

MARC VÀ XML

Ninh Xuân Phúc & Trần Thị Mộng Linh
Phòng Tài nguyên thông tin – Thư viện NH Khoa học Tự nhiên

Hiện nay, cùng với sự phát triển của công nghệ thông tin, người ta nói nhiều hơn về sự liên thông thư viện. "Chuẩn hóa – Hoá nhập – Phát triển" hiện đang là hướng phát triển chủ yếu của các thư viện. Muốn hóa nhập và phát triển theo hướng hiện nay thì các hệ thống phải "hợp" nhau. Hiện nay có rất nhiều hệ thống của các thư viện, bao gồm cả chuẩn hóa từ các trung tâm kỹ thuật. Trong bài viết này, chúng tôi chia sẻ về những nghiệp công tác trong lĩnh vực thư viện mà soái kiện thời học tập hiện tại là trong nội dung và ngoài và MARC và XML. Nội dung không nhằm trau dồi kiến thức tin học mà chỉ cung cấp những khái niệm cơ bản về ngôn ngữ của trung tâm nội dung dùng trong kỹ thuật lõi trội trình bày và trao đổi thông tin.

KHAI NIỀM MARC

Chuẩn biến mục MARC (Machine Readable Cataloging) – Biến mục có thể đọc bằng máy, là hình dạng cho phép máy tính lõi trội và truy xuất thông tin mục lúc bao gồm các biến ghi theo dạng MARC. Nghĩa là ngôn ngữ biến mục cần phải có sẵn thông tin trong biến ghi.

Vào những năm 1960, nhân viên Thư viện Quốc hội Hoa Kỳ (Library of Congress) đã phát triển một hình dạng tiêu chuẩn cho việc lõi trội các thông tin biến mục trên máy tính. Cho nên nay mỗi khoa lõi thông tin biến mục không loàiai nhau lõi trội Chứng nhận Thư viện Quốc hội Hoa Kỳ soái lõi thông tin biến ghi dạng MARC là hơn 50 triệu.

1. Tính chất của

Tính chất của biến ghi theo hiện ôi tiêu chuẩn là giao tiếp với nhau và có thể là một phần của một tiêu chuẩn. Nhìn nay cho phép người sử dụng mục lúc có thể tìm thấy tất cả các tài liệu liên quan đến cùng một tiêu đề.

2. Biến ghi MARC

Mỗi biến ghi MARC có thể chia thành bởi 3 thành phần:

- Cấu trúc biến ghi: nội dung xuất phát từ Tiêu chuẩn quốc gia Hoa Kỳ và trao đổi thông tin.
- Hình thức nội dung: bao gồm các mảng và các qui định do MARC format nêu ra. Chuẩn nêu ra các yêu cầu để tạo ra một biến ghi và cho phép máy tính xử lý nó.
- Nội dung nội dung của biến ghi là thông tin mà ta đưa vào qua quá trình sử dụng tiêu chuẩn biến mục như AACR2 và biến mục nemus LCSH.

3. Các tröông vànhán tröông

Trong heáthóng maiý tính, mot bieu ghi lamot tap hóp các tröông coi liên quan. Trong bieu ghi MARC, mot tröông bao gồm thông tin maihoa (vd. Ngày nhập vào heáthóng) và thông tin thô tách (vd. Môitäüvat lyihay Tiêu ñeàñeàmuic). Mot tröông coi nhán nhán daing, nhán này gồm 3 kyü töi. Vd. 250 – Thông tin län xuất bain.

4. Cấu trúc mot bieu ghi: gồm 3 phần chính

Nội bieu: tröông ñau tiên của bieu ghi, có 24 kyü töi.

Thô tách: do maiý tính cañ cõi vào bieu ghi thô tách tao ra. Nộinhán dieñ nhán tröông năo ñööic söïduing trong bieu ghi và sõi cho ñao. Phần này không hiện thi mà danh cho maiý tính quan lyi.

Các tröông coiñoadai thay ñoi: gồm 2 loai: Tröông ñieu khien vaströông döölieu

Tröông ñieu khien chöia maiñhóng tin ñööic söïduing trong quaütrình xöülyücaç bieu ghi.

Tröông döölieu chöia ñööic các thông tin thô tách của bieu ghi, chaing hañ nhö thanh phần moä taij các dañ muic chính và dañ muic boåsung, tieu ñeàñeàmuic, kyühiệu phần loai, v.v.

