MARC to ENC MARC: bringing the collection forward

Janet Kahkonen Smith
Roger L. Cunningham and
Stephen P. Sarapata

The authors

Janet Kahkonen Smith is the Digital Library Catalog Coordinator, **Roger L. Cunningham** is Senior Systems Development Engineer and **Stephen P. Sarapata** is Senior Software Development Engineer, all at Eisenhower National Clearing-house, The Ohio State University, Columbus, Ohio, USA.

Keywords

Standards, Administrative data processing, Archives

Abstract

This paper will describe the way in which the USMARC cataloging schema is used at the Eisenhower National Clearing-house (ENC). Discussion will include how ENC MARC extensions were developed for cataloging mathematics and science curriculum resources, and how the ENC workflow is integrated into the cataloging interface. The discussion will conclude with a historical look at the in-house data transfer from ENC MARC to the current production of IEEE LOM XML encoding for record sharing and OAI compliance, required under the NSDL project guidelines.

Electronic access

The Emerald Research Register for this journal is available at
www.emeraldinsight.com/researchregister

The current issue and full text archive of this journal is available at
www.emeraldinsight.com/0737-8831.htm

Library Hi Tech
Volume 22 · Number 1 · 2004 · pp. 28-39
Emerald Group Publishing Limited · ISSN 0737-8831
DOI 10.1108/07378830410524477

Directed under federal legislation of Public Law 101-589, Section 204, as part of the United States Department of Education's Excellence in Mathematics, Science, and Engineering Education Act of 1990, The Eisenhower National Clearing-house (ENC) supports classroom learning and the educational community. ENC acquires materials, creates and organizes catalog records for these resources, and provides public access to the collection via an online searchable database. Materials are collected from federal and state agencies, commercial publishers, professional organizations, local school districts, and individuals. ENC's goal is to facilitate educational reform by providing a collection of exemplary curriculum resources for mathematics and science K-12 learning. In addition to maintaining an online catalog of resources, ENC makes available for inspection and evaluation an on-site repository of physical materials.

Statement of work

ENC materials are described and cataloged with extensive content and bibliographic data by an integrated team of experienced mathematics and science professionals and cataloging librarians. The richly described resources, both digital (Internet-based) and non-digital (print, videotape, CD-ROM, graphic material, DVD, and multimedia kits) are entered into the ENC-developed cataloging database. The metadata schema chosen by ENC is based on the USMARC cataloging framework. Added-value field extensions, developed by ENC for educational metadata, have been incorporated into the cataloging framework.

Cataloging, using the long-standing standard USMARC and ENC MARC field extensions, has resulted in a collection of richly described curriculum resources. Each catalog record is divided into fields that are prefixed with a three-digit tag and a single character sub-tag

Received 1 September 2003
Revised 22 September 2003
Accepted 7 November 2003
The authors wish to thank Kimberly S. Lightle, PhD and Judith S. Ridgway for their support in the development of this paper.

Library Hi Tech
Volume 22 · Number 1 · 2004 · 28-39

that identifies the data that follow. There is a field for author, a field for publisher, a field for edition, etc. A 245 tag, subfield a, for example, is always identified as the title field. The MARC records are stored, read by a computer program and given instructions to display the information contained in each tag on the online catalog. Programs have been purchased and written to assist the computer to search and retrieve the MARC records by browsing certain fields.

Once the statement of work from the Department of Education was established, ENC began to plan how to locate, acquire, organize, describe, display, and maintain the collection of educational resources, and how to provide public access to them.

The statement of work and the MARC framework spelled out in general what kind of data would be included in the bibliographic record created for each resource. Additionally, ENC wanted teacher input into what they thought were important fields of information to include in each catalog record. What exactly did educators feel was important to them for search and retrieval purposes? Deciding what kind of data to incorporate into the catalog record after receiving feedback from staff and the educational community was ENC's next task.

The ENC cataloging framework was structured to hold bits of information about each curriculum resource and was developed with many considerations in mind. The purpose of the record was to identify the materials in the collection, make them respond to an online search and able to be retrieved by users. The discussion at ENC centered on what kind of information should appear in the catalog records beyond what was dictated by MARC, who would input the data into the cataloging tool, and how the record should look and act online. At this point in the development of the ENC project, the technology needed for searching and display of content-rich Web pages was very limited. The Internet at the time was also limited, primarily being used by universities and governmental bodies. Web page design and browser capability were limited to HTML use and, though this environment worked well for sharing information that seldom or never changed, it did not lend itself well to the advanced searching capability required at ENC.

It was determined that the basic bibliographic record should include the data found in standard library-based cataloging: title, author and publisher, to list some. In 1993 ENC decided to adopt the USMARC framework as a starting-point for cataloging because MARC most closely fit ENC's cataloging requirements: a traditional library cataloging schema that could be extended with additional data fields as ENC required. ENC felt that the full description of educational resources required the addition of field extensions to the MARC record, extensions that would give the ENC-cataloged record a particularly educational feel.

The original cataloging fields used were dependent on the fields built into the Dynix Marquis (n.d.) cataloging system, and followed the USMARC tagging schema (see Table I).

ENC soon felt the need to expand on the standard MARC fields above, because the MARC fields, while fairly comprehensive, did not address the educational scope of the ENC collection. Additionally, the educational data were what the ENC user community thought were important fields of information to include on the materials' records. In 1996, concurrent development of the database and the cataloging tool was begun, and the following field extensions in Table II were added to the cataloging tool.

