Abstract

Purpose – ABCD is a web-based open and free software suite for library management derived from the UNESCO CDS/ISIS software technology. The first version was launched officially in December 2009 with a MARC 21 bibliographic format database. This paper aims to detail the building of the UNIMARC bibliographic format database for ABCD.

Design/methodology/approach – ABCD provides tools for building any MARC-like or structured bibliographic format database; this feature is used to set up a UNIMARC bibliographic format database according to its structure as defined by IFLA. As a member of the ISIS family of software applications, ABCD provides a robust formatting language that allows the output of records according to the needs of the user.

Findings – ABCD provides a robust platform for building a UNIMARC bibliographic format database and many other facilities for editing records. But, it is not yet possible to have more than one active fixed length coded field in a worksheet.

Originality/value – CDS/ISIS users who have data in the UNIMARC format on old versions of CDS/ISIS software applications now have the opportunity to migrate their data into the new application without changing their structure. Also, a second international MARC format is provided to new CDS/ISIS users.

Keywords Library automation, Library software, CDS/ISIS, ABCD, UNIMARC, MARC, Standards, Bibliographic systems

Paper type Technical paper

Introduction

Automatisación de Bibliotecas y Centros de Documentación (ABCD, Automation of libraries and documentation centres) is an integrated library management free and open source software (de Smet, 2008, 2009a, 2010, 2011; de Smet and Dhamdhere, 2010) developed since 2007, by the Latin American and Caribbean Center on Health Sciences Information (BIREME, Sao Paulo, Brazil) with the financial support of the Consortium of Flemish Universities (VLIR, Belgium). It is a modernisation of the CDS/ISIS technology, a textual database management system (UNESCO, 1989), first developed and maintained by the International Labour Organisation (ILO) in the 1960s, and later by the United Nations Educational, Scientific and Culture Organisation (UNESCO).

This article is based on a communication presented at the 76th IFLA World Library and Information Congress at the UNIMARC Session, Gothenburg, August 2010 (see www.ifla.org/files/hq/papers/ifla76/148-megnigbeto-fr.pdf). The UNIMARC bibliographic format for ABCD was granted by Universiteit Antwerpen (University of Antwerp, Belgium), via the DOCBIBLAS 2 project, Services Agreement No. 2010/DOCBIBLASII/2.
ABCD is a client/server application of which functions are available throughout a browser; it is written in HTML, PHP, JavaScript, XML, to which is added the word wide web XML IsisScript Server (WXIS)[1], a web server for CDS/ISIS databases developed by BIREME. It provides functions for cataloguing, loans management, users management, acquisition management, selective dissemination of information, statistics generation, and open public access catalogue (OPAC). In order to meet adequately information professionals’ needs and to stand out in the professional community, ABCD should respect libraries automation standards; it should provide users with databases that are compelled to Machine Readable Cataloguing (MARC)[2] standards, making bibliographic data communication and exchange as well as International Standard Bibliographic Description (ISBD) display easier[3]. At the beginning of the software development process, developers have focused on MARC 21 standard implementation; therefore, the first release of the software (December 3, 2009) provides a MARC 21 bibliographic format database as well as a feature that allows the end-user to create any other format database, as structured as MARC 21. The purpose of this paper is the description of a solution rather than a real module developed for ABCD to cover the needs of incorporating and displaying data in UNIMARC format, a feature that is missing from the first version of the software.