5. Thủphản tích mot bieu ghi MARC:

Phiếu mục lục:

Brown, Vinson, 1912 –
Exploring Pacific Coast tidepools / by Vinson Brown and Ane Rovetta. – Rev.
and expanded ed. – Happy Camp, Calif. : Naturegraph Publishers, 1996.
v, 127 p. : ill. (some col.) ; 22 cm.
Includes bibliographical references (p. 119) and index.
ISBN 0879612177 (alk. Paper)

1. Tide pool animals – Pacific Coast (U.S.) – Identificaion. 2. Tide pool plants
– Pacific Coast (U.S.) – Identification. 3. Tide pool ecology – Pacific Coast
(U.S.) I. Rovetta, Ane.
DDC no: 574.979 20th ed.

Phiếu này ñööic moitäütheo maiMARC 21:

000	00952nam###2200265#a#4500
001	00012188151
003	CaOOAMICUS
005	9970910000000.0
008	960325s1996####caua#####b#####001#0#eng##

020 \$a0879612177 (alk. Paper)
 040 \$aLC\$beng\$cLC\$dLC
 043 \$an-us---\$ap-----
 050 00 \$aQH104.5.P32\$bB76 1996
 082 00 \$a574.979\$220
 100 1 \$aBrown, Vison,\$d1912-
 245 10 \$aExploring Pacific Coast tidepools /\$cby Vinson Brown, Ane Rovetta
 250 \$aRev. And expanded ed.
 260 \$aHappy Camp. Calif. :\$bNaturegraph Publishers,\$c1996
 300 \$av, 127 p. :\$bill. (some col.) ;\$c22 cm.
 504 \$aIncludes bibliographical references (p. 119) and index.
 650 0 \$aTide pool animals\$zPacific Coast (U.S.)\$xIdentification.
 650 0 \$aTide pool plants\$zPacific Coast (U.S.)\$xIdentification.
 650 0 \$aTide pool ecology\$zPacific Coast (U.S.)
 700 1 \$aRovetta, Ane.

Bieu ghi mai MARC tren nööc hieu nhö sau:

000 00952nam##2200265#a#4500

Nau bieu (000) goi 24 kyitöi, yinghoa cac vò trí:

vò trí 1-5 lañnoadai logic cua bieu ghi (00952) do may tao ra;

vò trí 6: tinh trang bieu ghi (n: moi);

vò trí 7: Loai bieu ghi (a: Tai lieu bang ngoi ngoi van ban, goi ca microform, microfilm, microfiche);

vò trí 8: cap tho töch (m: chuyen kha);

vò trí 9: Kieu kiem tra (#: kieu kiem tra khong cu the);

vò trí 10: khong xac nönh;

vò trí 11: soach thö luon lai2;

vò trí 12: maotrööng con, luon lai2;

vò trí 13-17: Nöa chæ DSDL (00265);

vò trí 18: Möc nöömahoa (#: maohoia nay nui);

vò trí 19: Hình thöc bieñ muic moita (a: Qui tac moita Anh-My);

vò trí 20: Yeu cau bieu ghi lieñ quan (#: khong yeu cau);

vò trí 21-24: Sö nöatö töch ainh xai, luon lai4500.

001 00012188151

Soakiem soat (001) nööc gain cho toachöc tao lap, söidung hoac phan phoi bieu ghi, nööc theahien trong trööng 003.

003 CaOOAMICUS

Nhañ dañg soakiem soat (003)

005 9970910000000.0

Ngay và thời gian thõi hiện thao tại gần nhất (005): gồm năm tháng ngày giờ phút giây

008 960325s1996####caua#####b####001#0#eng##

Các yếu tố/dữ liệu có/nhóm (008), ý nghĩa

1-6: Ngày nhận tin, do máy tối nồng gần (96/03/25)

7: Kieu năm xuất bản và tình trạng xuất bản (s: ché biết mỗi năm/năm có/theo).

8-11: Năm 1, năm này nêu gồm 4 chữ số/cứa vị trí 7 (1996)

12-15: Năm 2, thõi hiện nồng thời với việc lựa chọn mã cho vị trí 7 (####: không sử dụng).

