

# **MARC hay Dublin Core? :**

## **Việc chuyển đổi**

### **MARC-Dublin Core và**

### **Dublin Core-MARC**

**NGUYỄN MINH HIỆP, BA., MS.**  
GD Thư viện ĐH Khoa học Tự nhiên  
Đại học Quốc gia Thành phố Hồ Chí Minh

#### **Trao đổi biểu ghi thư tịch (bibliographic record) dạng liên biên (analog)**

Thư viện truyền thống lưu trữ và phục vụ tài liệu in ấn. Hệ thống mục lục thư viện giúp cho người sử dụng có thể tra cứu tài liệu lưu trữ trong thư viện luôn được cải tiến để đáp ứng yêu cầu phát triển số lượng tài nguyên thông tin cũng như nhu cầu đọc giả ngày càng cao. Từ hệ thống tra cứu bằng phiếu mục lục đến việc sử dụng mục lục trực tuyến, người cán bộ thư viện đã trải qua một tiến trình biên mục để xây dựng nên Hệ thống mục lục tra cứu trong thư viện: Từ phiếu mục lục được viết tay rồi đánh máy trên từng phiếu hay quay roneo phản ánh công việc biên mục riêng lẻ cho từng thư viện đến việc xây dựng hệ thống Mục lục liên hợp chung cho nhiều thư viện, người ta đã nghĩ đến vấn đề chuẩn hóa để có thể chia sẻ nhau công việc biên mục. Các thư viện trao đổi với nhau từng phiếu mục lục theo một tiêu chuẩn chung để xây dựng Mục lục liên hợp. Công việc biên mục mang một sắc thái khác là phải kiểm tra từng phiếu mục lục có đúng theo quy định của chuẩn thư tịch (bibliographic standards) không. Do đó công việc này mang một danh xưng là Kiểm soát thư tịch (bibliographic control) – Kiểm soát lý lịch của sách có được thể hiện đúng không. Cách trao đổi này so sánh từng thành phần trong phiếu mục lục. Tại Hoa Kỳ, Thư viện Quốc hội đảm trách việc phân phối phiếu mục lục.

Việc sử dụng máy tính trong hoạt động thư viện đã làm thay đổi rộng lớn công tác biên mục. Ban đầu thay vì Thư viện Quốc hội Hoa Kỳ cung cấp phiếu mục lục cho các thư viện thì khi sử dụng máy tính họ chỉ cung cấp các biểu ghi trên băng từ, các thư viện nhận băng từ, download dữ liệu, và tự in thành phiếu mục lục. Việc sử dụng máy tính như thế đòi hỏi phát sinh ra một chuẩn thư tịch mà máy tính đọc được, chuẩn này ban đầu được gọi là MARC I về sau cải tiến thành MARC II, rồi USMARC. Công việc trao đổi biểu ghi như thế này thật là thuận tiện cho việc in phiếu mục lục và về sau khi việc ứng dụng tin học phát triển, hệ thống mục lục trực tuyến thay thế hệ thống phiếu mục lục, chuẩn MARC tỏ ra rất phù hợp. Các quốc gia trên thế giới theo chân Hoa Kỳ phát minh ra hàng loạt MARC quốc gia. Về sau khi vấn đề xuất bản mang tính toàn cầu, có yêu cầu cần thống nhất chuẩn MARC cho nên UNIMARC của IFLA ra đời, tuy nhiên Hoa Kỳ, quốc gia có nguồn tài nguyên thông tin đồ sộ và hệ thống xuất bản bao trùm thế giới đã không chấp nhận UNIMARC do đó các quốc gia nói tiếng Anh khác dần dần bỏ MARC quốc gia để sử dụng USMARC. Để mang tính toàn cầu, USMARC kết hợp với CANMARC của Canada tạo nên MARC 21.

