

## Against the Grain

---

Manuscript 8495

---

### Biz of Digital – Developing and Growing a New Repository Service: Part 3 Expansion

Michelle Flinchbaugh

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>



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

---

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

# Biz of Digital — Developing and Growing a New Repository Service: Part 3 Expansion



by Column Editor: **Michelle Flinchbaugh** (Acquisitions and Digital Scholarship Services Librarian, Albin O. Kuhn Library & Gallery, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250; Phone: 410-455-6754; Fax: 410-455-1598) <flinchba@umbc.edu>

**Column Editor's Note:** This is Part 3 of a 3 part series on *Creating a New Repository Service*. Part 1: *Getting Started* appeared in the June 2019 issue (v.31#3). Part 2: *Procedures for Library Submissions* appeared in the September 2019 issue (v.31#4). This is Part 3, which completes the series. — MF

## Introduction

The **University of Maryland, Baltimore County (UMBC)**, a research-intensive institution with 546 full-time, 292 part-time

faculty, was utilizing the Maryland Shared Open Access Repository (MD-SOAR) DSpace platform to develop repository services. Few faculty would self-submit, so the library was

developing processes, procedures, and documentation for doing submissions in the library. A librarian shifting duties to work full-time on a new repository, the Digital Scholarship Services (DSS) Librarian, was eight months into a soft roll-out with minimal outreach only to individual faculty members, and she was the only staff working on the repository. Despite the limited outreach, and that the system and services hadn't been rolled-out to all of campus, a robust flow of submissions for her to process and enter had developed.

The system needed to be rolled-out to all of campus, and staffing needed to be added—both were done simultaneously at the beginning of the fall 2018 semester.

## Expanding Service to the Entire Campus

First, the implementation of the system was announced to all of campus. Then the DSS Librarian attempted to tell all of the faculty on campus about the new system. First, she sent an email to email list for academic department chairs asking to present at one of their meetings in the coming year. Most didn't respond, but she was scheduled into a few departments' meetings. She then began contacting department chairs for the remaining departments individually, and continued with this through an entire semester. She was able to set up many more presentations at departmental meetings, but still the vast majority of departments, 68%, didn't respond. For those departments where she wasn't invited to present, she emailed a flier to all of the faculty in the department. Later she began contacting campus centers about *ScholarWorks@UMBC*. Via these presentations and contacts, and ongoing processing of new **UMBC** publications via Google Scholar Alerts, at the end of the first year 91% of **UMBC's** academic departments had at least one work in *ScholarWorks@UMBC*. All of the major academic administrative units (the Provost's Office, Deans' Offices) and 18 centers also had works in *ScholarWorks@UMBC*.

## Hiring Help

The amount of time that the DSS Librarian had to process and enter new submissions was insufficient to handle incoming submissions as soon as she began receiving long lists of publications. She notified both her supervisor and other administrators that she was only able to add a little more than 20 items per month, and of the number of submissions she had, and requested a student and began developing and preparing processes and documentation for a student to do the data entry work, and also possibly some processing of submissions.

Column Name	Column Content
<b>Entered During the Submission Process</b>	
Title	The title of the work to be added
Collection	The name of the collection to add it to
What to submit	Link only, or link and pdf
File name	The name of a file provided by the author
URL to get PDF	The URL where the pdf to load can be found
Other Files to Load	Any supplements to be loaded with the work
Publishers URL to get Metadata (do not load pdfs from these sites)	The publisher's URL when the final published version can't be loaded
Type of Work	Journal articles preprints, technical reports, etc.
Conference, publisher or journal and citation (if not available elsewhere)	Any information not included elsewhere that needs to be included in the record
Series/Series Number	Series and series number
Publisher requirements to add during submission	Elements that publisher's sometimes require be included such as the citation, their DOI, a link to their version, etc.
Lab, Project, Facility or Class Keyword	e.g. UMBC Ebiquity Research Group, UMBC High Performance Computing Facility, HIST 490 Mao's China and After
Creative Commons License	If an item is on a creative commons license, its terms
<b>Entered After the Item is Submitted</b>	
<b>Administrator Privileges Required</b>	
Additional collections	If the item is interdisciplinary, additional departmental affiliations of authors and the status of authors to map the item to additional collections
Publisher's required statements	Copyright statement and things that the publisher requires we say about the item.
Creative Commons License	The systems assigns all items a version 3.0 U.S. Creative Commons license. When a work is on a different version, we include a link to the CC record for that license. THE CC license URL and title are put in place of the system assigned ones in the rights field
embargo length and expiration date	If there is an active embargo on the item, it's length and expiration date: e.g. 12 months, expires