16-18: Nôi xuất bản, sán xuất hoặc thõi hiện, gồm 2 hoặc 3 ký tự/theo bảng mã quốc gia theo chuẩn MARC 21 (cau: California)

19-22: Minh họa, tối đa 4 ký tự, vị trí không sử dụng chứa dấu trống (a: Cùm minh họa, ###: không sử dụng)

23: Nội töông nõi (#: không xác định)

24: Đăng tài liệu (#: không thuộc các đăng nát biết)

25-28: Tính chất nội dung, tối đa 4 ký tự (b: Thủ mục,###)

29: Xuất bản phim của chính phủ (#: không phải xuất bản của chính phủ)

30: Xuất bản phim hoa nghệ (0: Không phải hoa nghệ)

31: Tài liệu kyniem (0: không phải)

32: Bảng tra (1: cùm bảng tra)

33: Không xác định, luôn chứa dấu #

34: Thể loại (0: không phải viên töông)

35: Tiêu số (#: không phải tài liệu tiêu số)

36-38: Ngôn ngữ(eng: Tiếng Anh)

39: Biểu ghi nõõc số nõi (#: không số nõi)

40: Nguồn biến mục (#: Cò quan biến mục quốc gia).

020 \$a0879612177 (alk. Paper)

Số saich theo tiêu chuẩn quốc tế(020)

040 \$aLC\$beng\$cLC\$dLC

Nguồn biến mục (040), \$a: Cò quan biến mục gốc (LC), \$b: Nguồn ngõi biến mục (eng: Tiếng Anh), \$c: Cò quan chuyên tài (LC), \$d: Cò quan sốn chứa (LC)

043 \$an-us---\$ap-----

Mã khu vực nõa lỵ(043) lấy từ/bảng mã khu vực, gồm 7 ký tự, ký tự nào không dùng nõõc the/bảng dấu gạch ngang (-), \$a: Mã khu vực nõa lỵ(n-us: Hoa Kỳ)

050 00 \$aQH104.5.P32\$bB76 1996

Ký hiệu xếp giáicủa TVQH Hoa Kỳ cùi2 ché thò:

ché thò 1 (0: cùm trong bộ/sou tập của LC),

ché thò 2: Nguồn của ký hiệu xếp giá(0: Nõõc gần bôi LC)

\$a: Ché soaphân loai (QH104.5.)32

\$b: Soatai lieu (B67 1996)

082 00 \$a574.979\$220

Ký hiệu phân loại Dewey, có 2 chữ số

Chữ số 1: Dãy số bản (0: số bản này nút)

Chữ số 2: nguồn ký hiệu Dewey (0: nguồn gốc bối LC)

\$a: Phân loại (574.979)

\$2: Số tài liệu (20)

100 1 \$aBrown, Vison,\$d1912-

Dãy số chính – tên tác giả/cảnh (100), chữ số 1 dùng cho tên riêng

\$a: Họ tên cảnh (Brown, Vison)

\$d: Năm liên quan với tên cảnh (1912-)

245 10 \$aExploring Pacific Coast tidepools /\$cby Vinson Brown, Ane Rovetta

Thông tin về bản tin (245), có 2 chữ số

Chữ số 1: Bổ sung dãy số (1: Cố bộ bổ sung)

Chữ số 2: Các ký tự bat nhau không sắp xếp (0: số ký tự bat nhau không xôi lyi sap xep, nghĩa là sắp xếp tự do nhau trên)

\$a: Nhan nhè

\$c: Thông tin trach nhiệm

250 \$aRev. And expanded ed.

Lần xuất bản.

260 \$aHappy Camp. Calif. :\$bNaturegraph Publishers,\$c1996

Thông tin xuất bản: \$a – Nơi xuất bản, \$b – Nhà xuất bản, \$c – Năm xuất bản

300 \$av, 127 p. :\$bill. (some col.) ;\$c22 cm.

Mô tả và lý \$a – Khoi loong tai lieu (v, 127 p. : 5 trang nainh soi La maivai 127 trang nainh soi Alrap), \$b – Các chi tiết khác (ill. : minh hoa, some col. : mot so minh hoa mau), \$c – Khoi (22 cm.)

504 \$aIncludes bibliographical references (p. 119) and index.

Tham khảo thö töch van chæ muic (504)

650 0 \$aTide pool animals\$zPacific Coast (U.S.)\$xIdentification.

