



**Une base de données au format bibliographique
UNIMARC sous ABCD**

Eustache Mègnigbêto

Documentaliste

Bureau d'Etudes et de Recherches en Science de l'Information
(BERSI)

Cotonou, République du Bénin

E-mail : eustachem@gmail.com

Meeting: 148. UNIMARC

WORLD LIBRARY AND INFORMATION CONGRESS: 76TH IFLA GENERAL CONFERENCE AND ASSEMBLY
10-15 August 2010, Gothenburg, Sweden
<http://www.ifla.org/en/ifla76>

Résumé :

ABCD est un système intégré de gestion de bibliothèque (SIGB) libre et open source basé sur la technologie CDS/ISIS de l'UNESCO. Il est en cours de développement depuis 2007 par BIREME avec le soutien financier du VLIR. L'application s'utilise dans un environnement client/serveur et ses fonctions sont accessibles via le navigateur. ABCD offre les fonctions de catalogage, de bulletinage, de prêts, de gestion des utilisateurs, d'acquisition, de diffusion sélective de l'information, de génération de statistiques, et de catalogue accessible au public (OPAC). La version 1.0 du logiciel a été lancée le 3 décembre 2009 avec une base de données au format MARC 21. Pour accroître son utilisation parmi les professionnels du monde entier, un format bibliographique UNIMARC s'avère nécessaire. Le travail est en cours et les premiers résultats sont présentés.

Abstract :

ABCD is an open and free software suite for library management based on UNESCO CDS/ISIS software technology and being developed since 2007 by the BIREME, with financial support of the Consortium of Flemish Universities (VLIR, Belgium). ABCD is a web based system that integrates applications for online automation of database creation and management, using open protocols for web based interoperability, cataloguing of document collections, acquisitions, lending modules, and also statistics, periodicals control, creation, and OPAC. The first version of the software was launched officially on December 3rd, 2009 with a MARC 21 bibliographic format database. ABCD provides tools for building any MARC-like or structured bibliographic format database. To be more international and fulfill the desire of its ten of thousand of users, the new ISIS family software member should also provide a UNIMARC format database. The process is ongoing and first results are presented.

Introduction

ABCD, de la dénomination en espagnole *Automatisación de Bibliotecas y Centros de Documentación* (en français, *Automatisation des bibliothèques et centres de documentation*), est un système intégré de gestion de bibliothèque (SIGB). Son développement, en cours depuis 2007, est assuré par le *Centro Latino Americano e do Caribe de Informação em Ciências da Saúde* (BIREME, Centre d'Information sur les Sciences de la Santé pour l'Amérique Latine et les Caraïbes) avec le soutien financier du *Vlaamse Interuniversitaire Raad* (Conseil interuniversitaire flamand, VLIR, Belgique). ABCD est en fait une modernisation de la technologie CDS/ISIS originellement développée et maintenue par l'Organisation Internationale du Travail (OIT) dans les années 60 et reprise plus tard par l'Organisation des Nations Unies pour l'Education, la Science et la Culture (UNESCO). CDS/ISIS est un système de gestion de bases de données textuelles plus utilisé dans la gestion de la documentation, surtout dans les pays en développement (de Smet, 2009, p.3).

ABCD s'utilise dans un environnement client/serveur, et ses fonctions sont accessibles à travers un navigateur. Il est développé en HTML, PHP, JavaScript, XML, auquel s'ajoute le Word Wide Web XML IsisScript Server (WXIS), un serveur de bases de données de la famille CDS/ISIS développé par BIREME. ABCD offre les fonctions de catalogage, de bulletinage, de prêts, de gestion des utilisateurs, d'acquisition, de diffusion sélective de l'information, de génération de statistiques, et de catalogue accessible au public (Open public access catalog, OPAC). Cependant, pour répondre convenablement aux besoins des professionnels et s'imposer dans la communauté, ABCD se doit de respecter les normes en matière d'informatique documentaire; il devrait pouvoir offrir la description bibliographique au format MARC¹, facilitant ainsi la communication et l'échange de données bibliographiques avec d'autres systèmes d'information, et par la même occasion, facilitant la génération de la description bibliographique au format ISBD². Dès l'idée du logiciel, les développeurs avaient en ligne de mire le MARC 21 comme format international; ainsi, au lancement officiel, le 3 décembre 2009, la version 1.0 du logiciel était livrée avec une base de données au format bibliographique MARC 21, avec la possibilité pour tout utilisateur de construire tout type de base de données aussi structurée que le MARC 21. C'est ainsi que le format CEPAL, utilisé par des institutions documentaires latino-américaines en réseau, est construit. Le second format international de la famille MARC, UNIMARC, est l'objet du travail de l'équipe de développeurs.

