

# CẤU TRÚC CỦA KHỔ MẪU BIÊN MỤC ĐỌC MÁY MARC21

*PGS.TS. Đoàn Phan Tân*

Việc ứng dụng tin học trong xử lý thông tin càng thúc đẩy yêu cầu phải chuẩn hoá công tác biên mục. Tiêu chuẩn ISBD chưa hoàn toàn thích nghi với cách xử lý tin học đối với các mô tả thư mục. Các phương pháp xử lý tin học đòi hỏi dữ liệu phải được cấu trúc hoá. Để máy tính có thể nhận biết được các dữ liệu thư mục, các chỉ dẫn không những phải tuân theo những quy tắc chặt chẽ, mà còn phải được trình bày một cách chính xác theo một khổ mẫu thống nhất.

Khổ mẫu hay format là hình thức trình bày dữ liệu trên biểu ghi được cấu trúc hoá. Các dữ liệu trên biểu ghi được sắp xếp thành các trường, trường con, kết hợp với các mã số và các chỉ thị để điều hành sự sắp xếp này sao cho có thể nhận biết, trình bày và tìm kiếm dữ liệu bằng máy tính.

phân loại thập phân Dewey,... Các *Khổ mẫu biên mục đọc máy - MARC*

Năm 1966 lần đầu tiên Thư viện Quốc hội Mỹ xuất bản tập quy tắc của khổ mẫu MARC. MARC là từ viết tắt của cụm từ tiếng Anh: Machine Readable Cataloging có nghĩa là biên mục đọc được bằng máy. Khổ mẫu MARC là một mô tả có cấu trúc, dành riêng cho các dữ liệu thư mục được đưa vào máy tính điện tử. Nó là khổ mẫu cho phép máy tính lưu giữ và truy xuất thông tin.

Khổ mẫu MARC do Thư viện Quốc hội Mỹ xây dựng đầu tiên và sử dụng nên còn gọi là USMARC. Sau khi được chỉnh lý vào năm 1968, format MARC là cơ sở cho sự ra đời của một loạt các format quốc gia như: CAN.MARC của Canada, UK.MARC của Anh, AUS.MARC của Úc, INTERMARC của Pháp, IBERMARC của Tây Ban Nha. Mỗi hệ thống biên mục sử dụng format MARC đều có hướng dẫn riêng về hệ thống phân cách và hình thức trình bày các biểu ghi của mình.

Năm 1997 USMARC của Thư viện Quốc hội Mỹ kết hợp với CANMARC của Thư viện Quốc gia Canada tạo thành MARC21 và trở thành format chuẩn được nhiều phần mềm quản trị thư viện sử dụng.

Sau đây ta sẽ tìm hiểu cấu trúc của khổ mẫu MARC21.

Cấu trúc của khổ mẫu MARC là cấu trúc biểu ghi, trong đó các dữ liệu thư mục được sắp xếp trong các trường, có độ dài xác định, được mã hoá và trình bày theo một quy định chặt chẽ. Khổ mẫu MARC sử dụng các chữ số, chữ cái, các ký hiệu ngắn gọn đặt ngay trong biểu ghi thư mục để đánh dấu và nhận biết các loại thông tin khác nhau trong mỗi biểu ghi.

Mỗi biểu ghi của khổ mẫu MARC bao gồm các trường (fields). Ngoài các trường dành cho các yếu tố mô tả thư mục theo AACR2 như: nhan đề, thông tin về trách nhiệm, thông tin về xuất bản, từng thư, đặc trưng số lượng, phụ chú, tóm tắt, v.v... còn có các trường dành cho đề mục chủ đề, ký hiệu phân loại của Thư viện Quốc hội Mỹ, ký hiệu trường này lại có thể

chia nhỏ thành các trường con (subfields). Vì tên của trường thường khá dài nên trong biểu ghi MARC mỗi trường được biểu diễn bằng một nhãn gồm 3 chữ số.

Để tiện trình bày, người ta tập hợp các nhãn trường thành từng nhóm. Các trường có nhãn bắt đầu bằng số “0” thuộc nhóm trường “0XX”, các trường có nhãn bắt đầu bằng số “1” thuộc nhóm trường “1XX”, các trường có nhãn bắt đầu bằng số “2” thuộc nhóm trường “2XX”, v.v...

