Interim E-Books Report
May 2009

Presented by the E-Books Task Force
Mary Beth Weber (co-chair)
Gracemary Smulewitz (co-chair)
Fay Austin
Sara Harrington
Karen Hartman
Elizabeth Leister
Mei Ling Lo
Laura Mullen
Jane Otto
Cathy Pecoraro
Robert Warwick
Introduction

Careful consideration must be given to e-book purchases as funds are limited and RUL has a number of competing priorities to meet. To date, e-book purchases have largely been made by selectors on an individual basis, and Rutgers lacks an official collection development policy and workflow for these resources. There is a 1999 document available on the Collection Development portion of the RUL website titled “Evaluation Criteria for Electronic Resources.” While some of the criteria in this document are still valid, a policy for selecting, acquiring, and making available these resources is needed. The document will be prepared with the understanding that our policies will evolve over time as we gain experience with e-books and as our needs change. Conversely, standards for cataloging (including loading and processing of vendor supplied MARC record sets) need to be developed.

Charge

The Task Force charge follows:

“RUL/CDC E-books Task Force: Charge

The task force will review the electronic book marketplace. The charge is to evaluate available e-book platforms and services and make recommendations for RUL. Consider the following points in your evaluation:

• The quality and scope of available titles by discipline and audience level
• The relevance of available titles to RUL collections
• The sales and pricing models
• Archival options
• Integration with RU-ONLINE and IRIS

In order to assess these matters the task force will invite e-book providers to demonstrate their products to RUL librarians and staff. Be sure to consider input from faculty liaisons and staff whose workflow could be affected by e-books at RUL.

The task force should complete their final report by March 14, 2008 and put it on the agenda for discussion at the March 2008 CDC meeting. This should allow plenty of time to schedule vendor demos and to meet with e-book providers at ALA Midwinter.”

It should be noted that the original E-Books Task Force was formed in October 2007, and membership consisted of Karen Hartman, Mei Ling Lo, Laura Mullen, Mary Page (former Head of Acquisitions, and chair), and Cathy Pecoraro. Mary Beth Weber and Gracemary Smulewitz continued the task force’s work as co-chairs, beginning in October 2008. The task force was expanded to include Fay Austin, Sara Harrington, Elizabeth Leister, Jane Otto, and Robert Warwick.
**Collection Development Policies** - Selection criteria and a group specifically devoted to the selection and purchase of these resources are needed.

*Selection Criteria* - RUL needs to develop a standard set of criteria to apply to e-book offerings whether for a single title or packages. Consideration should be given to enhancements or limitations compared to the print version of a title, licensing, duplication with print or between packages, restrictions (number of simultaneous users, interlibrary loan, downloading). Evaluative criteria should also consider public services issues (instruction, reference service), technical services issues (cataloging, record maintenance), systems considerations (licensing, IP ranges, proxy servers, loading of record sets), pricing, digital rights management, currency and update, search interface, and customer service and support. This information could be incorporated into an online checklist for selectors. The existing documentation “Evaluation Criteria for Electronic Resources” could be expanded and updated to apply specifically to electronic books.

Duplication and overlap between packages should be avoided. Consideration should be given to the value of duplication, the workload required to make duplicate resources available, and the decision to limit selection to one format for a title (print or electronic, but not both if this can be avoided) must be given serious consideration.

Vendor considerations include the availability of slip plans for e-books and whether they can be integrated with an existing slip plan for print. Consideration should also be given to whether an option is provided to buy the first available format (print or electronic). Are titles profiled once when first available and regardless of format? Does the vendor treat e-books as another form of binding and include them in the general approval plan?

The impact of e-books on core collections is a factor. Will we need an approval plan if we get our core collections through packages? The ability of vendors to support labor intensive approval operations is a concern. If publishers generate revenue by selling e-book packages directly to libraries, this may affect approval plans in the future.

*Selection Strategy* - Selection strategy may be driven by particular needs (subject area, discount) and may vary with situation. The “critical mass” approach focuses initial e-books acquisition on building a mass of subject specific content that is large enough to guarantee intense usage. This approach lacks specificity, may be more cost effective, and typically has less restrictive digital rights management (DRM) policies for content. In contrast, single title purchases (also referred to as “pick and choose”) provides flexibility through targeted selection. It requires less initial investment, may incur greater overall costs in the long run since discounts applied to packages are not applicable, and DRM policies are often more restrictive.

