

Notes on Operations

Bringing Publisher Metadata Directly to the Library

Use of ONIX at the Library of Congress

Karl Debus-López, David Williamson, Caroline Saccucci, and Camilla Williams

The library community is discussing ways to use metadata created at the beginning of the bibliographic supply chain to reduce costs associated with cataloging and remove redundant work between publishers and libraries. The ONIX standard holds promise because many of the data elements found within ONIX can be mapped to the MARC standard. The Library of Congress (LC) has developed an ONIX-to-MARC Converter that is being used to create MARC bibliographic descriptions directly from publisher-supplied ONIX metadata for new publications received through its Electronic Cataloging in Publication Program. This paper presents background information on ONIX, provides detailed information on how the ONIX-to-MARC Converter functions, presents findings of a test of the ONIX-to-MARC Converter, and discusses the pros and cons of using ONIX in the daily work of a large cataloging operation. Use of the ONIX-to-MARC Converter has reduced the time needed to create bibliographic descriptions, facilitated the inclusion of enriched metadata to bibliographic records, and provided the LC cataloging staff with records that are comparable to high-quality copy cataloging records.

As library budgets are cut and cataloging operations shrink throughout the United States, many library administrators are urging maximization of use of metadata created early in the bibliographic supply chain to remove redundant work and reduce costs associated with cataloging. This sentiment is clearly expressed in the 2008 *On the Record: Report of the Library of Congress Working Group on the Future of Bibliographic Control*, which presents a vision for management of metadata in the twenty-first century. The Working Group on the Future of Bibliographic Control represented a cross-section of libraries and organizations that are principal stakeholders in the future information environment. Membership included leaders from academic research libraries, U.S. national libraries, public libraries, a law library, the Special Libraries Association, Google, OCLC, Microsoft, and the Coalition for Networked Information. Many librarians and others who are concerned about the nation's bibliographic future provided extensive input to the findings of *On the Record*.¹

Of particular relevance to this paper is Finding 1.1 of *On the Record*: "Increase the Efficiency of Bibliographic Record Production and Maintenance."² The finding notes, "Until very recently, bibliographic control has been an artisan

Karl Debus-López (kdeb@loc.gov) is Chief of the U.S. General Division and Acting Chief of the U.S. and Publisher Liaison Division, **David Williamson** (dawi@loc.gov) is Acquisitions and Bibliographic Access Directorate Cataloging Automation Specialist, **Caroline Saccucci** (csus@loc.gov) is CIP Program Specialist, and **Camilla Williams** (cewi@loc.gov) is Head of the Dewey Section, all at the Library of Congress, Washington, D.C.

Submitted February 10, 2012; tentatively accepted March 15, 2012, pending modest revision; revision submitted April 6, 2012, and accepted for publication April 8, 2012.

activity, as there was no alternative for providing access except to transcribe, by hand, data from the objects being described. Now, however, publishers and vendors are working in an electronic environment, and print material generally originates in electronic format.³ Moreover,

publishers can provide some elements of descriptive metadata in electronic format for much of their output and libraries need to capitalize on those metadata. Despite the fact that descriptive metadata are being created in other venues, libraries have so far taken minimal advantage of them. Given the explosion of material requiring some level of bibliographic control, the model of item-by-item full manual transcription can no longer be sustained. Libraries must find ways to make use of the data created by others in the supply chain, including data that can be derived from algorithmic analyses of digital materials.⁴

The report further recommended that all participants in the bibliographic record supply chain “make use of more bibliographic data available earlier in the supply chain” and “be more flexible in accepting bibliographic data from others (e.g., publishers, foreign libraries) that do not conform precisely to U.S. library standards.”⁵ Specific recommendations to the Library of Congress (LC) were to “fully automate the Cataloging in Publication (CIP) process” and “develop content and format guidelines for submission of ONIX data to the CIP program and require publishers participating in the program to comply with these guidelines.”⁶

After receipt of *On the Record* at the LC, the associate librarian for Library Services, Deanna Marcum, convened the Implementation

Working Group to review all of the recommendations and suggest projects and activities the LC could pursue to test and eventually implement some of *On the Record's* recommendations. In September 2009, the Implementation Working Group published its report.⁷ One of the highest priority activities identified by the Working Group was to establish an ONIX pilot to “determine if use of ONIX data is feasible and provides efficiencies.”⁸ Because this work was previously recognized by Marcum as a high priority for the LC, the pilot test had already begun in June 2009.

The purpose of this paper is to summarize the findings of the test and subsequent work done by the LC using ONIX data at the very beginning of the bibliographic record creation process—when the LC creates pre-publication metadata for publishers that participate in the Cataloging in Publication (CIP) Program. Features of the LC ONIX-to-MARC Converter will be described; test findings also will be presented, as will the benefits and problems of using ONIX records.

While the information presented is specific to the LC, the use of ONIX data by the LC has an impact on national and international library communities as LC records become a source of copy cataloging for other libraries to use. If the records created through the LC ONIX-to-MARC Converter are bibliographically sufficient, other libraries will have little need to manipulate them after they become available to the larger community by vendors and bibliographic utilities such as OCLC. If this occurs, the community can move forward on another goal of the *On the Record*—to reduce unnecessary work. As *On the Record* states, “Redundant work means wasted resources. Time and money are spent redoing work that has already been done, rather than creating new records for materials not yet cataloged. This leads to delays in providing access to materials, and to users

being unable to locate materials that, though owned, are not yet accessible.”⁹ The work done by the LC is also transferable to other libraries and utilities that may have access to ONIX data from their own institutional presses or from publishers generally.

The ONIX Standard

What is ONIX? According to EDITEUR, the organization responsible for coordinating the development of the ONIX standard, “ONIX stands for ONline Information eXchange; it is an XML-based family of international standards intended to support computer-to-computer communication between parties involved in creating, distributing, licensing or otherwise making available intellectual property in published form, whether physical or digital.”¹⁰ The ONIX for Books standard is used by the book industry for sending and receiving bibliographic data in support of the book supply chain from the publisher to distributors to retail stores. Figure 1 is an example of an ONIX record provided by a publisher.

ONIX 1.0 was released in 2000 in the United States and the United Kingdom as a way to provide standardized product data in a consistent format, particularly to online retailers. ONIX 2.0 followed quickly in 2001, which provided increased capabilities for transmitting richer product data. ONIX 2.1, released in 2004, became the standard currently in use by many U.S. publishers. It has remained stable since its release. The ONIX 2.x versions are both backwardly compatible with ONIX 1.0, but that eventually caused problems and confusion as increased capabilities conflicted with the original standard, leading to the development and release of ONIX 3.0 in 2009. ONIX 3.0 is not backwardly compatible with previous versions and, while the bulk of the standard is based on ONIX 2.1, some new elements

```

<product>
I.<a001>9781107020047</a001>
<a002>02</a002>
<a197>Cambridge University Press</a197>
<productidentifier>
<b221>03</b221>
<b244>9781107020047</b244>
</productidentifier>
<productidentifier>
<b221>15</b221>
<b244>9781107020047</b244>
</productidentifier>
<b246>02</b246>
<b012>BB</b012>
<n338 />
<title>
<b202>01</b202>
<b203>Economic Reform in India</b203>
<b029>Challenges, Prospects, and Lessons</b029>
</title>
<contributor>
<b034>1</b034>
<b035>B01</b035>
<b036>Nicholas Hope</b036>
<b037>Hope, Nicholas</b037>
<b039>Nicholas</b039>
<b040>Hope</b040>
<b044>Nicholas C. Hope has served as Director of the Stanford Center for International Development at Stanford University since 2007; from 2000 to 2007 he was the Center's Deputy Director. Dr Hope joined the world Bank in Washington DC in 1977 and held positions through 2000 as chief of the external debt division, chief of the industry and energy division for Indonesia and the Pacific Islands, country director for China and Mongolia, resident director in Indonesia, and director in the office of the vice president for Europe and Central Asia. He co-edited two volumes in 2003 and 2008 on Chinese policy reform. Dr Hope received his PhD from Princeton University.</b044>
</contributor>
<contributor>
<b034>2</b034>
<b035>B01</b035>
<b036>Anjini Kochar</b036>
<b037>Kochar, Anjini</b037>
<b039>Anjini</b039>
<b040>Kochar</b040>
<b044>Anjini Kochar is Senior Research Scholar and has served as India Program Director at the Stanford Center for International Development since 2000. From 1990 to 2000 she was an assistant professor in the department of economics at Stanford University. Her research has concentrated on poverty and rural development in India, and has appeared in the Journal of Political Economy, The Review of Economics and Statistics and The American Economic Review among other publications. Dr Kochar is also an associate editor of Economic Development and Cultural Change. Her PhD in economics was awarded by the University of Chicago.</b044>

