

# XỬ LÝ THÔNG TIN TRONG MÔI TRƯỜNG SỐ TẠI TRUNG TÂM THÔNG TIN – THƯ VIỆN, ĐẠI HỌC QUỐC GIA HÀ NỘI

*ThS. Hoàng Yến\**

*ThS. Nguyễn Bích Hạnh\*\**

**Tóm tắt:** Công tác xử lý thông tin trong kỷ nguyên số cần thiết phải thay đổi để có thể đối diện với những dạng thức số mới và những thách thức mới. Xây dựng một hệ thống thông tin thư mục toàn cầu, thiết lập các dữ liệu liên kết, các siêu dữ liệu, tạo nên những bộ sưu tập số có giá trị. Nhờ đó, các nguồn thông tin trở nên hấp dẫn, sinh động, có sự tương tác đa chiều và lôi cuốn người dùng tin hơn.

**Từ khóa:** Xử lý thông tin; Cơ sở dữ liệu; Thư viện số

## 1. Đặt vấn đề

Trong thời đại bùng nổ thông tin, số lượng các nguồn tin gia tăng theo cấp số nhân. Điều này gây rất nhiều khó khăn cho người dùng tin khi họ cần tìm kiếm, khai thác thông tin. Bên cạnh đó, cách thức người dùng tin sử dụng thư viện và khai thác các nguồn tài nguyên thông tin của thư viện cũng đã thay đổi và theo nhiều cách khác nhau: trực tiếp đến thư viện, qua OPAC, qua website, blog, qua tin nhắn SMS, chat trực tuyến ... Nhu cầu tin của họ cũng phát triển ngày càng phong phú và đa dạng, không chỉ là các nguồn thông tin truyền thống như sách, báo in mà còn là sách điện tử, các cơ sở dữ liệu điện tử, file âm thanh, hình ảnh, không gian 3 chiều, các bộ sưu tập số ... và họ còn mong muốn tìm thông tin với thời gian ngắn nhất mà hiệu quả nhất. Điều này đòi hỏi và thôi thúc các cán bộ xử lý thông tin phải tìm ra phương thức phù hợp để xử lý, bao gói thông tin giúp cho người dùng tin tiếp cận được với các nguồn tin với phương châm “bất cứ cái gì, bất cứ khi nào và bất cứ nơi đâu”, đáp ứng nhu cầu tin của họ nhằm giành lại “thị phần” đang bị cạnh tranh bởi Google Scholar, Google Books, Google Earth, Wikipedia ... trong thời đại số hiện nay.

## 2. Xử lý thông tin trong môi trường số

Để giải được bài toán này thì vấn đề cần phải thay đổi trong tư duy và cách làm trong việc xử lý thông tin là đặc biệt quan trọng; thay đổi để có thể đối diện với những dạng thức số mới và những thách thức mới trong xử lý thông tin trong thế giới số. Điều đó phản ánh một sự chuyển hướng nhằm xây dựng một hệ thống thông tin thư mục toàn cầu, thiết lập các dữ liệu liên kết, các siêu dữ liệu để tạo nên những bộ sưu tập số có giá trị cao. Thay đổi công tác xử lý thông tin trong kỷ nguyên số cần dựa trên những nguyên tắc định hướng:

---

\* Phòng Phân loại - Biên mục, Trung tâm Thông tin – Thư viện, Đại học Quốc gia Hà Nội

\*\* Phòng Phân loại - Biên mục, Trung tâm Thông tin – Thư viện, Đại học Quốc gia Hà Nội

*Đó là sự thuận tiện của người dùng tin.* Khi đưa ra quyết định mô tả hay tạo lập các điểm truy cập cho các nguồn tin, lựa chọn các hình thức có kiểm soát cho việc truy cập cần được thực hiện với ý thức hướng tới người dùng tin.

*Tính đại chúng.* Từ vựng sử dụng trong mô tả và truy cập cho các nguồn tin cần phù hợp với đại chúng.

*Tính nhất quán và chuẩn hóa.* Khi mô tả, xây dựng các điểm truy cập cho các nguồn tin cần được chuẩn hóa tối đa, đảm bảo sự thống nhất và nhất quán, từ đó giúp việc chia sẻ nguồn tin được dễ dàng và hiệu quả.

