ANALYSING TRANSACTION RECORDS OF INTEGRATED LIBRARY MANAGEMENT SYSTEMS TO ESTABLISH THE PROPORTION OF BIBLIOGRAPHIC RECORDS IMPORTED FROM FOREIGN LIBRARY CATALOGUES

ANDRÁS SIMON
Monguz Ltd. (Qulto Companies), Customer Manager

ABSTRACT
Libraries all over the world have built large digital catalogues for approximately sixty years. Databases of the Integrated Library Management Systems used by these institutions contain not only the bibliographic, authority and item records, but also a considerable part of the database transaction records that are logged during the day-to-day operations and application. The most important data are particularly the information about the creation and deletion of the records (creation and deletion date and time, datatype, creator user and institution) that need to be preserved for a long time as prescribed by the current international standards, first of all, by MARC21. The information about the institution and the time of creation are a compulsory part of a MARC record, thus these important transaction records are often migrated to a new database just in case of change in the Integrated Library Management System. The current Hungarian regulations for museums oblige these institutions to log and preserve each modification in their databases. Thus, the Integrated Collection Management Systems (ICMS) are used as online catalogues and digital inventory books both by libraries and museums; the databases of libraries using ICMS also preserve all the information about these modifications. In this way, and according to the prescription of Qulto – Huntéka software that was developed and supported by the Monguz Ltd. –, they are able to preserve the modification events so that from these data transactions one can gain much and useful information. As part of my research, I examined the proportion of the foreign databases imported from bibliographic records of library catalogues. The record importation was possible thanks to the shared cataloguing projects in Hungary that has been in existence for approximately fifteen years. The result of this examination is analysed and discussed in this paper.

Introduction

Data Mining from SQL Databases is an efficient way of getting information about library automation and the usage of Integrated Library Management Systems (ILMS). The system captures useful information about the institution, date, time, and other circumstances of record creation. Some of these important information units are compulsory and are...
part of the MARC record. The Hungarian regulations, first of all the NKÖM 20/2002\textsuperscript{1}, follow the international guidelines of CIDOC\textsuperscript{2}, that oblige museums to preserve the most important elements of the transaction records, and log the modifications as well.\textsuperscript{3} The Integrated Collection Management Systems (ICMS) were designed and developed to capture the memory of institutions, for both libraries and museums, as such, are able to preserve all the information about the creation, deletion, and the modifications. So, these transaction and bibliographic data units are preserved by the institutions for a long time, and can be extracted, analysed, and prepared for research.\textsuperscript{4}

As the main part of my Ph.D. studies at ELTE Institute of Library and Information Science, I researched the ICMS databases, especially the database transaction records for a couple of years. In this article, I try to find answers on how we can establish the proportion of bibliographic records imported from foreign library catalogues.

\textit{The concept of transaction record, role of MARC field 040 $a$}

Despite that the concept of data record is changing, and even the concept of the record itself seems to be obsolete, we still can talk about data records as the basic logical elements of the databases of Integrated Library Management Systems (ILMS).\textsuperscript{5} The main types of records are:

- Cataloguing – bibliographic, entity, authority and media records,
- Acquisition – order packages and items,
- Circulation – customers and loan items,
- OPAC – searches and results,
- Transaction records – generated by the application.

Maybe the most important transaction records are the ones which are generated by the ILMS when a catalogue record is created, deleted, or modified. These records should be preserved in the database according to the current international standards.

The creation of a catalogue record can happen in three ways:

- Local creation (typing),
- Daily record import,
- Conversion from former databases.

The information contained in transaction records should be according to the MARC description:\textsuperscript{6}

05 – Record status (c – Corrected or revised; d – Deleted; n – New …)
06 – Type of record (a – Language material …)
07 – Bibliographic level (a – Monographic component part; b – Serial component part)
005 – Date and Time of Latest Transaction
040 – Cataloguing Source $a$ (name of creator institute)
The field 040 $a - Cataloguing Source (name of creator institute) contains information about the way the record was produced in the ILMS. If it is created locally, the code of the library is entered automatically into the data field as default value. If it is imported directly from another database, the code of the foreign institution is entered into field 040 $a. If it is imported from a shared cataloguing system, the code of the original library is entered into the data field. The data migration from other databases should keep the original value of “Cataloguing Source field”. (Figure 1)

