Souping up the engine: making the most of the catalogue at The University of Auckland Library

Ksenija Obradovic
Cataloguing Manager
The University of Auckland Library
k.obradovic@auckland.ac.nz

Abstract:
It is widely recognised that libraries need to transform the way they present their collections to users. As library catalogues are still the main discovery tool for libraries’ resources, they need to change too. What can we do while we wait for the transformations possible with FRBR and RDA? This paper outlines The University of Auckland Library’s efforts to improve access to their resources.

Particular attention will be given to exploring the opportunities offered by modern technology. The paper will describe how different computer programs can be used to improve the quality of existing bibliographic data in the library catalogue, thus improving access to it. It will also look at strategies that can be employed for faster and more efficient creation of bibliographic data, including the automatic creation of data for both electronic and print resources. Emphasis will also be given to initiatives aimed at providing access to material that was not traditionally described in the catalogue, as well as creating gateways to specific parts of existing collections.

The paper will conclude with some speculations on the future of the library catalogue in an FRBR/RDA/Google/Amazon.com world.
Introduction

Recently, the value and future of libraries have been questioned and libraries have been pushed to reconsider their priorities. Libraries have been seen to be in competition with the Internet, and library catalogues in competition with search engines, such as Google, and Google sometimes comes out ahead. Libraries have lost their privileged position as the sole source of information for scholars and the general public. Users like tools like Google Scholar, they like tagging, giving their opinions, but most of all, they like to find what they are looking for quickly, almost immediately, and without hard thinking about what complicated searches they might have to perform to get the information they need. In that light, cataloguing and classification are themselves regarded as replaceable by keyword searching, relevance ranking, and user-referral mechanisms. (1)

In reality, the Internet and Google are not perfect either. Search engines can find only what they can crawl. They still cannot crawl everything that is on the Internet, and in addition to that, not all relevant literature in on the Internet yet. Google ranking is far more sophisticated than that of library catalogues, but often, when one searches for something, Google returns thousands of irrelevant hits. (2)

The role of academic libraries is to support learning and teaching, and in that environment, library catalogues are still the main source for information on libraries’ own resources. Academic libraries buy material with the specific intent of providing students and teachers with relevant literature. They also have important unique collections such as theses.

With this in mind, at the University of Auckland Library a lot of work has been done on improving access to the Library’s collections. Focus has been on improving utilization of the Voyager Library catalogue – by ensuring quality data for optimal access to collections already on Voyager and by making a number of existing collections more prominent by adding them to Voyager.

This paper will describe several projects that took advantage of new technology in a cataloguing environment to create bibliographic data more quickly and efficiently. It will also talk about data quality as the retrieval of the information directly depends upon it.

Technological advances and cataloguing

Until recently, Integrated Library Systems (ILS) were considered high technology. They enabled much faster cataloguing, good display of data and great searching opportunities. Unfortunately, they have not changed much since initially developed and the statement is no longer true. ILSs are not matching with expectations of either library patrons or librarians.

Library systems are big and complex and replacing or improving them is not an easy task. System vendors still have not developed more sophisticated systems, and libraries, no matter how big they are, do not have the time or money to invest in developing systems themselves.
To overcome the disadvantages of ILS cataloguing modules, computer specialists often write programs to augment them. For Voyager, the system the University of Auckland Library uses, Gary Strawn has created The Cataloguer's Toolkit, a set of programs which make cataloguing and managing existing data much easier. (3)

Other programs like MarcEdit, Marc Report, and Marc Wizard are standalone software packages that can be used independently of the ILS and can facilitate automated creation of data and database maintenance.

