

**Association for Library Collections and Technical Services  
(A division of the American Library Association)  
Cataloging and Classification Section**

**Committee on Cataloging: Description and Access**

**Task Force on ONIX International**

**Interim Report**

**Introduction**

The Task Force on ONIX International has been charged with but not limited to:

1. Evaluating the relationship between library metadata (AACR2, MARC 21) and the ONIX International standard for representing and communicating book industry product information in electronic form to determine how well ONIX International maps into AACR2 and MARC.
2. Identifying the issues surrounding the use of ONIX International metadata in AACR2 cataloging records. The Task Force shall refer to the four user tasks set forth in the IFLA Functional Requirements for Bibliographic Records and the core record requirements established by PCC in evaluating the ONIX International standard.
3. Assessing the consequences and impact of integrating records containing the ONIX International metadata into library databases, evaluating mechanisms for integration, and recommending appropriate measures for libraries.
4. Preparing rule revision proposals and discussion papers as needed.
5. Monitoring of projects and activities that use ONIX International.
6. Investigating the feasibility of informing others of library perspectives through a designated liaison to the developers of ONIX International.
7. Informing CC:DA about the development of ONIX International, including the preparation of a summary of the standard which shall include the following information:
  - 7.1 some background or history and community served
  - 7.2 description of metadata element set
  - 7.3 sample records if possible
  - 7.4 citations for more information, implementation projects, etc., including Web sites.

This report constitutes an interim report to CC:DA. It includes the information requested in item 7 of the Task Force's charge. The final report of the Task Force will be presented to CC:DA at the Midwinter Meeting in 2002.

### **Charge 7.1. Background, history, and community served**

**ONIX** stands for **ON**line **I**nformation **eX**change. It is a metadata standard developed by the publishing community as a standard means to exchange information about “book” product information electronically to wholesalers, e-tail and retail booksellers, other publishers, and anyone else involved in the supply chain. The American Association of Publishers (AAP) developed ONIX during 1999 in conjunction with the major wholesalers, online retailers and book information services. Intending to provide publishers a means of sharing product and supplier information usable on the Internet. ONIX was designed as a solution to two problems:

- the lack of consistency and standards in data exchange formats in use by book wholesalers and retailers and the need for a universal, international format in which all publishers could exchange information;
- the need for richer book data online since there is no physical book for the potential buyers to pickup and peruse on the Internet.

Based on EPICS (EDItEUR Product Information Communication Standards), the release of ONIX 1.0 in January 2000 is the culmination of the combined experience of the Book Industry Study Group (BISG) in the US and the Book Industry Communication (BIC) in the UK. ONIX release 1.2 was issued in December 2000. The most recent version, ONIX 1.2.1, released June 1, 2001, is known as **ONIX International**. Release 1.3 was postponed but was scheduled for publication in May 2001.

ONIX defines both a list of data fields and how to send that data in an “ONIX message.” ONIX specifies over 200 data elements with standard definitions. Each element is defined to ensure consistent use. Some elements are required (such as ISBN, author, title) and some are optional (such as book reviews, cover image). ONIX messages are written in XML using the ONIX DTD (Document Type Definition).

The ONIX standard is currently limited to exchanging information about print books. Efforts are underway to broaden the standard to include other formats, such as videos, serials, and electronic books, and to include additional information, such as the concept of digital rights. It is expected to be an evolving standard.

Publishers such as Cambridge University Press and Houghton Mifflin are currently using ONIX to send data to booksellers. Amazon.com, and Barnes & Noble are among the booksellers equipped to receive ONIX data from publishers. It is expected that the list of booksellers and publishers who have implemented ONIX will grow rapidly within the next few years. The Library of Congress and OCLC are both prepared to receive ONIX data, and hope to use it in the future as seed data for producing MARC records. Mappings from ONIX to MARC21 and to UNIMARC are both available.