Soon after deciding to use MARC, ENC chose another standard from the library field,

Table I Dynix USMARC tags used by ENC

Tag name	Tag	Tag name	Tag
Date	008	Abstract	520
ISBN	020	Grade level	521
ENC number	090	Equipment	538
ACQ number	092	Language	546
Title	245	Funding agency	599
Variant titles	246	Subjects (base set) (ENC)	650
Edition	250	Geographic focus	651
Publisher address	260	Audience	699
Specifications	300	Author/personal names	700
Series	440	Author/corporate names	710
Public notes	500	Author/conference names	711
Contents	505	Exemplary	770

Note: subfields are not shown

Library Hi Tech
Volume 22 · Number 1 · 2004 · 28-39

Table II ENC field extensions

Exemplary	770
Revision date	009
ENC record of supporting materials	596
Vendor contact	597
Standards	658
Pedagogical type (non-Web type)	695
Benchmark analysis	696
Empirical data	697
Review	698
Online address	856
Special ENC projects	996
Keywords	997
End of the record	999

the *Anglo-American Cataloging Rules* (*AACR*) (American Library Association, 1967). *AACR2* provides an extensive set of cataloging guidelines. The goal was to create a catalog record that would truly describe the curriculum resource. The record is a container for the highly granular data necessary for optimal search and discovery of the resource.

Technology

In 1994, after the initial research was done, ENC technology staff put together a client/server version of a product named *FileMaker* (FileMaker, 1994), which used a protocol developed by Apple Computer (AppleTalk) to communicate with the desktop computers at ENC. The system was slow with known database size constraints; however, this system allowed staff to move forward with the early creation of bibliographic records.

The library automation system chosen was a commercial system provided by Dynix Marquis Inc. It comprised a Sybase database with access via an interface running on OS/2 workstations. It soon became apparent that, while the Dynix Marquis system could handle a large number of records, as well as the fully defined set of MARC tags, it was lacking in both extensibility and accessibility. These were critical issues impeding the progress of staff in getting the cataloged records on to the public site.

During this time, some discussion revolved around the evaluation of a product based on the Z39.50 Information Retrieval Protocol. After a considerable amount of time experimenting with this product, ENC decided against its use, instead directing efforts to a new relational database product named Microsoft SQL Server, version 6 (MSSQL6), released by the Microsoft Corporation. This product was marketed as a small business alternative to the more complex and expensive relational database systems in use at the time. Three features made the MSSQL6 of interest to ENC: the databases and server were easy to maintain, the system made use of the American National Standards Institute (ANSI) SQL, and the ability of the system to use non-proprietary Open Database Connectivity (ODBC) Drivers. As a result of the evaluation of the system and the goals set for the organization, ENC decided on developing its own cataloging system with MSSQL6 as the database backend. The front end of the cataloging system would be developed using a product named PowerBuilder, version 4 (Pb4) from PowerSoft Corporation (n.d.), which was chosen due to the wide variety of database interaction objects available with it.

The next step at ENC was to analyze the workflow. The workflow encompassed many activities:

- acquiring materials;
- describing and indexing them;
- making a set of information for each record available on a public Web site in a way that could be displayed, searched, and retrieved; and
- maintaining a collection in-house of physical library resources.

The working queues, then, had to allow the flow of materials through a multi-step process. Breaking down the workflow, categories of work were identified: selection of materials, ordering and acquiring materials, descriptive (content-based) and bibliographic (traditional library) cataloging, abstracting, editing the fields of data, making the records "live" on the Web site, and maintaining an onsite repository of this collection of materials. ENC materials, therefore, would be required to move forward in a linear progression as they followed the workflow, allowing one staff member to work on one record at a time but also allowing any staff member to view the record at any time.

Writing the programming instructions for the workflow for single-item records (a book or a

Library Hi Tech
Volume 22 · Number 1 · 2004 · 28-39

videotape, for example) did not present a problem. For the technical team, the programming did become a challenge when a single record was to be a container for a resource with multiple pieces, i.e. a multimedia kit that included several textbooks, teacher guides, an accompanying CD-ROM, and a classroom poster. It was necessary that the predominant, most important piece in the kit should be made the primary item in the record and this would determine the GMD (General Material Designator) for the record. The GMD was important, because an online user can limit a search by the media type of the GMD; thus, an incorrectly assigned GMD can fail to retrieve all relevant records. Technical staff worked on how to proceed with programming the application tool, PowerBuilder 4, to meet ENC's workflow requirements and special cataloging needs. Staff named this internal cataloging tool QCat, after the multiple working queues.

Materials workflow

Establishing workflow queues as part of the cataloging tool became increasingly important as the steps in the processing of materials became more complicated. In order to collect and control the materials arriving at ENC, a workflow system had to be set up that would structure staff with very different tasks: materials selection, materials acquisition, record building, content editing, bibliographic cataloging, abstracting, and editing. Interspersed among these workflow queues were built-in quality control queues to help staff maintain the highest quality of attention and treatment possible of each record. The quality control queues were juxtaposed after the cataloging, abstracting and editing queues.

The processing of materials through QCat relied on an integrated team with very specific job duties. Owing to the subject-specific focus of the ENC collection, a team of math and science professionals, capable of understanding the nature and content of the materials being acquired, was necessary to develop the collection and abstract each resource. The acquisitions staff, responsible for contact with publishers and vendors in the ordering and follow-up stages, starts the initial processing of resources by building the record in QCat that will follow the material as it works its way through the working queues. A team of librarians catalog the bibliographic data, abstractors analyze the content of each resource, and editors proof the final records before they become "live" for public viewing.

The primary consideration for materials being reviewed for the collection was content validity and how useful the resource was to educators. The following are questions that the content specialists posed when reviewing potential materials:

- Who is the intended audience?
- Does the resource provide outstanding math or science?
- Is the content accurate?
- Is the material equitable with regard to culture, gender, race, religion, etc.?
- Are viewpoints balanced?
- Are the authors or publishers reputable?
- Is the physical format of good quality and usable in the classroom?
- How well-known are the sponsors of the material, the credentials of all contributors, and what are their affiliations?
- Is the material engaging and age-appropriate?
- Does the resource promote interaction and learning?
- Is the information current?
- Are the illustrations of good quality?
- Is opportunity presented for innovative teaching and learning in the classroom?