Structure d'une base de données ABCD

ABCD est une modernisation de CDS/ISIS développé par l'UNESCO pour la gestion des bases de données textuelles, mais utilisé en pratique dans la gestion des données bibliographiques. ABCD hérite de CDS/ISIS la structure de sa base de données qui comprend au minimum une douzaine de fichiers distincts physiquement, mais fonctionnellement unis. Quatre fichiers fondamentaux sont créés par le concepteur de la base de données; ce sont :

¹ Nous parlons de « format MARC » pour désigner tous les formats issus du ou consécutifs au MARC originel comme les MARC nationaux, INTERMAC, UNIMARC, et le MARC21.

² Cela découle de la « philosophie des données bibliographiques » relatives aux principes généraux d'un format tels qu'énoncés dans le manuel d'INTERMARC : « Il n'est pas question de passer de l'ISBD au format MARC mais l'inverse » (Groupe INTERMARC 1975, 13)

- 1) la Table de Définition des Champs qui est le fichier dans lequel le concepteur définit l'ensemble des champs de la base de données, ainsi que leurs caractéristiques comme l'étiquette, le nom, la longueur, la répétitivité, le type;
- 2) le Bordereau, un genre de "formulaire" où figurent les champs définis pour la base de données; il permet la création et la modification de notices;
- 3) la Table de Sélections des Champs : c'est dans ce fichier que le concepteur définit les champs de provenance des termes de recherche ainsi que la manière dont leur contenu est extrait;
- 4) le format : c'est un fichier où sont définies les sorties possibles du contenu des notices de la base de données. Il s'utilise pour envoyer les données à l'écran, à l'imprimante ou encore dans un fichier. Ainsi, une même base de données peut avoir plusieurs formats.

Si ces quatre types de fichiers sont créés par l'utilisateur, les autres sont automatiquement générés par le logiciel lui-même.

Les caractéristiques d'un champ d'une base de données ABCD

Dans le CDS/ISIS standard, un champ d'une base de données est caractérisé par une étiquette, un nom, une longueur, un type (numérique, alphabétique, alphanumérique), la répétitivité et l'existence ou non de sous-champs. Profitant des facilités du formulaire HTML et PHP qu'il intègre désormais, ABCD étend les caractéristiques d'un champ. Ainsi, on distingue comme caractéristiques nouvelles :

- 1) le type de champ : numéro automatique, champ fixe (comme la zone 008 du MARC 21), date (comme la zone 005 du MARC 21), Guide (comme le Leader du MARC 21), date (comme dans MARC), date ISO, ...
- 2) le type de saisie : il repose essentiellement sur les possibilités offertes par un formulaire HTML à travers ses types d'éléments. Ainsi, on distingue les types de saisie suivants : zone de texte, texte de longueur fixe, table, mot de passe, date, date ISO, liste déroulante simple, liste déroulante multiple, champ en lecture seule, case à cocher, bouton radio, fichier à télécharger, champ caché, ...
- 3) la liste de choix : il est possible que le contenu d'un champ ou d'un sous-champ soit à choisir dans une liste fermée; une telle liste peut provenir d'une base de données ABCD, d'un thésaurus ou encore d'un fichier texte externe.

Comme on peut le constater, les développeurs ont ciblé le MARC 21 dans leur travail. Ils ont ainsi donné la possibilité de créer des champs conformes à une "manipulation" facile du MARC 21, rendant agréable le travail aux catalogueurs.

