



University of
Massachusetts
Amherst

Adopting a Method to Evaluate Bibliographic Electronic Resource Title Sets of Metadata

Item Type	article;article
Authors	Eustis, Jennifer M.
DOI	https://doi.org/10.7275/410b-pt12
Download date	2025-05-24 07:08:31
Item License	http://creativecommons.org/licenses/by-nc-sa/4.0/
Link to Item	https://hdl.handle.net/20.500.14394/32164

University of Massachusetts Amherst

From the Selected Works of Jennifer Eustis

November, 2019

Adopting a Method to Evaluate Bibliographic Electronic Resource Title Sets of Metadata

Jennifer Eustis, *University of Massachusetts Amherst*

This work is licensed under a Creative Commons CC BY-NC-ND International License.



Available at: <https://works.bepress.com/jennifer-eustis/10/>

Adopting a Method to Evaluate Bibliographic Electronic Resource Title Sets of Metadata

By Jennifer Eustis, University of Massachusetts Amherst

Introduction

The challenges of managing electronic resources are well known. Guidelines adopting new practices were published by OCLC and the Library of Congress Standards Office in the early to late 1990s. Almost twenty years later, their management remains difficult. This is particularly the case with bibliographic title sets of metadata for electronic resources. Sets can be retrieved from multiple sources. Sometimes those record sources have different versions of the sets where the number of titles, quality, or documentation differs. As a result, the amount of work to prepare these sets to be batch loaded into library systems remains onerous.

At the University of Massachusetts Amherst (UMA), bibliographic title sets fall last in the workflow of electronic resources packages and databases. An electronic resource package and titles are enabled via discovery or third-party linkage services and holdings management software. This works well for journals and a majority of databases. Staff in acquisitions and discovery ensure that payment, agreements, and/or licenses are set. These steps are recorded in CORAL, our electronic resources management software. Cases exist where the resource is not available through discovery or one of our third-party applications. Some packages offer titles sets of metadata to be batch loaded into the system providing additional access points for users.

These various methods of providing access and discovery results in a complex system of integrated applications and workflows. Beyond a fracture experience for patrons, staff often have

difficulty understanding how these electronic packages come to be discoverable in either the catalog or discovery layer.

Adding to this, UMA has never had a formal evaluation process of bibliographic title sets.

Documentation is outdated and sometimes incorrect. Work was done by only a few individuals.

When the primary staff person left and other key people working on these resources moved on to other projects, this created significant gaps with knowledge transfer and how to manage these resources. Furthermore, many staff are either unaware or don't know the workflows surrounding electronic resources. By extension, their understanding of how, when, and the nature of bibliographic title sets of metadata can be minimal. Hence, staff, in particular, research subject specialists are often frustrated by what they feel should be done by just simply pressing a button or cataloging titles individually. This is why it was decided in the past year to adopt a method to evaluate bibliographic title sets. The aim is to develop a comprehensible narrative of their lifecycle before, during, and at the end of their batch loading process that can be shared with UMA administration and liaisons to provide a more holistic view of what it means to manage bibliographic title sets at UMA.

A Highlight of Existing Evaluation Methods

Though this project is in the early stages, it is important to see what other institutions or case studies are available on this topic. Our research thus far has found two distinct groups, qualitative and quantitative. The qualitative approach focuses on providing a checklist for staff to use for a bibliographic title set of metadata (Ashby, 2018). These checklists highlight the need for a staff person to check the quality of the metadata based on national and local practices, often

steps on how to preprocess MARC21 file, and sometimes steps to batch loading the set into the library system. The aim is to provide a set of criteria for any staff person to follow at this stage of the workflow for bibliographic title sets of metadata. On the other end of the spectrum is the quantitative approach (Thompson & Traill, 2017). This method applies a number score to each criterion. Each record in a file then receives final scores which are then output as visualizations and/or standard means. Both methods, the qualitative and quantitative, are used to determine the quality of the records, their completeness, and which record source is the more viable candidate to batch load into the system. This review which is still ongoing allows us to learn from these methods to find what could work here at UMA.

How to Adopt Those Methods

In addition to the literature review, understanding initiatives at UMA are key. The first is to increasingly adopt a data-driven method to evaluate processes, workflows, and strategies. For the Metadata unit at UMA, this means integrating the number of batch loaded title sets into the monthly statistics. For each bibliographic set of metadata, a local MARC field, 910, is used to record the batch loaders' initials, the date, and type of batch load (update or new). A script is run monthly that sends the data off to Tableau. Unfortunately, deletes and merges are not reflected in these statistics. Keeping a count of the numbers of items cataloged is an old practice. The problem is that numbers often don't tell the entire story of how much work goes into a record. This is particularly the case with bibliographic sets. The number of titles batch loaded reflects a later stage of bibliographic title sets of metadata, namely the maintenance stage. The selection of a record source and the preparation of the set remain a black hole. This is one of the reasons why many staff doesn't understand how and why bibliographic title sets are loaded.

