ABSTRACT

A meeting in April 2015 explored the potential withdrawal of valuable collections of microfilm held by the University of Maryland, College Park Libraries. This resulted in a project to identify OCLC record numbers (OCN) for addition to OCLC’s Chadwyck-Healey Early English Books Online (EEBO) KBART file. Initially, the project was an attempt to adapt cataloging workflows to a new environment in which the copy cataloging of e-resources takes place within discovery system tools rather than traditional cataloging utilities and MARC record set or individual record downloads into online catalogs. In the course of the project, it was discovered that the microfilm and e-version bibliographic records contained metadata which had not been utilized by OCLC to improve its link resolution and discovery services for digitized versions of the microfilm resources. This metadata may be advantageous to OCLC and to others in their work to transition from MARC to linked data on the Semantic Web. With MARC record field indexing and linked data implementations, this collection and others could better support scholarly research.

Collections, Discovery Tools, and Metadata Services

The University of Maryland, College Park Libraries’ (the Libraries; UM Libraries) collections include 3.45 million print books and 1.2 million eBooks, 17,000 electronic journals, and 352 electronic databases. In late 2011, the Libraries implemented WorldCat Local, OCLC’s single-search-box interface to the WorldCat database of cataloged resources and a central index of metadata provided by publishers, Abstracting and Indexing Services, institutional repositories, and so on. With WorldCat Local, and later, WorldCat Discovery, OCLC utilizes a knowledge base in managing e-resources discovery and access. Knowledge bases are “associated with link resolvers and electronic resource management systems” and “contain title-level metadata, linking syntax rules, publication ranges and other data.” KBART files are so named to represent files compliant with the NISO recommended practice, "Knowledge Bases and Related Tools (KBART).” KBART files, created and supplied by content providers, are used to transmit this title level metadata to knowledge base vendors and discovery service providers. Since OCLC enhances these files with OCLC numbers (OCN) in order to provide automated holdings maintenance on WorldCat bibliographic records, the Libraries’ Metadata Services Department (MSD) adopted a policy in 2012 to provide access to e-resources only via WorldCat when such files are available.

Beth Guay (baguay@umd.edu) is Continuing Resources Librarian, University of Maryland Libraries, University of Maryland, College Park.
Space Planning

Early on, the Libraries’ collection policies targeted duplicate copies of print monographs and print journals held electronically in trusted repositories, e.g., JSTOR, for deselection. By March 2014, the Libraries’ Collection Development Council discussed moving microfilm collections to the yet to be opened Severn Library, slated to “house lesser used materials ... in order to free up much needed space for users and the development of new collaborative learning spaces.” A year later, in April 2015, a meeting was called by the Assistant Head, Collection Development, to investigate microfilm collection retention decisions. This time the Libraries were considering the withdrawal of microfilm resources for which equivalent versions were held online. A caveat placed on the withdrawal of the microfilm by the collection managers was that prior to their withdrawal and subsequent deletion of the Libraries’ holdings on the WorldCat bibliographic records, the equivalent e-version resources should be made discoverable in WorldCat UMD (the Libraries’ WorldCat Discovery implementation) by the addition of the Libraries’ holdings on e-version bibliographic records corresponding to the microfilm version records.

Following the meeting, the Librarian for English, Latin American, & Latina/o Studies and Second Language Acquisition provided the Continuing and Electronic Resources Cataloger (C-ER Cataloger) with a list of eight valuable microfilm collections of resources and for each, the name of the comparable online collection (or e-collection) subscribed to. It was agreed that the C-ER Cataloger would investigate to determine if any of those microfilm collections could be withdrawn in compliance with the collection managers’ caveat. In other words, the C-ER Cataloger’s mission was to ensure a one-to-one correspondence of electronic and microfilm version bibliographic records for the equivalent versions of the resources.