**Figure 1: Fields in the submission spreadsheet that student assistant uses to input items into the IR**

*continued on page 84*

The challenge of having a student do data entry work was in how to convey information to the student on what to enter without entering all of the data on the item. The first attempt at this was a simple list of items to enter, along with a link to the item's metadata record on the publisher's site. Metadata values readily available on the publisher's site (author, abstract, keywords, etc.) weren't included in the list of items to enter since the student assistant could readily find, whereas information not available or that requires a judgement call was included in the list. Relying on copying and pasting from the publisher's metadata record where possible speeds up both the processing of items and their submission to the system.

However, using a list, the librarian noted that there were inconsistencies in what information she included and didn't include, some because of the nature of the work and others because of which versions of the work were available and what metadata was available, and yet others because she simply didn't prepare the items consistently. Unhappy with the list, she switched to spreadsheet with columns, and decided most information that could be found in metadata records and the work itself wouldn't be included in the spreadsheet, limiting the spreadsheet to links to works and metadata about works and information not readily available in metadata or on the work, or that required a judgement call. (See Figure 1.)

The spreadsheet was supplemented with detailed documentation on entering and completing items from a spreadsheet, available here: <https://wiki.umbc.edu/display/library/Entering+and+Completing+Items+from+a+Spreadsheet%2C+Full+Procedure> and a metadata chart to use as a short guide on what to look for and where to put it a record. (See Figure 2.)

To finish, the Dublin Core element was added to all lines in the "Where to put it column" to facilitate editing in administration which is done entirely by Dublin Core element.

In addition to providing these resources for the student assistant, after hire, the DSS Librarian spent a great deal of time training the student assistant, in steps, and checking and correcting work until it was done correctly. Initial training was only the submission processes utilizing the submission form. When that was mastered, the student was trained in utilizing administrative capabilities to edit metadata and map items to additional collections, then how to use the administrative capabilities to correct an error. In a final phase of training the student learned to add rights statements, change a Creative Commons license version, and embargo an item for automatic release when the embargo ends.

Once the student had worked through a substantive backlog of items to be entered, the DSS Librarian also trained the student to determine if an item is in-scope for the repository, to check rights and then enter those works that can be posted into a spreadsheet

Information to Look for	Where to put it
	<b>Page 1</b>
<b>Authors</b>	<b>Author(s) field.</b> Last name in the last name field. First name, middle name or initial in first name field.
<b>Contributors</b> —any role in the creation of the work except author and advisor which go elsewhere. Composer, illustrator, programmer, director, etc.	<b>Contributor(s) field.</b> Last name in the last name field. First name, middle name or initial in first name field.
<b>Advisors</b>	<b>Advisor(s) field.</b> Last name in the last name field. First name, middle name or initial in first name field.
<b>Title</b>	<b>Title field.</b> Title as it appears on the work.
<b>Other forms of the title</b> that appear on the work or elsewhere. The title with a non-searchable character replaced or omitted.	<b>Other title(s)</b>
<b>Date published, issued, or added</b>	<b>Date of issue.</b> Fill in as much detail as given on the work or In the metadata.
<b>Publisher.</b> An agency or company that played a role in producing the work via peer-review, copy-editing, and distributing it.	<b>Publisher.</b>
<b>Citation of publication.</b> If a work is not published, there is no citation of publication to include. Citations can usually be Copied and pasted from the publisher record, ensuring that it's complete, or created yourself.	<b>Citation:</b> Authors. Article title. Journal. Vol. #, issue #, year, DOI or link.
<b>Series Name and number</b>	<b>Series Name. Report or Paper Number.</b>
<b>Publisher's DOI</b>	<b>Identifier. Select DOI.</b>
<b>Link to publisher's record or PDF</b>	<b>External link field.</b>
<b>Type of work is it</b>	<b>Type.</b> Text for anything with typed characters.
<b>Format.</b> Take from spreadsheet.	<b>Format.</b> E.g. journal article pre-print.
<b>Number of pages.</b>	<b>Extent.</b> E.g. 8 pages.