MARC phản ánh toàn bộ những chi tiết của quy tắc biên mục Anh-Mỹ AACR2 nên khá phức tạp. Người sử dụng MARC cần phải được đào tạo chính quy, khi thành thạo họ cũng phải tốn rất nhiều thời gian để tạo nên một biểu ghi MARC, thường thì mất

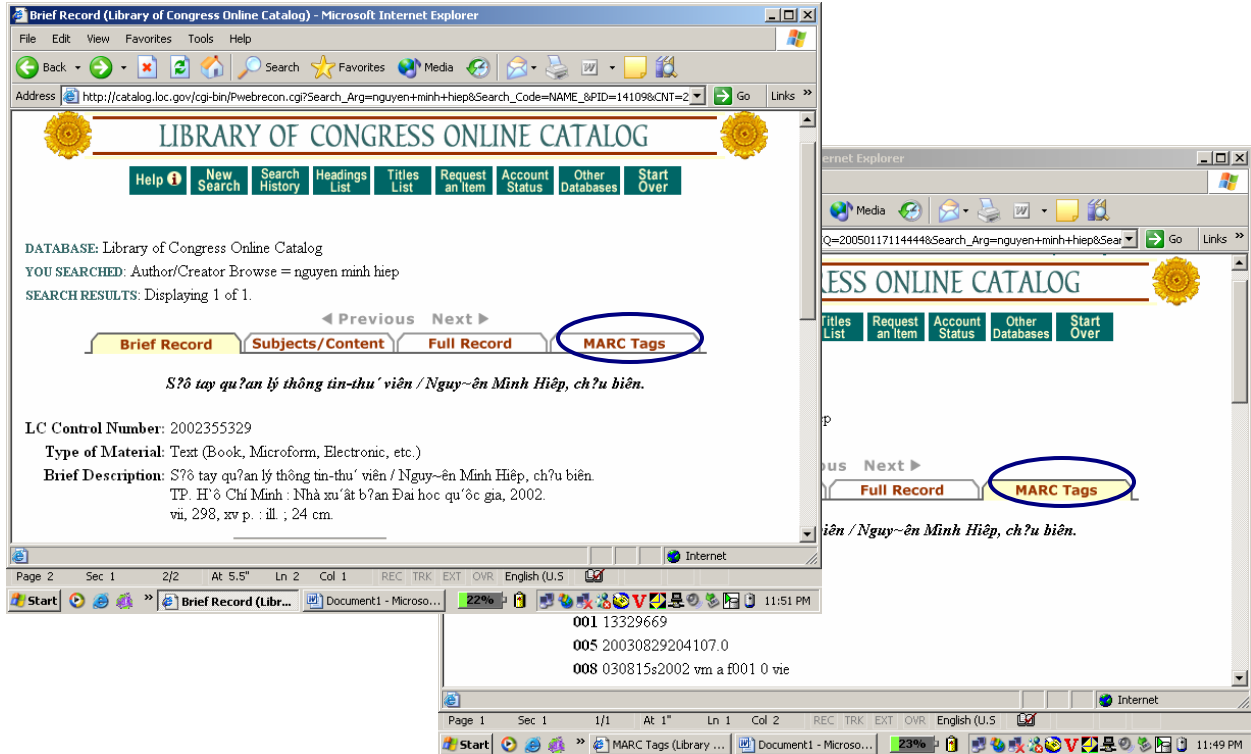
từ một đến hai tiếng đồng hồ để có thể biên mục được một biểu ghi MARC thực sự. Chúng ta có thể minh họa bằng một thống kê như sau: Vào năm 1997 Thư viện Quốc Hội Hoa Kỳ đã biên mục gần 300.000 biểu ghi MARC, tiêu tốn 25 triệu USD; OCLC có vào khoảng 34 triệu biểu ghi MARC phải cần đến 30.000 năm nhân công để thực hiện! (*“How to build a digital library”*/ Ian Witten. – New York: Morgan Kaufmann, 2003, trang 223). Cũng như phiếu mục lục, để trao đổi dữ liệu dạng thư tịch với nhau MARC phải được thể hiện các biểu ghi thật giống nhau từng “tag” một. Đây là cách trao đổi biểu ghi thư tịch trong thư viện truyền thông dạng liên biên.