One of the e-collections added to the WorldCat Knowledge Base (WCKB) by the Libraries was Gale’s, The Making of the Modern World, 1450-1850: Part I collection (MOMW). This collection is comprised of digitized versions of Gale’s microfilm resources in the series, The Goldsmiths’-Kress library of economic literature. A KBART file was derived from the Libraries’ MOMW MARC record set and uploaded to the WCKB sandbox where it supports the Libraries’ access to the e-version resources. The MOMW MARC record set had been reviewed and vetted by the Libraries prior to its purchase, and upon its purchase, Gale had set the Libraries’ holdings on the WorldCat bibliographic records representing the resources. With this information in mind, the C-ER Cataloger determined that the MOMW e-resource bibliographic records were comparable to those representative of the Libraries’ corresponding Goldsmiths’-Kress library of economic literature microfilm collection, thus meeting the collection managers’ criteria for deselection. The 3380 reels that could be withdrawn comprised a small but not insignificant allotment of physical space in the library.

Provision of discoverability of equivalent e-versions of resources held in other collections proved difficult. For example, the corresponding microfilm collections represented in the WCKB’s British Periodicals Collections I and II were held in the series, Early British periodicals and English literary...
periodicals. The Libraries had cataloged 186 individual serial titles in the microfilm series, *Early British periodicals* in 2002, but none in the series, *English literary periodicals*. Thus the objective would have been to ensure discoverability for the equivalent electronic versions of the Libraries’ 186 cataloged microfilm versions in the *Early British periodicals* series. At the time of this investigation, there were 580 British Periodicals I and II KBART file title entries; 390 of which had OCN. Whereas the OCN of The Making of the Modern World, 1450-1850: Part I WCKB collection were known entities, the OCN of the remaining e-collections had yet to be vetted. Thus the British Periodicals Collections I and II records were spot checked for evaluation. The quality of the 390 OCLC records ranged from excellent, e.g., OCLC record #297425799, to poor, e.g., #818401694 (see Figure 1, 2, 3, and 4). MARC record images in Figures 1-4 are sourced from OCLC’s Connexion cataloging client interface to the WorldCat bibliographic database. Figures 1 and 2 represent a microfilm version record and a comparable “excellent” quality record given for the resource in the WorldCat Knowledge Base, while Figures 3 and 4 represent a microfilm version and comparable “poor” quality record given for the resource in the WCKB. Note that the C-ER Cataloger’s definition of an excellent quality e-version record was one which provided metadata comparable to those of its equivalent microfilm version record; likewise, a poor quality record lacked comparable metadata. In other words, an excellent quality record was viewed as a guarantor of a discoverable resource, while a poor quality record was viewed as an obstacle to discovery. For this WCKB collection, the C-ER Cataloger determined that staff expertise with serial bibliographic records was required, and due to MSD staffing limitations, moved ahead to examine the other collections.
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>010</td>
<td>Microfilm 01106 reel 68-72</td>
</tr>
<tr>
<td>040</td>
<td>UMCP</td>
</tr>
<tr>
<td>090</td>
<td>m</td>
</tr>
<tr>
<td>049</td>
<td>SC008000002</td>
</tr>
<tr>
<td>047</td>
<td>300 N. Zeib Rd., Ann Arbor, MI 48106</td>
</tr>
<tr>
<td>042</td>
<td>pcc</td>
</tr>
<tr>
<td>043</td>
<td>e-uk—</td>
</tr>
</tbody>
</table>

**Figure 1.** Microfilm version record
Figure 2. Excellent quality e-version record — OCN in the KB file
<table>
<thead>
<tr>
<th>Type</th>
<th>Form</th>
<th>Cont</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>sL</td>
<td>Orig</td>
<td>EntW</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1883</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3.** Microfilm version record
In an investigation into OCLC's Chadwyck-Healey Early English Books Online (EEBO) KBART file, for which equivalent e-versions of microfilm resources in the series Early English books, 1475-1640 and Early English books, 1641-1700 are held, it was found that the availability of comparable e-version bibliographic records was optimal. In consultation with the MSD department head, a project to ensure the discoverability of equivalent e-versions of the Libraries' 5,062 cataloged microfilm resources in the series, Early English books, 1475-1640 was initiated. The C-ER Cataloger had hoped to follow with a similar effort for the Libraries' resources in the series Early English books, 1641-1700 (represented by 41,306 records in the Libraries' Integrated Library System).