**Figure 2: Guide for what to look for and where to put when entering items into IR**

for future entry. Substantive documentation was created covering this. The portion on scope covers theses and dissertations, CVs, obituaries, and abstracts with no full text, the requirement that an author must be affiliated with **UMBC** or the article about **UMBC** or someone affiliated with it. The section on checking rights is broken down by format, and covers Creative Commons licenses, open access, U.S. Federal Government publications. A final section covers determining which collections to add an item to. The full procedures is available here: <https://wiki.umbc.edu/pages/viewpage.action?title=Preparing+a+Spreadsheet+of+Items+to+Enter&spaceKey=library>.

Additional detail could be added to spell out some situations and decisions not covered, but this already extensive procedure is very challenging for someone just beginning this work. Training on it best divided up, a new staff person doing the steps they know, and the DSS Librarian doing the rest. This would

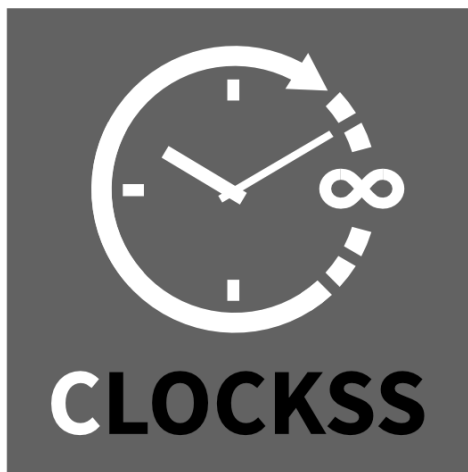
allow for training in more manageable phases and better mastery of the work.

### Still Needing more Help

To date, the DSS Librarian has barely scratched the surface of outreach and much more can be done, but the bulk of her time has been spent on a perpetual backlog of items to process and enter for faculty. The current backlog of works to check rights and enter is reaching nearly 1,200. With the DSS Librarian working full-time plus the half-time student assistant, the maximum monthly rate of processing has been 256 per month so that this constitutes nearly a 5 month backlog. Additionally, they continue processing Google Scholar Alerts, and a commitment was made to re-visit publications website annually to add new materials that were added to them. The student assistant won't be available during breaks, and graduates in a just a year and half. At the date of this writing, the DSS Librarian

*continued on page 85*

PRESERVED IN



# THE CLOCKSS ARCHIVE

- ▶ GLOBALLY DISTRIBUTED PRESERVATION
- ▶ COMMUNITY GOVERNANCE
- ▶ COMMITMENT TO OPEN ACCESS
- ▶ LOW PARTICIPATION FEES

AUSTRALIA CANADA GERMANY HONG KONG ITALY JAPAN UNITED KINGDOM UNITED STATES

## Come see us at table # 99 at the Charleston Vendor Showcase

For more information on how to support or participate in the archive contact us at [info@clockss.org](mailto:info@clockss.org).

CLOCKSS Archive is a dark archive that ensures the long-term survival of web-based scholarly publications, governed by and for its stakeholders.

The archive includes over

- 280 Participating Publishers
- 300 Library Supporters
- 25,000 journal titles
- 33,400,000 journal articles
- 188,000 ebooks
- 53 journals have been triggered as open access

CLOCKSS is the first archive to be re-certified by the Council of Research Libraries for our Trusted Repository Audit Checklist (TRAC). Our score was upgraded for Organizational Infrastructure to the top score of 5. We maintained our top score of 5 for Technologies, Technical Infrastructure, Security. Our total score of 14 out of 15 is the highest score of any of the archives that have been certified.

<https://www.clockss.org>

### Biz of Digital from page 84

has requested a full-time line to hire a staff person which was promised if/when there is enough work to justify doing so.