With confidence in the structure of the catalog record and the refining of the workflow queues, ENC was ready to focus on developing the collection. Parameters were set forth in the statement of work to build a library of science and math curriculum resources for K-12 classroom use and for professional development of teachers. The initial collection of traditional library formats, books, videotapes, CD-ROM, graphic materials, multimedia kits, and objects was processed easily and posed no problems with respect to repository management. With the growing presence of the Internet, and the availability of Web-based resources and downloadable software applications for classroom use, it was desirable for ENC to consider acquiring these newer formats of curriculum resources. Virtual resources, by

Library Hi Tech
Volume 22 · Number 1 · 2004 · 28-39

their very nature, are challenging with which to work and require some special considerations when acquiring, collecting and maintaining them in what had been a traditional physical repository. Thought was given to how these new formats could be optimally stored, retrieved, and examined by educators.

The workflow at ENC originates with the mathematics and science content specialists, who work under parameters set forth in the ENC Collection development policy. It is their primary responsibility to locate materials through a variety of avenues: publisher catalogs, online sources, professional workshops, educator suggestions, conferences, and through professional reading. On initial selection, the content specialist assigns the materials a general subject designation: mathematics, science, general education, or an interdisciplinary subject that combines math and science. Further along in the workflow, the abstractor will assign subject headings that are more specific. The content specialist assigns a pedagogical type and grade level. The resource is given a priority ranking (the urgency with which the material should make it through all the processing queues), from zero (generally unsolicited materials), to four or six (processing at a steady rate), to ten (high attention and processing rate). A priority ten ranking indicates that the material has been targeted for a special in-house project at ENC. Special projects have been set up for resources destined for Digital Dozen, ENC Focus, and other ENC projects. As the content specialists complete this phase of the work, the requests for materials go to the acquisition staff, whose responsibility it is to start the ordering process.

The acquisitions team is responsible for getting new resources to ENC. Requests for materials can come from any of the mathematics or science specialists on the staff. The acquisitions team researches Web sites, publisher catalogs, and product literature for each item request, then creates a shell record with a title, subtitle, series name, GMD, pricing, ISBN or other ordering numbers, and general subject identifier, to hold a place in the cataloging database until the item arrives. Once the order is successfully placed, dialogue is maintained with vendors via telephone, e-mail or facsimile as needed. Once the material is unpacked and verified against the order, it is barcoded, repackaged, and labeled. Additional data can be added to fields in the previously-built record as needed, for example, special handling instructions or ordering updates.

The resource is then forwarded via the online database to the content queue. Either the mathematics or science content team member examines and verifies that the item is what they requested. They are responsible for adding additional data to the record by assigning pedagogy, rechecking the designated subject field (M for mathematics, S for science, G for general education, I for integrated/ interdisciplinary approaches, and E for education technology), grade level, special project status and priority, and, in the case of Internet resources, making sure that the URL points to the appropriate resource. Once again, the priority number that the resource is given may determine how the material is selected and cataloged.

A cataloger selects the material from the queue and adds bibliographic data to the record. Some of the data are already filled in on the record by acquisition or content staff; catalogers review and modify these data as needed. Cataloging fields include title, subtitle, variant titles, edition, and series. Catalogers also assign authors, publishers, and other contributor roles for those responsible for the creation of the resource. Catalogers add publishing location, copyright information, and dates associated with the item. At ENC, funding data are important. In this locally-created field, catalogers assign those individual or groups responsible for funding sponsorship. This is an especially important field associated with ENC mission and goals; it shows to the Department of Education that ENC is fulfilling its governmental contract (see Figures 1-7 to view various screen shots of the ENC cataloging database).

Catalogers provide a detailed description of the resource that includes pagination, illustrative material and physical dimensions. The resource is also identified by any distinguishing properties it may have, i.e. running time for videotape, CD-ROM, and DVD; any equipment requirements or downloads; or downloadable equipment for the