In addition to packages or single title purchases, institutions have begun to explore the idea of patron initiated acquisitions. Under such a scenario, approval plans are used to identify appropriate titles that are available as e-books. Only titles that are used will be purchased. An argument used by those who support patron initiated acquisitions is that predicting use is difficult using traditional methods of collection development. The sentiment is that traditional collection development will lead to increasing patron dissatisfaction since patrons become aware of
resources not offered by the library that are available through Google, Amazon, Open WorldCat, etc. An additional consideration posed by a book vendor is that if a library purchases books that are not being used, they will not want to pay the 160 percent retail cost to purchase e-books that also will not be used. A major concern about this approach as expressed by Rutgers subject specialists is the fact that Rutgers must adhere to core curriculum standards as part of our accreditation. Patron initiated acquisitions will also not be readily accepted when our budget is limited and we adhere to core collection standards.

It may be necessary to provide access to more than one publisher’s platform since one vendor may not be able to meet all of our e-book needs. Purchases need to include consideration of a vendor’s ability to provide seamless access to multiple publisher platforms and the availability of federated searching when multiple platforms are used.

In all the preceding scenarios, RUL needs to determine the best approach. Are packages the best approach? It is hard to turn down packages because they are often discounted and include new academic titles which make up a significant portion of approval sales. However, packages frequently include titles that might not have been otherwise selected, and may not be heavily used. Packages may be a good option if the vendor enables us to swap out titles in exchange for others. Single title purchases are an attractive option, but may incur greater costs in the long run.

**Funding and Fees for E-books**- Our current selection budget does not include funds specifically earmarked for e-books. Selectors have used their own funds to purchase e-books (packages and single titles). E-books currently cost considerably more than their print counterparts. A decision must be made whether heavy and cross-campus demand can be satisfied with a single electronic copy (which ideally can serve simultaneous users), and if this is sufficient to justify the increased cost. Another issue to resolve is whether universally useful resources should be purchased with an individual selector’s funds or purchased by a subject team. Our fund code structure must be adjusted to manage e-books, and they should be easily distinguished from electronic serials and database purchases. Some of this can be facilitated through our approvals vendor and their online ordering interface.

E-books incur costs that are not associated with their print counterparts. Access fees and continuing costs are considerations when purchasing e-books. Perpetual access is not built into our current funding structure. This is complicated by the fact that perpetual access is defined differently by vendors and publishers. It may be a one-time cost or an annual fee. Related costs include “life of the edition” versus permanent access.

**Form a Group to set priorities for purchasing/licensing**- The Library Resource Council (LRC) may consider appointing a group devoted to the selection of e-books as well as other considerations such as cataloging and archiving. This could be a new group similar to the subject selector teams, this responsibility could be incorporated into the work of the existing subject teams, or a small subset of the E-Books Task Force could handle this work. The main point is that a qualified group of individuals takes responsibility for this work and that it is handled in a consistent manner. The purchase of e-books is costly and has major service and workload implications for the Libraries. E-book selection should conform to our general selection policies and the goals of the collection development librarians.
Internal support is needed before decisions are made. Statistics should be obtained from subject specialists before big purchases are made. Decisions should not be made in isolation or by limited groups of individuals. Policy changes should be planned and coordinated with subject specialists.

**Business Models**- Different business models should be explored before committing to an e-book purchase. Common considerations include ownership versus subscription and leasing content. Examples of different business models include:

- Subscription to copies of individual titles with access of single-user per copy and a premium for “ownership”.
- “Pay per view” in which the first viewing of an e-book is usually free with one click. Subsequent times may incur a vendor charge or the library must purchase the e-book.
- Flat monthly subscription rate for access to the entire database. This type of subscription is marketed to individuals, not libraries.
- Free browsing of the entire database with fees for printing and downloading. The vendor anticipates that libraries will establish accounts with a maximum amount to spend per user. The library receives 5 percent of the revenue generated from the library’s account.
- Annual FTE-based subscription model. The library pays for new content during the subscription year at renewal the following year.
- Short and long-term subscription to e-books.

Any vendor that we choose should be required to provide Counter compliant statistics, including the ability to compare our usage with other libraries using their e-books. Additionally, turnaway statistics are helpful to measure demand for titles and packages.

Vendor health should be examined before any purchases are made. E-books are a costly long-term investment.

**Access and Licensing**- Licensing, access and restrictions, and perpetual access should be clearly spelled out in agreements between RUL and e-book providers.