650 0 \$aTide pool plants\$zPacific Coast (U.S.)\$xIdentification.

650 0 \$aTide pool ecology\$zPacific Coast (U.S.)

Tieu nua nua muic (650) \$a: Nua muic chinh, \$b: Tieu phan muic, \$x: Phu nua chung, \$z: Phu nua nua lyi

700 1 \$aRovetta, Ane.

Dãy số bổ sung (700), số chữ số 1- xác định tên riêng, \$a: tên cảnh (Rovetta, Ane).

6. Một số vấn đề cần lưu ý

Minh họa trên cho thấy MARC format rất chất chơi chuẩn mực, nội bộ chính xác. Do vậy, ngày càng nhiều người, đặc biệt các người nói tiếng Anh và cả hệ thống thư viện dùng tiếng Anh sõi dưng nhầm lẫn

Tuy nhiên, chuẩn lỏng MARC format không phù hợp cho việc chuyển giao dữ liệu trong các thư viện nên vì tính chất phức tạp của nó. Daeng MARC ra đời nhằm phục vụ cho quá trình phân loại và in theo mục lục truyền thống, khi đó chủ đề là sinh vật và trao đổi thông tin qua maingroup. Sau này laryukieen phát triển của Dick R. Miller trình bày ở Hội nghị Hội Thư viện Hoa Kỳ tại Chicago năm 2000.

Các phần tử và thuộc tính của chúng liên lõi với nhau khiến cho việc quản lý thông tin khó khăn hơn.

Vd. 700 1 \$aBillings, John Shaw, \$d1813-1938,\$eCollector

Phân tích ví dụ này, ta thấy:

Phan töü \$a Billings : Höi

John Shaw: ten

\$d 1813-1938 : niet bij u

Thuộc tính:

700 Dañ müç bojsung

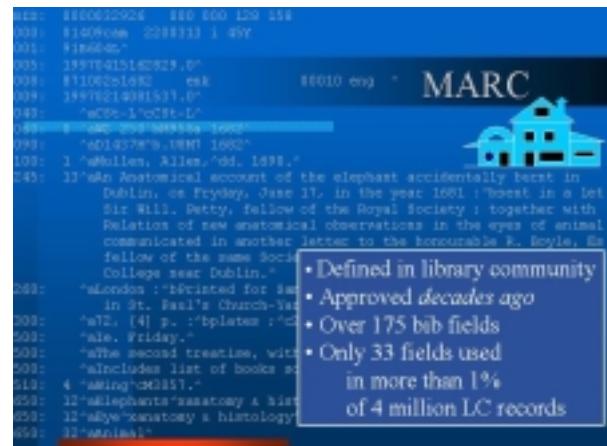
1 ten ñaiñ

\$eCollector : trách nhiệm liên quan

Nhain trööng 700, danh ñeà lòu thö tich
nhain vat, nhain dieñ caic phan töi teñ caiñnhain.
Trong thöc teí vung döi lieü nay chöia hoï, teñ,
ngay sinh, ngay mat. Caic yeú toá thuoc tính
ñooic ñat ôi caitrööic vañsau caic phan töi trööng
con 'e' ôi cuoi dong. Nhañ trööng 700 ñòng
thöi con mang thuoc tính moï taibøa sung trach
nhiem lieñ quan. Ñieùu nay gaÿ cain tröüñoi voi
viet quan trò thông tin.

Ông cho biết kết quả nghiên cứu hôn 40.000 biểu ghi trong hôn 4 triệu biểu ghi của TVQH chép sõi dung 33 tröông trong soá 175 tröông thò tách nhööc caú tao trong MARC. Taic giàu icon phän tích nhööng khöù khän khai, chaing hän veà boä kyü töi, cheá ñoä hien thò, caic biểu mẫu, nhañ tröönq, tuy meim deö, linh ñoäng nhööng phöic taip, trung laip, v.v.

Và Dick R. Miller kể tóm tắt: "Có thể hình dung MARC nhò mờ ngoài nhau tách rời, rỗng rã, nhöng cuôkyivàñööc söà chöa nhieù lañ. Nöìng trööc nhu caù hiêñ ñaiñ hoà, noigay cho ngööi ta cám qiaç lõöng lõi qiaç trung tu hay xáy môi".