How we work with electronic packages is often misunderstood which has led to an important initiative at UMA, the vendor scorecard. This project began about two years ago with the aim to score vendors based on several criteria that reflect UMA values and ethics. Criterion includes the availability of VPATs or voluntary product accessibility template, vendor transparency, or metadata. The Metadata unit adopted a checklist to determine whether a title or set of titles meet the minimum quality standards for UMA. This checklist formed the basis for how to think about evaluating bibliographic title sets of metadata.

The first iteration of an evaluation template was qualitative. It incorporated the checklist developed to minimally quality records. It took key elements from the vendor scorecard. These elements consist of documenting issues that aimed to provide a better understanding of how records sources differ and decisions into how title sets are treated. For example, the evaluation asked whether the title set was static or not, if MARC records were free and easy to find, or if there is accompanying documentation. Further, the evaluation asked about the presence of support and whether any work needed to be done before batch loading title sets into the system.

As this evaluation began to be used for title sets, ways to improve it became clear. This first iteration failed to distinguish 3 phases to the evaluation: the work involved in selecting a record source, the preparation needed to initial batch load the set into the system and then maintaining that set going forward. In addition to this, it was difficult to create a baseline from which to compare sets. One reason was that this checklist worked to evaluate a bibliographic title set from each record source rather than compare sets from different vendors. In short, this first evaluation

was a checklist to ensure that an appropriate record source was selected. And it didn't tell a meaningful story of how title sets are selected, ingested, and maintained.

This is why the second version of our evaluation is being developed. This second version draws from the quantitative approach explained in Thompson & Traill's article on leveraging python for the selection, ingest, and maintenance of title sets (Thompson & Traill, 2017). This idea of separating the 3 phases apart is also being adopted. The evaluation consists of the following sections. The General Information provides the name and CORAL reference number. The Pre-processing Evaluation phase aims to score record sources based on metadata quality and completeness, transparency of vendor practices in terms of support, documentation, any fees, record count between record sources and to the title list of what UMA has access. This phase addresses whether or not title sets need to be enhanced before ingestion and if additional scripts (either using Python or MarcEdit tasks) need to be created for this purpose. The Initial Load Evaluation phase score criteria aimed at determining steps needed to be done before loading a title set. This includes creating a new collection code in the system, perhaps removing records from the system, or even if the records can be loaded. The Post Initial Load Evaluation seeks to score the work needed to maintain the title set going forward. This last phase considers whether the title set is static or ongoing and the nature of what it means to be an ongoing set. For instance, if the record source delivers multiple types of files (updates, merges, new, deletes), the score is higher.

At first, the scoring was simply a 1 or 0 based on yes or no answers. This simplistic scoring was found to be inadequate rather quickly. When comparing a title set that couldn't be loaded with a

title set that could be, the visualizations of the total score of the 3 phases were pretty much the same. This result failed to illustrate the differences between the two sets. Another shortcoming was that visualizations derived from the total scores didn't illustrate any visible difference between sets that needed significant enhancement or none. Lastly, the visualizations combined scoring information related to the total scores for each phase and the estimated time to evaluate each phase.

To address the first and second issues, weighted scores have been applied to many of the questions in each phase. For example, having a correct and working URL is seen as crucial. Even if the title set records are complete and meet or exceed minimal UMA standards, if the URLs don't work, then the title set shouldn't be batch loaded. A good example of this scenario occurred with the title set of records for the epackage HistoryMakers. This set only came from OCLC Worldshare Collection Manager. The records were complete and met UMA's minimal quality metadata. All the records had a URL. But on further inspection, it turned out that this URL has a prefix that couldn't be reconciled in our EZproxy service. For this reason, the title set wasn't selected for ingest. As a result, more scoring weight is applied to a correct and working URL. The total time was separated from the scoring applied for each phase. And a question was added on whether or not the set could be batch loaded and weighted more than other questions.

This is not a systematic approach. Rather the weighting is determined by what we expect to see given the evaluation of the set. The goal going forward is to figure out the best ways to score rather than relying on this arbitrary approach.