**Licensing**- Information about terms and conditions of use should be easy to interpret and understood by both library staff and users. Permissions and restrictions should be explicitly stated in the license agreement. RUL should be informed in writing regarding what controls the vendor uses to monitor post-access uses of e-books. Information for transmitting permissions and restrictions should be provided with each resource. Licensing should also address some assurance of reasonably controlled price increases that are communicated to RUL in a timely manner.

The vendor or provider should be able to guarantee their right to license all components of e-books, including text, image and sound.¹ A provision should be included to protect RUL from any consequences that could result from unauthorized use of a resource, or use that violates what is stated in the license.
**DRM/Permissions**- Permissions should provide access to complete content and access to a reasonable number of simultaneous users.

**Perpetual access**- Vendors must explicitly state how they define perpetual access and how it will be guaranteed. The original format of the publication should be preserved.

**Restrictions on Access**- It is preferable to select e-book vendors that do not impose interlibrary loan restrictions or restrictions on how much content (pages, chapters, etc.) may be downloaded and/or printed. The option to download an entire e-book to one’s desktop is also a consideration. Downloaded e-books should enable users to mark up, highlight and annotate content for their personal use in external software such as PDF.

RUL should have the option to print on demand and bind e-books on an as needed basis.

Metadata and other descriptive information that facilitates search and resource discovery should be available to all indexing and access services provided by RUL.

**Archiving**- The license and permissions need to consider long-term access, retention, and migration strategies as technology evolves.

**E-Book Sources**- E-books are available from a variety of sources, which are outlined below. Each source also includes mention of benefits and drawbacks.

**Book Vendors**- Book vendors partner with publishers and aggregators, and serve as a middleman for purchasing e-books. They are able to coordinate electronic and print acquisitions, which enables libraries to maintain their current workflows for approvals and selection. They permit title by title acquisition and provide good collection development support. They typically offer MARC records for the titles they offer. One drawback to using book vendors is that their partnerships with publishers or other businesses can fail. Partnerships may also produce a second middleman.

**Publishers**- Publishers provide direct purchasing of e-books. Their prices are not always the best, and better prices are typically negotiated from aggregators. Using a publisher may require numerous licenses and invoices, which presents bookkeeping challenges. Many publishers lack the ability to support title by title purchases, and provide limited support for collection development.

**E-Book Aggregators**- E-book aggregators provide a single search silo. Their selection is limited to title by title. They provide collection development support and may offer MARC records. They offer one license and fewer invoices as compared to publishers. The extent of the content they offer may be limited or questionable. E-book aggregators may require purchase of specialized readers to access their content, which is an additional cost. Specialized readers may not be interoperable with other resources. Lastly, e-book aggregators often may not provide usage statistics, which are critical for collection analysis.
Subscription agents- Subscription agents enable libraries to aggregate e-books with other electronic content. The quality of the MARC records they provide may be questionable. Most subscription agents offer subscriptions to titles, rather than permitting purchase. Their content may be limited, and collection development support is typically limited or nonexistent. Subscription agents may not coordinate electronic and print versions of titles, and may generate multiple licenses accompanied by their own set of DRM issues.

Free e-books- Freely available e-books provide another means of collecting these titles. Some of the major initiatives include the Open Content Alliance, the Universal Library (a continuation of the Million Book Project), Project Gutenberg, and Google Books. Separate reports on each of these projects are included at the end of this report.

Issues related to free e-books include:

- Deciding which free e-books to add to our collection. Who makes this decision? Policies are needed for adding free e-books to our collection.
- Should RUL take advantage of the opportunity to participate in any of these initiatives? Who makes this decision? What are the advantages and drawbacks of participating?

Cataloging- Cataloging considerations include quality of vendor supplied MARC records and their workload implications, and restrictions on uploading records to OCLC. Technical Services personnel with cataloging responsibility have developed draft guidelines to be applied when vendor supplied record sets are loaded into IRIS. A copy of that document is in the appendix at the end of this report.

Restrictions on uploading records to OCLC- If content is leased, rather than purchased, it is not likely that these records can be uploaded to OCLC and access to this category of e-books will be restricted to our integrated library system (ILS). Restrictions should be explicitly stated in licenses.