```

Figure 1 Partial ONIX Record

in version 3.0 are not in version 2.1. While version 3.0 has been available for about three years, its adoption has been very slow with most ONIX distributors or receivers still using version 2.1. Until the book industry and other users of ONIX have a need for more of the features in ONIX 3, U.S. publishers have indicated they will stay with 2.1.¹¹

As noted, the ONIX standard was first developed in the United States and the United Kingdom as a product of the Association of American Publishers and EDItEUR. Development expanded to include the Book Industry Study Group (BISG) in the United States and Book Industry Communication (BIC) in the United Kingdom. Since then, fifteen other countries began participation in the development of the standard. Now development is overseen by the ONIX

International Steering Committee with EDItEUR responsible for coordinating all of the various country user groups. BISG, through its Book Industry Standards and Communication (BISAC) arm, assigned the BISAC Metadata Committee to be the user group in the United States responsible for participating in the development of the ONIX standard. In addition to the U.S. and U.K. user groups, additional groups in are in Australia, Belgium, China, Canada, Egypt, Finland, France, Germany, Italy, Japan, South Korea, the Netherlands, Norway, Russia, Spain, and Sweden. Additionally, ONIX work is being conducted in Bulgaria, Poland, and Turkey; although no official user groups are in place yet.¹² The particular focus of this paper is on the ONIX for Books standard. Other formats use ONIX, but they will not be addressed in this paper because the

LC has only implemented ONIX for Books within its workflows.

Literature Review

Two articles by different authors published within *Publishing Research Quarterly* in 2002 and 2004 begin with the statement "It is a proven fact that the more information customers have about a book, the more likely they are to buy it."¹³ The fact that both authors from the publishing industry could begin their analyses of the utility of the ONIX standard format with the same statement is an indication of a consensus on the benefit of sharing as much information about a book as broadly as possible. In the first article, Daly, executive director of the Book Industry Study Group at that time, stated that "ONIX was developed as a solution

to two modern problems: a) the need for richer book data online; and b) the widely varying format requirements of the major book wholesalers and retailers.¹⁴ While ONIX has been used from its release to describe elements of print books, very early in its development ONIX was deemed the “ideal standard to transmit metadata about e-books. In addition many of the retailers and wholesalers in the industry sell all forms of media. ONIX could provide a platform for the transmission of metadata for all types of information and entertainment products.”¹⁵

By 2004 Beky reported that, while small publishers still did not have the resources to convert their metadata to ONIX, the standard “has been adopted by all major U.S. publishing houses, which together produce approximately 50 percent of all trade book titles.”¹⁶ That number has continued to grow during the last eight years, and services to create ONIX records have been developed by companies such as Firebrand Technologies and NetRead to assist smaller publishers who do not have the capability to work in the ONIX environment. By 2010, the primary publishers’ trade journal, *Publishers Weekly*, was reporting on the importance of metadata to the publishing community. A leading consultant in digital publishing services reported “accurate metadata has become a marketing tool for publishers, a shopping guide for consumers, and an absolute necessity for distributors and retailers.”¹⁷ ONIX was seen as central to these marketing efforts. The statement of 2002 “that the more information customers have about a book, the more likely they are to buy it” had morphed into “accurate, rich metadata sells books” by 2010.¹⁸ To handle the creation of this metadata for publishers that do not have adequate resources, a cottage industry of metadata producers has developed, including companies such as Firebrand Technologies, NetRead, and, most important to the bibliographic record supply chain in

the library community, OCLC.

An understanding of the promise and value of ONIX exists not only within the publishing industry, but within the library community as well. As noted above, *On the Record* fully embraced the need to use publisher created metadata—and specifically, ONIX metadata—as broadly and as effectively as possible. But even before *On the Record* was published, the library community recognized the value in ONIX metadata. Within the library trade magazine *Library Journal*, Tennant noted in 2006 that

Publishers are increasingly supplying machine-readable metadata about the publications they put out—largely to enable their books to be sold to Amazon and other online booksellers. These records could provide much enriching information to our existing MARC data if the infrastructure were in place to normalize the records. Publishers often provide cover art, pull quotes from reviews, descriptive text, author biographies, and other useful material that MARC records typically lack. . . . How do I know this? I walk around with over 10,000 ONIX metadata records on my laptop that I downloaded from willing publishers. If we had a service to collect these records from publishers and make them available to catalogers, we could have access to many valuable facts about library materials.¹⁹

OCLC saw the value of ONIX as well. The availability of publisher data in ONIX provided an opportunity “to break down traditional silos between library and publisher supply chain metadata.”²⁰ OCLC created a Next Generation Cataloging pilot to improve the interoperability of publisher and

library data, adding value to both by leveraging the strengths of each. In 2009, the National Information Standards Organization (NISO) and OCLC solicited a white paper titled *Streamlining Book Metadata Workflow* from Informed Strategies to provide an industry overview of producers and stakeholders of bibliographic metadata.²¹ The first idea proposed by Luther, the author of the white paper, was to “use crosswalks between ONIX and MARC to facilitate the creation of CIP and to provide publishers with an XML feed of MARC data.”²² OCLC has moved forward on creating a crosswalk between ONIX and MARC, which is the foundation of their new OCLC Metadata Services for Publishers.²³ This fee-based service accepts titles from publishers in ONIX format and enhances them for publisher use. OCLC describes five principal benefits of use of this service by the publishing community:

1. Reduces cost and duplication of effort in bibliographic description, categorization, and name authority work.
2. More titles found = more titles sold.
3. Provides richer marketing data to support buying decisions for wholesalers, retailers, libraries, and end users
4. Adds and enhances data to support marketing, sales-analysis, and business-intelligence needs for multiple markets.
5. Opens additional channels for exposure of title metadata—for use in library workflows and to end users on the web.²⁴

In the United States, OCLC and the LC are the principal institutions working with an ONIX-to-MARC conversion program, although other organizations can map ONIX to MARC. The LC, through its management of the CIP Program and its corresponding relationship with more than 5,100

major U.S. publishers and imprints, and OCLC, with its strong penetration in the worldwide bibliographic environment (including strong connections with the publishing community), have the greatest ability to maximize use of ONIX metadata for the library community worldwide. Use of ONIX by the library community is relatively new and as such needs to be studied in more depth. Stalberg and Cronin suggest “with several concrete ONIX-MARC projects underway, analysis can now be done to determine the extent to which ONIX data are valuable for cataloging workflows.”²⁵ The information provided in this paper is a first step toward providing information on the LC’s use of ONIX and its impact on cataloging workflows. As the LC moves its ONIX-to-MARC Converter program into full production later this year, more analysis will be done and shared with the broader library community.

