its staffing, policies, and procedures, particularly with relation to the handling of faculty publications. A task force convened to study the issues, and its research and recommendations led to a focus on born-digital files and the scaling back of solicitation and staff support for faculty publications. As expected, the changes have led to a decline in the number of faculty publication added to the database, but student works are set to continue to provide a steady stream of additions to the IR. NC DOCKS has also embarked on a new track through a partnership with the Odum Institute at The University of North Carolina at Chapel Hill and is in the process of adding support for management of faculty research data sets. This move helps better align the Libraries with the needs of the university's faculty members. Through the changes outlined in this paper, NC DOCKS will continue to grow and to showcase faculty and student scholarship from The University of North Carolina at Greensboro.

Introduction

The North Carolina Digital Online Collection of Knowledge and Scholarship¹ (NC DOCKS) project began in 2008 as a collaboration among the University of North Carolina at Greensboro (UNCG), Appalachian State University, East Carolina University, the University of North Carolina at Pembroke, and the University of North Carolina Wilmington. The group's goal was to create an open access repository to showcase the scholarly output of the participating institutions. Since that time, the participant group has grown to include Western Carolina University and the North Carolina School of Science and Mathematics, and the shared database holds more than 8,000 items, including works by faculty members and students. While overarching database policies and directions are decided by the cooperative group, individual institutions have latitude to design their own local workflows and procedures. This article will discuss issues specific to the NC DOCKS implementation at UNCG, where the year 2011 brought a restructuring of institutional repository policies, procedures, and staffing.

Early Policies, Procedures, and Staffing

NC DOCKS at UNCG serves faculty publications, student electronic theses and dissertations, and

occasional other student works, such as award-winning papers. Item and metadata ingest into the database is handled by the University Libraries; faculty and student users are not permitted to enter or edit their own data in the system. While the bulk of the items in the database are scholarly articles served as full-text PDFs, the system is also set up to handle many other file types, including PowerPoint files, video, and other media. Faculty participation in NC DOCKS is entirely voluntary, there is no institutional mandate regarding the deposit of faculty scholarly work. However, students completing theses and dissertations are required to have their work deposited in NC DOCKS, a mandate that is enforced by the Graduate School.

From 2008 to mid-2011, staffing and support for NC DOCKS came from four library areas: Electronic Resources and Information Technology, Scholarly Communications, Acquisitions, and Cataloging. The system was built and maintained by the Electronic Resources and Information Technology department (ERIT). This unit handled development, upgrades, and maintenance for the system as well as provided leadership to the combined NC DOCKS user group. The Collections and Scholarly Resources Coordinator handled publicity and outreach, oversaw faculty item solicitation activities, and worked on copyright discovery for faculty publications. Acquisitions

staff members also contributed to solicitation efforts and copyright discovery, in addition to handling processing, reformatting, upload, and metadata creation for faculty publications. The Cataloging unit provided oversight to local policies and procedures and handled the workflow for electronic theses and dissertations (ETDs), involving item ingest, author profile creation, and metadata application.

Before the changes instituted in 2011, item submission policies allowed faculty members to submit articles and other works in any accessible format—and also permitted faculty members to simply send in single citations or citation lists. Acquisitions staff members worked from what was received, searching for articles by citation and/or reformatting and processing as necessary, to yield full-text searchable PDFs of written works. These workflows frequently included ordering items through interlibrary loan, scanning from print copies, reformatting a variety of file types, and sometimes spending significant time correcting optical character recognition (OCR) transcripts. In addition, staff faced challenges regarding formatting the display of images, non-English languages, symbols, diacritics, and scientific and mathematical formulas. But despite the intensive nature of the process, upload numbers climbed steadily from the inception of faculty publication processing in 2008, reached a high in 2010, and subsequently were cut short in 2011 as the climate of shrinking resources brought a juncture where the Libraries were no longer able to provide that continuing level of staff- and time-intensive support.