Figure 1 ENC cataloging maintenance manager

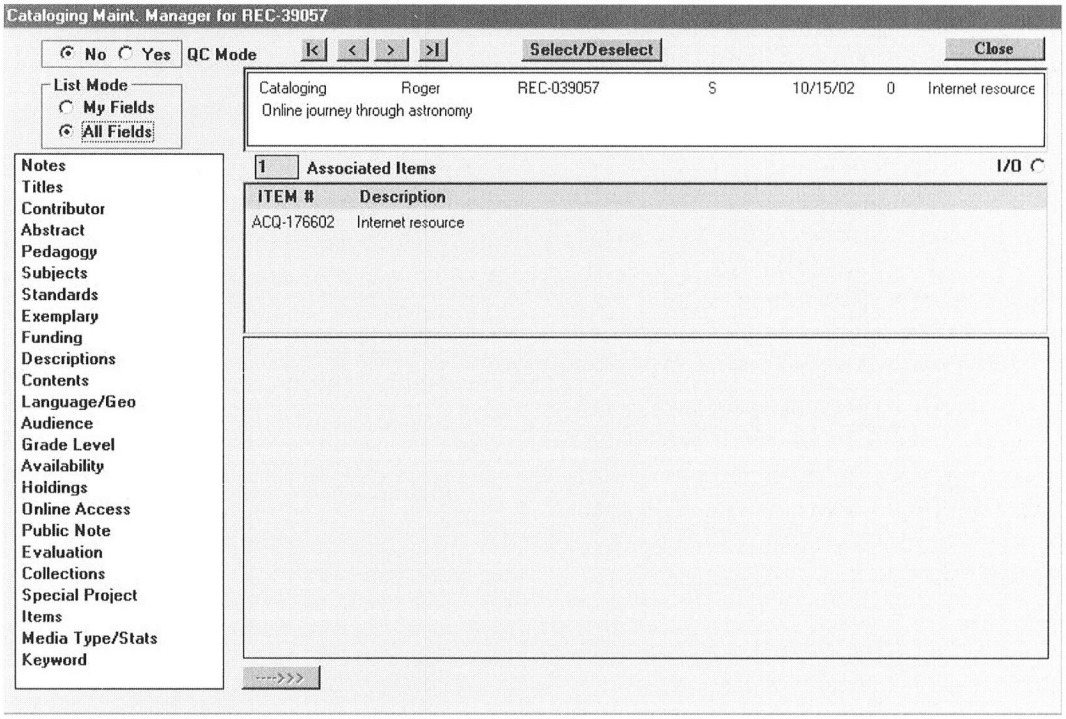
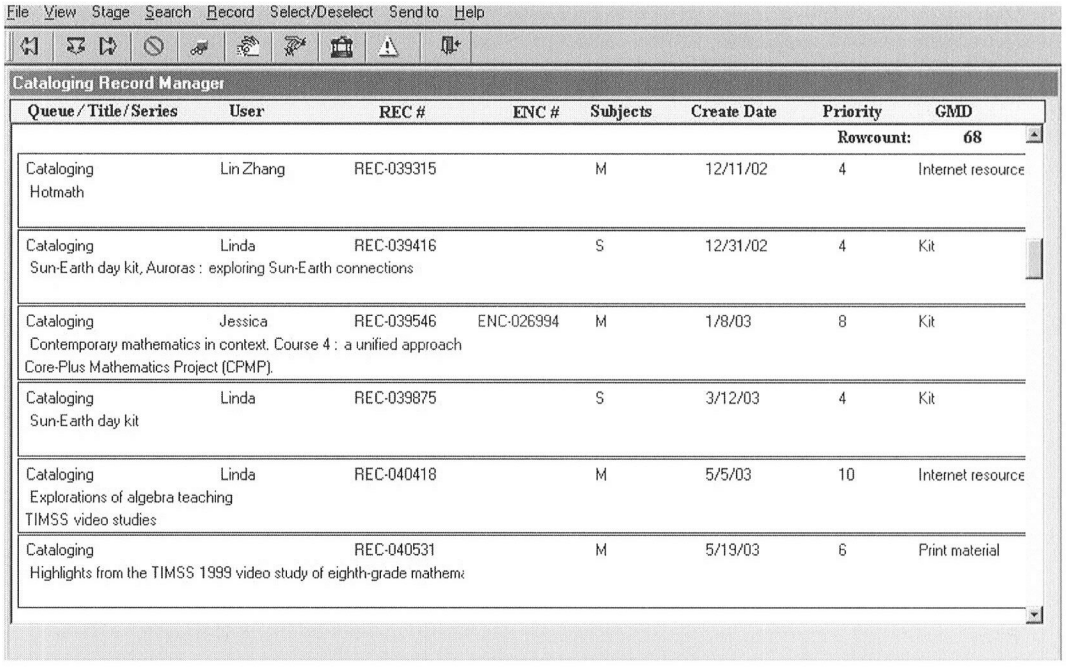
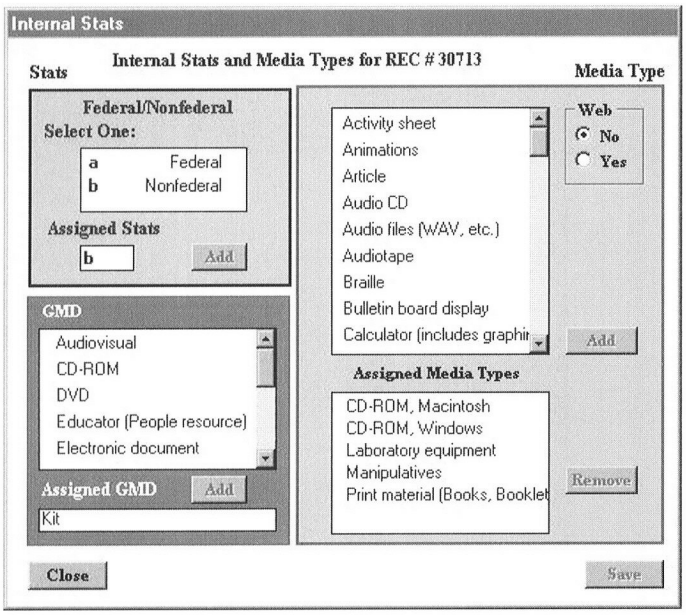


Figure 2 ENC cataloging record manager displaying cataloging queue and record owner



resource to be accessed, viewed, or displayed by the user. If a resource has a table of contents, catalogers add this to the contents field. The table of contents, which is a searchable field in the public catalog, provides valuable keyword search terms that aid in the discovery of the resource. Other cataloging fields include: language of the resource, geographic focus, availability notes, public notes, and supporting records (used to point to other materials in the

Figure 3 ENC internal statistics, general material designator (GMD), and media types



ENC collection that support the usage of the first record, i.e. a teacher guide for a set of videotapes, a calculator that support the use of a student workbook, etc.). Catalogers review the number of items in the record, the barcode, the GMD, the special projects, and then assign the media type, the technical format of the resource, and federal statistics. Once the cataloging is complete, the resource moves forward to the cataloging QC (quality control queue), then on into the abstracting queue.

The abstractors are responsible for helping the content team identify new resources to include in the collection. They assign appropriate subject headings and identifiers to the catalog record. In addition, they compose a 200-300-word descriptive abstract and select the curriculum resources that are the basis for all the ENC-based special projects. The abstractors identify standards, exemplary criteria, and state correlations, and add these fields of data to the cataloging database. From the abstractors, the record passes to abstract QC for review.