**Background: EEBO, Related Resources and Bibliographic Records**

Much has been written on EEBO's inception and continuing development as a collection of digital reproductions of microfilm reproductions of pre-1700 print resources, and on its scholarly value (Kitchuk, 2007; Martin, 2007; Gadd, 2009; Mak, 2013; Folger Shakespeare Library, 2015). Alfred Pollard and Gilbert Redgrave's A short-title catalogue of books printed in England, Scotland, & Ireland and of English books printed abroad, 1475-1640 ("STC"), and the “companion” volume, Donald Wing's Short-title catalogue of books printed in England, Scotland, Ireland, Wales, and British America, and of English books printed in other countries, 1641-1700 ("Wing"), respectively, were used in selecting the print resources for filming. Gadd (2009, 683) pinpointed the STC as “a catalogue of editions (or more accurately, editions and issues) not copies although, of course, the
information about any edition is derived primarily from the surviving copies ... Each entry gives the location of known copies ...”  

The “successor” to STC and Wing, the *English Short Title Catalog* (ESTC), “includes records for every item listed in STC, every item in Wing, every item in the *Eighteenth Century Short Title Catalogue* ... and newspapers and other serials which began publication before 1801” and is freely available online from the British Library.  

Gadd (2009, 685-686) offered this critique concerning EEBO’s bibliographic data and relationship to the ESTC:

> EEBO's relationship with the original STC and Wing is straightforward and clear; EEBO’s relationship with electronic ESTC, on the other hand, is less well-known. A series of agreements made between ESTC and University Microfilms/ProQuest between 1989 and 1997 allowed EEBO to draw directly on ESTC’s existing bibliographical data ... EEBO heavily edited ESTC’s data for its own purposes; certain categories of data were removed (e.g. collations, Stationer’s Register entrances), some information was amended (e.g., subject headings), and some was added (e.g. microfilm specific details). Second, there is no formal mechanism for synchronizing the data between the two resources. Occasionally, snapshots of data are sent by EEBO to ESTC but there is no guarantee that a correction or revision made to an ESTC entry will be replicated in the corresponding EEBO or vice-versa: neither ESTC nor EEBO will necessarily know when the other made a correction.

Gadd postured that “as both resources continue to amend and expand their bibliographical data for their own purposes, there is an increasing likelihood of significant discrepancy between the two resources.” He did not further address the quality of the bibliographic records describing the EEBO versions of the resources; perhaps he was unaware of the sources of the EEBO bibliographic data.

Microfilm version bibliographic records serve as the basis of the metadata describing the EEBO version resources. According to ProQuest, “MARC records (from which EEBO Bibliographic records derive) are produced for the microfilm collection Early English Books (EEB) after they are filmed.” OCLC’s cataloging database has served as one source of microfilm version records for titles in the series since the 1980s. In 1984, the Association of Research Libraries (1984, p. J-3) reported that one library had “input an indeterminate amount [of bibliographic records] into OCLC” for *Early English books, 1475-1640*, and that one had “input records for an indeterminate percentage of the set into OCLC” for resources in the series, *Early English books 1641-1700*. The cataloging sources of these microfilm resources have varied over time, from cooperative projects to UMI/ProQuest staff to individual libraries, however, adherence to standards has characterized the totality of the efforts invested. Joachim (1993, p. 111) described the cooperative effort begun in 1984 by the Indiana University Libraries, University of California, Riverside, University of Delaware, and the University of Utah to catalog microfilm version resources cataloged by *Wing*:  

---

14 Information about any edition is derived primarily from the surviving copies. Each entry gives the location of known copies.

15 The “successor” to STC and Wing, the *English Short Title Catalog* (ESTC), includes records for every item listed in STC, every item in Wing, every item in the *Eighteenth Century Short Title Catalogue* ... and newspapers and other serials which began publication before 1801.