### **Trao đổi siêu dữ liệu thư tịch (bibliographic metadata) dạng kỹ thuật số (digital)**

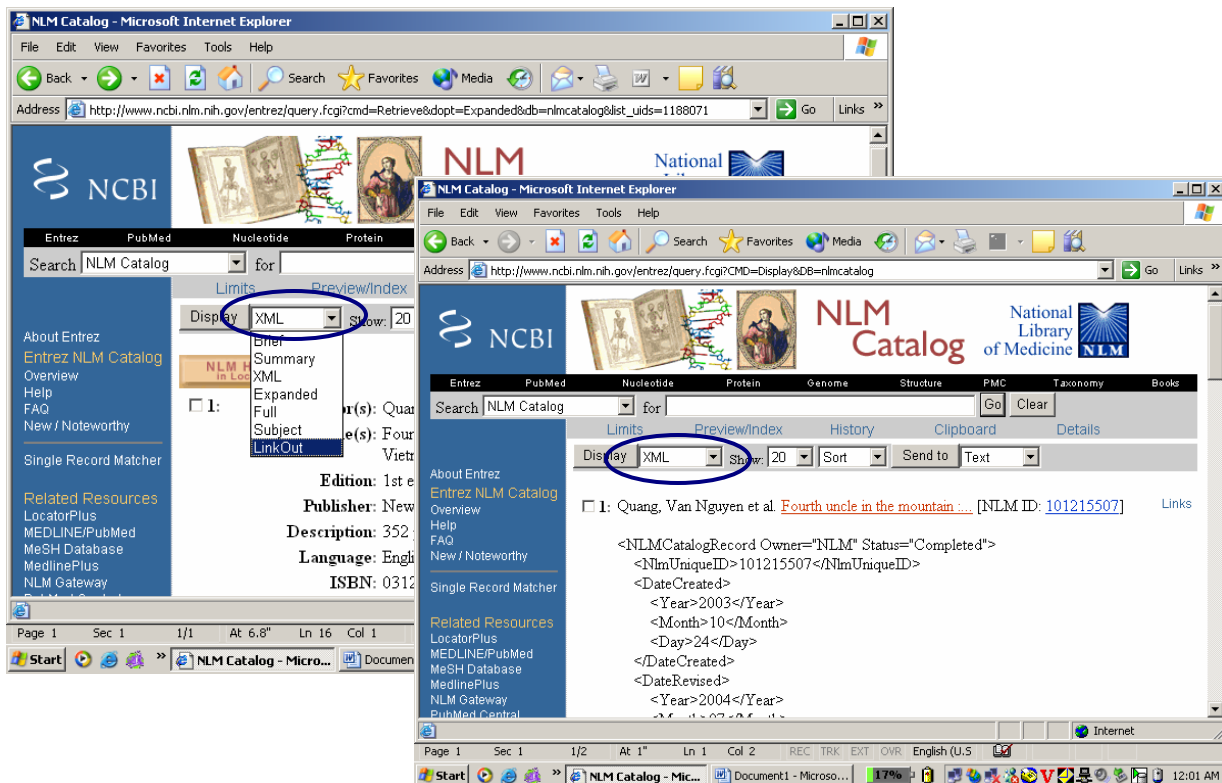
Kỹ thuật số ra đời cùng với công nghệ truyền thông và công nghệ đa phương tiện trên nền tảng web đã làm thay đổi hoàn toàn phương thức trao đổi và truyền dữ liệu. Công tác biên mục cũng thay đổi: dữ liệu thư tịch được đóng gói bằng ngôn ngữ XML. Phiếu mục lục và biểu ghi MARC được chuyển sang dạng metadata và trao đổi theo phương thức “gói thông tin”, phương thức này cho phép chúng ta không những chỉ trao đổi dạng thư tịch (chỉ lý lịch tài liệu) mà có thể trao đổi cả nội dung văn bản (toàn văn), thậm chí có thể trao đổi mọi hình thức đa phương tiện (âm thanh, hình ảnh, phim, vv...).