Setting

The LC is the world’s largest library with more than 151 million items, including more than 34.5 million cataloged books and other print materials in its collections.²⁶ In FY11, LC staff cataloged more than 363,000 new titles.²⁷ Of that total, 105,000 new titles were cataloged by the two divisions principally responsible for the U.S. national imprint collection within the Acquisitions and Bibliographic Access Directorate—the U.S. General (USGEN) and U.S. and Publisher Liaison (USPL) Divisions. Nearly all titles cataloged by the two divisions represent new monographic publications received from either the LC Copyright Office or the CIP Program. More than 51,000—almost half—of the new titles were processed through the CIP Program.²⁸ The rest were received from the Copyright Office. The bibliographic records created by the staff in the two divisions are

distributed to OCLC and through other means, making them readily available to researchers and the public, thereby saving libraries of all types the expense of duplicating this effort.

The CIP Program has been in existence since 1971. Its mission has remained the same: to provide cataloging data to libraries before publication, thus saving the libraries the cost of cataloging and supporting other library functions, such as acquisitions.²⁹ However, without the continued support of the U.S. publishing community, the program would have ceased long ago. The CIP metadata created by the LC represents the “accurate, rich metadata” mentioned by Reid as being essential for enhanced sales for publishers.³⁰ Currently, more than 5,100 publishers and imprints participate in the CIP Program. Their titles represent the cream of the crop of U.S. publications. Over 95 percent of titles received through the CIP Program are retained for the LC’s permanent collections. At the end of FY11, LC staff had cataloged more than 1.5 million books received through the CIP Program since 1971.³¹

Because the librarians in USGEN and USPL are working with new publications received either from the CIP Program or the Copyright Office, close to 80 percent of the cataloging done within the divisions is original work.³² The high percentage of original work is one of the reasons why the LC has been so interested in implementing the ONIX-to-MARC Converter program. The converter will (hopefully) reduce the amount of time spent on bibliographic description of new titles so staff can focus on subject analysis or work on special collections or more unique materials.

Early Use of ONIX at the Library of Congress

As early as 1996, four years before the creation of ONIX, Williamson and

Davis-Brown of the LC had noted “experiments involving Standard Generalized Markup Language (SGML) have demonstrated that bibliographic records can be created directly from electronic texts with little operator intervention. If a text were marked up to the MARC subfield level, a program could scan the text automatically and extract all of the data.”³³ This was an early prediction of how a standard like ONIX cross-walked with MARC could benefit the library community by reducing the amount of manual creation and input of data elements when creating bibliographic records.

The LC was an early adopter of use of ONIX data, having received ONIX data since April 2002. As the use and production of ONIX data increased, so did the number of available sources of ONIX data. Today the LC receives data directly from publishers as well as from data aggregators. Depending on the capabilities of the data supplier, ONIX data files are received containing daily, weekly, or monthly updates, with some suppliers providing occasional “full file” data files of all of their items available, as well as “delta files,” which contain only those changes made since the last file. In an average week, the LC will receive approximately 200 data files representing ONIX records from thousands of imprints and tens of thousands of individual items in the book supply chain. Many of these are update records containing changes that are of little interest to the LC, such as price changes or availability information. However, many records are of interest to various projects at the LC.

Individual ONIX records can contain a wealth of information. Elements such as author and contributor information, titles, editions, imprints, publishing dates, extent, and series may be available and can be mapped to MARC fields. Additionally, information not regularly included in a MARC record, such as summaries, tables of contents, and BISAC subject codes, may be present and can be mapped

to MARC fields. Many other fields may be present that are not used in MARC records but are used in the book supply chain. These include rights information, accompanying material information, related material information, author biographies, websites, awards, affiliations, sample texts, and much more that may be displayed on the website for an online retailer.

One of the first ONIX-related projects the LC embarked on was the ONIX-TOC (tables of contents) application to extract TOCs from ONIX records and automatically link the information in the 856 (electronic location and access) MARC field. The LC enhanced hundreds of thousands of bibliographic records through this mechanism. In FY11, 17,714 records were enhanced with TOC information received via the ONIX-to-MARC Converter program. Implementation costs have been very low. In 2006, Byrum and Williamson found that the cost of manually adding a typical TOC note was about \$40 per record while automatic addition of the note via the ONIX-TOC process was \$0.80 or less per record.³⁴ They further found that

The ONIX costs vary depending on the size of the data file received and how many new matches can be extracted from that file. The costs to set up the processing are about eight dollars. . . . Once the program is running unattended, the number of successful new TOC files created, determines the cost. If ten new TOC files are created that's about \$0.80, if one hundred are created, the cost drops to \$0.08, and if one thousand or more are processed, the cost is less than one cent per TOC for accomplishing extraction and linking.³⁵

The success of the TOC project was an early indication that

enhanced use of ONIX at the LC could potentially provide even greater financial benefits to the library while providing enriched data to the library community.

Following the success of the TOC Program, the LC implemented a Publisher Provided Summary Program. This Program allows publishers to voluntarily add summaries to their ECIP application and extracts summaries directly from the ONIX data. The ONIX summary information is linked in the 856 MARC field, while any summary from the ECIP application is input in the 520 (summary) MARC field in the bibliographic record. In FY11, 8,303 summaries were included within ECIP records—a 44 percent increase over FY10's 5,783 summaries. A wide range of publishers provide summaries, including children's publishers, university presses, religious publishers, and popular presses. The LC recently expanded the program to include juvenile fiction publishers. At the close of FY11, 32,504 summaries have been provided by publishers and added to the fully cataloged CIP bibliographic records.³⁶ Enriching bibliographic records through the inclusion of tables of contents and summaries has been found to assist the user by providing more terms for retrieval or relevant titles.³⁷

The ONIX-to-MARC Converter Program

The latest project to use ONIX data at the LC is also part of the Electronic Cataloging in Publication (ECIP) Program. Because ONIX data contain several of the same bibliographic elements needed in a preliminary ECIP record, a project began in June 2009 to look into using ONIX as the basis for an ECIP record, determine the quality of the ONIX data and its conversion, and see if the conversion would help decrease processing time. The LC developed an application to process

the ONIX files that are received to select any prepublication records in those files and create or update them in a database of prepublication ONIX records. This database resides on a server that is accessible to all catalogers who process ECIPs.

A second application also was needed to search the database and perform the ONIX to MARC conversion. Currently in the normal ECIP workflow, when a cataloger is ready to create an ECIP record, he or she clicks on a link in the web-based ECIP application form that starts an application created in the 1990s when the LC began processing CIP applications electronically. In this project, the link starts a different application that scans the ECIP form for the ISBN of the item. The application then searches the database of prepublication ONIX records, looking for a match. If no match is found, the new application calls up the old application and the cataloger processes the ECIP as before. If a match is found, however, the associated ONIX record is retrieved from the database. A skeletal MARC record template is hard-coded in the application, which goes through the ECIP form looking for needed elements, such as additional ISBNs, contact information, and place of publication (not in ONIX records), and then adds those to the skeletal template so that a MARC record begins to take shape within the application's memory.