16 Gadd (2009, 685-686) offered this critique concerning EEBO’s bibliographic data and relationship to the ESTC:

> EEBO’s relationship with the original STC and Wing is straightforward and clear; EEBO’s relationship with electronic ESTC, on the other hand, is less well-known. A series of agreements made between ESTC and University Microfilms/ProQuest between 1989 and 1997 allowed EEBO to draw directly on ESTC’s existing bibliographical data ... EEBO heavily edited ESTC’s data for its own purposes; certain categories of data were removed (e.g. collations, Stationer’s Register entrances), some information was amended (e.g., subject headings), and some was added (e.g. microfilm specific details). Second, there is no formal mechanism for synchronizing the data between the two resources. Occasionally, snapshots of data are sent by EEBO to ESTC but there is no guarantee that a correction or revision made to an ESTC entry will be replicated in the corresponding EEBO or vice-versa: neither ESTC nor EEBO will necessarily know when the other made a correction.

17 Gadd postured that “as both resources continue to amend and expand their bibliographical data for their own purposes, there is an increasing likelihood of significant discrepancy between the two resources.”

18 He did not further address the quality of the bibliographic records describing the EEBO versions of the resources; perhaps he was unaware of the sources of the EEBO bibliographic data.

19 Microfilm version bibliographic records serve as the basis of the metadata describing the EEBO version resources. According to ProQuest, “MARC records (from which EEBO Bibliographic records derive) are produced for the microfilm collection Early English Books (EEB) after they are filmed.”

20 OCLC’s cataloging database has served as one source of microfilm version records for titles in the series since the 1980s. In 1984, the Association of Research Libraries (1984, p. J-3) reported that one library had “input an indeterminate amount [of bibliographic records] into OCLC” for *Early English books, 1475-1640*, and that one had “input records for an indeterminate percentage of the set into OCLC” for resources in the series, *Early English books 1641-1700*. The cataloging sources of these microfilm resources have varied over time, from cooperative projects to UMI/ProQuest staff to individual libraries, however, adherence to standards has characterized the totality of the efforts invested. Joachim (1993, p. 111) described the cooperative effort begun in 1984 by the Indiana University Libraries, University of California, Riverside, University of Delaware, and the University of Utah to catalog microfilm version resources cataloged by *Wing*:  

---
In order to maintain standards and consistency among the five libraries, the project director prepared a “Wing STC Project manual.” The manual includes general information, information on authority work, a bibliography, a discussion of special cataloging problems and procedures, sample records, and database input guidelines.21

OCLC’s MARC records for the microfilm and EEBO version resources contain note fields identifying the locations of the print copies filmed and subsequently reproduced digitally by UMI/ProQuest. Gadd (2009, p. 686) emphasized the importance of this information to scholars in stating that “different copies from the same edition might vary, sometimes markedly.”22

As to Gadd’s (2009) critique concerning the lack of a formal synchronization mechanism and increasing likelihood of discrepancies between EEBO and ESTC, further examination of EEBO and ESTC bibliographic record displays such as those shown in Figures 5 and 6 suggest that the British Library is working with ProQuest to align their data. It appears a focus of the British Library may be to inform the scholar of the availability of the microfilm and electronic versions of the print resources. In its ESTC overview, the British Library states that “the existence of selected … printed and digital surrogates within products such as Early English Books Online … is … noted” in its records and that its records “act as an index to several major research microform series … including Early English Books, 1475-1640 … [and] Early English books, 1641-1700.”23
Figure 5. EEBO bibliographic record for the resource cited by STC 2nd edition entry 9164 and reproduced from the copy held at the Society of Antiquaries, London.
Figure 6. ESTC catalog record for STC 2nd edition, entry 9164 (http://estc.bl.uk/S3614). The code, “Lsa” given as “Loc. Of filmed copy” is the British Libraries’ MARC code for the Society of Antiquaries Library. 24
Finally, to add to this mix of print, microfilm, and EEBO digitized images, XML/SGML versions of
the resources are being created by the Text Creation Partnership (TCP), formed in 1999 by the
university libraries of Michigan and Oxford, ProQuest, and the Council on Library and Information
Resources, to provide full text search capability.25 Catalog records describing TCP versions are
available in WorldCat. According to the TCP, “the TCP does not have the resources to create new
catalog records for each text we produce (though you are welcome to do so, and if you are willing
to share them we would be very glad to know about it).”26