MarcEdit is a free MARC editing tool, developed in 2000 by Terry Reese, in response to a database clean-up project on the Oregon State University's electronic catalogue. It runs on Windows 95, Windows 98, Window ME, NT 4+, 2000, and XP. On NT systems (this includes Windows NT, 2000, XP) one has to have administrative permissions to install it. It requires 10.5 MB (setup) and 5 MB (application) of disk space. It has been continuously developed. More about it can be found at http://oregonstate.edu/~reeset/marcedit/html/index.php

Marc Wizard is a software package developed by Mitinet, Inc. Library Solution and marketed as a library resource suite. Since 1987 it has had a utility (MARC Magician) that can be used for importing, creating and editing MARC records. Marc Wizard provides sixteen global editing commands. It can work with most computers. There is more about it at http://www.mitinet.com/Downloads/Download-Wizard.aspx

MARC Report and MARC Global are programs developed by The MARC of Quality, a company that also provides cataloguing training and database services. MARC Report was written in 1995. It validates MARC records according to the latest MARC21 standards, and also does cataloguing crosschecks to identify MARC problems, such as wrong coding. MARC Report also includes a MARC editor and many integrated MARC utilities.

MARC Global was first copyrighted in 1997. It is a tool that can 'find and replace' MARC files.
MARC Report and MARC Global run on Windows 2000 Professional or XP Professional. Both are totally customisable. Further information about it can be found at http://www.marcofquality.com/soft/softindex.html

The start screen of MARC Global

The Cataloguing Department of the University of Auckland Library has made extensive use of MARC Report and MARC Global over the last few years. The programs were used to create a large number of records all at once, in cases where the records for the whole collection did not differ much, and to clone records for e-books from existing records for print editions. They were also used to quickly create bibliographic records in MARC format for collections being migrated to Voyager.

The same software was also used to improve the quality of data in existing records, in order to improve access to the Library material. Because MARC Report and MARC Global are not part of Voyager, records must be exported from the Library’s database, and re-imported after editing is finished - not the most efficient workflow.

With tools like MARC Global it has been very easy to convert non-MARC records into MARC. Over the years, various collections had had metadata created in InMagic DBTextworks, but it became obvious that it was better to have these collections on Voyager, as it is easier and faster to create metadata for new additions to these collections, as well as to take care of lending the material. Currently, a database of DVDs and videos developed by the Department of Film, TV & Media Studies is in the process of conversion from InMagic to Voyager, and there are plans for more projects of this kind in the future.
Two projects will be described where using the above programs helped create metadata in a much faster and more efficient way.

**Creating data for Artists’ and Galleries’ Files**

The Fine Arts Library, one of the University of Auckland branch libraries, has a collection of New Zealand artists’ and galleries’ files which contain ephemera from galleries, exhibition information, small catalogues and press cuttings. Approximately 350 galleries’ files and nearly 1700 artists’ files are kept in vertical files. Nicole Jackson, a librarian at the Fine Arts Library, spoke about the history and importance of this collection at the 2006 ARLIS/ANZ Biennial Conference. (4)

Artists’ and galleries’ files did not have any kind of metadata except two lists of names, one of artists, the other of galleries. The lists were made in MS Word format by the librarians at the Fine Arts library, and were used as the main reference tool. The librarians at the Fine Arts Library felt that this valuable resource was under-utilised, as it was uncatalogued and therefore essentially non-existent. Students and researchers were often unaware of it.

Having this collection on Voyager was recognised as important for several reasons. Firstly, Voyager access would make it more prominent. Students and researchers searching Voyager would find it as well as relevant books. Voyager would facilitate borrowing and provide statistics of usage.

Instead of creating MARC records for these files individually, one at a time, it was decided that records would be created automatically from the Word documents of artists’ and galleries’ names by converting each list into MARC records.

First, the list was prepared for conversion to MARC. Names were used to create the subject headings. They were also duplicated, with name and surname inverted, to produce a form that could be used for title.

| 245 00 | $aZena Abbot |
| 600 10 | $aAbbot, Zena. |

*Example of an artist’s name in the file prepared for conversion*

The file was then run through MARC Global software, which created MARC records from it. Other required MARC fields and codes were added with MARC Global. Much of the data was same for all the records. This included:

Subject Heading: 650 $a Art, New Zealand
Restriction on Access Note: 506 $a To be used in the Fine Arts Library only. Not available for interloan.
Content Note: 520 $a A collection of ephemeral material that may include exhibition invitations and small catalogues, brochures, reviews, photographs, articles, press releases, etc.
Example of a bibliographic record for Artist's File

When all the bibliographic records were finished, the file of nearly 1700 bibliographic records was split into more manageable chunks and uploaded into the Voyager local database. The records were uploaded manually, which allowed authority work to be done as each record was uploaded. Authority work was the most time-consuming part of the project. However, it is crucial to have all works by and about an artist listed under the same version of their name. Luckily, for more than three quarters of the names, authorities were already established on Voyager or Te Puna.