The application then goes through the ONIX record looking for the various fields that can be used, starting with any authors and contributors and creating the needed 100 (personal name main entry) or 700 (personal name added entry) MARC fields for those individuals. The converter only accepts personal names because only personal names appear in the ONIX files. The ONIX record contains information identifying the order in which any authors and contributors are listed as well as the function that person had in relation to the publication; this

allows the application to determine who is the person to put in the 100 field (if needed) and who goes in a 700 field and in what order. Some one hundred or so functions, such as authors, editors, compilers, arrangers, adaptors, illustrators, actors, composers, etc., are defined in the ONIX standard. The application takes each name and prepares it for inclusion in the statement of responsibility (SOR). For authors, the names are simply held in direct order for inclusion in the SOR. For other contributors, the name of the function as provided in the ONIX standard list for contributor functions is added before the name in square brackets because the application needs to inform the cataloger that the person is not an author but it does not know the exact SOR wording in the galley mock-up. For example, a name identified as an editor would be processed to show “[edited by] John Smith” in the SOR.

When the application searches the ONIX record for the title, it extracts the title, any subtitle, and, remembering the authors and contributors and their function, it creates an SOR and puts them all together to create the 245 (title statement) field. Additionally, because the catalogers discovered that the ONIX data provided did not always match what came on the galley mock-up provided by the publisher, the applications displays the title page and copyright page from the galley mock-up as well as the proposed 245 field. That way the cataloger can compare them and note if any adjustments to the title or SOR are needed. If the cataloger notices a serious problem, such as the title on the ECIP form not matching the title from the ONIX data, the cataloger can opt to use the old application to process the ECIP. If it looks like the correct title (with minor variations), the cataloger notes any differences that need to be looked at after the ONIX conversion has been completed. The cataloger then clicks a button and the application then

creates the 245 field data to be added to the record being constructed and then extracts information for the 250 (edition statement) field, if available. Again, a display provides the edition information found for the cataloger to compare against the title page and copyright page from the galley mock-up, and the cataloger notes any differences.

Another click on the same button and 250 field data are created, the information for the 260 field is extracted, and another display comes up for comparison. This time, however, when the cataloger clicks the button, the application asks the cataloger to supply the place of publication for \$a of the MARC 260 (publication, distribution, etc.) field and the equivalent place code in the 008 (fixed-length data elements) field because the place is not provided in the ONIX data. Finally, information for any series is extracted, if available, and the final display comes up for comparison. When the cataloger clicks the button this time, the application goes through the ONIX record and extracts the extras that will be added to the record, such as a link to a cover image, a summary, or a TOC; those are converted into their appropriate MARC fields and added to the record to be constructed.

Regarding the TOC, the information extracted from the ONIX record is not manipulated to try to make it conform to *Anglo-American Cataloguing Rules*, 2nd ed. (AACR2) specifications.³⁸ Instead, the data are provided more or less as given in the ONIX record. If elements are separated one per line in the ONIX record, then the application will insert two hyphens between the elements, but if the elements are strung together in one long data string, it is given as found in the 505 field. In this case, the publisher usually separates the elements with punctuation and the presence of “Chapter” or “Part” to distinguish each separate element in the TOC. The first indicator value “8” is given to indicate

that no display constant will be generated by an online public access catalog or other MARC display system, and the data are then given.

For summaries, the wording is presented exactly as found in the ONIX record. The entire summary is quoted and at the end of the field “—Provided by publisher” is given to indicate the source of the quoted summary. The only manipulation the application will perform is to convert certain symbols or punctuation marks from an HTML equivalent or a Unicode value to something that can be reproduced on a keyboard, e.g., the symbol for “less than” can be represented in HTML as “<” and the application simply converts it to “<.”

The final task for the application to do is to ask the cataloger to input his or her cataloging code to put into an internal work-tracking field. The application then puts all other data elements in their correct MARC fields and writes the MARC record to a work file, calls up the cataloging client, and sends keystroke commands to the client window causing the client to import the MARC record from the workfile. The cataloger then is presented with the MARC record in the client where he or she can review it, making any of the changes noted during the conversion process.

Use of the ONIX-to-MARC Converter program to create ECIP bibliographic records has grown exponentially in the past two and a half years. In FY09, 532 ECIPs were added through use of the ONIX-to-MARC Converter program. In FY10, 2,810 records were added through the program, and in FY11, 499 records were added—an increase of 202 percent from the previous year.³⁹ The LC is on the cusp of putting its ONIX-to-MARC Converter into production later this year once a systems upgrade is completed. At that point, all of the LC catalogers responsible for processing ECIP galleys will be fully trained on use of the ONIX-to-MARC Converter.

ONIX-to-MARC Pilot Test

The ONIX-to-MARC test that evaluated use of the ONIX-to-MARC Converter process described in the above section began in June 2009 and ended in August 2009. To fully understand the results of the test, one needs to know about the existing system used to create ECIP prepublication metadata.

Text Capture and Electronic Conversion Program

The text capture and electronic conversion (TCEC) software program is currently known as “On the MARC” and has been the traditional method for LC staff to process an electronic galley in an ECIP application. Before submitting an ECIP application, the publisher attaches an ASCII text file that includes the title page, copyright page, series page, TOC, and chapters with each section tagged to enable TCEC to function properly. The LC prefers the full text of the galley to assist catalogers in subject analysis; however, the publisher may submit just the core text if a useful summary is included in the application. Publishers are obligated to send more text if the cataloger cannot make a determination of the proper subject analysis with only this core text.

TCEC shows up as a link at the bottom of the ECIP application page, commonly referred to as the “data view.” When the cataloger clicks on that link, a new window is presented in a split screen with the contents of the text file exactly as they appear in the galley in the window above and a work screen below. On the left side of the screen are MARC field tags for the areas of the bibliographic record that can be processed with TCEC. The field tags are arranged in numerical order, although many of them can be manipulated out of order. The ISBN and any other qualifier are captured from the data view and are automatically inserted into a 020 (ISBN) field at

the top of the work screen. If multiple ISBNs are included in the application, they are converted to multiple 020 fields.

The cataloger then begins to select the text in the galley view screen for the title, subtitle, and statement of responsibility. The cataloger usually needs to rearrange some of the text to include *International Standard Bibliographic Description for Single Volume and Multi-Volume Monographic Publications (ISBD)* punctuation.⁴⁰ If a title and a subtitle are present, a space colon space (i.e., “:_:”) must be between those elements. The statement of responsibility must have a space forward-slash space (i.e., “_/”) between it and the previous element. After that forward slash, commas and semicolons can be added with appropriate spacing as well. If a parallel title is present, the equal sign must be properly spaced between the elements.

Once the elements are in the right order with ISBD punctuation, the cataloger highlights the text to select it and then clicks on the 245 field tag. This generates a MARC-coded field with all indicators and subfield coding. The cataloger can choose to select a name for the main entry by highlighting a name and selecting the 100 tag. This changes the field tag coding to 100/245, and the 245 first indicator will be changed from 0 to 1 to reflect title added entry. Added entry fields (700s) can now be added. TCEC works best with access points for personal name because it includes a field tag only for 100 and 700 and inverts the name according to cataloging rules.

Because most words in the title should not be capitalized according to AACR2, every word after the first word occurring before the forward slash is automatically lowercased and any word after the forward slash is presented as-is. After text appears in the work screen, the cataloger can change any capitalization. This can be done easily by rolling over the letter or letters with the mouse or by highlighting the text

and selecting from a dropdown menu to uppercase or lowercase the letters. All the fields can be manipulated in a similar way. The cataloger also can add any additional subfield coding or any notes or other text to the record.

The TOC also can be included in the bibliographic record because TCEC will do most of the work of stripping chapter and numerical designations, leaving the text of the chapter headings. The cataloger usually needs to do some manipulation of the text and fix any capitalization errors. Catalogers are encouraged to include the TOC if the process will take no more than five minutes.