The UM Libraries’ EEBO Project

The OCLC EEBO KBART file, which contained 129,544 title entries when downloaded, 58,518 of
which lacked OCN, was combined with a file extracted from the 5,062 MARC records that
represented the microfilm resources. The merged file was to be used as a tool in identifying the
OCN of the equivalent e-versions of the microfilm resources held. The plan was to add the e-
version OCN to the EEBO KBART file via OCLC’s OCN correction form.27

Significant time was spent developing and documenting procedures by which staff could perform
the work of identifying OCN for addition to the EEBO file. The basic procedures are as follows: (1)
via the OCLC Connexion cataloging client, search and retrieve the e-version record using the
microfilm version record data; (2) use titles and/or OCN of the microfilm version record to
identify the comparable EEBO resource in the KBART file; (3) view the EEBO resource record
using the URL in the file; and (4) record the OCN of the matching e-version record in the
appropriate row/column of the file.28

Subsequently, two MSD staff members were recruited to assist in the effort. In early November
and mid-December, 2015, training sessions were held with both staff, followed by an individual
session with each. Before the year’s end, each staff member had successfully completed an
assigned number of “titles” for review. Importantly, from the initial investigative work, a KBART
file with 50 OCN was compiled and submitted to OCLC. Confirmation from OCLC Customer
Support was given that the file would be loaded. Due to the ongoing developmental status of
OCLC’s services, the OCN were not loaded into the WCKB until June 2016. However, a second file
sent in April 2016 was loaded in June as well. The number of OCN added to the WorldCat
Knowledge Base from the project’s inception through 2016 was small due to staffing issues. The
average staff time to complete a microfilm/equivalent e-version title entry in the KBART file was
13 minutes.29

As the project progressed, staff following the procedures confirmed that some OCN in the EEBO
KBART file were incorrect. Most often, the “errors” stemmed from the attribution of TCP or
German language of cataloging record OCN to the EEBO version resources. These TCP and German
language of cataloging records correctly corresponded to matching EEBO version resources,
however, TCP version records refer to XML/SGML encoded text editions; secondly, OCLC attempts
to prefer English language of cataloging records over others in its knowledge base.30
Other OCN errors seriously detract from the WCKB’s EEBO file’s value. For example, WorldCat record number 606541404 describes the “fourth edition very much enlarged” of “A Most exact catalogue of the Lords spirituall and temporall, as peers of the realme, in the higher House of Parliament, according to their dignities, offices, and degrees: some other called thither for their assistance, & officers of their attendances ...” yet this OCN in the WorldCat Knowledge Base’s EEBO KBART file links to an EEBO record describing the “third edition much enlarged.” See Figure 8 illustrating the WorldCat UMD record which links to an EEBO resource record describing the “third edition much enlarged.” Note that the OCLC record (as seen in the Connexion client view of the record in Figure 9) is cited by STC (2nd ed.) 7746.3 while the EEBO version record linked to is cited by STC (2nd ed.) 7746.2. To make matters worse, the author determined that the corresponding image associated with the EEBO catalog record cited by STC 7746.2 and displayed at the site corresponded to neither resource cited as STC 7742.2 and STC 7746.3. These were both printed in 1628, but the image provided at the EEBO site was of a resource printed in 1640 (see Figure 10).

Figure 8. WorldCat UMD record OCN 606541404 linking to the wrong version of a resource in EEBO.
Figure 9. Connexion client view of OCN 606541404
Further investigation identified errors of misappropriation of OCN in the KBART file to EEBO version records describing copies of editions filmed at locations other than those noted in the corresponding OCLC records. For example, the EEBO resource, “By the King. A proclamation for the adiournement of part of Trinitie terme,” identified in the WCKB as associated with OCN 71492075, links the scholar to a resource described by the EEBO version record as the copy filmed at the British Library. OCLC record 71492075 however indicates that the copy it describes was the copy filmed at the Henry E. Huntington Library and Art Gallery. See Figures 11-13.