#### Workflows

One workflow for loading ETDs, and second bifurcated workflow, first checking rights and finding info and filing it into a spreadsheet, and then using the spreadsheet to enter items was developed. These two workflows handle 90% of items going into the repository. However, some items don't fit well into the spreadsheet because of their nature, and require utilizing Dublin Core elements not normally utilized, or entering multiple values into Dublin Core element where there is usually only one. Other items may be serials, multivolume sets, art, video, symposia with video of multiple presentation given by different people. Modified spreadsheets were developed, or will be developed to be used in these instances. At some point in the future, items that are particularly time consuming to enter manually (for example, works with more than ten authors) may also be loaded, depending on our ability to develop automated methods of reformatting data accurately.

#### Future Plans

With the large backlog of work, methods for making work more efficient are a high priority. In the short term, Macro Express can quickly automate some data entry tasks, making the entry of new submissions less time-consuming. The DSS Librarian has read articles on how other libraries have automated the submission processes using citation managers and spreadsheets to batch load new submissions, and in the future will investigate if what other libraries have done will work at **UMBC** and with MD-SOAR.

Another high priority is further extending outreach. Additional outreach needs to be done to professional programs, and faculty and programs located at distant locations in either Baltimore City or at **University of Maryland System's** Shady Grove campus. Outreach also needs to be done to lecture series, campus awards, and student publications and research forums. Finally, outreach to new faculty six months to a year after they come to **UMBC** needs to be put in place. Finally, an annual email needs to go to faculty affiliated with each department and center reminding them to send materials.

Works in the repository also need to be promoted. A plan to have the system automatically tweet all new items has stalled. **UMBC**

Library's Committee of Social Media and Outreach will promote items on social media, but interesting items have to be identified and sent to them. Other means of promotion are also possible, but it's been difficult to find time for this with a perpetual large waitlist of materials to process and add.

The Digital Service Librarian has also been working the MD-SOAR Governance Group toward enhancements to solve various inefficiencies and problems related to the system configuration. A current effort is being made to standardize how metadata indicates that an item is a preprint or postprint and to select enhancements to move forward. A new extended submission form is desirable, and additionally tweaks to the indexing and display, and field configuration would be of value.

Another area needing work is resolving inconsistencies in metadata as procedures have changed over time. In some instances, we've learned, in others, reached agreements not previously realized, so there have been inconsistencies between how we entered records six months ago and how we enter them now, and at some point hope to do a large scale batch edit of our metadata to make records consistent. Of particular importance is putting in place some type of authority control on the names

*continued on page 87*



## Squirreling Away from page 86

we serve.<sup>2</sup> The most successful libraries in navigating through change are likely the ones that have the greatest ability to be flexible given the changes that our profession is being asked to take on. Here are three areas where flexibility has been key in the way we have faced change.

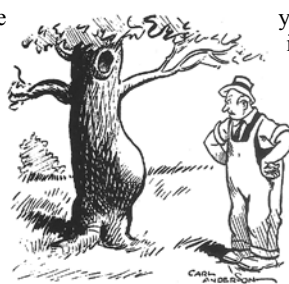
First, flexibility in the services we provide is one of the most important aspects to consider with library change management. There are certain aspects of our work and the services we provide where we need to make sure that we are aligned with our community needs. It seems that much of the literature about marketing in libraries (and elsewhere) is about how to successfully increase adoption or use of a service or product that you are providing vs. providing what is desired. One instance that we have had to change is our desire of managing course reserves in our new library. Taking into consideration the lack of space and the desire not to hire evening staff to work onsite, we opted to change course and not bring back course reserves to our library portfolio. We aspired to move items into electronic format if possible — but as you know, this can be difficult with course adoption texts, let alone textbooks. We had to balance between what we felt we could do with what the needs are in our community and moved accordingly. It would be nice if we could provide this, but it did not meet our abilities and resources.

Second, flexibility with collections is certainly a key aspect to change management. Over the past eight years in our library, we have had flat or declining collections budgets. Even in a flat year, that means cuts to cover inflation for the resources that you are planning to keep. The flexibility involved here is to ensure that your cuts are balanced and that redundancies are eliminated first. While you can work on a straight cost per use model, this will likely not tell the entire story. Resources targeted for faculty and researchers will likely be far less commonly used than ones geared to student use. Additionally, there might be resources that had been staples in the library collection since before you arrived — and represent the core holdings of many libraries. But if they are not being used, then that is an opportunity for you to make changes that reflect your reality at the library.