As both the back-end of the QCat database and the front end of the cataloging interface were being concurrently developed and refined, MSSQL version 6.5 was released. A decision was made to delay a server upgrade until the second service pack was released by Microsoft. During this intervention, a new version 5 of PowerBuilder became available. Once service pack number two was released, both the database and the partially developed client interface were converted. The server and

Figure 4 ENC funding screen

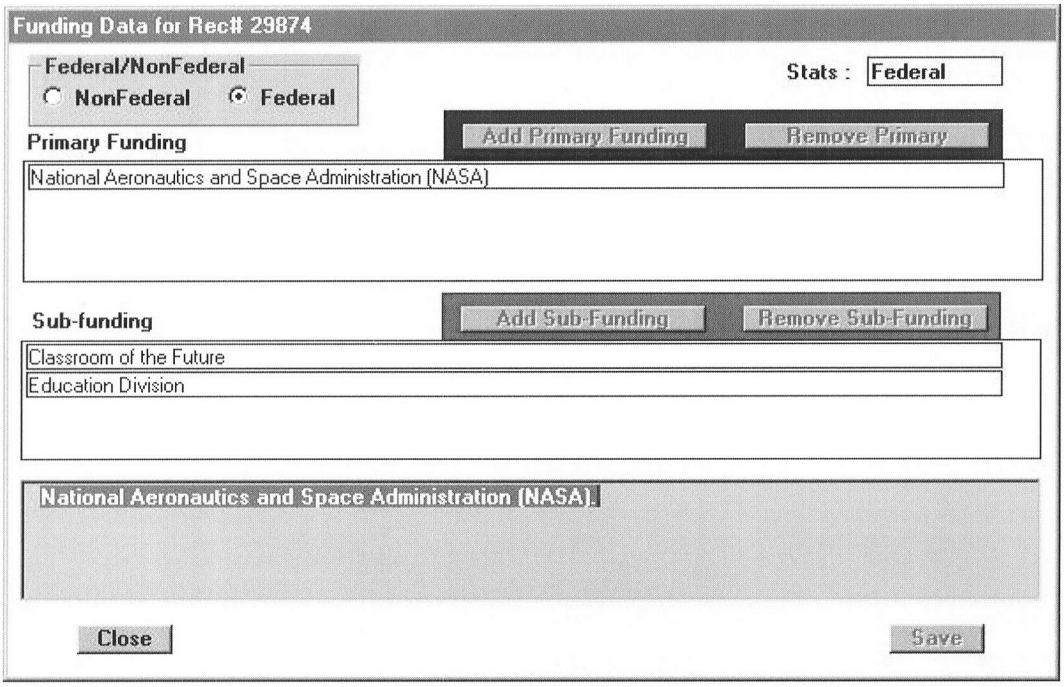


Figure 5 ENC specifications and equipment screen

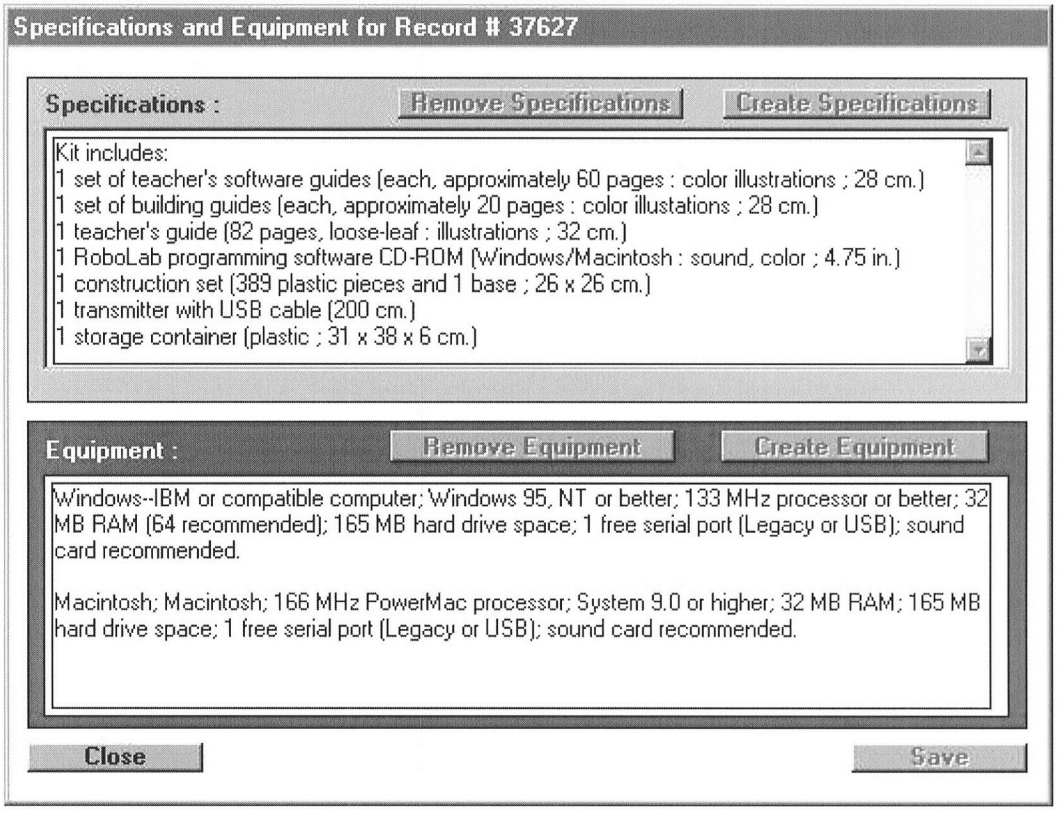


Figure 6 ENC contributor screen displaying contributor roles, names, and publisher location