CDS/ISIS and MARC formats

Many initiatives had been taken to provide the CDS/ISIS system with MARC bibliographic and authorities databases. Cabral (1989) described Probase, a UNIMARC database based on the CDS/ISIS system. Bokos (1993) dealt with the implementation of a UNIMARC database format with CDS/ISIS in the National Library of Greece. A workshop held in 1993 in Budapest and Barcelona was entirely devoted to CDS/ISIS and UNIMARC (Plassard and Holdt, 1994); it had been an opportunity to CDS/ISIS users to share experiences in MARC data handling with this application. Hopkinson (1994a) produced a CDS/ISIS database to facilitate the exchange of data between institutions participating in the International Information System on Cultural Development. The database was based on and was fully compatible with the Common Communication Format. Hopkinson (1994b) dealt with the capabilities of the CDS/ISIS system to manage UNIMARC data. Hai and Foong (1996) described the development of ONLIS, a customised user interface for computerised MARC cataloguing and information based on the CDS/ISIS MS-DOS version. Hopkinson (1996) underlined the weaknesses and the strengths of the CDS/ISIS system in handling MARC data. Hopkinson (1997) dealt with the building of a UNIMARC format database with the CDS/ISIS system and stated that there is a little difference while implementing others MARC. Hopkinson (1999) described the Unibase database, a UNIMARC bibliographic format database for CDS/ISIS by the Instituto da Biblioteca Nacional e do Livro, Portugal. Buxton and Hopkinson (2001) provided the structure of two MARC format databases namely UNIMARC and the UNESCO’s Common communication format bibliographic (CCF). Freire (2002) built the BIBLIObase database for the Windows version of CDS/ISIS on a UNIMARC structure[4]. However, these initiatives didn’t result in files or applications distributed at central level by UNESCO.

Web based Library Information System (WebLIS) is another application based on the CDS/ISIS technology. As its names implies, its functions are available in a client/server environment via a web-browser. It provides a MARC-like structured
database that is neither MARC 21 nor UNIMARC. ISISMarc is basically a data entry interface for CDS/ISIS databases that works on Windows operating system; it was meant to replace the standard CDS/ISIS Windows version data entry window with a more powerful environment. The application provides both MARC 21 and UNIMARC bibliographic formats databases. ISISMarc is particularly adapted for handling MARC-like formats namely MARC21 and UNIMARC; however, it can be used with any other CDS/ISIS database formats very easily. The application has been listed on the web site of the Library of Congress as MARC 21 compliant system (www.loc.gov/marc/marcsysvend.html). WebLIS and ISISMarc are both distributed by UNESCO and are available from the CDS/ISIS home page.

ABCD and other integrated library management software
CDS/ISIS is a textual data base management system, used for documentation management, mainly in developing countries (Hopkinson, 1997; de Smet, 2008, 2009b). Müller (2009) studied a dozen of free and open integrated library systems (ILS)[5]; he identified four main that are CDS/ISIS[6], PMB, Koha, and Evergreen. Müller (2011) defined three categories of criteria for choosing a free and open ILS that are software licensing, community and functionalities. Among twenty ILS[7], he ranked Koha first, Evergreen second and PMB third on the criteria he defined. Both Koha and Evergreen are listed as compliant systems by the Library of Congress, but PMB is not.

Unlike the standard CDS/ISIS, PMB, Koha and Evergreen are ILS. PMB does not allow direct cataloguing in any MARC-like format (Müller, 2009; Mègnigbèto et al., 2011). Koha provides two MARC formats namely MARC 21 and UNIMARC; but, one should be chosen while installing the software. Evergreen provides only a MARC 21 database (Müller, 2009, 2011). In the tradition line of its ancestor CDS/ISIS, ABCD accepts the coexistence of several databases within a single instance; therefore, a MARC 21 and a UNIMARC databases may coexist within the system. It is the responsibility of the information system the application is installed in to determine which database format it should use.

Technical description of an ABCD database
ABCD database files
As in the standard CDS/SIS, four main types of files are created by the user while setting up a new database in ABCD:

1. The Field Definition Table (FDT) in which the database designer defines all the necessary fields and set their characteristics like tag, label, length, repeatability, subfield identifiers and type.

2. The worksheet, a sort of form where appear defined fields; it enables record creation and edition.

3. The Field Selection Table (FST) defining fields where research terms are extracted from, and the techniques used for the extraction.