Comparaison structurelle entre MARC 21 et UNIMARC

Cette comparaison est volontairement sélective et porte uniquement sur deux différences :

- 1) MARC 21 dispose d'une seule zone codée de longueur fixe (008) dont les valeurs de chacune des positions sont codées. Dépendant du type de matériel à décrire et précisé à la position 6 du guide, les valeurs des positions de la zone 008 varient. A contrario, UNIMARC a prévu une zone codée de longueur fixe (100) pour tous les types de matériel, et plusieurs autres zones codées de longueur fixe (105 à 141), chacune pour un type de matériel spécifique. Rappelons que l'équipe de développement avait pour priorité, dans la version 1.0 du logiciel, une base de données MARC 21. C'est

pourquoi tous les types de champs nécessaires à la réalisation de ce format ont été créés; l'hypothèse générale qui sous-tendait cette démarche méthodologique était qu'une fois le MARC 21 mis en œuvre, il deviendrait facile de construire tout autre format de base de données à lui semblable. Au total, il y a au moins 20 champs codés de longueur fixe dans l'UNIMARC bibliographique.

- 2) Par ailleurs, la zone codée de longueur fixe (008) du MARC 21 n'admet pas de sous-zones alors que dans UNIMARC, l'ensemble de la vingtaine de zones codées de longueur fixe qui équivalent au champ 008 du MARC 21 admettent chacun une sous-zone de code *a*.

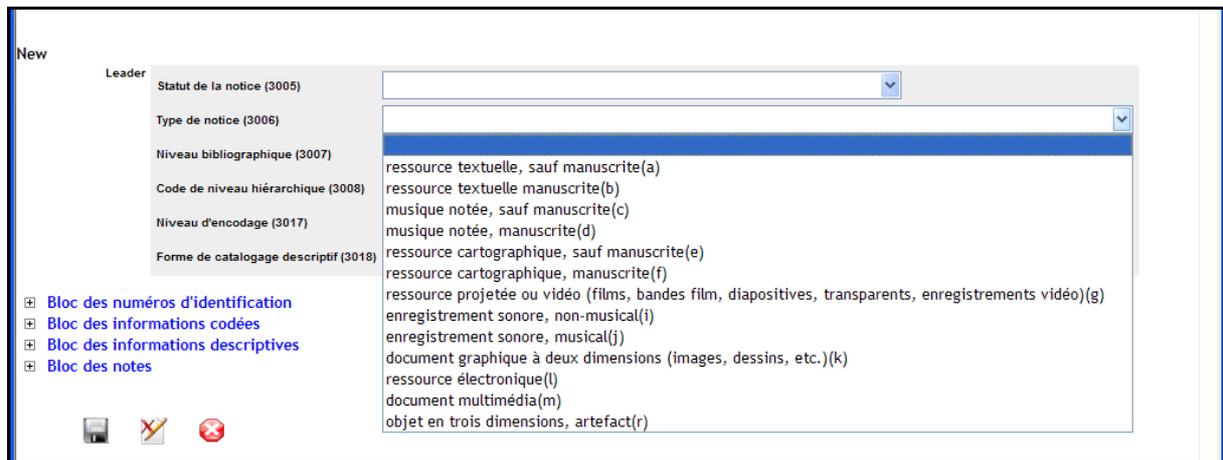
Un champ de type fixe dans ABCD est automatiquement relié à la position 6 du champ Guide de la base de données dans laquelle il est défini. Ce type de champ n'est pas directement éditable par l'utilisateur. ABCD offre une facilité qui active automatiquement, pour ce champ, une fenêtre fille dont les champs présents dépendent de la valeur de la position 6 du Guide sélectionnée par l'utilisateur. Dans MARC 21 en effet, la zone 008 (en correspondance de laquelle ce type de champ est prévu dans ABCD), a 40 positions dont celles allant de 18 à 34 incluses sont spécifiques au type de matériel décrit et les autres communes à tous les types. Du fait qu'un champ de type fixe de ABCD est automatiquement lié à la position 6 du Guide, il est impossible d'avoir deux champs de ce type présents dans une même base de données. Par conséquent, il n'est pas encore possible de créer la vingtaine de zones codées de longueur fixe de l'UNIMARC dans la base de données bibliographique UNIMARC sous ABCD.

Ces deux éléments de différence constituent donc des handicaps à la mise en œuvre d'UNIMARC sous ABCD dans sa version actuelle. La résolution technique est en cours d'étude par l'équipe de développement. En attendant, une solution provisoire pourrait être appliquée au second élément à travers le format d'affichage des données; en effet, le langage de formatage de CDS/ISIS permet d'afficher outre des données extraites de notices, des chaînes de caractères voulus par l'utilisateur à l'endroit désiré. Dans le cas d'espèce, cette solution consiste à faire précéder le contenu de chaque champ codé de longueur fixe du code de sous-zone³. La limite de cette solution, cependant, réside dans le fait que les données elles-mêmes telles que stockées dans les notices ne sont pas affectées.