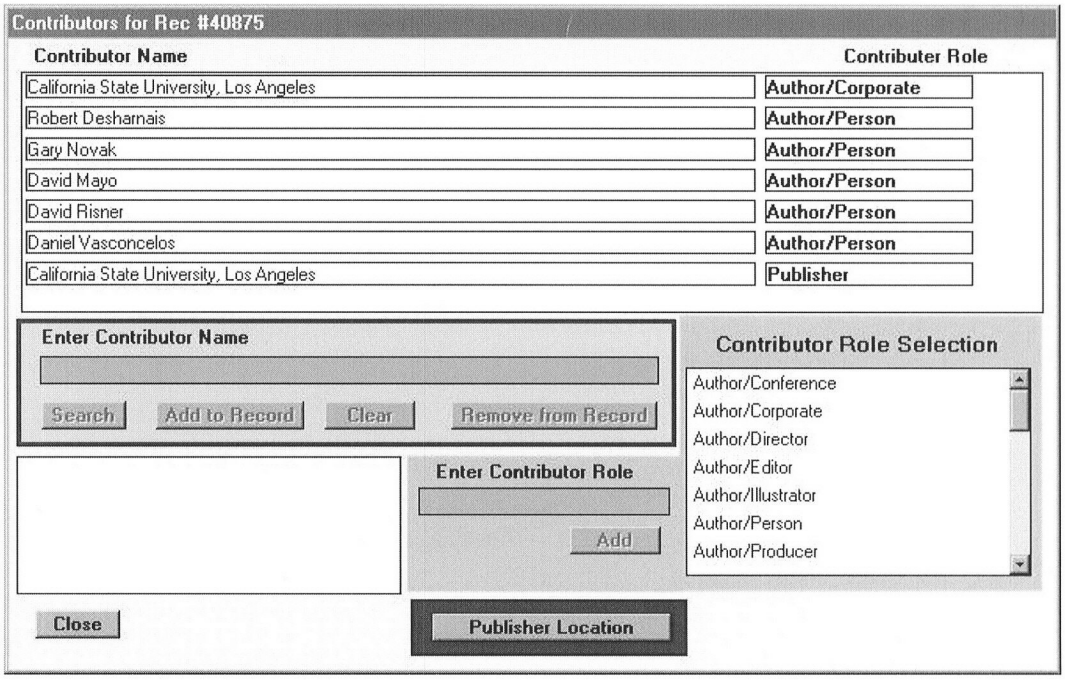
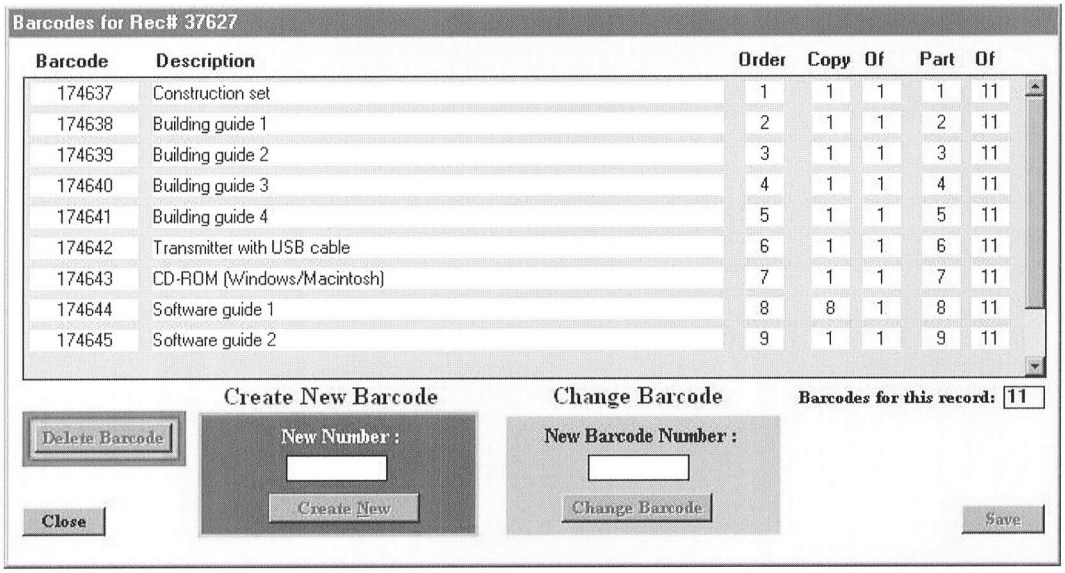


Figure 7 ENC items screen



database conversions were straightforward, and proceeded quickly. The client interface presented some problems: the compilation of PowerBuilder 4 to PowerBuilder 5 resulted in over 1,100 errors due to the development software changes. Not all the changes were negative; in fact, some very positive features had been added; primarily, the ability to use non-visual data stores. Correcting the errors as well as modifying the code to make use of some of the new features took several months.

Once the code was completed, tested, and ready for distribution, a week was set aside for the migration of the more than 5,500 completed catalog records. This proved to be the most difficult part of the development process. FileMaker stored data as a large flat table, which divided into 42 SQL relational database tables. With the conversion completed, it took two weeks to train in the new version and also debug the client application.

With use, it became apparent that the building of records as a collection of items was not an efficient way to process and maintain the data. The client application was slow and cumbersome, being burdened with the task of maintaining the relationships between several items and their records. To resolve this, steps were taken to modify the processing, so that the client interface could move as many of the data as possible from the items to the record itself. There was a limit, however, to how many data could be moved without a major modification to the client interface. The problem in the interface was due to the fact that it was designed to adhere to strict business rules with regard to the relationship between items and records. Problems also arose from the use of inheritance in the application, resulting in complications when changing the properties of the base objects.