| 000 | 01158nam 2200253 i 4500 |
| 001 | MOKK-AW0005517816 |
| 005 | 20190707131418.1 |
| 008 | 080404s2006 enkab b 001 0 eng d |
| 020 | $a978 0 500 34223 7 (hbk.) |
| 020 | $a0 500 34223 7 (hbk.) |
| 035 | $abibDEK00681166 |
| 040 | $aDEENK |
| 080 | $a72(593)=20 |
| 100 | 0 | $aNithi Sathāpitānon $d(1947-) |
| 245 | 1 0 | $aArchitecture of Thailand : $ba guide to traditional and contemporary forms / $cNithi Sthapitanonda and Brian Mertens ; photographs by Michael Freeman ... [et al.] |
| 260 | $aLondon : $bThames & Hudson, $c2006 |
| 300 | $a256 p. : $bill., főként színes, 1 térk. ; $c28 cm |
| 504 | $aBibliogr. hivatk (p. 253) és indexet tartalmaz |
| 505 | 0 | $aAn architecture of layers -- Thai houses -- House elements -- Religious architecture -- The grand palace -- Temple and palace elements -- Temple and palace ornament - Immigrant and foreign-influenced architecture -- Thai architectural forms today |
| 650 | 7 | $aÉpítészet $2lcsh//hun |
| 650 | 0 | $aArchitecture $zThailand |
| 700 | 1 | $aMertens, Brian |
| 700 | 1 | $aFreeman, Michael $d(1945-) |
| 850 | $aD1 $bD1 $cD108 $dD108 $wbibDEK00681166 $x111010111 |

Figure 1. A record containing original source of information in national shared cataloguing system

Describing the research

Most of the technical information units in the databases of the library systems of our customers are not public. The readers can see only the record elements which are pub-
lished in the OPAC. The librarians can look at the records of their own databases – as much as the data are visible on the user interface. The transaction records, like other technical data, are searchable for the system librarians or for the system managers, who can examine their own systems, but most of the transaction records are hidden from them too. Hence most of the system managers and system librarians do not know the data structure and business logic of their system, they cannot get all the relevant information even from their own database.

I am a Customer Manager of Monguz Ltd. (Qulto Companies), a Hungarian firm developing integrated collection management systems for libraries and museums in Eastern-Central Europe. As an employee of Qulto Company, it is possible for me to enter the databases of our clients, to do my work, for instance, repairing the damaged databases, preparing reports, and modifying data according to the requests of our customers. According to the customer contracts we, as workers of Qulto Ltd., are allowed to get statistical information from the ICMS databases of our partners. However, it is useful both for vendors and for libraries to have relevant information about the catalogues.

Our company has about 400 customers: libraries and museums. From them I have chosen 39 Hungarian libraries for researching the proportion of bibliographic records coming from external data sources. Every selected institution uses ICMS at least for five years. They represent the Hungarian libraries according to institution type and geographical location. I have prepared an SQL script to detect the content of 040 $a MARC field. This script was run on all the examined databases, and the results were collected into data tables, in which each row contained the data for one anonymized institution. The attributes of the selected libraries are:

- They use the Qulto – Huntéka ILMS application,
- They are located in the territory of Hungary,
- They have relevant amount of transaction records,
- They were using the ILMS continuously in the past 5 years.

All of them have the possibility to import bibliographic records from different external databases. Mostly they use the catalogs of the National Széchényi Library (NSZL – Országos Széchényi Könyvtár – OSZK) and the Hungarian National Shared Catalogue System (Hungarian abbreviation: MOKKA – Magyar Országos Közös Katalógus). For books in foreign languages, they use mostly the Library of Congress Catalog, and GVK – Gemeinsamer Verbundkatalog. (The most widely spoken foreign language in Hungary, after English, is German, that is why the proportion of German books in Hungarian libraries is relatively high.) It could be interesting to differentiate two cases: importing a record from a shared cataloguing system, or importing directly from a library database. Nevertheless, there is neither in MARC field 040 $a, nor in other data fields any information about it. However, taking into consideration the high diversity of the libraries, recorded in the data source MARC field in the examined library catalogues,
we can assume that the widely known Hungarian National Shared Catalogue, having 55 member libraries and more than 9 400 000 bibliographic records, is used by most of the librarians in case of record import.

In the library catalogues I have chosen bibliographic records to examine their origin:
- Bibliographic level: monographic,
- Having content in the field: 040 $a
- Created by the librarian either through typing or importing.