XU HỘIING MÔI - XML

Bieu ghi MARC coi thei chia hon 800 troieng moataidoi lieu chi tieu neaninh daeng thei nhong phan lom cao troieng nay lai khong can thiet va sau nhieu nho khong soi dung trong cao thô vien nien tot hien nai. Trong thô vien nien tot quai trinh toachoi cao kho doi lieu luon doa tren cao he quan tri CSDL voi chiec nang giup ngoai soi dung nhanh choing tim kiem va truy cap thong tin troi tiep thong qua mang may tinh, khong dung nen cao theitho muc in san. Han chean oai cau bieu ghi MARC veaphoeng phap trao noi doi lieu lau noic khou noic va viet boi con ngoai, moataicong nha c moi soatriong voi noadai coaninh, han chea trong viet hoatroi cao ngoai ngoitheo chuan UNICODE khi can soi dung nen 04 bytes cho moi kyuttoi.

Xu hoiing hien nay tren thegioi, naie biet lai oicao quoct gia chua hinh thanh cao kho doi lieu thong tin – thô vien theo daeng MARC, lai chuyen qua soi dung **chuan bien muc MARC voi ngoi ngoi moataidoi lieu XML**. Doi lieu noic troi tiep trao noi neanlien thong tren mang Internet totcait website, khong can soi dung nen giao thoi lieu thô vien quoct teZ39.50.

1. XML lai gi?

XML (eXtensible Markup Language) ngoi ngoi nainh dau moi roing coi nguoin goi totngoin ngoi nionh daeng sieu van ban HTML (Hyper Text Markup Language), caihai ngoi ngoi nay neun bat nguoin totchuan ngoi ngoi nionh daeng van ban tong quat coicau truc SGML (Structured General Markup Language).

XML lai ngoi ngoi nhoic nionh nghia boi toa chiec mang toan cau (World Wide Web Consortium), thôiing nhoic viet tat theo cach chiec chiec lai W3C. Nay lai toachoi quoct teaninh ra cauchuan cu Web vanInternet.

Mot van ban XML hinh thanh totcait thei(tag) voi ten goi phan tot(element). Khaic voi ngoi HTML, soi loeung van ten goi cao phan tot trong XML lai khong han chei XML lai ngoi ngoi totng quat dung nionh nghia doi lieu thong qua cao thei Trong HTML cao thei nhoic nionh nghia van qui nionh troi. Trong khi noi voi XML ta coi thei tuy yinonh nghia moi thei Nhieu vay coi thei coi XML nhieu tap cha cuu ngoi ngoi HTML. Doa van mot soiqui tac, XML toi ton tai vanphat trien tot thanh cao ngoi ngoi nionh nghia khaic.

Niem quan troieng nhat lai XML cho phep deidang xoi lyi chuyen tai vantrao noi doi lieu gioi rat nhieu ong dung van tai lieu ngoi dung voi cao nionh daeng khaic nhau. Neu naiquen voi may tinh, han ta biет rang coi rat nhieu nionh daeng file khaic nhau. Viet chuyen noi doi lieu gioi chung quai lauan gai ma duoc coi khong it trinh ong dung hoatroi. Ví dui nhieu file .DOC (Van ban Word), .XLS (Ban tinh Excel), .DBF (Lap trinh FoxPro, .MDB (Lap trinh Access), .TXT (File van ban), .RTF (Rich Text Format) van moi nay lai .HTML. Chie rieng cao file van ban thoai nai nui gai khou chou, neu ban nhan nhoic mot file Word 2000 ma may tinh cuu ban con dung Word 7.0, coi gaing lam cung chie nhoic phan van ban con cao noi dung khaic thôiing bi bien daeng.

Trong XML, dữ liệu và trình diễn không lỏng ôi làng vàn bain vàn coi theo de dang cau hình cung nhõ thay nõi chung bang caic trình soan thaib thõong neub khõi trong tay trình soan thaib XML chuyên nghiệp. Dữ liệu vàn caic thei trong XML khõi maõ hoia, khõi nõi hoai bain quyen.