After considerable discussion, it was determined that record processing was a linear process, and that catalog records should move forward only with few exceptions. The client interface was updated to the latest version, and the data entry windows were simplified. Once the server was upgraded to MSSQL2000, the client software could be upgraded to PowerBuilder 7.

By this time, the basic work-processing queues were well-defined and were capable of withstanding future modifications when necessary. By building stand-alone windows, it eliminated the dependence on inheritance as well as greatly improving the ability to perform any alterations. If a window needed modification, the code could be changed directly without concern for ancestors; also, if the modification was extensive, a new window could be built to simply replace the old one. The size of the code was reduced by a fourth, the need for the high memory overhead of the original code was reduced, and development

Library Hi Tech
Volume 22 · Number 1 · 2004 · 28-39

time was reduced by one third. This new version of the ENC cataloging tool was appropriately renamed NeoCat (ENC, 2002).

NeoCat has been in use at ENC since 2001, and contains over 25,000 catalog records. Problems have been kept to a minimum, while several new features have been added: a new spell-check system has been installed, fields have been modified and added, and controlled vocabulary lists enhanced and refined. Export functions have been expanded, so that records are generated as ENC XML files, Open Archives Initiative (OAI) files based on the OAI Protocol for Metadata Harvesting, version 2.0, and Learning Object Metadata files based on the IEEE Draft Standard (IEEE Learning Technology Standards Committee, 2002).

Converting data files from the internal to the public site

The ENC catalog was released for online public use in 1997. The Web site featured search and display features. Searches were performed by the Fulcrum SearchServer search engine (Fulcrum Technologies, n.d.). HTML was generated on demand from the MARC and ENC MARC extension-encoded flat files through PERL CGI scripts. MARC tagging and ENC MARC-extensions come into use when the daily creation of the single-tab delimited flat file is divided into individual records for the search engine, which creates and displays records for public view.

The Microsoft Windows database server that holds the large MARC and ENC MARC-encoded flat text file moves the data to a UNIX server. This large flat file contains the complete set of data for each record that is created or modified during the day by ENC staff. Each record within the file contains a MARC field identifier or an ENC MARC field extension.

Several preprocessing steps must be applied to the new and updated ENC records before they can appear on the Web site. The MARC file is divided into separate files, one for each record created. The files are named using an ENC record number. The large MARC file becomes obsolete once the individual files are stored on the UNIX server. The Fulcrum SearchServer indexing engine was chosen by ENC in 1996 as the indexing engine that would drive Web-based searches of the ENC catalog. Fulcrum also resides on the UNIX server. The Fulcrum engine is capable of searching the individual large text fields like those found in the abstract. These large text fields cannot be entered directly into the Fulcrum database; they are processed a second time instead, their fields being tagged with specific encoding character sequences. These tagged fields are written to individual files named using the ENC record number.

As the separate files are developed and the large text fields identified, the records are passed into a preprocessor that builds SQL statements and automatically inserts the data into the Fulcrum database for full indexing. This is where the online searching takes place. The Fulcrum search engine allows ENC users to apply a simple or advanced search query against the database and receive a return list of ENC records that matches their search string. The resource identifier is passed on to a PERL CGI program which dynamically generates HTML.

At the beginning this procedure worked well but, as the usage of the ENC site increased, a more efficient way of generating HTML was necessary. The Vignette product was introduced in 1999 and dealt well with the issue of being capable of rendering completed Web pages. Vignette could also cache completed Web pages so that re-rendering them was not necessary until a time when some of the content on the page had changed (Vignette Corporation, n.d.).

Vignette has the ability to read and parse XML, a much faster process than reading individual MARC-encoded files. In 2000 an additional step was introduced into the indexing process: the generation of XML records after Fulcrum indexing. The MARC-encoded files are used as input to the XML generator, which has a one-to-one correlation of XML files to MARC files. The container tags used in XML are the MARC tags used by ENC. As a user queries the ENC collection, Vignette reads the appropriate XML file, generates and displays HTML, and also caches the HTML to eliminate the need to re-render the exact record when next selected.

Library Hi Tech
Volume 22 · Number 1 · 2004 · 28-39

ENC extensions and searchability

The new build of the ENC cataloging tool, QCat, once strictly MARC-based, was extended in 1996 to include ENC-developed fields (see Table II). ENC's collection, currently containing over 25,000 records, has 6,050 resources funded or published by governmental agencies. The statement of work from the US Department of Education charged ENC to tag each cataloged record in such a way as to show its "federal-ness". ENC extensions for funding, special projects, and statistics contain the data that signify the materials' federal creation or funding status. These fields were a significant addition to the catalog record for two reasons: they provide proof of the US government's presence in education and they lend credence to the ENC collection.

Other ENC-developed cataloging fields hold data for vendor/distributor contact information, so that each record contains all the information needed for educators to order the materials for classroom use. Data in the acquisition fields contain vendor address, telephone number, e-mail address, ISBN, ordering numbers, availability terms and pricing for each resource in the collection, with the exception of Internet resources. Also added were cataloging fields that could hold exemplary data for the curriculum resource.

Decisions were made about how the fields included in the catalog record should be populated. Most importantly, it was necessary that the ENC user community would successfully retrieve the resources for which they searched. In some fields, controlled vocabulary lists were deemed necessary to facilitate improved user searchability. The controlled vocabulary for subject headings was developed by a collaboration of focus groups of K-12 educators and ENC subject specialists. In both the simple and advanced searches, the user has the opportunity to fill in a search box with one or more topics or subject words. Although these searches pull from a controlled vocabulary list, i.e. words must be on ENC's list of subject terms, most subject terms familiar to educators are included in the list.

There are specific ways, however, that ENC staff use terms in order to be consistent and to most accurately describe teaching materials. To know precisely what terms ENC uses, the browse-by-subject function takes users through the different levels of ENC subject terms to find resources that match the specific term they request. For example, the first level of terms is general (mathematics); the next level is slightly more specific (algebra); the third level is quite specific (coefficients).