Once all the descriptive elements of the bibliographic record are in the work screen, the cataloger clicks the send button to convert the TCEC record into an integrated library system (ILS) record. The process is instantaneous and some local fields are automatically added. The cataloger can now proceed to do any authority work and subject analysis, adding appropriate subject headings and LC classification and shelving. The Dewey decimal number is applied to the record at the last stage. The bibliographic record is then ready to be sent back to the publisher as CIP data that can be printed on the copyright page. Figure 2 illustrates a completed TCEC record that is ready to be loaded into the ILS.

ONIX Process

While TCEC has worked effectively for cataloging ECIPs, the above description makes clear that it can be cumbersome to manipulate, punctuate, and highlight text. This is where the ONIX-to-MARC Converter is so beneficial. The data are automatically preselected, so no text manipulation is required at this point. As noted above, a compare mode allows the cataloger to note any discrepancies between the ONIX data and the galley to fix once the record is in the ILS, but the process is more akin to proofreading.

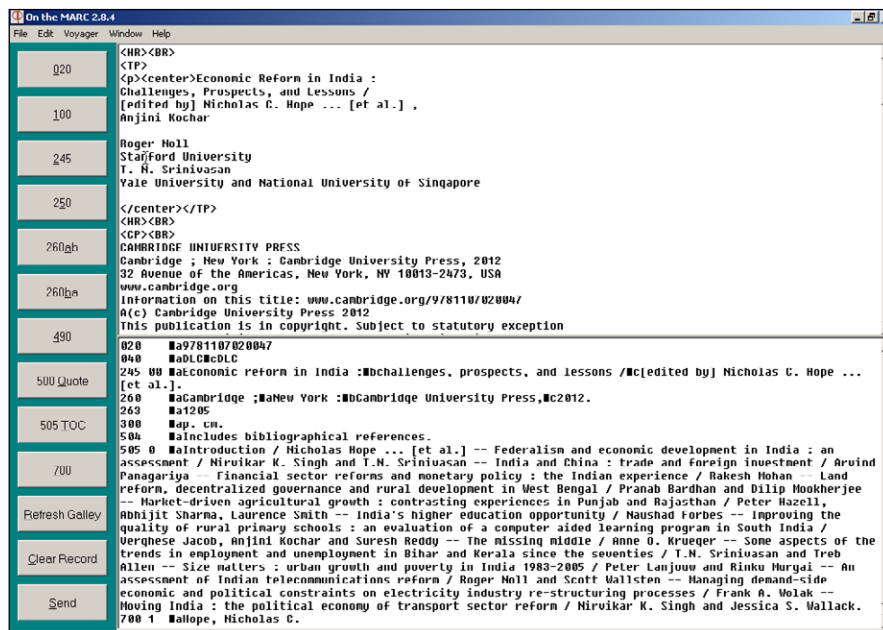


Figure 2. Completed TCEC Record

Figure 3 illustrates the compare mode at the title level. The compiled title and statement responsibility as they will appear in the MARC record are shown at the top of the page. The rest of the page shows the galley view so that the cataloger can note any discrepancies between the ONIX data and the galley.

Once the ONIX data have been converted to the MARC record (see figure 4), the cataloger can begin to manipulate the ONIX-generated data. Any discrepancies that the cataloger noted in compare mode can now be modified. The cataloger always double-checks the galley to make sure the bibliographic data accurately describe the galley. Any capitalization errors must be corrected manually because the ILS has no highlighting short cuts. The cataloger also now can add any additional fields or text.

ONIX Pilot Test Evaluation Process

The ONIX Pilot test was a test of both the utility of the ONIX records

as a source for accurate bibliographic description and a comparison of the TCEC process to the new ONIX-to-MARC Converter process to determine which was most effective to use. The pilot began in June 2009 with two CIP program specialists who performed the descriptive cataloging and tested ONIX data from two publishers, Cambridge University Press and Wiley. The pilot evaluated several aspects:

- The availability of ONIX data for items in the CIP stream
- The usefulness of the data in cataloging
- Any problems or unexpected results from converting the data from ONIX to MARC
- Time comparison between TCEC and the ONIX-to-MARC Converter process

As the project progressed and to evaluate more data, the ONIX/MARC Conversion Committee added the imprints of Harper Collins, Palgrave Macmillan, and Oxford University Press to the pilot. In July 2009, the

pilot team was joined by the National Library of Medicine (NLM). The findings of the test were shared with Acquisitions and Bibliographic Access (ABA) directorate director Beacher Wiggins in August 2009. Based on the positive results described below, Wiggins decided to continue the pilot.⁴¹ The pilot project team also was expanded to include two more LC testers in December 2009.

In October 2010 the pilot was again expanded. Each monographic section that processed ECIP applications within the ABA Directorate identified at least one staff member to work with the ONIX-to-MARC Converter. This increased the number of testers by twelve and the number of publishers was expanded to include all that might provide data in ONIX format (more than 5,100 imprints).

Availability of ONIX records

The test uncovered inconsistencies in the quantity of the ONIX data received from the publishers. For example, the results for Wiley showed that only 64 of 274, approximately 23 percent, of the ECIP applications submitted by Wiley and its imprint, Jossey-Bass, were in the ONIX database. In addition to this low hit rate, the Wiley ONIX files were problematic. Some of the problems encountered included the following:

- They did not always contain summaries or TOCs
- The complete number of authors and editors was different from the galley
- The titles, series, publisher, and publishing dates did not match the galley

Because of these disappointing results and problems, the committee decided not to continue with Wiley in the ONIX pilot project. David Williamson, ABA cataloging automation specialist, spoke with representatives

from Wiley about the discrepancies between their ONIX data and their actual galleys. Since then, the discrepancies have diminished and, with the subsequent expansion of ONIX to all of the U.S. imprint monographic cataloging units in ABA, Wiley and Jossey-Bass were returned to the pilot.

The results for Cambridge University Press were more promising. During the initial test period, fifty-three of eighty-eight (approximately 60 percent of Cambridge titles) were received with ONIX data. The content of the ONIX records tended to be accurate and very few changes were made to the bibliographic records. The files generally included summaries and TOCs. The Cambridge results also confirmed that the use of ONIX data is efficient and effective when processing ECIP applications.

Benefits of Using ONIX

Several benefits result from using publisher-supplied ONIX data in cataloging ECIP applications. Use of ONIX data is faster and more ergonomic for catalogers. Very little keying and usually little data manipulation are required when working with ONIX records as compared with regular TCEC descriptive cataloging. Although TCEC eliminated a large need for keying information into the bibliographic record, ONIX data are readily available by the publisher and largely formatted when the record is imported into the LC's ILS.

Use of ONIX has streamlined LC operations because the staffs in USPL and USGEN are able to utilize the ONIX information much like they use copy cataloging records available from other sources. Most, if not all, of the information needed for the descriptive elements of the title are available within the ONIX record.