Chúng ta có thể hình dung hai phương thức trao đổi như sau: Trao đổi phiếu mục lục và biểu ghi MARC là trao đổi theo từng mặt hàng (thư viện truyền thông); còn trao đổi theo gói thông tin là trao đổi theo container, đơn vị trao đổi là container chứ không phải mặt hàng (thư viện điện tử - thư viện số). Cả hai đều đi đến kết quả chung là hàng hóa đã được trao đổi. Chính vì thế khi có cuộc bùng nổ thông tin điện tử và thư viện số ra đời (1994), người ta đặt vấn đề metadata (trình bày lý lịch của tài liệu như phiếu mục lục và biểu ghi MARC dưới dạng web) có cần thiết phải mô tả quá chi tiết và công phu như MARC không? Năm 1995, chuẩn Dublin Core ra đời đáp ứng yêu cầu dễ dàng tổ chức thông tin (biên mục và chỉ mục) để trao đổi theo phương thức mới: chỉ có 15 thành phần nòng cốt so với 800 trường mô tả của MARC. Dublin Core ra đời từ thành phố Dublin, bang Ohio, Hoa Kỳ, đã nhanh chóng phổ biến khắp thế giới. Năm 2001 được tổ chức Tiêu chuẩn Quốc gia Hoa kỳ - ANSI công nhận và phê chuẩn. Đến nay đã có 12 lần hội thảo để phát triển, lần họp mới nhất là vào ngày 11/12/2004 tại Thượng Hải, Trung Quốc.

Biểu ghi thư tịch MARC bao gồm quá nhiều chi tiết phức tạp nhưng chỉ trao đổi biểu ghi dạng lý lịch tài liệu (bibliography – thư tịch). Muốn trao đổi toàn văn và đa phương tiện thì biểu ghi thư tịch MARC phải chuyển sang dạng siêu dữ liệu thư tịch MARC với ngôn ngữ đóng gói XML, trở thành MARC-XML. Một nghịch lý là nếu trao đổi theo dạng XML này thì dữ liệu không đòi hỏi phải chi tiết đến độ không cần thiết như những biểu ghi MARC chuyển sang MARC-XML. Đó là lý do trên thế giới có nhiều thư viện chỉ dùng hình thức XML để tổ chức dữ liệu mà không quan tâm đến MARC, chẳng hạn như hệ thống thư viện lớn thứ hai nước Mỹ, Thư viện Quốc gia Y học Hoa Kỳ sử dụng phương thức này. Hình 1 và Hình 2 minh họa những biểu ghi thư tịch của Thư viện Quốc hội Hoa Kỳ sử dụng MARC và của Thư viện Quốc gia Y học Hoa Kỳ không sử dụng MARC.



Hình 1: Biểu ghi của Thư viện Quốc hội Hoa Kỳ sử dụng MARC



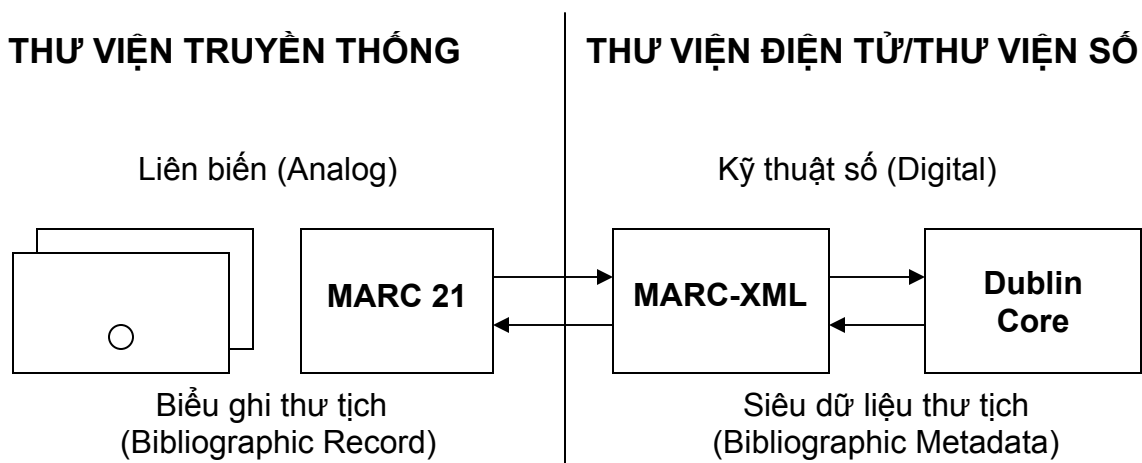
Hình 2: Biểu ghi của Thư viện Quốc gia Y học Hoa Kỳ không sử dụng MARC