The ENC homepage offers two search strategies: a simple search which provides a search based on a question, word, or phrase, and can be restricted to curriculum resource by its media type, grade level, or cost, and the advanced search which offers a Boolean-type search strategy at the first level: searches can be limited by "contain any of the words", "all of the words", "none of the words", or a particular phrase. Recently, the Fulcrum search engine was replaced with a search engine named Autonomy (Autonomy, n.d.). Autonomy is programmed to perform searches in the following fields: the subject lists, abstract, and table of contents, title, series, language, publisher, author, funding agency, and ENC number. Resources that are relevant to a given search are assessed according to the way they are represented and interpreted in the metadata that describe them, and then displayed to the user. ENC staff realize that ensuring relevance in a search relies to a large extent on the expertise of the staff and how well they understand indexing, cataloging and how the search and retrieval function works.

Handling the metadata today

While ENC has adopted the MARCXML record, the data lend themselves to a straight conversion to both OAI XML format and IEEE LOM XML format.

The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) version 2.0 came into use in 2002. This XML format has become the standard form for submitting records via the OAI-PMH to the National STEM (Science, Technology, Engineering, and Mathematics) Digital Library (NSDL). ENC has contributed all its digital resources from all its collections to the NSDL for harvesting. OAI specifies the *Dublin Core Metadata Element Set* (Dublin Core, n.d.).

Library Hi Tech
Volume 22 · Number 1 · 2004 · 28-39

Part of the daily maintenance of the ENC catalogs is the generation of OAI XML records. These OAI records are stored on a server and their status is maintained through an SQL database. The database is used to ensure the proper distribution of records requested through the OAI harvest.

The work on the IEEE Learning Object Metadata (LOM) began in 1998; it became IEEE LOM 1484.12.1-2002 in June 2002. For standardization and compatibility in the digital library community, ENC has adopted the IEEE LOM format for digital resources.

Conclusion

The use of USMARC tags and the ability for MARC to be enhanced with ENC added-value educational extensions have resulted in a collection of richly-described curriculum resources. A working model was established with input from the educational community. Each resource in the collection is cataloged completely and objectively by mathematics and science educators and library professionals based on this model. The working model itself has been reviewed and modified over time.

Through the years, ENC has improved the discovery of current and relevant curriculum resources which benefit the user. The workflow of resources has been optimized through the cataloging system and streamlining the process of entering data. Applying these efforts with new technology has resulted in content-rich data that are easily transportable to the formats in use today and those of tomorrow.

The ENC collections are available to the public through the World Wide Web and support both simple and advanced searching capabilities, providing ease of use in locating curriculum resources.

Visit us at the following sites:

- http://enc.org
- http://gsdl.enc.org
- http://icontechlit.enc.org
- http://thelearningmatrix.enc.org

References

American Library Association (1967), *Anglo-American Cataloging Rules*, American Library Association, Chicago, IL.

Autonomy (n.d.), *Autonomy Server*, Version 3, Autonomy, San Francisco, CA.

Dublin Core (n.d.), *Dublin Core Metadata Element Set*, Version 1.1: Reference Description, available at: http://dublincore/documents/dces/

Dynix Marquis (n.d.), *Marquis Automated Library System*, Version 3.1, Provo, UT.

Eisenhower National Clearing-house (ENC) (2002), *NeoCat*, ENC, Columbus, OH, March.

FileMaker (1994), *FileMaker*, FileMaker, Santa Clara, CA.

Fulcrum Technologies (n.d.), *Fulcrum SearchServer*, Version 3.5, Ottawa, Canada.

IEEE Learning Technology Standards Committee (2002), IEEE 1484.12.1-2002 LOM, available at: http://ltsc.ieee.org/

Powersoft Corporation (n.d.), *PowerBuilder*, version 4, Concord, MA.

Vignette Corporation (n.d.), *Vignette Content Management*, Version 6, Austin, TX.

Further reading

ENC Online (n.d.), Eisenhower National Clearing-house, available at: http://enc.org

Gender and Science Digital Library (n.d.), Eisenhower National Clearing-house, available at: http://gsdl.enc.org/

Innovative Curriculum Online Network Digital Library (n.d.), Eisenhower National Clearing-house, available at: http://icontechlit.enc.org/

International Business Machines (n.d.), *OS/2*, Version 2.1, Armonk, New York, NY.

(The) Learning Matrix (n.d.), Eisenhower National Clearing-house, available at: http://thelearningmatrix.enc.org/

Lynch, C.A. (1997), "The Z39.50 information retrieval standard", *D-Lib Magazine*, April, available at: www.dlib.org/dlib/april97/04lynch.html (accessed 22 May 2003).

Microsoft Corporation (n.d.), *Microsoft SQL Server*, Version 6.0, Redmond, Washington, DC.

Microsoft Corporation (n.d.), *Microsoft SQL Server*, Version 6.5, Redmond, Washington, DC.

Microsoft Corporation (n.d.), *Microsoft SQL Server*, Version 7.0, Redmond, Washington, DC.

Microsoft Corporation (n.d.), *Microsoft SQL Server 2000*, Redmond, Washington, DC.

National Science Digital Library (n.d.), available at: http://nsdl.org/

Open Archives Initiative (n.d.), available at: www.openarchives.org/

Sullivan, T. (n.d.), *AccessFlorida A Z39.50 Control System Programmers Reference and System Guide*, Florida Center for Library Automation, Gainsville, FL.

Sybase Inc. (n.d.), *PowerBuilder*, Version 5, Emeryville, CA.

Sybase Inc. (n.d.), *PowerBuilder*, Version 7, Emeryville, CA.

Sybase Inc. (n.d.), *Sybase SQL Server*, release 4.9.1, Emeryville, CA.