Despite this arbitrariness, the results seem more promising than the first version of the evaluation template. Figures 1 and 2 illustrate the first foray into the visualizations from the weighted scoring and evaluation times. In Figure 1, the score for HistoryMakers is extremely high. This was the title set that we were unable to load because of the URLs. Our scoring works so that lower scores mean an easier set to pre-process, load, and maintain. Larger scores indicate either a set that is not selected for ingest or a very challenging title set for all three phases.

There are many questions with this approach that are worth asking. First, is assigning arbitrary weights to criteria an acceptable practice? If not, what method of reasoning would apply to the weights? Is it possible to create a baseline so that scores can be identified with the level of difficulty encountered with each set? Currently, the total score for all three phases is 100. Does this make sense? In Figure 1, the Initial Load evaluation score for HistoryMakers is fairly low and not that much different visually from the other title sets. This is due to how HistoryMakers was scored as illustrated in Figure 3. This title set didn't get the total maximum score of 10 because of 3 questions in this section. In this case, would it make sense to weight these questions differently or review the questions in this section?

Figure 1: Score for Each Phase and Total Overall Score

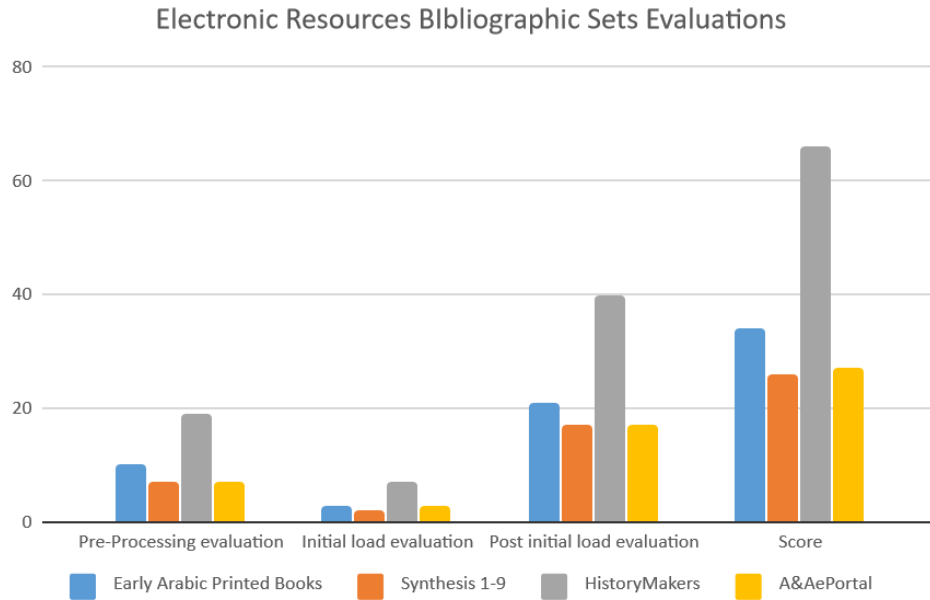


Figure 2: Time for Each Phase and Total Time For Each Title Set

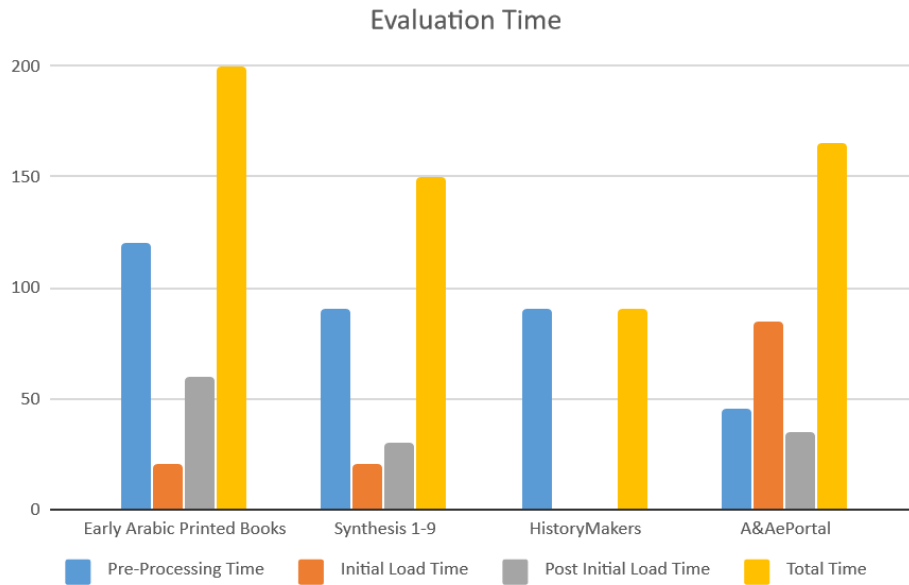


Figure 3: HistoryMakers' Scores for Initial Load Evaluation

Initial Load Evaluation

Criteria	Response	Response Guide	Scoring	Score (Max Score: 10)
Does a new collection and load code need to be created?	Yes	Yes/No.	<ul style="list-style-type: none"> • Yes = 1 • No = 0 	1
Does new documentation need to be created?	Yes	Yes/No.	<ul style="list-style-type: none"> • Yes = 1 • No = 0 	1
Does the Record Loading shared network drive need to be set up for these MARC records?	Yes	Yes/No.	<ul style="list-style-type: none"> • Yes = 1 • No = 0 	1
Are Aleph records being overlaid?	No	Yes/No	<ul style="list-style-type: none"> • Yes = 1 • No = 0 	0
Do Aleph records need to be deleted?	No	Yes/No	<ul style="list-style-type: none"> • Yes = 1 • No = 0 	0
Do the MARC records have to be loaded at off hours?	No	Yes/No. Off hours means between after 8pm and before 4am.	<ul style="list-style-type: none"> • Yes = 1 • No = 0 	0
Can the MARC records be loaded?	No	Yes/No.	<ul style="list-style-type: none"> • Yes = 0 • No = 4 	4
Total:				7

How to Construct an Evaluation Narrative

This work is still very much at its beginning stages. As seen from the section above, there are some questions to be answered before addressing the one on a constructive narrative. With that in mind, our work on refining the second version of the evaluation template is a step in the right direction. The work going forward will focus on the scores, an approach to weighting, and resulting visualizations. Then, it will try to answer the question of whether these visualizations illustrate a meaningful narrative that can be understood by staff. Part of the work going forward are points already mentioned in reviewing the weighting process and the questions for each phase. The other important aspect is to define what is meant by a meaningful narrative and how visualizations can convey that in a way that aligns with UMA's values and ethics.