## Tồn tại hai chuẩn biên mục

Cuối thập niên 1960, điểm mốc đầu tiên của sự phát triển ngành thông tin - thư viện – bắt đầu tin học hóa, chuẩn MARC ra đời và từ đó ngày càng phản ánh đầy đủ từng chi tiết của AACR2, đáp ứng yêu cầu của ngành thư viện truyền thống. Thư viện số bắt đầu xuất hiện vào năm 1994, đánh dấu điểm mốc thứ hai của sự phát triển đó. Một năm sau chuẩn Dublin Core ra đời đáp ứng yêu cầu biên mục tài liệu điện tử trong giai đoạn việc ứng dụng công nghệ thông tin trong ngành thông tin - thư viện đạt đến đỉnh cao để hình thành thư viện điện tử, thư viện số.

Chuẩn MARC được phát triển công phu, kiểm soát chặt chẽ, chi ly và bao hàm đến độ khá phức tạp, được tạo nên bởi những nhà biên mục học chuyên nghiệp chủ yếu để sử dụng trong thư viện truyền thống. Chuẩn Dublin Core chủ trương đơn giản hóa để có thể áp dụng rộng rãi cho tài liệu thư viện số đối với những người không cần được huấn luyện biên mục thư viện. Hai chuẩn này không những chú ý đến giá trị đặc thù của mình mà còn lưu tâm đến những triết lý căn bản đối nghịch nhau một cách tuyệt đối. MARC mang tính chất truyền thống, trong khi Dublin Core mang tính hiện đại phù hợp với việc sử dụng công nghệ mới. Do đó ngày càng được sử dụng rộng rãi. Cho đến khi Dublin Core được chính thức công nhận là một Chuẩn biên mục vào năm 2001, ngành biên mục tồn tại hai tiêu chuẩn biên mục: MARC 21 và Dublin Core.

Từ đó sự đối nghịch không những không còn tồn tại mà ngày càng nhiều chuyên gia tìm những giải pháp để chuyển đổi nhau: MARC sang Dublin Core và Dublin Core sang MARC, chẳng hạn như tại website của Văn phòng Chuẩn MARC và phát triển mạng lưới của Thư viện Quốc hội Hoa Kỳ có hướng dẫn các phương thức chuyển đổi. Thậm chí còn có nhiều phần mềm chuyển đổi được phổ biến rộng rãi, chẳng hạn như những phần mềm MarcEdit của Đại học Bang Oregon, Hoa Kỳ mà ta có thể download để dùng miễn phí tại website: <http://oregonstate.edu/~reaset/marcedit/html/downloads.html>.



Hình 3: Sự thay đổi biên mục và phương thức chuyển đổi

## Việc chuyển đổi

- **Dublin Core sang MARC:** Việc chuyển đổi 15 thành phần của Bộ thành phần Dublin Core sang Những Thành phần dữ liệu thư tịch của MARC khiến nhiều trường của MARC 21 được ghép chung vào một thành phần Dublin Core riêng lẻ. Điều này không thành vấn đề đối Dublin Core. Bởi vì cách cấu tạo của Dublin Core khiến mỗi thành phần mang tính cốt lõi (core) nên có thể chứa nhiều giá trị. Chẳng hạn như thành phần *Chủ đề (Subject)* có thể chứa tiêu đề đề mục, từ khóa, và số phân loại. Ví dụ:

```
<meta name="DC. Subject"
  content="Thư viện học">
```

```
< meta name="DC. Subject"
  scheme="DDC"
  content="020">
```

Trong việc chuyển đổi sang MARC, chẳng hạn như thành phần *Chủ đề (Subject)* được chuyển đổi như sau:

- Scheme=LCSH: 650 #0\$a (Subject added entry--Topical term)
  - Scheme=MeSH: 650 #2\$a (Subject added entry--Topical term)
  - Scheme=LCC: 050 ##\$a (Library of Congress Call Number/Classification number)
  - Scheme=DDC: 082 ##\$a (Dewey Decimal Call Number/Classification number)
  - Scheme=UDC: 080 ##\$a (Universal Decimal Classification Number)
  - Scheme=(other): 650 #7\$a with \$2=code from MARC Code List for Relators, Sources, Description Conventions
- **MARC sang Dublin Core:** Vì dữ liệu của MARC phong phú hơn Dublin Core nhiều nên trong việc chuyển đổi MARC sang Dublin Core, một trường riêng lẻ của MARC được ghép với một thành phần của Dublin Core, tuy nhiên thành phần Dublin Core có thể được lặp lại. Ví dụ đối với những trường liên quan đến chủ đề.