Tháng 12/1997, phiên bain hau tiên XML 1.0 (Extensible Markup Language – Nguõi ngõi nãinh daú môiroõng) ra nõi vaølaøchuan nôn gian cuà SGML. Từñõi nhieu công ty phan mềm nãøi chaïy nua öng dung XML vaø nhieu lõnh vöc. Hàng trám ngoi nõi nõinh daeng chuyên dung döia trên XML nãøra nõi. Nien hinh moi soatuy bien ngoi nõi nõinh daeng döia trên XML cho thay söc mainh cuà XML:

- BITS – Banking Industry Technology Secretariat : Ngoi ngoi van phong ve kyi thuât nghiep vuøng ngan hang
- IFX – Financial Exchange : Trao nõi döilieu tai chinh
- BIPS – Banking Internet Payment System : Heäthoõng thanh toan qua Internet cuà nghiep vuøng ngan hang
- TIM – Telecommunication Interchange Markup : Nõinh daeng trao nõi vien thõong
- CBL – Common Business Library : Thô vien kinh doanh phoøthoõng
- ebXML – XML kinh doanh nien töi
- PDML – Product Data Markup Language : Ngoi ngoi nõinh daeng döilieu san pham
- FIX – Financial Information eXchange Protocol : Giao thõi trao nõi thõong tin tai chinh
- CML – Chemical Markup Language : Ngoi ngoi nõinh daeng trong lõnh vöc hoia hoic, cho pheip bieu dien caic công thõi hoia hoic, hoia tri phan töiöidaeng nõa hoia,
- V.V.

Döilieu thõong tin – thô vien roïraøng cung laømoø lõnh vöc tiem naing. Dung chuan XML:

- coitheataø bieu ghi thô tõch mot lan vaøxuat bain chung theo caic daeng khaic nhau;
- hién thi bieu ghi thô tõch tröc tiep trên trình duyệt Web, search engines (công cui tìm kiém) vaøcaic heäthoõng thô vien tiem naing khaic markhoõng caøn chuyen nõi;
- bieu ghi thô tõch coitheanõoic chuyên nõi qua lai giøa XML vaøMARC markhoõng bì ton that.
- nhieu van nẽatoøn tai trong nõinh daeng MARC nõoic khaic phuic, ví duï nhõ viet kiem soat tieu nẽachuan.

Nam 1995, TVQH Myø bat näu nghiên coiù tính khai thi cuà viet dung SGML (Standard Generalized Markup Language – Chuân ngoi ngoi nõinh daeng van bain toøng quat coicau truc) nea maõ hoia nõinh daeng MARC 21. Sau nõi phiên bain MARC DTDs (Document Type Definitions) nea nghia loaii van bain MARC nea nõinh nghia döilieu MARC 21 trong daeng thõi SGML nõoic phat hanh nam 1998. Cung nam nay TVQH Myø công boø phan mềm chuyên nõi giøa MARC 21 vaøSGML.

2. Chuyển biến ghi MARC sang XML

Nhiều nghiên cứu cuôc tranh luận về việc mã hóa biến thi thô tách theo chuẩn MARC nên các biến ghi mà không rõ ràng hơn và có thể hoàn toàn trong môi trường Internet. Tại giai K. T. Lam, Nhiều học Khoa học và Kỹ thuật Hồng Kông nghiên cứu ứng dụng XML để cài đặt một liên kết giữa tiêu đề chính và các tiêu đề không rõ ràng thiết lập trình bày kết quả sau:

Một trong những vấn đề lớn trong mục lục MARC là cách trình bày và kết nối tiêu đề tại các nghiên cứu đang tên khác nhau, không ghi bằng nhiều ngôn ngữ. Có một số cách để giải quyết vấn đề này là: đặt tên bằng tiếng thường không rõ ràng phiên âm bằng nhiều ngôn ngữ khác nhau.

Trong biến ghi MARC chỉ cho phép chọn một hình thức tên trong CSDL thô tách. Như câu hiện tại là phải theo tên không rõ ràng là cái tên gọi của cùng một tác giả. Sau đây là một biến ghi tiêu đề tại các giá trị là một nhà báo, nhà văn Hồng Kim Dung từ TVQH Mỹ

```
....  
100 1 $a Jin, Yong,$d1924-  
400 1 $aChin, Yung,$d1924-  
400 1 $aZha, Liangyong,$d1924-  
400 1 $a Cha, Louis,$d1924-  
400 1 $a Cha, Liang-yung,$d1924-  
400 0 $a Kim-Dung,$d1924-  
400 1 $a Kim, Dung,$d1924-  
400 0 $aJinyong,$d1924-  
400 1 $a Yong, Jin,$d1924-  
400 1 $a Kin, Yo ,,$d1924-  
....
```

Biến ghi này gặp phải những vấn đề sau:

- Chỉ bao gồm những hình thức tên theo chối cái La tinh, còn tên tiếng Hoa không ghi rõ ràng.
- Trong ví dụ này, "Jin, Yong" không rõ ràng là tên của Kim Dung không rõ ràng là thông tin tung hô (400). Tuy nhiên, những quốc gia khác lại muốn chọn hình thức tên khác làm tiêu đề là tiếng Trung Quốc sẽ chọn hình thức tên tiếng Hoa làm tiêu đề.