4. The format where are specified commands that output records for displaying on the screen or printing to the printer or to a file.

Whereas these four types of files are defined by the end user, most of the other required files (about ten) are generated automatically by the application itself.
Characteristics of an ABCD database field
In the standard CDS/ISIS, a database field is identified by a tag, a label, a length, a type (numeric, alphabetic, alphanumeric, pattern), its repeatability, and the existence of subfields. ABCD profits from HTML and PHP form features and extends field characteristics; therefore, new available field characteristics are:

- **Field type.** Automatic numbering, simple field, field with subfields, subfield attached to a prior defined field, fixed field (as field 008 in MARC 21), date (as field 005 in MARC 21), Leader (as Leader in MARC 21), ISO format date, etc.

- **Data input type.** It is mainly based on HTML and PHP form features through their elements, and Javascript; hence, the data input types are: date selection from a Javascript calendar, text area, fixed length field, table, password, date, ISO date, simple and multiple selection throughout a drop-down list, read only, check box, radio, file uploading, hidden field . . .

- **Selection list.** It is possible to enable the cataloguer to choose from a drop-down list the current field (or subfield) content from a predefined closed list; such contents may come from a database, a thesaurus or an external text file.

Structural comparison between MARC 21 and UNIMARC bibliographic formats
The comparison below is deliberately selective and focuses only on fixed length coded field:

1. The MARC 21 bibliographic format has only one fixed length coded field (008) with 40 positions of which values are obviously coded. Positions 18 to 34 are specific to the type of material being described, and others are common to all types; a contrario, the UNIMARC bibliographic format has one fixed length coded field (100) for all material types, and many other fixed length coded fields (105 to 141), each for a specific material type (IFLA, 2008).

2. The fixed length coded field (008) in MARC 21 does not admit subfields whereas in UNIMARC, all fixed length coded fields (100 and 105 to 141) admits each the subfield a (IFLA, 2008).

Technical limitations for implementing UNIMARC in ABCD
The two elements stressed on in the comparison above (one fixed length coded field versus multiple length coded fields and absence versus presence of subfields in fixed length coded fields) make the difference between the MARC 21 and the UNIMARC formats as fixed length coded fields are concerned, and hindered the implementation of the UNIMARC format with the current version of ABCD. Before a technical solution is provided by developers, the later may be bypassed provisionally thanks to the CDS/ISIS formatting language. Indeed, CDS/ISIS provides formatting commands which allow displaying not only content of fields but also strings the user would like to attach to it, as if they were contained in the database (UNESCO, 1989). In this specific case, the solution consists in prefixing, while displaying it, the content of the fixed length coded field with the subfield identifier[8]. The limit of this solution resides in the fact that, the field content is not affected; but on displaying, it appears as compliant to UNIMARC.
Technically in ABCD, a fixed length coded field is automatically linked to position 6 of the Leader of the database in which it is defined. Such a field is not directly editable by the cataloguer; indeed, ABCD provides a special feature that automatically activates a child window containing a sub-form according to the material being described and indicated in position 6 of the Leader. Therefore, it is impossible to have more than one active fixed length coded field in the same database. Hence, it is not yet possible to define and activate the about 20 fixed length coded fields the UNIMARC bibliographic format requires. So, only field 100, which is mandatory for any resource being described, is defined. It is why all types of fields required by this format has been implemented; the general hypothesis underlining this methodological approach was certainly that, once MARC 21 is implemented, it should be easier to build any other MARC-like database format.

**Mnemonic facilities in a database worksheet**

In UNIMARC as in MARC 21, positions in the fixed length coded fields (100 and 105 to 141 in UNIMARC), in the Leader, in some subfields and even in indicators (when present) are coded. Values to be entered to these positions should take into account the available information according to the material being described. Like PMB, Evergreen, Koha and many other web-based ILS, ABCD uses mnemonic method, as illustrated in Figures 1 and 2, making the job to the cataloguer easier. Cataloguers should not mind the code, but should choose an expressive value and the application should understand the associated code.