Thành phần DC	Từ hạn định DC	Trường MARC
Chủ đề (Subject)	LCSH	600, 610, 611, 630, 650
Chủ đề (Subject)	MeSH	600, 610, 611, 630, 650
Chủ đề (Subject)	LCC	050
Chủ đề (Subject)	DDC	082
Chủ đề (Subject)	UDC	080

- **Lưu ý:**
  - Việc chuyển đổi MARC sang Dublin Core và Dublin Core sang MARC được thực hiện bằng những phần mềm chuyển đổi.
  - Trong ngữ cảnh thông tin điện tử, các siêu dữ liệu thư tịch chuyển đổi đều có giá trị như nhau.

- Một phần mềm quản lý thư viện trong tương lai sẽ xuất ra cả hai dạng thức MARC và Dublin Core. Việc nhập vào MARC hay Dublin Core không gây trở ngại gì cho kết xuất đó; cho dù việc nhập vào là do bản thân thư viện hay do trao đổi dữ liệu với các thư viện khác.
- Vấn đề là chúng ta chọn MARC hay Dublin Core để biên mục. Cái nào tiện lợi cho chúng ta hơn?

## Kết luận

Lịch sử biên mục chính là lịch sử phát triển thư viện. Chúng ta có thể chia lịch sử biên mục theo hai điểm mốc phát triển: (1) Tin học hóa – cuối thập niên 1960 và (2) Thư viện số – 1994.

Hình 3 minh họa một bức tranh tổng thể của lịch sử biên mục. Ranh giới giữa hai điểm mốc chính là sự phân chia rành mạch giữa phương thức cũ và công nghệ mới của Thư viện truyền thống và Thư viện điện tử / Thư viện số; hay nói một cách khác là sự phân chia giữa giá trị cũ và giá trị mới.

Đôi khi giá trị cũ mang một ý nghĩa truyền thống mà ta luôn muốn lưu giữ, nhưng chính giá trị mới mới quyết định sự phát triển. Giá trị mới gắn liền với công nghệ mới là cứu cánh của chúng ta. Đứng nhằm lẫn giữa **GIÁ TRỊ MỚI** với **PHÁT HIỆN MỚI** trên nền **GIÁ TRỊ CŨ**.

MARC là một giá trị cũ được làm mới để thích ứng với công nghệ mới, trong khi Dublin Core là một giá trị mới. Hiện nay vì nhiều lý do khách quan và chủ quan, tồn tại như hai chuẩn biên mục và cả hai đều có thể chuyển đổi cho nhau. Vấn đề chọn chuẩn nào để sử dụng không nằm trong tiêu chí giá trị của mỗi chuẩn, mà ở chỗ dùng chuẩn nào thuận tiện cho điều kiện của chúng ta.

Chúng ta có thể hình dung MARC như một ngôi nhà to lớn đầy đủ tiện nghi được thiết kế để đáp ứng yêu cầu của thế kỷ trước, đến nay cần sử dụng một số thiết bị hiện đại thì phải sửa chữa cho phù hợp. Dublin Core như một ngôi nhà nhỏ, gọn được thiết kế để ứng dụng những thiết bị hiện đại. Chúng ta chưa có ngôi nhà nào hết, chúng ta chọn nhà nào?

## TÀI LIỆU THAM KHẢO

1. *Dublin Core/MARC/GILS Crosswalk* / Library of Congress. Network Development and MARC Standards Office.
2. LESK, Michael. *Practical Digital Libraries: Books, Bytes, and Bucks*. – San Francisco: Morgan Kaufmann, 2000.
3. *MARC to Dublin Core Crosswalk* / Library of Congress. Network Development and MARC Standards Office.
4. NGUYỄN MINH HIỆP. *Thư viện số : Bài giảng Powerpoint*. – 2004.
5. WITTEN, Ian H. và BAIBRIDGE, David. *How to Build a Digital Library*. – New York : Morgan Kaufmann, 2003.
6. <http://www.glib.hcmuns.edu.vn/greenstone/DLib.htm>