Just as important, and of more significance to the users of records created by LC staff, is that the ONIX records have additional data elements

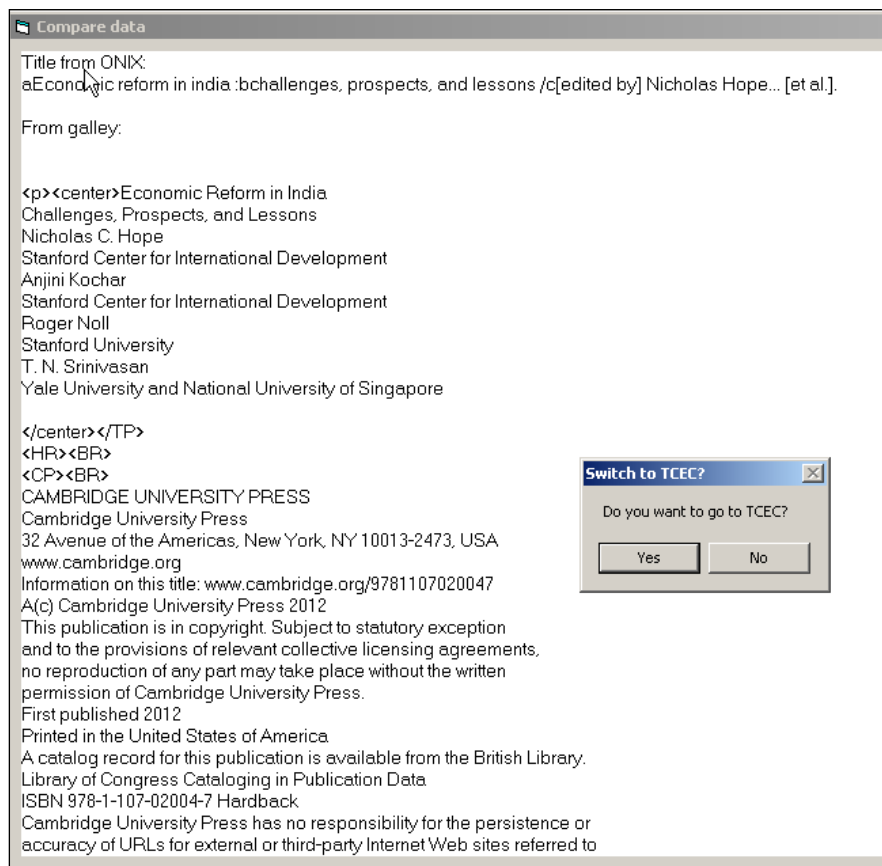


Figure 3. ONIX Compare Mode at the Title Level

that provide more information and access points, such as TOCs, summaries, and BISAC terms. Indeed, one of the great advantages of ONIX data is the rich keyword access provided in the often-lengthy summary statements and TOCs. As previously noted, these fields provide added value to researchers. A summary often is included with the ECIP application and is captured in the ONIX-to-MARC Converter, resulting in two summaries. If they are duplicative, the cataloger will delete one of them; if they are different, the cataloger usually will keep both. Sometimes only either an ONIX or an ECIP application summary is provided; occasionally neither is provided. For ONIX-provided TOCs, the field includes the following disclaimer: "Machine generated contents note." With this disclaimer, the TOC can include additional data that otherwise

would not be included, such as part and subchapter titles; catalogers generally do not review these TOCs for accuracy. With TCEC, the TOCs must be manipulated manually, so having this extra data already included is advantageous. Because the LC utilizes ONIX data in bibliographic records, those libraries that download these records into their ILS also will benefit from the publisher supplied ONIX data.

ONIX Problems Encountered

While the LC has determined an overall benefit to using the ONIX-to-MARC Converter program and intends to put it into full production in early 2013, a number of problems surfaced during the testing of the converter. While these problems do not occur frequently enough to prevent the LC from moving forward on

Tag	11	12	Subfield Data
010			#\$ 2012005153
020			#\$ 9781107020047 (hardback)
040			#\$ DLC \$c DLC
042			#\$ pcc
084			#\$ BUS008000 \$2 bisacsh
245	0	0	#\$ Economic reform in india : \$b challenges, prospects, and lessons / \$c [edited by] Nicholas Hope... [et al.]
280			\$a New York : \$b Cambridge University Press, \$c 2012.
263			#\$ 1205
300			\$a p. cm.
520			#\$ *The essays in this volume are written by leading economists working on the Indian economy. They collectively emphasize the importance of policies and institutions for sustained growth and poverty reduction, stressing that the success of sector-specific policies is vitally dependent on the nature of markets and the functioning of institutions such as those charged with regulating and overseeing critical sectors. Individual contributions assess the role of Indian government policy in several key sectors and emphasize the policies required to ensure improvements in these sectors. The first section discusses aspects of the macro economy; the second deals with agriculture and social sectors; the third with jobs and how labor markets function in agriculture, industry and services; and the fourth with infrastructure services, specifically those of electricity, telecommunications, and transport. The essays are drawn from the most influential papers presented in recent years on Indian economic policy at the Stanford Center for International Development" \$c Provided by publisher.
505	8		#\$ Machine generated contents note: 1. Introduction Nicholas C. Hope, Anjini Kochar, Roger Noll and T. N. Srinivasan; Part I. The Macro Economy; 2. Federalism and economic development in India: an assessment Nirvikar K. Singh and T. N. Srinivasan; 3. India and China: trade and foreign investment Arvind Panagariya; 4. Financial sector reforms and monetary policy: the Indian experience Rakesh Mohan; Part II. Institutional Reforms: Agriculture and Education: 5. Land reform, decentralized governance and rural development in West Bengal Pranab Bardhan and Dilip Mookherjee; 6. Market-driven agricultural growth: contrasting experiences in Punjab and Rajasthan Peter Hazell, Abhijit Sharma and Laurence Smith; 7. India's higher education opportunity Naushad Forbes; 8. Improving the quality of rural primary schools: an evaluation of a computer aided learning program in south India Verghese Jacob, Anjini Kochar and Suresh Reddy; Part III. Employment, Industrial Structure and Poverty: 9. The missing middle Anne O. Krueger; 10. Some aspects of the trends in employment and unemployment in Bihar and Kerala since the seventies T. N. Srinivasan and Treb Allen; 11. Size matters: urban growth and poverty in India 1983-2005 Peter Larjouw and Rinku Murgai; Part IV. Infrastructure: Electricity and Transportation: 12. An assessment of Indian telecommunications reform Roger Noll and Scott Wallsten; 13. Managing demand-side economic and political constraints on electricity industry restructuring process Frank A. Wolak; 14. Moving India: the political economy of transport sector reform Nirvikar K. Singh and Jessica S. Wallack.
650	7		\$a BUSINESS & ECONOMICS / Development / Economic Development. \$2 bisacsh
700	1		#\$ Hope, Nicholas.

Figure 4. ONIX Record Converted to MARC

its plan to use ONIX metadata more extensively, documenting them for the record is important.

Catalogers have found that about 60 percent of the time, no ONIX data for a particular ECIP is present even though, according to the BISG Product Metadata Best Practices for Data Senders, data should be provided at least six months before publication.⁴² As mentioned earlier, the ONIX program searches a database of ONIX records with a matching ISBN. If it does not find a corresponding ONIX record, the normal TCEC screen will immediately open, incurring no lost cataloging time.

Discrepancies between the ONIX data and the galley may be present. Words in the titles, subtitles, statements of responsibility, and series can be missing or different. The date of publication or even the name of the publisher in the ONIX-generated data are often different from the projected date of publication or associated imprint in the ECIP application or the galley. Sometimes the series statement appears as the title proper, differs

between ONIX and galley, or is simply missing in the ONIX data. The series numbering may be included in the ONIX data but appears nowhere in the galley. These kinds of errors occur about 40 percent of the time and are relatively easy to fix.

