

Against the Grain

Manuscript 8190

Let's Get Technical – Taking the CIP Program into the Future with PrePub Book Link

Caroline Saccucci

Stacey Marien

Alayne Mundt

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 for additional information.

Let's Get Technical — Taking the CIP Program into the Future with PrePub Book Link

by **Caroline Saccucci** (CIP and Dewey Program Manager, Library of Congress) <csus@loc.gov>

Column Editors: **Stacey Marien** (Acquisitions Librarian, American University Library) <smarien@american.edu>

and **Alayne Mundt** (Resource Description Librarian, American University Library) <mundt@american.edu>

Column Editor Note: *In this issue's column, we profile how the **Library of Congress'** CIP program upgraded to a new platform. **Caroline Saccucci**, CIP and Dewey Program Manager describes the history of the CIP program and the steps that went into deciding and implementing a new platform for the publishers to use. — SM & AM*

Introduction

The **Library of Congress** Cataloging in Publication (CIP) Program has relied on technology built in house between 1999 and 2003 to manage the flow of applications for CIP data. In 2017, Library management determined it was the right time to fully upgrade the platform used to submit and process CIP applications. As the CIP and Dewey Program Manager, I am the product owner of the database and therefore the primary customer of the development project to replace the aging system. As product owner, I have the responsibility and the authority to make decisions on most aspects of the project, in conjunction with my division chief, who is the system owner, and my team of subject matter experts. This article will discuss the reasons for the platform upgrade, the development process, and our hopes for the future with the new system.

Since 1971 the CIP Program has partnered with U.S. publishers to create prepublication metadata for forthcoming books. This metadata is printed in the books and distributed via subscription to bibliographic utilities such as **OCLC** and to publishers and vendors. Today approximately 4,500 publishers participate in the CIP Program. In FY2017, the **Library of Congress** cataloged over 60,000 CIP titles, representing a significant percentage of academic, scholarly, and trade monographs published in the United States. Because these titles have already been cataloged by the **Library of Congress**, local libraries can quickly add these titles to their collections. CIP publishers are required to send a complimentary copy of their published books to the CIP Program, thereby saving millions of tax payer dollars which would have been spent purchasing these titles with appropriated funds.

When the CIP Program started in 1971, the

entire process was paper based and depended on the U.S. Postal Service. Publishers used to fill out a form by hand or with a typewriter. The form with marked up manuscript galley was mailed as a package to the Library. CIP Program staff would create an initial bibliographic record based on the information in the application and then send the package to the cataloging divisions to complete the cataloging. The CIP metadata would be mailed back to the publisher for inclusion in the printed book; publishers would need to retype the metadata onto the copyright page because it came in hardcopy only, not electronic format.

The Problem

In 2003, the CIP Program launched an online CIP system. Publishers could now electronically submit their applications with galleys as attachments in ASCII format. This meant, however, that publishers needed to convert their Word documents to ASCII and then tag their files with html-like codes to indicate title page, copyright page, series title page, table of contents, first chapter, second chapter, etc. Very little data in the application was prepopulated, and each application was time consuming to complete and submit. After the cataloging process was completed, the CIP data was emailed to the publisher, but the mailer often garbled diacritics in the CIP data.

The online CIP system was named **ECIP Traffic Manager** for internal use because CIP applications were submitted electronically, and it managed the traffic or flow of CIP applications. Catalogers could view the galleys as html web pages; the tagging created specific pages and chapter headings for easier navigation. Internal staff created a special cataloging software called **OnTheMarc** to convert application data to a MARC record.

In 2010 the CIP Program was able take advantage of ONIX metadata in the cataloging workflow by linking the ISBN in the CIP application to the ISBN in the ONIX product record and mapping that record's ONIX to MARC fields.

While this process worked for internal catalogers, it was not available to our 32 partner institutions, even though their university presses produced ONIX. **ECIP Traffic Manager** routed CIP applications to a cataloging team, such as the History and Military Science Section, but it

was not able to be assigned to a specific cataloger. With cataloging teams having large numbers of CIP applications to process, it became necessary for some section heads to create spreadsheets to manage their CIP work. Because this system was implemented before many security protocols had been developed, **ECIP Traffic Manager** would have to be updated or replaced.

The Process

Library Management decided to replace the aging system. As product owner, I collaborated with my chief/system owner and staff from the Office of the Chief Information Officer (OCIO) to create a statement of work and award a contract in September 2017. OCIO provided project management, user experience design, and server integration support. The CIP team worked with the user experience designer for the best possible design for a new system. The contractor provided the developers, a business analyst, and a dedicated tester. The CIP Team provided user roles, legacy application mappings, ONIX mappings, and other documentation as needed; all documentation was posted on the team Confluence site. The work was divided into two-week sprints, and user stories were created in Jira for all the functionality. Both LC and contractor staff created user stories, and the team tested the results of each sprint. The team met weekly for two-hour development meetings and two-hour design meetings. The team also met twice a week in the morning for 15-minute standup meetings to give up-to-the-minute progress. The team made data migration decisions, and the project manager mapped legacy application fields to new **PPBL** application fields.