³ Un code de sous-zone est composé de deux caractères : un séparateur (^, |, ou \$), et un code sur un caractère alphabétique ou numérique (Comité français de l'UNIMARC), dans le cas d'espèce le code *a*.

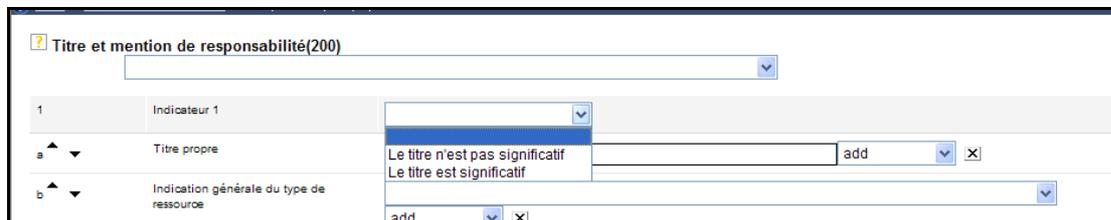
La facilité mnémotechnique du bordereau de saisie

Dans UNIMARC, les positions dans les zones codées de longueur fixe (100 et 105 à 141), du Guide ou dans certaines sous-zones ou même des indicateurs (lorsqu'ils sont présents) sont codées. Les valeurs à entrer à ces positions doivent tenir compte de l'information disponible selon le type de document à décrire. ABCD utilise la méthode mnémotechnique de la liste des choix du formulaire HTML pour faciliter le travail à l'utilisateur.



The screenshot shows a web-based form for entering UNIMARC data. On the left, there is a 'Leader' section with several dropdown menus for fields like 'Statut de la notice (3005)', 'Type de notice (3006)', 'Niveau bibliographique (3007)', 'Code de niveau hiérarchique (3008)', 'Niveau d'encodage (3017)', and 'Forme de catalogage descriptif (3018)'. Below these are four expandable sections: 'Bloc des numéros d'identification', 'Bloc des informations codées', 'Bloc des informations descriptives', and 'Bloc des notes'. The main area of the form is a large dropdown menu for position 6 of the Guide, which is currently open. It lists the following options with their mnemonic codes in parentheses: 'ressource textuelle, sauf manuscrite(a)', 'ressource textuelle manuscrite(b)', 'musique notée, sauf manuscrite(c)', 'musique notée, manuscrite(d)', 'ressource cartographique, sauf manuscrite(e)', 'ressource cartographique, manuscrite(f)', 'ressource projetée ou vidéo (films, bandes film, diapositives, transparents, enregistrements vidéo)(g)', 'enregistrement sonore, non-musical(i)', 'enregistrement sonore, musical(j)', 'document graphique à deux dimensions (images, dessins, etc.)(k)', 'ressource électronique(l)', 'document multimédia(m)', and 'objet en trois dimensions, artefact(r)'. At the bottom left of the form, there are three small icons: a printer, a pencil, and a red 'X'.

Figure 1. Extrait du bordereau de saisie. Liste de choix pour la position 6 du Guide.



The screenshot shows a web-based form for entering UNIMARC data. At the top, there is a dropdown menu for 'Titre et mention de responsabilité(200)'. Below it, there is a table with two rows. The first row is for 'Indicateur 1' and has a dropdown menu that is currently open, showing two options: 'Le titre n'est pas significatif' and 'Le titre est significatif'. The second row is for 'Titre propre' and has a text input field with the value 'add' and a dropdown menu with 'add' and 'x'. The third row is for 'Indication générale du type de ressource' and has a text input field with the value 'add' and a dropdown menu with 'add' and 'x'.

Figure 2. Extrait du bordereau de saisie. Liste de choix pour le premier indicateur de la zone 200.

Le format d'affichage

CDS/ISIS est doté d'un langage de formatage qui donne à tout utilisateur (qui en connaît la syntaxe) la possibilité d'écrire la sortie désirée des informations contenues dans une notice. ABCD fonctionne, entre autres, avec WXIS qui, lui, est un serveur Web pour les bases de données CDS/ISIS. WXIS est basé, lui, sur une autre application appelée CISIS qui intègre entièrement le langage de formatage du CDS/ISIS qu'il a rendu plus puissant et étendu au point où l'on peut le comparer à un véritable langage de programmation. Ainsi, l'on peut générer à partir des données en ABCD un affichage en ISBD (ce qui n'est pas l'objet de cet article) ou en format UNIMARC en vrac.