More complicated discrepancies occur when the number of authors or editors differs between the ONIX and galley versions. This can result in having to recatalog the ECIP. If the ONIX data list three authors, the record will be presented with author main entry and two author added entries. However, if the galley presents four authors, then the record has to be converted to a title main entry with an added entry for the first-named author added; the names of the additional authors in the statement of responsibility and the corresponding added entries also must be deleted. Fortunately, this kind of error is relatively rare, occurring perhaps 10 percent of the time. Sometimes an author is incorrectly presented as a main entry or as an added entry (e.g., an editor is coded as an author) because the ONIX data were incorrectly coded

at the outset, a more likely occurrence encountered 20 percent of the time.

Special characters and diacritics frequently occur in cataloging. Diacritics in the ONIX data usually convert to the MARC record without a problem. The ONIX diacritics and special characters are in Unicode or HTML coding, which can cause problems in the conversion process. When these diacritics and special characters do not properly convert, the ILS may not allow the bibliographic record to be saved to the local database. More often than not, these errors appear in the summaries and machine-generated tables of contents and can be difficult to locate. If a number of these errors are present, the cataloger has to decide whether to look for and fix all the errors or to delete the whole field. These types of errors have dramatically reduced over time and, with the next upgrade of LC's ILS, the number will further reduce as fixes are added to the application.

Time Comparison Study

As part of the ONIX Pilot Project, the CIP Program Specialists performed a time comparison study to document the time needed to catalog records with ONIX data to the TCEC software used to create bibliographic records from electronic galleys received from publishers. The purpose was to determine how much faster and efficient descriptive cataloging with ONIX data is compared to descriptive cataloging with the TCEC software. This study began in week ten of the ONIX Pilot Project and was completed in week fifteen. The study did not include time needed to search and establish or update authority records.

ONIX Results

The time study showed that from fifteen seconds to five minutes were required to create descriptive elements for an ECIP galley. Of the 130

ECIP galleys completed during the five-week period, one hundred (80 percent) required little editing and were completed in one minute or less. The remaining thirty had various problems previously identified by the ONIX/MARC Conversion Committee, such as

- title, author, and contributor that differed between ONIX data and ECIP galley;
- subtitles, statements of responsibility, and series statements that appeared in neither the ONIX data nor the galley;
- HTML coding in 520 and 505 fields;
- symbols, usually an opening apostrophe (i.e.,'), in the 520 and 505 fields that prevented the bibliographic record from being saved to the ILS; and
- significant amount of text that require proper capitalization.

TCEC Results

The results of the time study showed that the times for the ECIP galleys cataloged using the TCEC software ranged from thirty seconds to four minutes. Approximately seventy ECIP galleys were processed using TCEC during this five-week period. Fifty of the seventy (80 percent) ECIP galleys processed using the TCEC software were completed in two minutes or less. These were ECIP galleys that included primarily authors and editors, titles, summaries, and tables of contents that required very little, if any, editing.

Those ECIP galleys with processing times of more than two minutes usually required considerable editing. Some of the problems encountered were

- large contents notes that required formatting, upper case, and additional editing;
- significant amount of capitalization in the galley;

- complicated titles, statements of responsibility, and punctuation;
- formatting of added entries; and
- inclusion of an ISSN.

The results of the time study suggested that descriptive cataloging of ECIP galleys with ONIX data is significantly faster than with the TCEC process. While the study did not compare the processes using the same ECIP records, the time study indicated that the ONIX process might be twice as fast as the TCEC process for the vast majority of ECIP galleys received.

Resolving Problems

Communication has been the key to resolving the kinds of problems described above. As catalogers discover new errors, they report them to the cataloging automation specialist. Most errors found in ONIX records can be broken down into two major categories: a mismatch between the galley and ONIX in the data elements (such as the words in the title and author names) and Unicode conversion errors. Catalogers are responsible for catching and correcting both types of errors, although they notify the cataloging automation specialist about the problems. At the beginning of the pilot, significantly more errors of both types were encountered. Over time, the cataloging automation specialist has fixed a large number of errors caused by Unicode. He reports back to publishers on ONIX-related problems discovered by LC staff.

One example demonstrates how effective communication between the CIP program staff at the LC resolved a problem concerning ISBNs. The unique product identifier in the ONIX record is the ISBN for that particular item. When the testing first began, only that ISBN would carry over to the MARC record, even though a number of ISBNs might be included in the ECIP application. As mentioned

above, if multiple ISBNs are included in the application, they are to be converted to multiple 020 fields. Publishers brought to the LC's attention via change requests that the multiple ISBNs were not in the CIP data. Coauthor Williamson provided a solution for this problem so that all the ISBNs for a given title would carry over to the MARC record in an ONIX-to-MARC conversion. A CIP program specialist sent revised CIP data to the publisher, who then had complete CIP data to print in the published books.

Lastly, CIP Publisher Liaison staff members at the LC play a vital role in the CIP Program. They are the primary point of contact between publishers and the LC. At the beginning of the pilot, publisher liaisons frequently requested that the galley be properly coded to allow ONIX to function. Since the pilot began, publishers have sent change requests to have the lengthy summaries removed from the CIP data due to limited space on the copyright page. This has no effect on the bibliographic record because CIP data are sent as an e-mail message, which can be altered as needed. The publisher liaisons simply delete the summary text in the message and the bibliographic record stays intact. Researchers do not lose the rich keyword access in these summaries and publishers have the CIP data they need for their books.

Conclusion

The LC ONIX-to-MARC Converter has been implemented successfully at the LC. The converter has allowed LC cataloging staff to directly use publisher-supplied metadata for the creation of most descriptive data elements required for a monographic record. The use of the ONIX-to-MARC Converter has benefited the LC by reducing the amount of time spent on creating records for titles received through the CIP Program while

providing more enriched records for the greater library community. When the ONIX-to-MARC Converter is put into full production later in 2012, the LC anticipates that more than half of the records created for publishers through the CIP Program will have as their foundation the original ONIX record received from the publisher. At that point, even greater benefits will be realized by the LC and the library community.

However, as Stalberg and Cronin remark, more research needs to be done to determine how use of ONIX will affect the library supply chain.⁴³ The LC and OCLC, which have done the most to implement use of ONIX in their workflows, are unique because they have direct links to publishers that other libraries rarely have. The LC has a need to ingest and manipulate ONIX data as the data are directly tied to the galleys received as part of the CIP Program. OCLC has moved into the ONIX marketplace by providing a fee-based service to publishers who want their ONIX data enhanced with additional data elements, such as LC Subject Headings.

One question that will need to be answered is whether other libraries can use ONIX as effectively as the LC and OCLC. Will use of ONIX by the library community be largely confined to the LC and OCLC with work done largely behind the scenes from which the larger community benefits? Because ONIX files with information on postpublication items are available for use by the library community, can other libraries use these files as a reliable source of copy? Perhaps the postpublication data can be used to assist in removing backlogs of materials in specific areas.

How can ONIX data created by foreign publishers be used by the U.S. library community? Can the LC partner with other institutions to take advantage of ONIX metadata from foreign publishers? The Library and Archives Canada (LAC) has

concurrently implemented a metadata pilot titled "Repurposing User-Generated Metadata Pathfinder" (RUGM Pathfinder) to utilize ONIX data in the cataloging of materials published by Canadian publishers and thereby reduce redundancies in cataloging.⁴⁴ EDItEUR has developed the "Linked Heritage" project, which aims "to extend and enrich the content and metadata holdings of the Europeana digital library."⁴⁵ Some cross-sector collaboration between the LC and these institutions could be beneficial. The possibility that postpublication ONIX data for foreign publishers might be used as a source of copy for libraries exists. With all of the talk of the promise of ONIX, seeing if other libraries can implement it in their workflows will be interesting. This area is ripe for research, particularly in the large academic and research library community.