After saving the record to Voyager, a holding record was created. Two new locations were created on Voyager that would display as ‘Artists' Files' and ‘Galleries' Files'. There is no classification.

Example of a holding record

In the Voyager database an item record needs to be created as well, in which the barcode number is entered. Folders need to be barcoded to enable easier issuing and for statistical purposes. This part of the work has been left for later, as existing envelopes are going to be replaced, and at that point, barcodes will be added and item records created.
Creating records for these two collections was very fast. It took less than two hours to prepare the list of 1700 artists’ names for conversion, and another hour to prepare the list of galleries’ names. Conversion itself was done in a few minutes. Adding global fields took another two hours for artists’ files, and 30 minutes for galleries files. The time difference is due to cataloguers changing their minds about wording and fields they wanted to add, not the size of the files. It actually does not matter how big the file is, MARC Global creates MARC records in only a few minutes. Authority work on artists’ and galleries’ names took the longest time. For the artists’ files, it took about 30 hours of authority work, and for galleries’ file another 10 hours. This would have been required even if the records had been created one by one.

Since these two collections were put on Voyager, their usage has increased considerably. How much prominence was given to them is best shown by the statistics of usage:

2004      5 files used  
2005      3 files used  
2006      72 files used  
2007 (so far) 175 files used.

Creating records for e-books for the eHRAF collection

eHRAF (Human Relations Area Files) consists of two separate databases: eHRAF Collection of Archaeology and eHRAF Collection of Ethnography. (5) Each database contains full-text books, documents, and journal articles. The database vendors do not provide MARC records for books in the collection.

Several years ago, to enable patrons to find e-books from eHRAF from the Library catalogue, URLs were added to existing MARC records for print books. However, this is not an ideal solution. Managing e-books is much easier if every version has a separate record. New library web portals, which will replace standard OPACs, are being developed with FRBR, and FRBR has been developed on the premise that cataloguers have followed the rules and created separate records for each manifestation of a book.

There were 343 records for printed books from eHRAF collection with an URL link to the electronic eHRAF book added in the Voyager catalogue. These records were exported from Voyager and then modified using the MARC Global software. On Voyager, in the records for the print versions (Voyager exports the copy of a record, leaving the master record on the database) the 856 field was stripped off with the Gary Strawn’s Cataloguer’s Toolkit.

Following OCLC recommendations for creating records for e-books, books from eHRAF collection were described as electronic versions of already published material. (6)

Of the records in the file downloaded from Voyager were striped fields with data relevant only for print version (e.g. 010 LC Number, 035 System Control Number, etc.) After that, fields relevant for the e-version were added, including:

245 ‡h [electronic resource]
533 †a Electronic reproduction. †b New Haven, Conn. : †c HRAF. †n Available via World Wide Web. †n Access may be limited to HRAF affiliated libraries.
655 7 †a Electronic books. ‡2 local
710 2 †a Human Relations Area Files, Inc.

Bibliographic data for the original work remained unchanged.

Example of a bibliographic record for the printed version of a book from the eHRAF collection
Example of a bibliographic record for the e-book from the eHRAF collection

Time saved by creating the records this way was substantial - it took only one hour to finish the whole job.

### Importance of data quality

Each MARC record carries a lot of information, some of which is used by human and some only by automated systems, such as limiting, display, etc. Any mistake in the record, or lack of data, can compromise searching and finding. For these reasons a library collection is only as good as its bibliographic records.