Theo bản MARC21 đầy đủ cho dữ liệu thư mục, MARC21 có 10 khối trường chính:

- 0XX: Trường điều khiển, các chỉ số nhận dạng và phân loại
- 1XX: Tiêu đề mô tả chính (tên cá nhân, tên tập thể)
- 2XX: Nhan đề và thông tin liên quan đến nhan đề.
- 3XX: Mô tả vật lý
- 4XX: Thông tin về từng thư
- 5XX: Phụ chú
- 6XX: Các tiêu đề mô tả theo chủ đề
- 7XX: Các tiêu đề mô tả bổ sung khác chủ đề, từng thư
- 8XX: Tiêu đề mô tả bổ sung về từng thư
- 9XX: Thông tin nội bộ

MARC21 có khoảng trên 200 trường (không kể trường con). Qua thống kê người ta thấy rằng chỉ có khoảng 10% số nhãn được sử dụng thường xuyên trong các biểu ghi MARC, 90% còn lại ít khi được sử dụng.

Các nhãn trường thường được sử dụng là:

- 010 Số kiểm tra của Thư viện Quốc hội Mỹ
- 020 Chỉ số sách quốc tế (ISBN)
- 022 Chỉ số ấn phẩm định kỳ quốc tế (ISSN)
- 043 Mã vùng địa lý
- 050 Chỉ số phân loại của Thư viện Quốc hội Mỹ (LCC)
- 082 Chỉ số phân loại thập phân Dewey (DDC)
- 100 Tiêu đề chính, tên cá nhân (tác giả)
- 245 Nhan đề và thông tin về trách nhiệm
- 250 Thông tin về lần xuất bản
- 260 Thông tin về xuất bản, phát hành
- 300 Mô tả vật lý (đặc trưng số lượng)

|     |                                      |
|-----|--------------------------------------|
| 440 | Thông tin về từng thư                |
| 520 | Chú giải hay tóm tắt                 |
| 650 | Điểm truy nhập chủ đề, đề mục chủ đề |
| 700 | Điểm truy nhập bổ sung, tên cá nhân  |

Danh sách tất cả các nhãn trường được in trong 2 tập *MARC 21 Format for Bibliographic Data* do Văn phòng Phát triển mạng lưới và tiêu chuẩn MAC của Thư viện Quốc hội Mỹ xuất bản.

Để hỗ trợ cho công tác biên mục Văn phòng này còn cho xuất bản các tài liệu bổ trợ sau:

- MARC21 format for authority data.
- MARC21 format for classification data.
- MARC21 format for holdings data.
- MARC21 format for community information.
- MARC21 code list for geographic areas.
- MARC code list for countries.
- MARC code list for languages.
- MARC code list for organizations.

Các tài liệu trên phối hợp với nhau để tạo ra một biểu ghi thư mục hoàn chỉnh, kiểm soát tính nhất quán của các tiêu đề mô tả, các thuật ngữ chủ đề, hỗ trợ phân loại chính xác, tạo tham chiếu qua lại giữa các tiêu đề và chủ đề, giữa các ký hiệu hay mục lục phân loại, theo dõi và thông báo vốn tài liệu hiện có gắn liền với công tác bổ sung và lưu thông.

Ngoài nhãn trường, trong biểu ghi MARC một số trường còn được xác định bằng các chỉ thị (indicators). Chỉ thị được mã hoá bằng hai chữ số (từ 0 đến 9) đi theo sau nhãn trường. Có trường chỉ dùng chữ số thứ nhất hoặc thứ hai, có trường dùng cả hai. Trong các trường không cần chỉ thị người ta bỏ trống.

Ví dụ:

|     |    |                                                                                                           |
|-----|----|-----------------------------------------------------------------------------------------------------------|
| 245 | 14 | \$a The emperor's new clothes / \$c adapted from Hans Christian Andersen and illustrated by Janet Stevens |
|-----|----|-----------------------------------------------------------------------------------------------------------|

Trong ví dụ trên, 245 là nhãn trường nhan đề và thông tin về trách nhiệm, 14 là chỉ thị. Chỉ thị thứ nhất "1" có nghĩa là trong phiếu mục lục phải in từ "Title" vào trước nhan đề. Còn chỉ thị thứ hai "4" có nghĩa là bỏ qua 4 ký tự đầu của nhan đề (T, h, e và dấu trống). Khi đó nhan đề "The emperor's new clothes" sẽ được ghi ra là "emperor's new clothes".

Trong biểu ghi MARC, trường con được nhận biết bởi mã trường con (subfield codes), đó là một ký tự bằng chữ in thường, đặt sau dấu ngăn cách \$ (delimiter).

Ví dụ:

|     |           |            |            |
|-----|-----------|------------|------------|
| 300 | \$a 675p. | : \$b ill. | ; \$c 24cm |
|-----|-----------|------------|------------|

Trong đó 300 là nhãn trường mô tả vật lý của cuốn sách, bao gồm các trường con \$a (Số trang), \$b (Thông tin minh hoạ), \$c (Khổ, cỡ). Trường này không cần chỉ thị.