Để hiện thực hóa các vấn đề trên thì việc áp dụng các công cụ chuẩn – các quy tắc trong công tác xử lý thông tin số đặc biệt quan trọng. Công tác xử lý thông tin được thực hiện chính xác theo các chuẩn không chỉ giúp cho hoạt động kiểm soát thư mục và tiếp cận các nguồn tin ở thư viện nhanh chóng, hiệu quả mà còn giúp cho các thư viện trong và ngoài nước có thể trao đổi dữ liệu một cách dễ dàng. Các chuẩn phổ biến được áp dụng trong công tác xử lý thông tin hiện nay ở Việt Nam cũng như trên thế giới là:

- Chuẩn cấu trúc dữ liệu/trao đổi dữ liệu: Khổ mẫu thư mục (MARC21 Bibliographic), Khổ mẫu siêu dữ liệu Dublin Core, Tiêu chuẩn mô tả đối tượng siêu dữ liệu (MODS), Tiêu chuẩn truyền và mã hóa siêu dữ liệu (METS)

- Chuẩn nội dung dữ liệu: Quy tắc biên mục Anh – Mỹ (AACR2), Mô tả và truy cập tài nguyên (RDA)

- Chuẩn mô hình thư mục: FRBR - Yêu cầu chức năng cho biểu ghi thư mục

- Chuẩn kiểm soát tên (tên cá nhân, tên tập thể, tên địa lý): Khổ mẫu kiểm soát nhất quán (MARC21 Authority)

- Chuẩn kiểm soát chủ đề: Khổ mẫu kiểm soát nhất quán, SKOS (từ vựng có cấu trúc được kiểm soát trong mạng ngữ nghĩa)

- Chuẩn đánh giá dữ liệu: các bộ tiêu đề chủ đề như: LCSH, MeSH, Sears List, LCNAF

Từng bước, từng bước một với sự nỗ lực các cán bộ xử lý thông tin của Trung tâm đã dần hiện thực hóa các công việc trên, thể hiện trong sự thay đổi xử lý thông tin. Công việc xử lý thông tin, bao gồm: Xử lý hình thức (là mô tả những yếu tố hình thức của tài liệu, bao gồm: nhan đề, tác giả, nơi xuất bản, năm xuất bản, nhà xuất bản, số trang, số tập ...) và xử lý chủ đề (là xử lý chi tiết những yếu tố về nội dung như: làm từ khoá, định chủ đề, tóm tắt, định ký hiệu phân loại, cập nhật cơ sở dữ liệu) trong môi trường số không chỉ dừng lại ở việc mô tả, tạo các điểm truy cập cho tài liệu mà nó còn phải chứa những chỉ dẫn, giải thích trong các điểm truy cập, mô tả cho tất cả tài nguyên và liên biến. Mang đến hiệu quả là sử dụng cho việc truy cập và mô tả tài nguyên thông tin trong thế giới số. Đó là những điểm truy cập được thiết lập dựa trên các tiêu đề chuẩn - một công cụ chuyên chở thông tin liên quan đến các hình thức của những tên (bao gồm tên người, tên tập thể, hội nghị, địa danh) và những tiêu đề chủ đề được dùng làm những điểm truy cập trong các biểu ghi thư mục. Các hình thức của những tên và những tiêu đề chủ đề này cùng tiêu phân mục được dùng như là những tham chiếu dẫn đến hình thức chuẩn và các mối liên quan giữa những hình thức này. Các tiêu đề chuẩn này giúp cho các cán bộ thư viện cung cấp cho người dùng tin sự truy cập đồng nhất đến các tài liệu

thư viện thông qua việc sử dụng mục lục thư viện (OPAC) cũng như cung cấp những thông tin rõ ràng về tác giả, nhan đề và tiêu đề chủ đề. Thiết lập các tham chiếu thích hợp trực tiếp chỉ dẫn người dùng tin đến những hình thức chuẩn của tên, nhan đề và tiêu đề chủ đề với 4 tham chiếu chính:

Tham chiếu *xem (See)* là tham chiếu trực tiếp chỉ dẫn người dùng tới tên bút danh, biệt hiệu của tác giả hay tới hình thức thích hợp của tên.

Tham chiếu *xem thêm (See also)* là liên kết các biểu ghi thư mục và cung cấp những dẫn từ một tiêu đề này đến tiêu đề khác để chỉ ra các mối quan hệ. Chúng cũng liên kết các nhan đề với các nhan đề mới hơn, đặc biệt là các tài liệu nhiều tập/bộ.

Tham chiếu *thông tin (information)* được các cán bộ sử dụng khi có thông tin quan trọng cần được thêm vào, yêu cầu giải thích nhiều hơn *xem/xem thêm*. Nó được các cán bộ thiết lập nhằm cung cấp những hướng dẫn cho người dùng tin đi từ một điểm truy cập tới một số các điểm truy cập khác.

Tham chiếu *ghi chú* là để giải thích cho những biệt danh, biệt hiệu, từ viết tắt/từ cấu tạo bằng những chữ đầu của một nhóm từ, đa tên.