## Charge 7.2. ONIX Metadata Element Set

### General Description

ONIX is a rich metadata scheme that is comprised of 235 elements of information that fall into 24 categories. The set includes descriptive, administrative and structural metadata elements. The level of granularity of information is finer than that has been developed by MARC/AACR2. Although most of the descriptive metadata elements map to MARC, many of the administrative and structural elements in ONIX do not have an equivalence in MARC. Examples of these non-equivalent data elements include:

- Weight [of the book] (Descriptive)
- Outof PrintDate (Administrative)
- Supplier and Trade data (only 9 of 30 Administrative elements map)
- Downloadable copyright notice (Administrative)
- PriceComposite (Administrative)
- ONIX message elements (Structural)

ONIX standards allow two levels of implementation: Level 1, or ONIX-Lite, is designed to support publishers who have not developed a database for information management. It does not require reference to the XML DTD (release 1.2, 5/11/2000 p. 5). Level 2 contains all of the elements of Level 1 plus elements that provide for more in-depth description. In version 1.0 of ONIX (01/2001), 42 of the 148 elements were considered "crucial." With release 1.2 (11/2000) the total number of elements rose from 148 to 236, 82 of which comprise Level 1. Release 1.2.1 (06/01) contains only minor corrections and a few new values for existing elements, but no new elements. Unfortunately, what complicates the situation is that different countries have defined different subsets of elements belonging to Level 1 (release 1.2, 24 /11/ 2000 p. 6). 45 of the elements are core to all countries (denoted by \*); US publishers have added 4 elements (\$); UK and other European countries add 2 other elements (#)

### Distribution of Categories of Metadata (number of elements in each category)

Categories	Level 1 # of elements	Level 2 # of elements
Reference number, type Source	5	8
Product numbers	3	8
Product form	2	11
Series	3	5
Set	4	6
Title	2	7

Authorship	5	16
Conference	0	6
Edition	3	3
Language	1	2
Extent and other content	2	7
Subject	2	23
Audience	4	12
Publisher	2	8
Publishing Dates	1	4
Territorial Rights	2	7
Dimensions	4	7
Descriptions and other supporting text	2	11
Links to image/audio/video files	3	14
Prizes	0	5
Related Products	4	18
Supplier and trade data	14	30
Sales promotion information	2	5
Message header data elements	12	18
<b>Total</b>	<b>82</b>	<b>236</b>

Crosswalks from ONIX to UNIMARC have been created, (<http://lcweb.loc.gov/marc/onix2marc.html> [ONIX 1.0 mapped]; see also Alan Danskin report at <http://bic.org.uk/reporton.doc> and Bob Pearson's mapping at [http://222.editeur.org/ONIX\\_MARC\\_Mapping\\_External.doc](http://222.editeur.org/ONIX_MARC_Mapping_External.doc)) that support publisher data to be added to national utility databases (OCLC/RLG/BL). Although the mapping is not one-to-one, the presence of publisher records provides a basis for creating and sharing bibliographic information using the MARC record structure based on publisher information. Many of the ONIX fields reflect the level of inventory control, rights, and publication or release information that the book industry requires that the library does not. Much of the information that does not currently map to MARC would not be missed by catalogers. In truth, some of the elements that are mapped to MARC are only possible if Level 3 Cataloging is being performed. Mainstream cataloging for publisher trade publications normally only uses Level 2. Hence, information about an author's affiliations (<b046> to 100/700 \$u) would not appear in a library's bibliographic

record. The end-user, however, would appreciate the links to the reviews, abstracts, and prizes won by the titles that are available in the publisher's records. Now that the use of the URL has expanded beyond the 856 field, these are links that libraries may want to consider retaining. The differences in granularity between the two metadata schema means that although fields can be mapped from ONIX into UNIMARC, they can not "reconverted."

Mapping the fields, however, is only one aspect of sharing these records. The form of names and the controlled vocabulary for subjects, audience, etc. differ from those used by the library community. The level of subject access for publishers and bookstores is far more general than the library access. Personal names may be inverted or in natural word order.

### **Charge 7.3. ONIX Record Examples**

These records are created for business purposes and therefore the information contained within proprietary databases precludes outside access. However, there are a few examples that are available on the Web. Alan Danskin at the British Library has provided examples of records that were originally created in ONIX and converted in UNIMARC (<http://www.bic.org.uk/reporton.doc>, appendix, pp. 4-9).

Guidelines for Online Information Exchange (ONIX) January 2000 1.0 (<http://www.publishers.org/home/onix/onix.pdf> includes examples of ONIX records (pp. 53-54 of AAP pdf document)

There is an example of standard 1.1, level 2 in XHTML (<http://www.editeur.org/samples/950-731-260-9A.xml>) (requires downloading to view).