Ví dụ:

|     |               |                   |          |
|-----|---------------|-------------------|----------|
| 260 | \$a Newyork : | \$bChelsea House, | \$c 1986 |
|-----|---------------|-------------------|----------|

Trong đó 260 là nhãn trường xuất bản, bao gồm các trường con: \$a (Nơi xuất bản), \$b (Nhà xuất bản), \$c (Năm xuất bản).

Các nhãn trường, chỉ thị và mã trường con là các dấu hiệu để nhận biết và điều khiển cách bố trí các trường và trường con, do các chương trình quản trị CSDL quy ước khi xây dựng. Khi một biểu ghi thư mục đã được đánh dấu một cách chính xác và được lưu trữ dưới dạng một tệp dữ liệu trên máy tính thì chương trình quản trị CSDL sẽ đánh dấu và tạo khuôn dạng cho các thông tin này để in ra thành một bản thư mục, một phiếu mục lục hay hiển thị trên màn hình. Các chương trình này còn cung cấp công cụ tìm kiếm các thông tin thoả mãn yêu cầu đặt ra, dựa trên các điểm truy nhập nằm trong các trường của biểu ghi MARC.

Ngoài các trường dữ liệu thư mục là phần chính của biểu ghi thư mục, mỗi biểu ghi MARC còn có các thành phần cố định sau, xuất hiện ở đầu mỗi biểu ghi:

a. *Đầu biểu (Leader)*. Đầu biểu gồm 24 ký tự đầu tiên của biểu ghi, ứng với 24 vị trí, mỗi vị trí này được gán cho một ý nghĩa xác định và thể hiện bằng một mã (chữ in thường, chữ số hoặc khoảng trống) cho ta biết các thông tin về trạng thái và các thuộc tính của biểu ghi như: độ dài biểu ghi, loại hình tài liệu, cấp thư mục, mức độ mã hoá, quy tắc mô tả được sử dụng (ISBD, AACR),... Nhiều thông tin trong đầu biểu ghi là để dành cho máy tính sử dụng để nhận dạng biểu ghi.

Ví dụ: Giả sử một biểu ghi MARC có đầu biểu:

01401cam#2200265#a#4500

thì ý nghĩa cơ bản của nó là:

| Vị trí | Ký tự | Ý nghĩa ký tự của vị trí                  |
|--------|-------|-------------------------------------------|
| 0- 4   | 01041 | Độ dài biểu ghi: 1041 ký tự               |
| 5      | c     | Trạng thái biểu ghi: đã sửa chữa          |
| 6      | a     | Dạng tài liệu: văn bản in                 |
| 7      | m     | Cấp thư mục: sách chuyên khảo             |
| .....  |       |                                           |
| 18     | a     | Quy tắc mô tả được sử dụng: AACR2         |
| 19     | #     | Không yêu cầu các biểu ghi liên kết       |
| 20-23  | 4500  | Sơ đồ các mục trong thư mục, luôn là 4500 |

b. *Danh mục (Directory)*. Tiếp theo đầu biểu là một loạt các tiểu dẫn (bằng các chữ số) có độ dài xác định gọi là danh mục. Danh mục này cho biết các nhãn trường có trong biểu ghi, độ dài của trường và vị trí bắt đầu của trường trong biểu ghi. Danh mục được tạo ra bởi máy tính từ biểu ghi thư mục, dựa trên các thông tin đã nhập. Nó không phải là phần MARC hiển thị cho người biên mục hay người sử dụng mục lục. Nó chỉ được sử dụng bởi người lập trình và máy tính để thực hiện việc trao đổi các biểu ghi thư mục giữa các hệ thống.

Ví dụ: Đoạn tiểu dẫn sau đây trong danh mục

. . . 245003600354250001200390260003700402 . . .

có nghĩa là: trong biểu ghi có trường với nhãn 245, có độ dài là 36 ký tự và bắt đầu từ vị trí thứ 354. Tiếp theo là trường có nhãn 250, có độ dài 12 ký tự và bắt đầu ở vị trí thứ 390 ( $36 + 354 = 390$ ). Tiếp theo nữa là trường có nhãn 260, có độ dài 12 ký tự, bắt đầu từ vị trí 402 ( $402 = 12 + 390$ ). Như vậy đoạn tiểu dẫn trên trong danh mục có cấu trúc và ý nghĩa như sau:

| Nhãn    | Độ dài | Bắt đầu từ vị trí |
|---------|--------|-------------------|
| . . . . |        |                   |
| 245     | 0036   | 00354             |
| 250     | 0012   | 00390             |
| 260     | 0037   | 00402             |
| . . . . |        |                   |

Tóm lại mỗi biểu ghi MARC phải bao gồm các thành phần cơ bản sau:

- Đầu biểu gồm 24 ký tự.
- Một danh mục các trường dữ liệu mà với mỗi trường dữ liệu phải bao gồm một nhãn với 3 chữ số, độ dài của trường dữ liệu và vị trí của ký tự đầu tiên.
- Các dữ liệu tương ứng với các trường trong danh mục, chứa các dữ liệu thư mục cần xử lý với độ dài thay đổi.