Also in question is whether greater use of ONIX by the LC, OCLC, and perhaps other libraries or bibliographic utilities in the future will, in fact, eliminate redundancies and waste as hoped for by the authors of *On the Record*.⁴⁶ Even with the enriched data provided by ONIX, will libraries continue to manipulate and tweak the records, thereby accruing costs to their institutions?

Many questions need to be explored in the coming years as libraries begin using and relying more on ONIX metadata supplied by publishers. This is a first foray into how ONIX can be used in a practical setting. As the LC continues to work with ONIX in the years to come, it will share its findings with the library community.

References and Notes

1. Working Group on the Future of Bibliographic Control, *On the Record: Report of the Library of Congress Working Group on the Future of Bibliographic Control*, Jan. 8, 2008, www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf (accessed Mar. 24, 2012).
2. *Ibid.*, 13.
3. *Ibid.*
4. *Ibid.*
5. *Ibid.*, 14.
6. OTR Report Implementation Working Group, Regina Reynolds and Bruce Knarr, co-chairs, *On the Record Report Recommendations the Library of Congress Should Pursue Over the Next Four Years: Report to the Associate Librarian for Library Services*, Sept. 15, 2009, www.loc.gov/bibliographic-future/news/OTR_rep_response_final_091509.pdf (accessed Mar. 14, 2012).
7. *Ibid.*, 15.
8. OTR Report Implementation Working Group, *On the Record*, 4.
9. Working Group on the Future of Bibliographic Control, *On the Record*, 14.
10. EDItEUR, ONIX, FAQs, www.editeur.org/74/FAQs/#q1 (accessed Feb. 5, 2012).
11. See EDItEUR, ONIX, About Release 3.0, www.editeur.org/12/About-Release-3.0 (accessed Mar. 23, 2012) for a summary of the key areas of change and discussions at the BISAC Metadata Committee through March 2012, as reported by co-author David Williamson, who attends as the Library of Congress representative to the committee.
12. EDItEUR, ONIX, FAQs, www.editeur.org/74/FAQs/#q3 (accessed Feb. 5, 2012); EDItEUR, "Minutes of the ONIX for Books ISC meeting 12 October 2011, Frankfurt Book Fair," www.editeur.org/16/Maintenance-and-support/#Steering%20committee (accessed Feb. 5, 2012).
13. Frank Daly, "ONIX: The Metadata Standard for the Information and Entertainment Industries," *Publishing Research Quarterly* 18, no. 2 (Summer 2002): 28; Endre L. Beky, "ONIX: Is There a Return on Investment for All Publishers?" *Publishing Research Quarterly* 20, no. 2 (Summer 2004): 3.
14. Daly, "ONIX," 29.
15. *Ibid.*, 30.
16. Beky, "ONIX," 5.
17. Calvin Reid, "Accurate Metadata Sells Books," *Publishers Weekly* (Jul. 5, 2010), www.publishersweekly.com

- .com/pw/by-topic/industry-news/publishing-and-marketing/article/43740-accurate-metadata-sells-books.html (accessed Apr. 8, 2012).
18. Ibid., 5.
 19. Roy Tennant, "The New Cataloger," *Library Journal* 131, no. 7 (Apr. 15, 2006): 32.
 20. Karen Calhoun and Renee Register, "Next Generation Cataloging," *Journal of Library Administration* 49, no. 6 (Aug./Sep. 2009): 652.
 21. Judy Luther, *Streamlining Book Metadata Workflow: A White Paper prepared for the National Information Standards Organization (NISO) and OCLC Online Computer Library Center, Inc.*, (June 30, 2009), www.niso.org/publications/white_papers/StreamlineBookMetadataWorkflowWhitePaper.pdf (accessed Mar. 14, 2012).
 22. Ibid., 15.
 23. Carol Jean Godby, *Mapping ONIX to MARC: Report and Crosswalk Produced by OCLC Research* (Dublin, Ohio: OCLC Research, 2010), www.oclc.org/research/publications/library/2010/2010-14.pdf (report) and www.oclc.org/research/publications/library/2010/2010-14a.xls (crosswalk) (accessed Feb. 5, 2012).
 24. OCLC, *OCLC Metadata Services for Publishers* (Dublin, Ohio: OCLC, 2010), 2. publishers.oclc.org/en/213918usb_services_for_publishers.pdf (accessed Mar. 14, 2012).
 25. Erin Stalberg and Christopher Cronin, "Assessing the Cost and Value of Bibliographic Control," *Library Resources & Technical Services* 55, no. 3 (Jul. 2011): 131.
 26. Library of Congress, Fascinating Facts, www.loc.gov/about/facts.html (accessed Feb. 5, 2012)
 27. Library of Congress, "Library of Congress Acquisitions and Bibliographic Access Directorate Summary Annual Report, Year Ended September 30, 2011," 3, www.loc.gov/catdir/cps/aba11.pdf (accessed Mar. 14, 2012).
 28. Library of Congress, "U.S. and Publisher Liaison Division Acquisitions and Bibliographic Access Directorate Annual Report, Fiscal Year 2011," Oct. 19, 2011, www.loc.gov/catdir/cps/annrepuspl2011.pdf (accessed Mar. 26, 2012).
 29. Library of Congress, About CIP, www.loc.gov/publish/cip/about (accessed Feb. 5, 2012).
 30. Reid, "Accurate Metadata Sells Books," 5.
 31. Library of Congress, "U.S. and Publisher Liaison Division," 6.
 32. Ibid, 19.
 33. Beth Davis-Brown and David Williamson, "Cataloging at the Library of Congress in the Digital Age," *Cataloging & Classification Quarterly* 22, no. 3/4 (1996): 192.
 34. John D. Byrum Jr. and David Williamson, "Enriching Traditional Cataloging for Improved Access to Information: Library of Congress Table of Contents Projects," *Information Technology & Libraries* 25, no. 1 (Mar. 2006): 9.
 35. Ibid.
 36. Library of Congress, "U.S. and Publisher Liaison Division," 6.
 37. Byrum and Williamson, "Enriching Traditional Cataloging."
 38. *Anglo-American Cataloguing Rules*, 2d ed., 2002 rev., 2005 update (Ottawa: Canadian Library Assn.; London: Library Assn. Publishing; Chicago: ALA, 2005).
 39. Library of Congress, "U.S. and Publisher Liaison," 4.
 40. International Standard Bibliographic Description for Single Volume and Multi-Volume Monographic Publications, Recommended by the Working Group on the International Standard Bibliographic Description Set up at the International Meeting of Cataloguing Experts, Copenhagen, 1969 (London: IFLA Committee on Cataloguing, 1971).
 41. Caroline Saccucci and Camilla Williams, "Time Comparison Study of Cataloging Records of Those Done Using ONIX Data and Those Records Cataloged by the TCEC Software" (internal report, Library of Congress, Aug. 2009).
 42. BISG, *Product Metadata Best Practices*, www.bisg.org/what-we-do-21-8-product-metadata-best-practices.php (accessed Feb. 9, 2012).
 43. Stalberg and Cronin, "Assessing the Cost and Value of Bibliographic Control," 131.
 44. Library and Archives Canada, Pathfinder Working Group, "Repurposing User-Generated Metadata Pathfinder: Interim Report," Mar. 31, 2010, rev. Apr. 13, 2010, www.collectionscanada.gc.ca/obj/012004/f2/012004-2057.01-e.pdf (accessed Mar. 27, 2012).
 45. EDItEUR, Collaborations, "Linked Heritage", www.editeur.org/112/Linked-Heritage (accessed Mar. 27, 2012).
 46. Working Group on the Future of Bibliographic Control, *On the Record*.