



67th IFLA Council and General Conference

August 16-25, 2001

Code Number: 001-188(WS)-E
Division Number: IV
Professional Group: Division Bibliographic Control
Joint Meeting with: UBCIM
Meeting Number: 188
Simultaneous Interpretation: No

OCLC's UNIMARC/MARC 21 Conversion

Jay Weitz

Consulting Database Specialist
OCLC Online Computer Library Center, Dublin, Ohio, USA

Abstract:

OCLC's involvement with UNIMARC/Bibliographic is outlined briefly, beginning with the February 1995 decision to pursue international growth, through the development of UNIMARC and MARC 21 conversion utilities during 1995 and 1996, and implementation in various phases between March 1997 and early 1998. Some major conceptual and practical differences between UNIMARC and MARC 21 are also mentioned.

OCLC's involvement with UNIMARC/Bibliographic stems from two circumstances that converged in February 1995. The OCLC Board of Trustees decided to pursue OCLC's international growth through a number of product enhancements, including the development of a UNIMARC capability. And OCLC entered into an agreement with the National Library of the Czech Republic to load UNIMARC bibliographic records from the Czech National Bibliography into WorldCat, the OCLC Online Union Catalog.

After examining existing utilities to convert records between UNIMARC and USMARC (now known as MARC 21), OCLC decided to build its own conversion software. The project began in May of 1995 to write specifications for the two-way conversion, translate these requirements into a software utility, and finally to fold the utility into four phases of OCLC services: UNIMARC output through OCLC Export; UNIMARC output through OCLC-MARC subscription and tape services; UNIMARC Batchload capability for WorldCat; and UNIMARC output from the OCLC CatME for Windows system and other micro cataloging products.

Because UNIMARC developed from a significantly different perspective than MARC 21, virtually every field and most subfields within each field had to undergo conversion. This includes character set and code list translations, re-designation of tags, subfield reassignments, indicator modifications, and in numerous cases, manipulation of data within certain fields and subfields. Furthermore, several conceptual differences between UNIMARC and MARC 21 proved to be challenges. UNIMARC de-emphasizes the notion of "main

entry" by placing all intellectual responsibility in its 7-- area, whereas MARC 21 explicitly separates main entries into the 1XX area, for instance. UNIMARC raises Family Name entries to a status equivalent to that of personal, corporate, and uniform title headings, whereas in both AACR2 and MARC 21, family name headings are rare except as subject entries, and are designated by an indicator in the Personal Name fields. MARC 21 makes more extensive use of general notes designated as 500 fields, where UNIMARC divides its "general" notes into over a dozen different fields corresponding to other areas of the bibliographic record.

At the same time, OCLC tried to make its application as general as possible, keeping in mind the wide latitude UNIMARC allows for individual implementations. To that end, UNIMARC team members met with representatives of three European national libraries during the course of development to get a taste of how the format was actually implemented in real situations. This allowed us to grapple with such questions as data loss, local variations among implementations of the same UNIMARC standard, practices that are supported in one format and not the other (most notably, embedded fields and certain forms of record linking), and the impact of different cataloging rules on individual implementations of UNIMARC.

Subfield-by-subfield draft specifications for the conversion in both directions between UNIMARC and MARC 21 were begun in June 1995 and completed in January 1996. At that time, programmers began coding and testing the UNIMARC to MARC 21 conversion. That effort was substantially completed in July 1996 when OCLC sent a file of converted records to the National Library of the Czech Republic for their judgment. From that time through March 1997, OCLC fine-tuned the software through testing with UNIMARC records from the Czech National Library as well as data from a number of other European UNIMARC implementations. The first converted Czech records were loaded into WorldCat in March 1997. The UNIMARC capability was made available in OCLC Export in late March 1997. Late in calendar 1997, we made available the option of receiving UNIMARC records via the OCLC-MARC Subscription Service. When the new OCLC Cataloging Micro Enhancer for Windows (OCLC CatME) software was released early in 1998, it included a UNIMARC export capability.

Although specifications have been written for many of the changes that have resulted from the three UNIMARC Updates published in 1996, 1998, and 2000, as well as MARC 21 changes during the same period, these have not yet been implemented by OCLC. OCLC does not currently have plans to support the UNIMARC/Authorities, UNIMARC/Classification, or UNIMARC/Holdings formats. Because of our current global strategy, however, it is reasonable to expect that UNIMARC will play an increasing part in OCLC's future thinking and planning.