Pour la communication et l'échange de données bibliographiques entre systèmes, les données de la famille MARC s'exportent au format ISO 2709, un format lisible "uniquement" par ordinateur. Pour une lecture directe, le format *étiquette contenu* est plus indiqué. Ce type de format est offert par les grandes bibliothèques comme les bibliothèques nationales (par

exemple la Library of Congress, la Bibliothèque Nationale de France, la British Library, Biblioteca Nacional de Portugal) sur leur site web. Il permet de saisir d'un coup d'œil les zones présentes dans une notice, et par la même occasion permet la vérification des règles de catalogage ou de saisie, et une comparaison avec des descriptions bibliographiques de référence. Nous avons écrit un format sous ABCD pour afficher les données UNIMARC sous cette forme. Le format lui-même n'est pas l'objet de cet article. Nous nous contentons de présenter ici quelques références bibliographiques en UNIMARC sous ABCD. Toutes les zones du Guide au bloc des notes sont présentes, sauf évidemment les zones codées de longueur fixe (105 à 143) réservées aux types de matériel spécifique, compte tenu des explications ci-dessus.

```
LDR -----nam0 22  ## 450
005 201005171205.25
010 __$a2-7384-7768-2
100 __$a 20100511d1999  m y0frey03  zz
101 0_ $afre
200 1_1 $aLa démocratie au Bénin$e1988-1993$ebilan et perspectives$fPhilippe Noudjenoume
210 1_ $aParis$cL'harmattan$d1999
215 __$a395 p.$d21 cm
225 $aCollection Sociétés africaines & diasporas$fdirigée par Babacar Sall
```

Figure 3. Description bibliographique sous ABCD avec les étiquettes des zones UNIMARC: *La démocratie au Bénin : 1988-1993 par Philippe Noudjènoumè.*

```
LDR -----nam# 22  ## 450
005 201005171206.35
010 __$a2-7384-4138-6
100 __$a      d      m y0frey03  ba
101 0_ $afre
200 1_ $aLe Bénin$eles hommes et leur milieu$eguide pratique$fBernard Passot
210 1_ $aParis$cL'harmattan$d1996
215 $a335 p.$ccouv. ill. en coul.$d21 cm
300 $aIndex
```

Figure 4. Description bibliographique sous ABCD avec les étiquettes des zones UNIMARC: *Le Bénin : les hommes et leur milieu par Bernard Passot.*

Conclusion

Tout d'abord, faisons un ajustement terminologique : nous avons, en effet, traité d'un logiciel documentaire qui a son langage et également d'un format de communication de données, l'UNIMARC, qui a également le sien; nous avons utilisé les termes *ABCD champ* et *sous-champ* qui équivalent respectivement dans UNIMARC aux termes *zone* et *sous-zone*.

ABCD est un Système intégré de gestion de bibliothèque (SIGB) jeune sur le marché professionnel. Nous nous attelons à le mettre aux normes pour la satisfaction des utilisateurs de la technologie CDS/ISIS. Notre objectif immédiat est le format bibliographique UNIMARC qui n'est pas encore livré avec le logiciel. Ensuite viendra la base de données des autorités pour alimenter les différents accès de l'UNIMARC bibliographique. Le travail continue et durera encore plusieurs mois.

Références bibliographiques

de Smet, Egbert. 2009. *The abc of ABCD : the reference manual : version 1.0*. URL : http://bvsmodelo.bvsalud.org/download/abcd/manuais/ABCofABCD_atualizado.pdf (Consulté le 12 février 2010).

Comité français UNIMARC. *UNIMARC/B 3 : Structure du format*. In : Manuel UNIMARC : format bibliographique. URL : http://multimedia.bnf.fr/unimarc_b_trad/Bsection3-5-2006.pdf (Consulté le 26 avril 2010)

Groupe INTERMARC. 1975. *INTERMARC (M) ; format bibliographique d'échange pour les monographies : manuel*. Paris : Groupe INTERMARC.