Availability of MARC records- The availability of vendor supplied MARC records and their quality will have workload implications. Some vendors provide free records while others provide them at a cost. If e-book titles duplicate print titles in our collection or overlap with other e-content, vendor records have little value for us. Vendor records should be full-level and will ideally conform to standards set forth in MARC Record Guide for Monograph Aggregator Vendors (http://www.loc.gov/catdir/pcc/sca/FinalVendorGuide.pdf).

Cataloging e-books is typically not a one-time deal. Different approaches to vendor supplied records will be used depending on the size and nature of collections or record sets. There are three possible approaches: batchloading of record sets, title by title cataloging, and merged format records. The first and third approaches require preprocessing standards and scripts executed by Systems to add desired elements (URLs with Rutgers specific notes, for example) and to strip unwanted fields and information. Electronic titles that duplicate print titles in our collection will need to be added to the records for the print in accordance to our merged format record policies. Title by title cataloging is best for small record sets (100 titles or less, for example). URL maintenance may be required, and content may need to be refreshed. Since
each record set will require different preprocessing, it is desirable to create a sample record for collections for future reference and to record policy decisions.

**Publicity-** How and where e-books are available must be determined. E-books must also be actively promoted to our user community. It is pointless to acquire and make available costly materials if no one is aware of their availability.

RUL needs to explore search and discovery options that extend beyond the ILS. A new section on the RUL website for e-books, similar to the Indexes and Databases page, is a possibility. IPAC’s new resources list includes e-books. New e-books could also be announced on the “News” portion of the RUL website introductory page. Information can be disseminated via departmental liaisons, as well as through other sources such as the Human Resources Weekly Digest.

**Evaluation-** Regular and ongoing evaluation of e-books is necessary to determine usage, effectiveness in terms of subject and content, cost effectiveness, etc. It is not practical to purchase and make available e-books without further evaluation and assessment.

*Counter compliancy*- Ideally, e-book products that RUL acquires will be Counter compliant so that we can run statistical reports. Vendors should be able to provide statistics that to enable us to compare our usage with peer institutions that are using their products. Statistics will also help us to analyze peer institutions’ collection development practices and subject strengths.

*Regular and ongoing evaluation*- Evaluation should also include comparison of similar titles and packages with an eye towards eliminating redundancy and overlap.

Obtaining feedback from users within three to six months of e-book purchases is critical to plan renewals and future purchases accordingly. Pop up surveys, surveys targeted to specific groups, or a survey on our website is one option. Departmental liaisons can also obtain feedback from their respective departments.

**Conclusion-** The E-Books Task Force has done preliminary research and some e-book packages have been ordered. Next steps include development of formal collection development policies specifically for e-books. Consideration must be given to specific business models, and one approach may not be appropriate for all our needs. A group to select e-books and to develop policies is needed. Free e-books present different concerns, and this issue needs to be more fully explored. As we, as an institution gain expertise with e-books selection, licensing, cataloging, etc., our policies will be more fully developed and applicable to a wide range of situations.
Appendix
Free E-books Initiatives and Projects

1. Open Content Alliance http://www.opencontentalliance.org/about/

The Open Content Alliance (OCA) is a collaborative effort of a group of cultural, technology, nonprofit, and governmental organizations from around the world that helps build a permanent archive of multilingual digitized text and multimedia material. An archive of contributed material is available on the Internet Archive website and through Yahoo! and other search engines and sites.

The OCA encourages access to and reuse of collections in the archive, while respecting the content owners and contributors. Contributors to the OCA have agreed to the principles set forth in the Call for Participation.

The Open Content Alliance is administered by the Internet Archive, a 501c3 non-profit library. Newest contributors including the Boston Library Consortium can be found here: http://www.opencontentalliance.org/contributors/

2. The “Universal Library” (a continuation of the Million Book Project)

The Million Book Project (or the Universal Library), led by Carnegie Mellon University School of Computer Science and University Libraries, is on the mission to create a Universal Library which will foster creativity and free access to all human knowledge. As a first step in realizing this mission, it is proposed to create the Universal Library with a free-to-read, searchable collection of one million books, available to everyone over the Internet. Within 10 years, it is our expectation that the collection will grow to 10 Million books. The result will be a unique resource accessible to anyone in the world 24x7, without regard to nationality or socioeconomic background. As of today, they have completed the scanning of more than 1 Million Books and have made accessible the entire database from http://www.ulib.org.

3. Project Gutenberg (also referred to as PG)

PG is the oldest e-book provider (started in 1971 with the keying in of “The Declaration of Independence”), and now has approx. 32,000 e-books available for download through its website. Volunteers from all over the world copy or scan books that are available in the public domain. PG has strict guidelines, especially about copyright verification.