Providing good quality bibliographic records is a tedious and time-consuming job. Current cataloguing practices are very much limited by lack of time and money. Cataloguing departments are often faced with insufficient resources for proper cataloguing; they usually do not have enough staff, there is a continuous demand to have items shelf-ready in the shortest
possible time, backlogs are growing, and an atmosphere in which the value and need for any cataloguing is questioned, is not helpful either.

Numerous libraries turn to “shortcuts” – they create very short records to which they add holding records – just enough to be able to issue the material. Or they omit data they think is unnecessary for retrieval, and which takes time to add. Although such strategies may look useful at the time, they always turn out costly in the long run. Whenever data has to be transferred from one system to other (and this happens relatively often) such records always cause problems. Even in their original database, they are often confusing for both patrons and librarians.

Many years ago, libraries began collaborating by sharing their resources. The main idea was to save local cataloguing time and effort by minimising the number of records to be created and corrected in each local library. Although very practical and admired by all cataloguers, this idea has been realized only in part. Most libraries do not upgrade records on the shared databases.

The idea needs to be revisited. Cataloguers need to re-create shared cataloguing behaviours, practices and institutions.

Taking part in big cataloguing communities, such as OCLC, has proven to be important for libraries. OCLC is the biggest source of bibliographic records, and a source of good quality bibliographic records. Not having to create bibliographic records is a great time saver. However, not all records on big databases are of good quality. Some are CIP records; some are inadequate for various other reasons. Many records need to be corrected. Usually members of big cataloguing communities download records at the time of ordering. Later, when the material arrive at the library and when they do copy cataloguing, libraries upgrade records only on the local database and not on the shared database.

Collaboration in copy cataloguing between cataloguing agencies can be proven to save local cataloguing time and effort when the result is records that fewer libraries have to correct in order to adopt them. It looks like spending a lot of time on editing, but any time spent initially will turn into a gain later. This is the same benefit that libraries derive from shared original cataloguing. The more libraries contributing, the greater the benefit will be. Libraries need to be aware of common interests, needs, and responsibilities in a global rather than provincial information community.

As libraries increasingly rely on copy cataloguing and outsourcing for their cataloguing data, and more and more libraries enter into resource-sharing agreements and virtual catalogues, MARC database quality control has never been so important.

Cataloguers at the University of Auckland Library take advantage of MARC Report and Cataloguer’s Toolkit. MARC Report is used to analyse the quality of bibliographic records and to determine mistakes in MARC tagging, coding and punctuation in bibliographic records outside the Voyager database, and Cataloguer’s Toolkit for editing bibliographic records that are part of Voyager database. Of course, it is not possible to correct the whole database at once, but such process has proved its worth on many occasions.
MARC Report software is particularly helpful when checking the quality of bibliographic records supplied by vendors, such as records for e-books. The Library has huge e-book collections. As of August 2007, there are nearly 50 collections and over 280,000 bibliographic records for e-books in the local catalogue. Having records for e-books has proved very important for library patrons and more and more e-book vendors are keen to provide them. However, the quality of records that e-book vendors produce varies to a great extent. Even vendors who generally produce good records sometimes have records with mistakes that can compromise searching or cause records to be rejected when batch loading to Voyager.

The main problem with these records is their sheer number. In some cases, the Library has been provided with several thousands of records to be batch-loaded to the system. The large number of e-book records makes it impossible to carry out bibliographical checks. Files are run through the MARC Report programme to identify faults in MARC tagging before loading the records onto Voyager.

### Improved Access to Theses and Dissertations

Library systems are not perfect, but neither do libraries make full use of their potential. Frequently, libraries buy software, set it up, and leave it as something that cannot be altered. However, some things can be improved to enable better functioning of library systems and easier finding of library material.

In every academic library, theses are one of the most important collections, but it was always very difficult to find them in the University of Auckland Library Voyager catalogue. Information about theses is recorded in the 502 field but that field cannot be limited when searching. Theses were buried among other material and searching for them was not intuitive. In 2006, the Library began digitising theses, and this was an excellent opportunity to re-examine access to theses via Voyager and to try to improve it. A solution was found in the development of a special gateway to theses collections.