Figure 11. The WCKB associates OCN 71492075 with the EEBO resource, “By the King. A proclamation for the adiournement of part of Trinitie terme,” described by the EEBO website as the copy filmed at the British Library.
Figure 12. The EEBO resource record linked from OCN 71492075 by the OCLC EEBO KBART file indicates the copy filmed was held by the British Library.
Figure 13. OCN 71492075 indicates it describes a copy of the resource, “By the King: a proclamation for the adiournement of part of Trinitie terme,” filmed at the Henry E. Huntington Library and Art Gallery.

**Evaluation**

The UM Libraries’ EEBO project procedures revealed that match points of equivalent microfilm and e-version records were the names of the institutions holding the filmed copies and the STC citations to the resources. STC citations are carried in the MARC 510 fields of the bibliographic records in two subfields:

1. in subfield “a,” the names of citing works, given in a brief form, e.g., “STC” to represent Pollard and Redgrave’s *Short-title catalogue*; and

2. in subfield “c,” the location (e.g., page number or volume) within the citing works, e.g. “8626.”

Figure 14 displays a Connexion Client view of OCN 33150534, cited as STC 9170, and Figure 15 shows the same record in the WorldCat display view. Unfortunately, the MARC 510 fields are neither indexed by OCLC nor displayed in WorldCat. OCLC could enable the identification and
collocation of records for equivalent print, microfilm and electronic versions by indexing the MARC 510 fields and subfields.\textsuperscript{34}

\textbf{Figure 14.} Microfilm version record OCN 33150534, cited as STC 9170.
Figure 15. WorldCat.org view of OCN 33150534, STC 9170 (http://www.worldcat.org/oclc/33150534). The underlying MARC 510 field metadata is not displayed.

Investigation by the author revealed that TCP version records supply these metadata elements in duplicate in different MARC fields; one a free text note field, the other a number/code field, 024. The 024 field is defined to carry a “standard number or code published on an item which cannot be accommodated in another field (e.g., field 020 (International Standard Book Number)).”\(^\text{35}\) It should be noted that use of the 024 field to carry a number that is *not* published on the item is not in accordance with the field’s definition. The TCP records use the 024 field with a first indicator value “8,” conveying that the number is an unspecified type of standard number or code.\(^\text{36}\)

Subfield “a” of the 024 field, which carries the STC numbers in the TCP version records, *is* indexed by OCLC. In the TCP version records, however, these elements are ensonced within strings of text, e.g., “(stc) STC (2nd ed.) 9170.”\(^\text{37}\) A search on standard number, “9170,” in WorldCat will therefore fail to retrieve the appropriate record. See Figure 16 for an example of a TCP version record of a resource cited as STC 9170.
In respect to the MARC field definitions, should there be a need to retrieve bibliographic records representing TCP versions of resources via STC citations, these numbers should be entered in “a” subfields, and the brief abbreviated names of the citing source, e.g., “STC (2nd ed.),” “Wing,” etc. in the “2” subfield which is defined to carry the “Source of number or code.” Should OCLC choose to index the MARC 510 fields as described above, the Text Creation Partnership records would be missed.

Figure 16. Text Creation Partnership version OCN 832931179, STC 9170

Indexing of the MARC 510 fields/subfields by OCLC combined with use of other MARC field/subfield values, such as language of cataloging, to limit results to desired OCN could support elimination of EEBO KBART file OCN errors and identification of thousands of new OCN for addition to this and perhaps other similar files. As a point of reference, according to OCLC’s “MARC Usage in WorldCat” webpages, as of January 1, 2016, there were 6,382,317 instances of MARC 510 “a” subfields and 4,082,280 instances of the “c” subfields. It should be noted, however, there are five first indicator values available for use in MARC 510 fields and only one of them is used to convey the information that the location in the source data is given in the field. Also worth noting, 024 data at the “MARC Usage in WorldCat” webpages shows that there were 4,633,776 occurrences of subfield 2 of the 024 field, and 43,711,819 occurrences of subfield “a.”