E-books are available in 21+ languages (PG. has mirror sites in several countries, including Canada, Australia). PG maintains a test database of MARC records created by volunteer catalogers. The database is constantly being updated.

Many of the e-books created by PG are available through NetLibrary. NetLibrary maintains a collection of free-e-books and its co-operation with PG is noteworthy.
PG also features recorded books (audiobooks) and several book collections through the related project: Project Gutenberg Consortia Center which increases the number of e-books available for download to 75,000+. (See http://gutenberg.us/Collections.htm)

Texts are copied in a plain vanilla ASCII format, while other formats such as HTML are also available for download. These options provide flexibility for the user since no special proprietary reader is required.

PG is committed to providing free e-books. E-books are freely available and can be downloaded without restrictions. They have recently introduced downloads for mobile devices.

Libraries can selectively provide URLs on older catalog records for public domain titles. This can be done either as a link to the PG website, or as a link to the individual title located there. Using links to PG for RUL titles in the public domain provides flexibility for the student or user who may have discovered the title through NetLibrary on campus, but does not realize that the title is also available for free use and download. Titles are in several languages and the project is truly international in scope and philosophy.

PG is involved in cooperative ventures with groups such as the Creative Commons, the World Public Library, NetLibrary, etc. suggest that the project will continue to grow and to be maintained.

Some of the disadvantages of PG include the fact that offerings are limited to public domain titles. Though, increasingly, authors are giving permission for their copyrighted material to be copied and made available via this project. Changes/challenges to copyright law may restrict what is available in the public domain.

4. Google Books

Available books include “fiction, non-fiction, reference, scholarly, textbooks, children's books, scientific, medical, professional, educational, and other books of all kinds” in English and other languages. Browsable categories are: fiction, literature, science fiction, fantasy, romance, mystery, fairy tales, short stories, poetry, nonfiction, philosophy, economics, political science, linguistics, mathematics, physics, chemistry, biology, random subjects (naturalists, icons, real estate business--United States, landscape painting, semantics, princesses, friendship fiction, wit and humor). Seven million books with full-text searchable (only public domain titles are downloadable) are available.

Each book includes an “About this book” page with basic bibliographic data (title, author, publication date, length and subject). Additional information is available for some books, such as key terms and phrases, references to the book from scholarly publications or other books, chapter titles and a list of related books. There are links to bookstores and libraries for every book.

The full view option permits readers to view any page (contingent on permission/copyright status); users may download/print/savable in pdf if the book is in the public domain. (Note: Not all public domain works have the “Download” button, but Google is working to include it on all
public domain works). The limited preview option is limited a number of pages based on permissions. The snippet view option provides a “card catalog” description plus a few “snippets” (sentences showing search term in context). The no preview available option provides basic “card catalog” information about the book. Sample screen shots are available at \textit{http://books.google.com/intl/en/googlebooks/screenshots.html}.

In addition to straight-ahead digitization, Google provides, for each book, a reference page with book reviews, web references, integrated Google maps, etc. Examples are available at \textit{http://books.google.com/books?id=yNFn1OpnkBkC}. There are zoom and navigation features, as well as configurable page views.

Google Book Search retrieves books based on ISBN, LCCN, and OCLC record numbers to enable links to specific editions. Google Book Search URLs follow a pattern which facilitates automated linking to specific parts of a book (front and back covers, title page, copyright page, table of contents, index).

Search results may not be limited to downloadable books. However, searches may be limited to books with “full view” availability, using the Advanced Search feature of Google Book Search. Users must go into each individual record to determine if it is downloadable. A project to identify and link to downloadable titles could be doable, but currently there is no notification service to announce when new titles are available. A subscription service is planned once the agreement has been approved.

Users can make available Google Book Search (with free searching) directly from their own sites via Google Book Search APIs. Dynamic links information is available at \textit{http://code.google.com/apis/books/docs/dynamic-links.html}.

As part of its Libraries Project, Google is exploring the expansion of its program to “include special collections from libraries both in the U.S. and other countries. If you want to let us know about your library's special collection, please feel free to contact us and include the size of your collection, specialization or unique content, how much of your content is already digital, and what languages it includes. You can also visit this page \textit{(http://books.google.com/support/bin/answer.py?answer=43741&topic=9082)} to learn more about who we're working with, what we're digitizing, and how library books appear in the program.”