Dưới đây là ví dụ về một biểu ghi MARC:

```
LEADER 00718nam//2200217/a/4500
001    ///93154367
003    DLC
005    19951019113246.0
008    940827s1993///at/ac///b///001/0/eng//
010    Sa///93154367
020    $a1863731695 (pbk.) :$c$22.95
043    $au-at---
050 00 $aHQ1391.A8$bS28 1993
082 00 $a320.994082$220
```

100            \$aSawer, Marian,\$d1946-  
245 12        \$aA woman's place :\$bwoman and politics in  
                  Australia /\$cMarian Sawyer and Marian Simms  
250            \$a2nd ed.  
260            \$aSt. Leonards, N.S.W. :bAllen & Unwin,\$c1993  
300            \$axiii, 345 p. :\$bill., ports. ;\$c22 cm.  
500            \$aPrevious ed.: Sydney : Allen & Unwin, 1984.  
500            \$aIncludes index.  
504            \$aBibliography: p. 309-328.  
650 0        \$aWomen in politics\$zAustralia.  
650 0        \$aWomen legislators\$zAustralia.  
700 1        \$aSimms, Marian.

Trong biểu ghi trên, 001 là nhãn trường số kiểm soát, 003 là nhận dạng số kiểm soát, 005 là ngày và thời gian thực hiện giao tác gần nhất, 008 là các thông tin chung được mã hoá về biểu ghi.

Trong biểu ghi MARC, các trường cách nhau bởi dấu phân cách trường. Trong MARC21, dấu phân cách trường ký hiệu là ^. Dấu phân cách trường đặt ở cuối mỗi trường. Người ta dùng ký hiệu \ để đánh dấu kết thúc mỗi biểu ghi.

Cấu trúc biểu ghi của khổ mẫu MARC21 là cấu trúc điển hình của khổ mẫu biên mục hiện đại, với ba thành phần cơ bản: đầu biểu, danh mục các trường, các dữ liệu tương ứng với các trường trong danh mục. Hệ thống chỉ định nội dung với các nhãn trường, chỉ thị và mã trường con, các ký tự phân cách được quy định rất chi tiết và cụ thể, tạo thuận lợi cho máy tính nhận biết, sắp xếp và tìm kiếm dữ liệu. MARC21 phù hợp với tiêu chuẩn biên mục hiện đại AACR2 nên tạo khả năng biên mục linh hoạt cho nhiều loại hình tài liệu trong thư viện. MARC21 được xây dựng dựa trên chuẩn ISO.2709 nên việc trao đổi thông tin được thực hiện dễ dàng.

Cấu trúc biểu ghi của khổ mẫu MARC tạo ra nhiều khả năng cho máy tính lựa chọn và sắp xếp các dữ liệu thư mục:

- Cho phép người sử dụng dễ dàng truy cập tới các biểu ghi.
- In ra các thông báo sách mới, các ấn phẩm thư mục, các mục lục dưới dạng thức khác nhau, các nhãn trên gáy sách.
- Trao đổi dữ liệu thư mục với các thư viện khác trong nước và trên thế giới.

Khổ mẫu MARC có ý nghĩa quan trọng trong biên mục tự động. Thực chất của biên mục tự động là sử dụng một phần mềm tư liệu hoặc sử dụng phân hệ biên mục của phần mềm quản trị thư viện để tạo lập các biểu ghi cho một CSDL thư mục và tạo ra các mục lục thích hợp.

Trong biên mục tự động việc tạo lập biểu ghi thường là xử lý tiền máy và nhập dữ liệu, do con người thực hiện. Người ta nhập dữ liệu qua các khổ mẫu hiển thị trên màn hình theo kiểu xử lý văn bản. Còn việc tổ chức và sắp xếp biểu ghi và biên soạn các mục lục thì do máy tính thực hiện. Máy tính có thể in các phiếu mục lục và chế bản cho các ấn phẩm thư mục. Đó là sản phẩm đầu ra của CSDL thư mục.