The Outcome

The new system, named **PrePub Book Link (PPBL)**, is built on a **ServiceNow** platform. It is hosted in the cloud and is therefore accessible to all users inside and outside the firewall. **PPBL** will provide an array of enhanced functionality and features for publishers, catalogers, and other internal users who work with CIP applications.

Publishers will be able to create and manage user accounts for single and multiple publishers and imprints. In the legacy system, imprints were not linked to each other, so the CIP Program had no real overview of how the publishing industry was connected. **PPBL** will allow them to have one log-in for all the publishers and imprints they need. With the legacy system, each account had one user ID and password for all users, and users locked

continued on page 89



Let's Get Technical from page 88

themselves and others out; in PPBL, each user will have an individual account with password self-service.

Publishers will be able to attach galleys as Microsoft Word or PDF documents and either in a single file attachment or in multiple files. They will no longer need to convert their files to ASCII and tag them. Instead, they will indicate the content they have attached, e.g., title page, copyright page, table of contents, etc. The advantage to LC staff is that the document will look more like the forthcoming book. With a PDF, catalogers will finally see images and graphics, which is especially useful for cataloging art books, graphic novels, and works for juvenile audience.

Publishers will have the ability to use their own ONIX data to pre-populate much of the CIP application. If publishers create ONIX metadata early enough for the CIP process, they will save themselves time when filling out the CIP request. Metadata such as author(s), editor(s), title, subtitle, series, and summaries, are excellent examples of the kind of data required in the application for CIP data that are also contained in the ONIX metadata. The CIP Program already receives many feeds of ONIX from various publishers and aggregators, and we want to take better advantage of it. The CIP Program collaborated closely with the **Book Industry Standard Group (BISG)** to map ONIX 2.1 to CIP application fields and MARC fields.

All email correspondence to and from publishers will be recorded within PPBL, so any staff member can view the history. Outbound email to the publisher was recorded in ECIP Traffic Manger, but all incoming messages landed in the staff member's Outlook account; no one else could see the response, leading to delays and unnecessary follow-up emails.

The CIP data will be issued to the publisher via an emailed link, and that link will enable the publisher to copy and paste the data without concern for garbled diacritics. PPBL will enable publishers to identify any staff who may need access to this CIP data. The CIP data will be generated directly from the MARC record, and publishers will be able to indicate whether they want the age/grade levels and/or summaries to appear in the CIP data. (This hide/show feature would apply to the CIP data as it would appear in the published book and have no effect on the MARC record.)

Library of Congress staff will be able to manage and track their own work. Individual CIP applications can be assigned to individual staff members, and section heads will have better oversight of the pending CIP applications in their team. A new MARC Editor was developed for PPBL to enable the application data to be converted into a MARC record and edited before import into the Library's ILS. Catalogers will be able to continue cataloging the CIP immediately upon import.

IDEAS WITH IMPACT

UNIVERSITY of CALIFORNIA PRESS



Ask us about our special discounts for small colleges
for details, contact customerservice@ucpress.edu



UNIVERSITY
of CALIFORNIA
PRESS

journals.ucpress.edu

PPBL will have more advanced reporting functionality, so a section head can see their team account or an individual staff member's account. Staff members will be able to filter the pending CIP applications by subject, so they can see what is in scope for them. There will be other filters available as well, such as publisher, ISBN, projected print date or projected date of publication. Speaking of projected print date, this is a new data element required in the CIP application to help the CIP Program prioritize the workflow. The projected date of publication will be used for the new claiming functionality, which will generate claims from PPBL, not from an email address in the MARC record.

Our partner institutions currently must download custom software onto their PCs and provide LC with delete authority on their servers to follow the FTP process for their completed records to the Library. Not surprisingly, this proved cumbersome for some institutions. With PPBL, there will no more special software and no need for ftp and server access; instead, partners will be able to upload each MARC record from their cataloging client into PPBL.

Lessons Learned

This development project required an incredible number of meetings. In addition to the meetings listed above, there were many ad hoc meetings to make decisions on tiny pieces of the project, such as creating routing

rules or helper text for the CIP request form. Communication was paramount to every aspect of this project. Because the developers were working remotely, we relied heavily on WebEx and recorded most of the meetings. We utilized Jabber for quick in-house conversations, and we updated Confluence with decisions, documentation, and important links. PrePub Book Link was officially announced at **ALA 2018**, but we communicated with our users via email, conference presentations, and press releases. Finally, we were reminded that development projects always take longer than anticipated. The original period of performance was June 1 but was extended to later in the fall 2018.

The Future

Closer to the implementation date, the CIP team will create a full training plan and provide user documentation. The CIP Program will provide hands-on training for internal staff and training webinars for publishers and partners. The CIP website will be updated, and the online publishers' manual will be completely rewritten. Publisher data and all pending CIP applications will be migrated; there will likely be a one or two-week work stoppage during this migration period. The CIP Program is excited for the future with PrePub Book Link! For more information, please visit the CIP Program web page at <http://www.loc.gov/publish/cip/>. 