Examining the libraries one after another, I counted the values of MARC field 040 $a, separating the data table rows of the catalogue records by having the local institution code and the different values. I distributed the libraries into three groups:

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School Libraries</td>
<td>H</td>
<td>6</td>
</tr>
<tr>
<td>Municipal Libraries</td>
<td>M</td>
<td>16</td>
</tr>
<tr>
<td>Special Libraries</td>
<td>S</td>
<td>17</td>
</tr>
</tbody>
</table>

Results of the research

Amount of imported and locally created records

On Figure 2 I represent the amount of imported and locally created records.

The results were grouped by library type. The first column of each group shows the count of the bibliographic records summarized from the databases of the libraries. The second column of the group shows the count of the locally catalogued ones out of them.

The third column of each group shows the count of the bibliographic records describing a monographic document summarized from the databases of the library type, and the fourth one shows the count of the locally catalogued ones out of them.

Figure 2. Amount of imported and locally created records

413
Before the research, one of my hypotheses was that after the database of the Hungarian MOKKA system had been filled with records, the proportion of the imported records became higher. This hypothesis is verified in the case of municipal libraries, by the result of the research. It is represented in Figure 3. We can see on the figure that in a selected large municipal library, the proportion of the imported records became gradually higher year by year after especially 2007, and the proportion in 2017 was the highest.

![Figure 3. Proportion of imported records in large municipal libraries](image)

Another hypothesis was that it requires time to search, select and import a foreign bibliographic record. In the case of books written in a foreign language, there are a couple of duties for the librarians even in the case of record import:

- Select the appropriate foreign database and find the record,
- Rewrite some data elements of book description, like subjects or physical description,
- Correct the bibliographic description according to some special Hungarian regulations.

Sometimes it needs so much time that it seems better to create the bibliographic record locally.

This hypothesis was verified too. Comparing Figure 2 to Figure 4 and Figure 5, we can establish that the municipal libraries having mostly Hungarian books often import the bibliographic records. However, high school libraries and special libraries having more foreign books create their bibliographic records mostly locally.

On the diagrams of Figure 4 and Figure 5, numbers represent the examined libraries and letters give us information about the library type:
The diagrams show us the relation between library type, the proportion of foreign books, and the locally catalogued bibliographic records.
Another hypothesis was that the small libraries having little staff and small catalogues are more motivated to import records than the bigger ones. I supposed that the larger libraries which have specialists for bibliographic description use the solutions of shared cataloguing systems less willingly.

After all, this hypothesis was not verified by my research, as we can see in Figure 6. The height of the columns shows the proportion of locally catalogued records. This diagram is sorted by the catalogue size, which is represented by the number of bibliographic records written under the columns. We can establish that there is no connection between the size of the catalogue and the proportion of imported records.

![Figure 6. Proportion of imported books, sorted by catalogue size](image)

I hope I could prove in my paper the trust that it is worth to examine the transaction records of the library databases, because we can gain important information working by these results.

I examined the databases of customers of Monguz Ltd. from several other aspects:

- Connection between the size of the catalogue and the size of the database,
- Change of the size of the catalogue and full text records according to library type and the year the record was created,
ANALYSING TRANSACTION RECORDS OF INTEGRATED LIBRARY MANAGEMENT...

- Time spent by the librarians cataloguing, according to the way it was created (imported or not),
- Age of loaned items (difference between the loan date and the year of publication),
- Proportion of the special record types (bibliographic, entity, authority and media) according to type and size of the library.

I have searched the databases for years, and I hope, having my special opportunity, I can prepare for further examinations too. In my opinion, it is important to do this work, especially because plenty of the transaction information are not preserved for a long time, and usually are annihilated in the case of migration, or are simply erased because of the lack of space on the data carriers.

Notes and bibliography

6. MARC 21 Bibliographic 01X-09X - Numbers and Codes-General Information. Source: https://www.loc.gov/marc/bibliographic/bd01x09x.html [20 February 2022.]
7. Prepared by the author.
10–14. Figures are made by the author.

András Simon finished his studies in 1993 in History and Librarianship at the Faculty of Arts, graduating from ELTE. He has been working in Hungarian libraries since 1992 with the scope of library automation. Between 1999 and 2002 he was the system librarian of the Hungarian Shared Cataloguing System. Since 2002 he worked as a subcontractor and later as an employee for the MTA SZTAKI, then for Monguz Ltd. from Szeged in the field of design, development, customer support and distribution of the Huntéka integrated library management system. He is currently a Ph.D. student at ELTE Institute of Library and Information Science since 2019.