**Notes:** Drop down list for the position 6 of the Leader. If the cataloguer chooses *Notated music, except manuscript (c)*, ABCD will store the code *c* into the position 6 of the Leader.

**Figure 1.**
Extract from the worksheet

**Figure 2.**
Extract from the worksheet

**Note:** Drop-down list for the field 200's first indicator
The ISBD and MARC tags display formats

CDS/ISIS provides a formatting language that enables any user (who knows its syntax) to indicate into a file, called format, the desired output of information contained in a record. WXIS, the CDS/ISIS database web server, is based on another CDS/ISIS family application called CISIS[9] that integrates entirely the CDS/ISIS formatting language, extends it and makes it more powerful, so that it may be compared to a programming language. Therefore, a record content may be displayed according to the needs of the user; so, the content of a record may be dumped or displayed according to ISBD recommendations or any other cataloguing rules.

For communication and exchange between systems purpose, bibliographic MARC data are exported to ISO 2709, a machine readable format; but for a direct readable format by human beings, the format “Tag:content” – that we call hereafter MARC tags and Hopkinson (1997) called “diagnostic format” – as well as other formats are more indicated. Records in MARC tags format are available on the web site of many national libraries (Library of Congress, Bibliothèque nationale de France, British Library, Biblioteca Nacional de Portugal, and many other national libraries). The MARC tags display allows knowing at a glance the fields which are present in a record, checking cataloguing or typing rules, and comparing bibliographic description to an authoritative one. The ISBD is a set of recommendations by IFLA for establishing catalogue records. It sets the presentation of a record by giving the list of mandatory elements and the order they should appear in, as well as the punctuation introducing them (Cazabon, 1999; Groupe INTERMARC, 1975). “The ISBD is really a format (...) mainly devoted to the written presentation of bibliographic records, hence to printed bibliographies or to catalogue cards” (Cazabon, 1999).

Two formats have been written to display UNIMARC database records according to these schemes. We present here some bibliographic references according to them (Figures 3-6). In the MARC tags format, all fields from the Label to the notes block are displayed, except the fixed length coded fields (105 to 143) reserved to specific material type, due to the explanations given above.

<table>
<thead>
<tr>
<th>LDR</th>
<th>IND</th>
<th>-----</th>
<th>#22----- 450</th>
</tr>
</thead>
<tbody>
<tr>
<td>005</td>
<td></td>
<td>201012272156.01</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>##</td>
<td>Sa20100912d1962 k y0frey01 01 ba</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>#</td>
<td>SaferSceng</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>##</td>
<td>SaFR</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>#</td>
<td>SaManuel des fontes moulées$\texte imprimé$S$dThe gray iron casting$handbook$S$trad. et adap. par le Centre d’information des fontes moulées$S$zeng</td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>#</td>
<td>SaParis$S$12, avenue Raphaël – Paris 6$S$Ed. techniques des industries de la fonderie$S$d1965</td>
<td></td>
</tr>
<tr>
<td>215</td>
<td>##</td>
<td>Sa705 p.Scill.$S$22 cm</td>
<td></td>
</tr>
<tr>
<td>710</td>
<td>02</td>
<td>SaCentre d’information des fontes moulées</td>
<td></td>
</tr>
</tbody>
</table>

Note: Manuel des fontes moulées by Centre d’information des fontes moulées
Conclusion
ABCD is a new suite of library automation software based on the CDS/ISIS technology. The standard CDS/ISIS system presents some inadequacies while handling MARC data. However, with the new technological environment, the ABCD application has overcome all these weaknesses and gives more flexibility in record edition, mainly as