Nếu tên tiếng Hoa của tác giả Kim Dung vào biến ghi MARC, tại giai K. T. Lam nêu trên sử dụng hai cách, không rõ ràng A và B trong môi trường UCS/Unicode (UCS – Universal Character Set – Bộ ký tự toàn cầu) như sau:

Kiểu A

001 oca00560270

....

100 1 \$6880-01\$a Jin, Yong,\$d1924-

880 1 \$6100-01\$a ,,\$d1924- [ghi chුව Sau trööng con \$a lai Tiếng Hoa]

400 1 \$aChin, Yung,\$d1924-

400 1 \$6880-02\$a Zha, Liangyong,\$d1924-

880 1 \$6400-02\$a ,,\$d1924- [ghi chුව Sau trööng con \$a lai Tiếng Hoa]

400 1 \$a Cha, Louis,\$d1924-

400 1 \$a Cha, Liang-yung,\$d1924-

....

Kieu B

001 oca00560270

....

100 1 \$a Jin, Yong,\$d1924-

400 1 \$aChin, Yung,\$d1924-

400 1 \$a Zha, Liangyong,\$d1924-

400 1 \$a ,,\$d1924- [ghi chුව Sau trööng con \$a lai Tiếng Hoa]

400 1 \$a Cha, Louis,\$d1924-

400 1 \$a Cha, Liang-yung,\$d1924-

700 1 \$a ,,\$d1924- [ghi chුව Sau trööng con \$a lai Tiếng Hoa]

....

Lúc này khi so sánh trööng song song cùa theo 880 và trööng con 6 trong kieu A, chúng ta có thể thấy

- Tên tiếng Hoa tööng nhööng vôi tên neaphien am "Jin, Yong"
- Duy trì kết nối cùa daeng ngoan ngööigoi và các biến thai cùa noi trong các tên tiếng neakhoang thiết lập. Ví dụ: [tiếng Hoa] và "Zha, Liangyong" có mối liên kết thông qua trööng con \$6 nhööic khai bao trong caihai nhanh tööng 400 (thông tin tung thö) và 880 (đãn müc tung thö):

400 1 \$6880-02\$a Zha, Liangyong,\$d1924-

880 1 \$6400-02\$a ,,\$d1924- [ghi chුව Sau trööng con \$a lai Tiếng Hoa]

Ở kieu B, nhanh tööng 700 (Các đản müc boi sung) nhööic dung nea tao [tiếng Hoa] tööng nhööong vôi neamüc thiết lập "Jin, Yong" (hoặc cung coitheanhap Chin, Yung và theo 100 và "Jin, Yong" và theo 700). Tuy nhiên, Kieu B không duy trì nhööic cheñoilien kết song song nein nhööng hình thöic tên không nhööic thiết lập tieu nea. Ví dụ, [tiếng Hoa] và "Zha, Liangyong" không liên kết. Ông này chúng ta coithean i từ Kieu A sang Kieu B không không ni ngööic lai nhööic.