Conclusion

UMA is just at the start of its journey to evaluate bibliographic titles sets in a meaningful and constructive way. Thanks to the current literature, some methods can be adopted here and then

refined for UMA's practices, culture, value, and ethics. This project takes time and a final evaluation template is still being formulated. However, this work will help UMA staff understand better the challenges and decisions surrounding the selection, ingestion, and maintenance of bibliographic title sets of metadata.

References

Ashby, S. D. (2018). Managing Electronic Resource Collections and Batch Loads. Retrieved from https://www.ohiug.org/uploads/6/1/3/5/61351715/2018_managing.pdf.

CARLI (2013). Batch Loading Bibliographic Records for Electronic Resources. Retrieved from <https://www.carli.illinois.edu/sites/files/i-share/documentation/eresbatch.pdf>.

Coughlin, D. M. (2014). A Framework for Evaluating Electronic Resources. Retrieved from <https://etda.libraries.psu.edu/catalog/23831>.

Nagra, K. A. (2009). The Evaluation of Use of Electronic Resources and Services in Academic Libraries: A Study of E-metrics and Related Methods for Measurement and Assessment. *Journal of the Library Administration & Management Section*, 6(1), 28-41. Retrieved from https://academicworks.cuny.edu/cgi/viewcontent.cgi?article=1002&context=bm_pubs.

OCLC (2019). Cataloging electronic resources: OCLC-MARC coding guidelines. Retrieved from https://help.oclc.org/Metadata_Services/Connexion/Related_documentation/Cataloging_electronic_resources_OCLC_MARC_coding_guidelines.

Okogwu, F. I. and Dike, V.W. (2017). Evaluation Practices of Electronic Resources in University Libraries in South East Nigeria. *Library Philosophy and Practices* (1656). Retrieved from <https://digitalcommons.unl.edu/libphilprac/1656>.

Phillips, K. (2017). A Data-Driven Approach to Electronic Resources Collections. *Library Journal* (August 23, 2017). Retrieved from <https://www.libraryjournal.com/?detailStory=a-data-driven-approach-to-electronic-resources-collections>.

Rutgers University Libraries (2011). Evaluating Bibliographic Record Sets for Electronic Resources. Retrieved from https://www.libraries.rutgers.edu/rul/staff/technical_services/cataloging/eval_bib-record_sets.pdf.

Sapon-White, R. (2014). E-Book Cataloging Workflow at Oregon State University. *Association for Library Collections & Technical Services* (58:2). Retrieved from <https://journals.ala.org/index.php/lrts/article/view/5335/6515>.

Thompson, K. and Traill, S. (2017). Leveraging Python to improve ebook metadata selection, ingest, and management. *Code4Lib Journal* (38). Retrieved from <https://journal.code4lib.org/articles/12828>.