510 field indexing to support identification of OCN for addition to the EEBO KBART file may require the participation of the content provider, ProQuest. The 510 field elements are indexed in
its Early English Books Online collection. ProQuest could add these data to its EEBO KBART file in support of OCN matching. The KBART Recommended Practice allows content providers “to include any extra data fields after the last KBART utilized position.”

Finally, it should be noted that reconciliation of errors in the WCKB EEBO file pertaining to the locations of the filmed copies as noted in OCLC records but found to be different at the EEBO site would require more complex steps than 510 field matching. Furthermore, catalogers working on the EEBO project were not instructed to check the images at the EEBO website but only to confirm the STC citation match points in the EEBO version records. A closer examination of EEBO in light of the findings in this paper of an EEBO record linked to a resource printed 12 years later is an area calling for further study. In respect of the needs of scholars as eloquently described by Gadd (2009), the WorldCat Knowledge Base OCN must improve its accuracy in terms of access provision via WorldCat Discovery.

**MARC 510 Elements: Opportunities for Linked Data Applications?**

OCLC is actively engaged in research and collaboration with the greater library community to transition its metadata to linked data, however, MARC 510 metadata is lacking in its linked data record display views (see Figure 14 in a Connexion client view of a record and Figure 17 in the WorldCat linked data display view). On the other hand, in its work to transfer its English Short Title Catalog, a “MARC based ... vendor-supplied ILS” to “ESTC21” a “native linked data resource,” it appears the British Library combines the MARC 510 subfield values, e.g., “Bristol, B7384” as a resource property value (Figures 18 and 19). “Bristol, B7384” represents entry number 7384 in Roger P. Bristol’s *Supplement to Charles Evans’ American bibliography* (see Figure 20, WorldCat OCLC record number 88701). As presented in Figure 19 (Stahmer, 2014), “Bristol, B7384” may be comprehensible to a well-versed scholar, librarian or archivist, but not to a computer.

Hillmann, Dunsire, and Phipps (2013) posited that “it would be useful if all managers of schemas and other standards were to develop element sets and value vocabulary representations that match the source semantics at the finest granularity and make them available along with maps of the internal ontologies.” Could a Semantic Web implementation of MARC 510 metadata at the finest granularity, with resource identifiers representing citing works such as “Bristol” and with property values such as “7384” representing locations within citing works, offer benefits to scholarship? It has been demonstrated in this paper that the consistent match points across bibliographic records representing equivalent versions of these resources has been the metadata contained in MARC 510 fields. Ultimately, a linked data implementation of the MARC bibliographic 510 field should lead the scholar to every known print copy comprising every edition, according to Gadd’s definition of an edition, above, and to the institutional holdings of equivalent microform, digitized images, or digitized full-text versions, giving the scholar the path to the resources of interest. OCLC, the British Library, members of the TCP, and other stakeholders may want to consider further exploration of use case scenarios to determine or rule out additional benefits of transforming MARC 510 field metadata to linked data.
Figure 17. Linked data view of OCLC #33150534, [http://www.worldcat.org/oclc/33150534](http://www.worldcat.org/oclc/33150534)
Figure 18. MARC 510 field data in ESTC
**Figure 19.** MARC 510 metadata in structured data view in ESTC21
CONCLUSION

At the current pace, given available staffing and the number of EEBO resources lacking OCN, the time and effort spent by the Libraries’ Metadata Services Department staff toward the goal of adding OCN to the OCLC EEBO KBART file, though well spent, will be years in the making. A collective effort in this endeavor by the WCKB community of users is welcomed by this author. A combined effort by OCLC and ProQuest to improve discovery and link resolution services for these valuable scholarly resources could increase their discoverability exponentially, allowing MSD staff to spend more time creating and enhancing the metadata that will lead researchers to the uncatalogued EEBO resources they seek. As to the transition of MARC 510 field metadata to linked data, OCLC, the British Library, members of the TCP, and other stakeholders should consider their options before moving forward without it.
ACKNOWLEDGEMENT

The author wishes to thank Karen Coyle for reading and advising on earlier versions of this paper; Becky Culbertson, Nathan Putnam, and Patricia Herron for supporting the project; and Joshua Westgard for converting the data to get the project underway. Special thanks are due to staff members of the UM Libraries, Donna King, Roselin Becker, Erica Hemsley, Yeo-Hee Koh, and Tanisha Lee, and to Freeda Brook, Luther College, for their work on the project.