On October 28, 2008, a settlement was reached between Google, the Authors Guild, and the Association of American Publishers. It is awaiting approval by a United States district court judge (the hearing will take place sometime after May 5, 2009), and an approved settlement is expected to bring the following changes to Book Search:

- the Book Rights Registry, an independent, not-for-profit organization, will be established to represent authors, publishers and other rightsholders, and to help locate rightsholders, ensuring that they receive the money their works earned under the agreement.
- Institutional subscriptions to libraries, universities, and other organizations. Public and university libraries in the United States will also be able to offer terminals where readers can access the full text of millions of out-of-print books for free.
More in-copyright but out-of-print books will be available for preview, reading and purchase in the United States.
Fee-based full online access will be available to millions of books
Academic researchers will be able to apply through an institution to run computational database queries through the index without actually reading individual books.

An Institutional Subscription Database (ISD) will be established, to include books in the “in-copyright, out of print” category. Authorized users within an institution (e.g., faculty, students, staff) can view full text of ISD books for the duration of the subscription (note: not perpetual access). Authorized users will be able to annotate books, share links to books for course use, and have limited copy/paste and print capability that makes printing the entire book difficult, but doable. Google can subsidize purchase of institutional subscriptions by participating and cooperating libraries.

Subscription pricing will be set by Google and the BBR (subject to binding arbitration if they cannot agree), balancing “realization of revenue at market rates” and “realization of broad access to the books by the public, including institutions of higher education.” A number of factors will be included in determination of price, and the initial pricing strategy will probably be in effect for two to three years.

Under the settlement, Google would create four categories of partner libraries, each with its own privileges and responsibilities, including (in some cases) security obligations: fully participating libraries, cooperating libraries, public domain libraries, and other libraries. For a detailed description of rights and responsibilities, see Jonathan Band’s A Guide for the Perplexed: Libraries and the Google Library Project Settlement: http://www.arl.org/bm~doc/google-settlement-13nov08.pdf.

“In order to protect copyrighted books, we only allow Google Book Search users to view a limited portion of the book. Enforcing these limits requires us to keep track of our users' page views. We don't associate your searches or the pages you view with any personally identifiable information about you, such as your name and address, unless you login with your Google Account. However, to enforce limits on user page views, we do associate some non-personally identifiable information (IP address, cookies) with the books and pages that you've viewed regardless of login. We keep this data for a limited period of time to enforce viewing limits. As always, we encourage you to read our Privacy Policy to be fully informed about how your confidentiality is protected.” The Privacy Policy is available at http://www.google.com/privacypolicy.html.

Questions

Is there a way, possibly with ISAWG assistance, to identify titles available in Google Book Search (and other open content projects)? If so, would we link only when full text is available, or only when full-text is downloadable?
Would we link to non-matching editions?
Should we explore “alternatives to linking at the record level” (project expansion to special collections, Google Book Search links from Libraries website)?
How interested are we in the institutional subscription, should it become available? Would we want to participate in the scanning, or explore consortial participation?

Sources
Google Book Search Settlement Agreement (http://books.google.com/googlebooks/agreement/)
About Google Book Search (http://books.google.com/intl/en/googlebooks/about.html) and its linked pages
ALA’s Google Book Settlement website: http://wo.ala.org/gbs/
Evaluating Bibliographic Record Sets for Electronic Resources

[Draft]

May 2009

The following are factors to consider when evaluating record sets for e-resource batch loads. This is not a linear process and every question will not be relevant for every record set; however this document can be used as a checklist to ensure that all issues have been taken into account in making a decision.

There are five major aspects to consider:
1. Number of records
2. Quality of the records
3. Matching on existing records
4. Preprocessing scripts
5. Sending new records to OCLC

1. Number of records

- **How many records are in the set and is it worthwhile to modify them?** If the expectation is we hold most of the titles in print form, you may be able to obtain a spreadsheet from the vendor with titles and URLs which can be added individually.

- **What is the expected outcome for the record load?**
  - Will this be an ongoing process or is it a one-time batch load? If ongoing, record quality will be a primary concern.
  - Is the batchload intended to add electronic access to existing print records? If so, which fields will be added to the existing records (often just an 856, and possibly an RU-ONLINE holding). Thus, record quality would be a lesser factor, and matchpoint will be a major factor.
  - Is the batchload meant to add new content? In this case new records will be added. Thus, record quality will be a major factor.