Clifford Morgan at WileyUK has provide records for our study purposes. Wiley has adopted ONIX as their standard and continues to use the BIC set of subject codes. ONIX allows the implementor to use either BASIC or BIC codes. The records are at the end in the Appendix to Charge 7.3.

### **Conclusions**

A devil's advocate may argue that it is often easier to start from scratch than to convert a publisher record into a standard UNIMARC record with AACR2 serving as the guide to formulate the content, NAF for names, and LCSH to provide the subject headings. Granularity as well as content may reduce the usability of these records. However, we may be gaining information about the publication that would be useful for a patron to evaluate the title: increased information concerning the reception of the book, more information about the author, more detailed information about its currency.

## Charge 7.4. Citations for more information, implementation projects, etc.

### I. General and Background Information

**ONIX International homepage** (Developed and maintained by EDItEUR jointly with Book Industry Communication and the Book Industry Study Group)

<http://www.editeur.org/onix.html>

Includes:

- Link to Download of ONIX Release 1.2.1 (latest release as of 5/30/01) Guidelines and XML DTD
- Links to downloads of previous releases
- ONIX FAQs (frequently asked questions about ONIX)
- Link to subscription page for ONIX\_IMPLEMENT (the e-mail listserv for ONIX implementers)

### **PowerPoint presentations on ONIX International**

<http://www.bic.org.uk/onixsem.html>

Links to 7 presentations (PowerPoint, Word, or .pdf) presented at the seminar *ONIX International: How Better Product Information Sells More Books* (London, 14 November 2000):

- What is ONIX (David Martin)
- Amazon and ONIX (Mo Jacobs)
- Whitaker and ONIX (Michael Healy)
- BookData and ONIX (Peter Mathews)
- Cambridge University Press and ONIX
- Harper Collins and ONIX (Graham Bell)
- Libraries and ONIX (Alan Danskin)

### II. Technical Information

**ONIX to MARC 21 Mapping** (Network Development and MARC Standards Office, Library of Congress)

<http://lcweb.loc.gov/marc/onix2marc.html>

Includes mapping table in ONIX data element order, and record builder for creating MARC 21 records from ONIX data.

**ONIX to UNIMARC Mapping** (by Alan Danskin of The British Library)

<http://www.editeur.org/onixmarc.html>

**EPICS Version 3: EDItEUR Product Information Communication Standards**

<http://www.editeur.org/epics.html>

The data dictionary upon which ONIX International is based.

### III. **Implementations of ONIX International**

**ONIX International Implementers Information** (The Book Industry Study Group, Inc.)

<http://www.bisg.org/onix.html>

Project details and contact information at organizations that are implementing ONIX International.

### IV. **Related Projects**

**AAP Open Ebook Publishing Standards Initiative** (Association of American Publishers)

- General information about the project:  
<http://www.publishers.org/home/ebookstudyinfo.htm>
- .pdf documents on Digital Rights Management, Metadata, and Numbering:  
<http://www.publishers.org/home/ebookstudy.htm>

**Book Industry Communication**

<http://www.bic.org.uk/>

BIC, set up and sponsored by The Publishers Association, The Booksellers Association, The Library Association and The British Library, develops and promotes standards for electronic commerce and communication in the book and serials industry.

### V. **Similar or Supporting Standards, Specifications, Initiatives, etc.**

**PRISM (Publishing Requirements for Industry Standard Metadata)**

<http://www.primstandard.org/news/2001/0401.asp>

Press release announcing the release of Version 1.0 of PRISM and links to PRISM home page and documentation.

**BIC BASIC Standards for Product Information**

<http://www.bic.org.uk/bbinfo.html>

Another subset of the EPICS data dictionary

**Open eBook Publication Structure**

<http://www.openebook.org/specification.htm>

Includes the specification (Version 1.0) plus links to FAQ and sample records.

**ANSI/NISO Z39.82-2001** (Published April 9, 2001)

Available for free download in .pdf format at:

<http://www.niso.org/>

The ANSI/NISO standard on Title Pages for Conference Publications. This standard describes data elements that publishers, authors, and editors should use to create title pages or chief sources of information for conference publications in all subjects, languages, and formats and will be extremely helpful in assuring the communication of conference information to interested readers.