REFERENCES


3. In 2015, the Libraries implemented WorldCat Discovery, intended to be a replacement for WorldCat Local.


WorldCat Knowledge Base collections may be created by libraries and uploaded to the Knowledge Base. Details on the process are available at http://www.oclc.org/support/services/collection-manager/documentation.en.html#knowledgebase, accessed February 18, 2017.


15. “About EEBO.”

16. Details on the ESTC are provided by the British Library at http://www.bl.uk/reshelp/findhelprestype/catblhold/estccontent/estccontent.html, viewed March 12, 2017


   http://www.bl.uk/reshelp/findhelpertype/catblhold/estccontent/estccontent.html

24. The British Libraries ESTC codes for filmed copy locations are difficult to translate. See Meaghan J. Brown's finding aid, “STC Location Code Transcription” wherein she offers details on STC and ESTC location codes and the problem her finding aid addresses. Brown explains, “… it is currently possible to search the ESTC for items using MARC codes, but not the location codes familiar from the STC,” accessed February 18, 2017. http://www.meaghan-brown.com/stc-location-codes/

   http://www.textcreationpartnership.org/home/

   http://www.textcreationpartnership.org/catalog-records/

27. OCLC’s form is available at https://www.oclc.org/content/dam/support/knowledge-base/ocn_report.xlsx, accessed October 18, 2016.

28. See Appendix 1 for the Procedures

29. With streamlined KBART search features introduced by a Metadata Services Department colleague, it’s expected this time may be reduced moving forward.

30. A June 9, 2015 email from an OCLC staff member to the KB-L@oclc.org listserv reported on OCLC’s efforts to match OCN in its KBART files to English language of cataloging records, when available.

31. UM Libraries’ staff use this metadata in the equivalent OCLC microfilm and e-version and EEBO resource records as match points. Staff do not verify that the images linked to the EEBO version records correspond to those in the aforementioned bibliographic records. It is hoped that ProQuest will investigate the case described in this paper in which the EEBO resource differs from its corresponding record.


33. As of January 29, 2017, the MARC 510 field has not been indexed by OCLC. See http://www.oclc.org/support/help/SearchingWorldCatIndexes/#05_FieldsAndSubfields/5xx_fields.htm

34. E.g., OCLC indexes “internet resources” using a combination of MARC data elements. These are laid out in “Searching WorldCat Indexes” at http://www.oclc.org/support/help/SearchingWorldCatIndexes/#06_Format_Document_Type_Codes/Format_Document_type_codes.htm. MARC 21 Bibliographic at
https://www.loc.gov/marc/bibliographic/bdleader.html provides the Leader position 06 code for “Language material.” MARC Code List for Languages (http://www.loc.gov/marc/languages/) contains the language codes contained in the language of cataloging field/subfield (MARC 040 field, subfield “b”).


36. Ibid.

37. OCLC. Searching WorldCat Indexes, accessed February 18, 2017. http://www.oclc.org/support/help/SearchingWorldCatIndexes/#05_FieldsAndSubfields/0xx_fields.htm#TocPath%3DFields%2520and%2520subfields%7C2


39. An Oct. 18, 2016 review of OCLC's all-collections-list, available at https://www.oclc.org/content/dam/support/knowledge-base/all-collections-list.xlsx indicates that 38.5% percent of the 129,498 resources on the EEBO KBART file have OCLC number coverage.


44. The image of the linked data view of Figure 14 was captured on February 18, 2017.


49. See Reference 14 above.

50. A discussion and invitation to collaborate on this work took place in late 2016 on the OCLC WorldCat KB listserv (see http://listserv.oclc.org/scripts/wa.exe?SUBED1=kb-l&A=1). To date, the Preus Library, Luther College, will be working with the Libraries on this project.