A link to the Thesis Collections gateway was put on the Library home page, which leads to a pre-limited search of University of Auckland theses.

Limits are set up on two parameters in the bibliographic records: code “t” in the Leader Line and the word “Auckland” in 502 Thesis Note. This restricts a search to the University of Auckland theses. From the Theses Page, users can choose a Keyword, Date, Title or Author search. These searches can also be limited to a particular degree.

It was interesting to discover that most bibliographic records for theses had code “a” for printed material, instead of “t” for manuscript, although according to the cataloguing rules, theses are supposed to be treated as unprinted material. All thesis records had to be fixed to enable the limits to work. They were exported from Voyager and coding was globally changed using MARC Global. As the records were already exported from Voyager, they were run through the MARC Report program to determine other mistakes in MARC coding and tagging. After fixing the mistakes, records were re-loaded to Voyager.
Set-up for a canned search of the Library catalogue

A further improvement was creating links to the online thesis booking form from Voyager holdings. One copy of theses is kept in the General Library in non-public access collection and users have to fill in a form. Linking this form from Voyager facilitated the whole process.

Creating the Theses Gateway did not require any additional software. It was not an expensive exercise by any means. It only needed was a bit of enthusiasm and collaboration between cataloguing and IT departments. And the result has been very satisfactory for users.

Conclusion

The world of technology is developing rapidly, and that is affecting libraries. The need to improve library catalogues is obvious; yet, at the moment libraries are not offered the means of transforming them. Libraries are faced with many questions and very few answers. It is impossible to predict the direction of future developments. However, some things will never change - metadata quality remains essential to effective information retrieval.
The development of FRBR is something that promises much better display of data stored in library catalogues. FRBR (Functional Requirements for Bibliographic Records) is a conceptual model in which are recognised relationships between works described in the library catalogue. Instead of a meaningless list of hits, in FRBRized displays, search results are grouped according to these relationships. Results are presented in a more intuitive way that is easier for users to understand and navigate. However, all projects involving FRBR have shown that results depend directly on the quality of data in bibliographic records.

It is recognised that cataloguing standards need to be changed to reflect modern needs. RDA has promised to be user-focused and applicable to all types of content and resources, but even before it was published, it was widely criticised as being inadequate.

Whatever the new rules for description are, the important thing is to maintain cataloguing standards and to follow the rules for description. Any new developments that use existing data in library catalogues have been designed with the rules in mind. This is also the case with FRBR. If there are no elements in the bibliographic record that are used to establish relationships, these relationships cannot be established. If there are several versions of an author's name in the library catalogue, results will not be able to be clustered.

One day, in the near or distant future, we may get library systems that are much better. Or catalogues may be replaced with something totally new. For now, they are the main representation of libraries resources. Although ILSs are not as developed and advanced, as we would like them to be, there are still things that can be done to improve the use of the catalogues.

The University of Auckland Library believes that empowering cataloguing staff with new technology is critical to efficiently providing access to a wide range of information sources. The Cataloguing Department utilizes technology to automate and manage many of its functions and to streamline its procedures.

Modern technology offers many opportunities for cataloguers to improve workflows. Advances in computers and information technology have created huge potentials for cataloguing staff to increase efficiency and accuracy, reorganize the work force, change workflows, rationalize costs, and provide better services to patrons.

Better collaboration between libraries can also significantly cut time spent on cataloguing. The realm of information has crossed the boundaries of local communities, and so collaboration should be on an international level. Library catalogues, whether they are searchable by Google or not, are part of worldwide information world. Libraries need to be aware that they are part of the global communities at the same time as they serve their own patrons.

In spite of the preceding criticism it is important to recognise the continuing value of the library catalogue, in particular that the presence of a work in the collection of a major research library implies that the work has been deemed to be of scholarly value. We must build on our existing strengths and on the work of previous generations.

Do not throw away your catalogues yet, try to improve them!
Endnotes

3. http://www.library.northwestern.edu/ctkv/

Bibliography