2. Matching on existing records

- Is there a match point for the title (if we have one, see below)? In order to automatically link each record to our existing bib record for the title (if we have one; see below), there must be an identifier, typically an OCLC record number in the 001 or 035, or an ISBN (020) or ISSN (022). The LC control number (010) may also be an option.
- If there is no match point, consider the pros and cons of allowing all the records to sit side-by-side with our existing records.
- How likely is it that we would need to match existing IRIS records?
  - Estimate how many (approximate percentage) by searching a sample.
  - How often will the records in the incoming record set describe the same format as the
format described in the existing SIRSI records?
• Will the available matchpoint(s) enable the records to correctly merge?

Search a subset of the records in OCLC to see if the 035s correspond to the print, the electronic version, the microform, or a combination of the above. Records loaded to OCLC will match on the OCLC number regardless of whether the record describes the appropriate format. For example, a problem would arise if the control number (035) is for the OCLC print record, but the vendor has modified that record to make it describe the electronic resource. This would cause our holding to attach to the print master record if the titles were sent to OCLC.

3. Sending new records (the ones that did not match) to OCLC

• Are there license restrictions (i.e. does record provider hold a copyright on the records)? Check the license.
• Is this a static package loaded one time, or a package with ongoing additions, changes, or deletions?
  o Do we have perpetual access? If no, lean toward not sending.
  o Is the content leased rather than purchased? If yes, don’t send.
  o Are the records leased rather than purchased? If yes, don’t send.
  o Is the bib record likely to go away? (For example, for packages subject to change, the vendor might pull an entire set of records and replace it with a new set, rather than updating each record individually; this is more likely with serials.) If yes, don’t send.
  o Is the record set finite, or will new records be added periodically? If it’s not finite, contact the vendor to determine if there’s an alerting service for additions to the package.
  o If holdings are subject to change, is it prohibitively expensive to update our holdings in OCLC? (This is more likely with serials.)

NOTE: The trigger to send to records/holdings to OCLC is changing the Date Cataloged to the current date. This same trigger sends the record to LTI. If the records should not go to OCLC, but they should go to LTI, contact Bob Warwick.

4. Evaluate the record quality

• Were the records cataloged by the Library of Congress or by a PCC library? If not, is it a trusted cataloging source?
• Has authority work been done on the headings? What is the quality of the authority work?
• Examine the first few records in full, carefully, for idiosyncrasies or anomalies
• For a larger subset of the records (not all), review the following fields to determine if cataloging is in accordance with our practices:
  o 245 |h
  o 530
  o 533
  o 534
5. Determine pre-processing needs (fields to be deleted, retained, evaluated, or edited) when adding new records

NOTE: All changes will be automatic, and must be valid for every record.

- Evaluate:
  040 |dNjR (if meaningful change on every record)
  260 (is it for the original publisher or the e-resource provider? Is it consistent with the 530/533 and 245 |h?)
  530
  533
  534
  590 (for rare books; otherwise, delete)

- Delete:
  037 [?]
  050 if second indicator ‘4’
  090
  506
  655 value of Electronic books, Electronic serials, Electronic resources, etc.
  690 value of Electronic books, Electronic serials, Electronic resources, etc.
  710 for the vendor
  84x
  85x except 856
  86x
  87x
  882
  886
  887
  Obsolete fields (e.g., 256; refer to MARC21) except 440
  Any local field (generally contains a ‘9’, e.g., 938)

- Retain identifiers and control numbers, including 015, 016 fixed fields, access points other than those listed above, even where they would not be added to our original cataloging as long as they are appropriate for the item in hand, including:
  02x
  041
  043
  044
  050 00
082
440 (even though obsolete)
77x
78x
800
810
811
830
880

- Edit:
  856 |z (value should be changed in accordance with our policy, typically “Access from campus or login via Rutgers account”)

Be sure your preprocessing includes adding an RU-ONLINE holding where appropriate. (Every record with a link to the full text should have one RU-ONLINE holding, and only one.)

The option exists to send records to a review file (matched records, unmatched records, or the entire set).

If the batchloading process unavoidably produces duplicates, determine if
  - Duplicates will be resolved in each case
  - Duplication will be resolved only as encountered
Notes


Resources

Morris, Carolyn, “eBook Workflows Presentation, Rutgers University, March 4, 2009” [PPT presentation given to the E-Books Task Force].