Year	2008	2009	2010	2011 ²
Faculty publications added	133	629	1736	1171

Table 1. UNCG Faculty Publications Added to NC DOCKS, 2008-2011

In mid-2011, a combination of factors moved the Libraries to reconsider local policies and procedures relating to NC DOCKS, especially in the area of handling of faculty publications. Facing a climate of budget cuts in the UNC System and

looking to align staff support with library goals and objects, in the summer of 2011 the Libraries undertook a staffing realignment. This reorganization shifted several staff members out of technical services and into Special Collections and University Archives and Electronic Resources and Information Technology. Among the staff members moved out of technical services were two positions from Acquisitions, and these positions had been almost entirely responsible for the processing and uploading of faculty publications into the NC DOCKS database. The job changes for these staff members necessitated a complete rethinking of the NC DOCKS faculty publications processing workflow.

Task Force and Findings

The IR Content Processing/NC DOCKS Task Force was convened in the summer of 2011. The group was charged to examine best practices for the processing of IR content, including “difficult” content such as items including symbols, diacritics, non-English languages, charts, and pictures; to recommend appropriate staffing levels; to review available processing software; and to establish guidelines and best practices for submissions and communications with faculty contributors. The task force examined local IR policies and procedures for both faculty publications and student works, but, as the main area of need, faculty publications were the primary issue of study. Processing of faculty publications was suspended for the remainder of 2011 as the task force set about its work. The group examined existing practices and procedures, reviewed relevant literature, and contacted peer institutions with IR programs.

Based on its findings, the task force recommended a shift to focus on procurement of born-digital documents, moving away from ordering items via ILL and scanning physical copies. The task force also discouraged the continuance of the policy of accepting articles in formats that would require significant processing, instead suggesting a focus on HTML files, Word documents, and self-archiving-supported PDFs. Additionally, the task force recommended a shift to focusing on current or recent scholarship over older titles and provided suggested staffing scenarios, based on

staff hours available. A number of these changes were implemented as suggested, which will be discussed in the following sections.

Item Submission Policies

In the current climate, it is no longer feasible for the library to continue to provide processing and reformatting support at pre-restructuring levels. Though the library still accepts citation lists from faculty members, under the new policies the focus is on born digital content. Items are no longer ordered through ILL and are no longer scanned from print copies. Faculty articles continue to be served as PDFs, but can only be accepted in formats that do not require significant reformatting, such as most HTML files, Word documents, and PDFs where the publisher allows archiving of the publisher’s PDF version. Non-article items (slides and other media) are still accepted, but these formats continue to make up a very small percentage of items in the database.

Staffing Changes

After the restructuring, IR duties in ERIT, Scholarly Communications, and Cataloging remain relatively unchanged. The Acquisitions unit is no longer involved in the work of NC DOCKS, and a new half-time student position has been created to work on faculty publications uploads, metadata, and other tasks as assigned. This position handles the tasks previously centered in Acquisitions, but hours devoted to faculty publications-related duties have been reduced. As is the case with many student positions, turnover has been an issue, and training time costs are high due to the specialized nature of this work. These factors, combined with reduced staff hours, altered item submission policies, and a reduction in item solicitation efforts result in faculty publication upload numbers for 2012 being down to a little over 20% of what they were in 2010.

The Importance of Student Works

While the pre-restructuring faculty publications workflow was based out of Acquisitions, Cataloging handled and continues to handle the ETD process. The NC DOCKS restructuring process did not alter ETD workflows. Students completing

Month	Number of items
January	1
February	42
March	21
April	38
May	48
June	57
July	23
August	0
September	26
October	82
November	50
Total	388