Dùng Kieu A hay Kieu B nếu có ý nghĩa thuận lợi và bất lợi riêng. Tuy nhiên, nếu dùng XML để định danh dữ liệu tiêu chuẩn, không rắc rối hay hoán toán không khác phức. Vì vậy, chúng ta có thể tái sử dụng.

```
<?xml version="1.0" encoding="UTF-8"?>
<marc name="authority" cdate="19980625" udate="19980625" rcn="ABrG">
.....
<fd id="0" name="001" ind1="" ind2="" label="Control Number">
    <sf name="">oca00560270</sf>
</fd>
<fd id="1.1" script="cjk.chinese" name="100" ind1="1" ind2="b" label="Author">
    <sf name="a">    </sf><sf name="d">1924-</sf>
</fd>
<fd id="1.2" script="latin.pinyin" name="100" ind1="1" ind2="b" label="Author">
    <sf name="a">Jin, Yong,</sf><sf name="d">1924-</sf>
</fd>
<fd id="1.3" script="latin.wadegiles" name="100" ind1="1" ind2="b" label="Author">
    <sf name="a">Chin, Yung,</sf><sf name="d">1924-</sf>
</fd>
<fd id="2.1" script="cjk.chinese" name="400" ind1="1" ind2="b" label="See From Author">
    <sf name="a">    ,</sf><sf name="d">1924-</sf>
</fd>
<fd id="2.2" script="latin.pinyin" name="400" ind1="1" ind2="b" label="See From Author">
    <sf name="a">Zha, Liangyong,</sf><sf name="d">1924-</sf>
</fd>
<fd id="2.3" script="latin.wadegiles" name="400" ind1="1" ind2="b" label="See From Author">
    <sf name="a">Cha, Liang-yung,</sf><sf name="d">1924-</sf>
</fd>
<fd id="3" script="latin.english" name="400" ind1="1" ind2="b" label="See From Author">
    <sf name="a">Cha, Louis,</sf><sf name="d">1924-</sf>
</fd>
.....
</marc>
```

Dùng bảng định kiểu XSL (eXtension Style Sheet – Bảng định kiểu mô hình) để triển khai ra kiểu dữ liệu mà một hệ thống thông tin có thể chuyển thành cách Kieu A và Kieu B mà không gặp phải khó khăn nào. Mỗi phần <fd id ...> sẽ </fd> là thông tin về một hình thức tiêu đề trong nó bao gồm cả cách chia nhận rõ ràng và rõ ràng con MARC, v.v., có thể ghi lại không tại cách cách hình thức tên bảng mỗi hình thức ngoài ngoặc của một cách giao tiếp cần biểu ghi rõ ràng kỹ thuật hình thức tiêu đề tên cách cách tên.

Minh họa nêu trên cho thấy khai nang khác phuć nhöng bat cap ve ket noi hoac chuyen giao döilieu con ton tai trong nönh dang MARC.

Thiet nghö, söi phat trien nhanh choa töng coi cuia công nghe thông tin ñai van nang chi phoi hoat nöong thông tin – thô vien. Trong qua trinh phat trien, MARC ñai nöong vai tro quan trọng, giao quyết nööc nhiều van ñeai ñinhöng bööc ñau tin hoac hoa thô vien, tao ra hieü qua xai hoai khöng nöihap öng nhanh choing nhu cau lœu tröivastim kiem thông tin. Khoa hoac thông tin khöng döng lai ma tiep tuic phat trien, nhöng nööc niem cuia MARC format se ñööc khać phuć. Công nghe XML nang dan thay thea MARC. Nieu này khöng coi ngua MARC bù loai boi hoan toan ma ñööc phat trien len möc ñoäcao hon, hien ñai hon, ñoila soi dung **chuan bien muc MARC voi ngoi ngoi moat doi lieu XML**.

TÀI LIỆU THAM KHẢO

HOANG LE MINH. *Giai phap ky thuat hoa tröi lieu thong tin – thô vien. Soatay quan ly thong tin – thô vien* / Nguyen Minh Hiep chuu bien. – TP. HCM : Ñai hoac Quoc gia, 2002.

K. T. Lam. *Moving from MARC to XML* <http://ihome.ust.hk/~lblkt/xml/marc2xml.html>

MILLER, DICK R.. *XML and MARC : A Choice or Replacement?* <http://elane.stanford.edu/laneauth/ALAChicago2000.html>

MORTIMER, MARY. *Kien thöc cõi bain vea MARC 21.* – H. : Cty Nam Hoang, 2001.

NGUYEŃ PHÖÖNG LAN. *XML : Nein taing vaøing dung* / Nguyen Phööng Lan, Hoang Nöic Hai. – TP. HCM : Giai Duic, 2001.

"Coi thea hinh dung MARC nhö mot ngoi nhau tieu nghi, roäng rai, nhöng cuikyova ñööc söa chöa nöieu lai. Nöing trööit nhu cau hien ñai hoa, noigaij cho ngööi ta cám gaij lööng löt giööa trung tu hay xaij möi"

Dick R. Miller