Figure 4.
ISBD format

Centre d'information des fontes moulées

Note: Manuel des fontes moulées by Centre d’information des fontes moulées

Figure 5.
Marc tags format

LDR IND ----- #22----- 450
001 14426353X
005 201012171640.44
010 ### 0Sa978-0-19-537348-6$bv. 1 : pbk.
010 ### 0Sa0-19-537348-0$bv. 1 : pbk.
100 ### Sa ^a20100512d2010 u y0frey0103 ba.
101 0# Saeng.
102 ### SaGB.
200 1# SaAfrica and the West$sea documentary history$sh1SiFrom the slave trade to conquest, 1441-1905$sfWilliam H. Worger, Nancy L. Clark, Edward A. Alpers.
205 ### Sa2nd ed.
210 ### SaOxfordSoxford University Press$sd2010.
320 ### SaNotes bibliogr. Index.
700 ###1 SaWorger$sbWilliam H.
701 ###1 S$3033693641SaClark$sbNancy LS$f1950----$4070.

Note: Africa and the West by William H. Worger, Nancy L. Clark, Edward A. Alpers

Figure 6.
ISBD format

Africa and the West : a documentary history$swilliam H. Worger, Nancy L. Clark, Edward A. Alpers. – 2nd ed. – Oxford : Oxford University Press, 2010
Worger, William H.
Clark, Nancy L.
Alpers, Edward Alters

Note: Africa and the West by William H. Worger, Nancy L. Clark, Edward A. Alpers
fixed length coded field is concerned. The first version of ABCD provided a MARC 21 bibliographic format database; the building of the UNIMARC database is completed and the resulting files will be available in the version 2.0 of the software mid 2012. These two MARC format databases will co-exist within the distribution of the software; therefore, end users would have to choose the suitable one for their information system. Consequently, there is no need or requirement to convert existing MARC 21 data to UNIMARC.

Notes
1. World wide web XML Isis Script is a web server for ISIS technology databases developed by BIREME. URL: http://bvsmodelo.bvsalud.org/php/level.php?lang=pt&component=28&item=2
2. We use the term MARC to designate all formats derived from the original MARC, as national MARC formats, MARC 21, INTERMARC, UNIMARC, Common Communication Format, etc.
3. This derives from the «philosophy of bibliographic data» relative to the general principles of a format as stated in the Intermarc manual: «It is excluded to go from ISBD to MARC but rather the opposite (Groupe INTERMARC, 1975)
4. BIBLIObase is listed as a MARC specialized tool on the web site of the Library of Congress as MARC 21 (www.loc.gov/marc/marctools.html)
6. In fact, CDS/ISIS is not an ILS.
7. They are (in alphabetic order): Avanti 1.0, BiblioteQ, Emilda, EspaBiblio, Evergreen, Gnuteca, InfoCID, Jayuya, Koha, LearningAccessILS, NewGenLib, oBiblio, OPALS, OpenAmaph醚que, OpenBiblio, Open MarcoPolo, PhpMyLibrary, PMB, Senayan and Weblis.
8. A subfield code is composed of two characters: a delimiter (^, |, or $), and a subfield identifier which is a one alphabetic or numeric character code, in the specific case, the letter a (Comite´ franc¸ois de l'UNIMARC, n.d.); in this case, the code a
9. The CISIS Interface is a library of functions, written in the C programming language, designed to allow the development of CDS/ISIS database applications (without calling the CDS/ISIS software). CISIS applications are fully compatible with CDS/ISIS, including multi-user applications.

References


Further reading

About the author
Eustache Mégnigbêto is an information scientist from Benin. He has a Master’s in Information Science and specializes in library automation. He is an expert on the UNESCO CDS/ISIS bibliographic software and represented Africa at the Third World Congress of ISIS Users held in Rio de Janeiro, Brazil, in 2008. He has been embedded in a team of information professionals to implement the MARC 21 format under ABCD and has taken the leadership for the implementation of UNIMARC under the same software. Eustache Mégnigbêto has published about 30 articles, reports, and communications at international scientific events. Eustache Mégnigbêto can be contacted at: eustachem@gmail.com