Table 2. Faculty Publications Uploaded in January through November of 2012³

theses and dissertations are required to have their work loaded into NC DOCKS, though embargo options are available for students who do not want their work immediately available. New ETDs are added to the database three times a year, after graduations in May, August, and December. ETD numbers will grow over and beyond usual graduation rates in the coming months, as an outsourced retrospective digitization project will contribute approximately 3,000 additional back file theses to the database. Previous IR research by Morrow and Mower noted that some content can be considered “low-hanging fruit,” which is desirable for its ease of entry into IRs⁴. Student works such as ETDs fall into this category at UNCG, as this content generally requires no additional time for solicitation or processing before items can be added to the database. In addition to ETDs, UNCG has recently started adding other student works: award-winning papers and other works of significant merit. Where number of faculty works added to the database declined in 2011 and 2012, the number

of ETDs uploaded grew in 2011. The 2012 count is incomplete, as at the time of this writing, ETDs from December 2012 graduates have not yet been received by the Libraries.

Year	2008	2009	2010	2011	2012 ⁵
ETDs uploaded	198	194	210	230	154

Table 3. ETDs Uploaded into NC DOCKS, by Year

Future Directions

The changes in NC DOCKS staffing and policies between 2011 and 2012 have helped set the stage for additional changes to the system and its services. Recent NC DOCKS user group discussions involving all participating schools have yielded the decision to add a new component to the system. At the outset of the collaboration in 2008, the decision was made to focus on full-text scholarly materials in their completed forms, specifically excluding in-process materials and data sets. But in an attempt to align library services with the needs of faculty members on NC DOCKS campuses, NC DOCKS will soon start allowing faculty members to add completed and in-process data sets in a variety of file types. The campuses have seen an increasing need for support for faculty research data management, and as a result of this, NC DOCKS has partnered with the Odum Institute at the University of North Carolina at Chapel Hill on the integration of NC DOCKS with a product called Dataverse. Dataverse was created specifically to work with data sets, and through the partnership with the Odum Institute, participating faculty members will have assistance with accessibility, migration, preservation, and replication of their files. At UNCG, the Libraries will provide support to faculty members on creating data management plans (DMPs) and choosing an appropriate repository, as some data sets may be more appropriate for externally-based subject-specific repositories. If NC DOCKS is the chosen repository, the Libraries will provide assistance with loading data sets and creating metadata. To end users, these data set records will have the same NC DOCKS look and feel, but behind the scenes the data sets will live in an instance of Dataverse instead of on the NC DOCKS

servers. Faculty members will have the option to share their data sets publically or to restrict access to certain users. This venture is not live yet; it is still in development.

Conclusion

There have been many changes in the library landscape since the start of the NC DOCKS project in 2008. Budget cuts and a staffing realignment forced a reevaluation of UNCG's NC DOCKS policies and procedures in 2011. As a result the Libraries have scaled back on solicitation of faculty publications and on the level of service available for processing these documents, directly leading to a marked decrease in the number of faculty publications added to the database. But the addition of student works, especially ETDs, has not slowed and will grow considerably in the coming year with the addition of approximately 3,000 items from a retrospective digitization project. While faculty publication numbers will likely continue to grow at a slow pace for the foreseeable future, student works can be relied upon for a steady stream of additions to the IR. NC DOCKS has also embarked on a new track through a partnership with the Odum Institute at the University of North Carolina at Chapel Hill, and is in the process of adding support for management of faculty research data sets. This move helps better align the Libraries with the needs of the university's faculty members. Through the changes outlined in this paper, NC DOCKS will continue to grow and to showcase faculty and student scholarship from the University of North Carolina at Greensboro.

¹ <http://libres.uncg.edu/ir/uncg/>

² This count represents only items uploaded between January and early July, as faculty publication processing and upload was suspended for the remaining months of the year.

³ This count updates and corrects data presented during the Charleston Conference, as new information has become available since the conference.

⁴ Morrow, A., & Mower, A. (2009). University scholarly knowledge inventory system: A workflow system for institutional repositories. *Cataloging and Classification Quarterly*, 47(3-4), 286-296.

⁵ This figure represents only ETDs from graduations in May and August. December numbers are unavailable at the time of this writing.