Third, flexibility with people might be the most important aspect of navigating through change in any organization. I believe that change is an extremely personal construct that will impact different people in vastly unique ways. In our staff of just under twenty, we had some people who saw relatively little change in their day to day life with our transformation — as well as some who had to learn nearly a completely new job. But long before this change, I preached flexibility in the workplace for one simple reason. My premise is that if I am flexible with my team, they in turn may be flexible with our users and the community we

serve. Conversely, if I were rigid or rule bound with my team, it would be difficult to expect them to be flexible with our community. Additionally, a great number of the ways that we may be flexible with our teams can create a better working environment. Giving your team the freedom and flexibility to navigate through these changes as they see fit enables the library group to better serve your community. While there will be aspects of library change that are fairly rigid, especially with space and budget constraints, creating a flexible environment will pay dividends for you and your team.

So just like that cute puppy, kitten, dog, cat or other pet you bring home from an adoption event, there is a great deal of joy that will come



your way. But the more rigid you are in bringing this new being into your family, the more likely you will be disappointed and troubled with the results. Pets can be a great deal of work, but all of it is worth it when they are curled up with you when you are working on your late articles — right? 🐾

**Corey Seeman** is the Director, Kresge Library Services at the **Ross School of Business** at the **University of Michigan, Ann Arbor**. He is also the new editor for this column that intends to provide an eclectic exploration of business and management topics relative to the intersection of publishing, librarianship and the information industry. No business degree required! He may be reached at <cseeman@umich.edu> or via twitter at @cseeman.

### Endnotes

1. **Sutton, Sarah**. "Flexibility in the Face of Change." *Library Resources & Technical Services*, vol. 57, no. 2, Apr. 2013, pp. 77–86. EBSCOhost, doi:10.5860/lrts.57n1.77.
2. **Kesselman, Martin A**. "Hot Tech Trends in Libraries: Flexibility and Changeability Is the New Sustainability." *Library Hi Tech News*, vol. 35, no. 8, Sept. 2018, pp. 1–5. EBSCOhost, doi:10.1108/LHTN-09-2018-0062.

## Biz of Digital from page 85

of **UMBC** authors. The system automatically creates author pages, but creates an additional author page for every different form of a name. A decision to put authority control in place for **UMBC** author names would collate their works all on one author page.

Repository training for faculty was developed and scheduled, but with very little interest in self-submission on campus, no one RSVP'd and the training session was cancelled. The **University of Maryland**, College Park, has had some success with a workshop on authors' rights, and important rights issues directly pertinent to the repository can be addressed in that type of workshop, so in the future, the DSS Librarian hopes to develop and offer such a workshop.

Moving from a part-time student assistant with diminishing returns on training to a full-time staff person is also highly desirable. That would facilitate much more work getting done at a higher quality. This would also allow the DSS Librarian to shift from spending most of her time doing staff-level production work (or correcting the student assistant's work), to doing further outreach and promotion of the repository and the works in it, refining procedures and documentation, and batch

clean-up of the data in the system. At present, the Digital Scholarship Service Librarian can only take on a single project each summer. More time would also allow her to develop and offer workshops and do more outreach. Additionally, she could develop other digital scholarship services in consultation with library management and write grant proposals toward obtaining funding for startup costs.

### Conclusions

The gradual implementation was good preparation for extending the repository service to all of campus, and moving to a production mode where many more items would be processed and added each month. Basic procedures were in place, allowing a shift in focus from legal and technical issues to people, from how to do work to working with staff to increase production. It also gave the novice DSS Librarian time to learn. In 18 months since the implementation of *ScholarWorks@UMBC*, 1,683 items have been added, bringing the total number of works to 2,773, and there have been 4,145 visits to *ScholarWorks@UMBC*. This is a strong start, but in time **UMBC** could potentially be adding 4,000+ items per year to the repository. 🌱