Cuối cùng là các tham chiếu cho các tiêu đề chủ đề sử dụng: từ rộng hơn, từ hẹp hơn, từ có liên quan. Hệ thống tham chiếu này được thiết lập để giúp cho người dùng tin sử dụng mục lục chủ đề được hướng dẫn đến tiêu đề chủ đề đã được thiết lập, cho dù lúc đầu họ sử dụng bất cứ từ hay cụm từ nào để truy tìm tin; được hướng dẫn đến những chủ đề có liên quan đến chủ đề mà người dùng tin đang tìm; được hướng dẫn đến những chủ đề chi tiết hơn, cụ thể hơn và chính xác hơn. Toàn bộ hệ thống tham chiếu cho các tiêu đề giúp cho người dùng tin sử dụng mục lục có cơ hội khám phá được hết tất cả các *tiêu đề* đã được thiết lập cho những chủ đề có liên quan chặt chẽ với nhau và nhờ vậy giúp cho những người dùng tin nhất là những người làm công tác nghiên cứu về bất cứ lĩnh vực nào có thể truy cập được tất cả những tài liệu hữu ích, mà họ có thể sử dụng được trong các bộ sưu tập khổng lồ.

Với công tác xử lý chủ đề khi áp dụng Khung phân loại DDC trong môi trường số thì sự sắp đặt theo thứ tự và những chức năng tìm tin của khung phân loại DDC cũng đã thay đổi. Đó là nhu cầu về việc phân loại những bộ sưu tập lớn và ngày một phát triển của nguồn tin điện tử/số cho những người dùng tin cá nhân trong môi trường truy cập thông tin với không gian cá nhân. Như vậy, đòi hỏi khung phân loại DDC phải trở thành một hệ thống mềm dẻo hơn, có thể vận hành như một hệ quản trị cơ sở dữ liệu quan hệ đối tượng về các môn loại. Điều này có thể đạt được bằng cách áp dụng một cách hệ thống các kỹ thuật khai thác dữ liệu và phát triển tri thức vào sơ đồ phân loại với chức năng như một cơ sở dữ liệu của các môn loại, như một khung phân loại các bộ sưu tập của thư viện và như một công cụ tìm tin. Với một khung phân loại DDC mềm dẻo hơn, nó sẽ rất hữu ích cho giảng dạy, học tập, ghi nhớ tri thức và mối quan hệ giữa các lĩnh vực khác nhau về môn loại và các nguồn tin. Từ các chỉ số phân loại được tạo lập trong các cơ sở dữ liệu, các cán bộ xử lý chủ đề sẽ tra tìm các chỉ số phân loại, bao gói thông tin theo các chủ đề và các dạng tài liệu (sách tham khảo, sách tra cứu, luận án, luận văn, bài báo ...) để tạo ra cơ sở dữ liệu Hướng dẫn theo chủ đề. Các đường link liên kết điện

tử hữu ích được tạo lập giữa các cơ sở dữ liệu theo chủ đề. Bạn đọc chỉ cần bấm chuột vào các liên kết đó để có được các thông tin về tài liệu có trong thư viện hoặc ngoài thư viện, có thể liên kết tới các nguồn học liệu mở cũng như những cơ sở dữ liệu điện tử/số. Đồng thời những liên kết điện tử này cũng có thể được các cán bộ liên kết đến các Hướng dẫn của các thư viện khác với cùng chủ đề đang được đề cập hoặc các chủ đề liên quan với chủ đề chính.

Với việc xử lý nội dung tài liệu bằng các tiêu đề chủ đề, cơ sở dữ liệu đã cung cấp phương pháp truy cập nội dung tài liệu theo cách đa chiều, giúp người dùng tin thấy được nội dung đa dạng của tài liệu. Bởi hệ thống tham chiếu được thiết lập giúp cho người sử dụng mục lục chủ đề: được hướng dẫn đến tiêu đề chủ đề đã được thiết lập, dù cho họ khởi đầu việc truy tìm tin bằng bất cứ từ hay cụm từ nào; được hướng dẫn đến những chủ đề có liên quan đến chủ đề mà họ đang tìm; và được hướng dẫn đến những chủ đề chi tiết hơn, cụ thể hơn và chính xác hơn. Một lần nữa nguyên tắc “người dùng tin là trọng tâm, là tiêu điểm” của công tác xử lý thông tin vẫn tiếp tục được xem như là kim chỉ nam và được khẳng định một lần nữa. Tức là tiện lợi cho người dùng tin phải được xếp trước tiện ích của người xử lý thông tin. Như vậy, có rất nhiều thứ đã thay đổi trong các thư viện hơn 100 năm qua, thập kỷ gần đây đã chứng kiến sự chuyển dịch mạnh mẽ với quy mô lớn từ mục lục truyền thống (mục lục phiếu) sang mục lục truy cập trực tuyến, thư viện truyền thống sang thư viện số ... Tuy nhiên, một điều không hề thay đổi chính là sự tận tâm của các thư viện và đặc biệt là các cán bộ thư viện làm công việc xử lý thông tin để cung cấp những công cụ tìm kiếm thông tin cũng như xây dựng những cơ sở dữ liệu, những bộ sưu tập số ... cho tất cả người dùng tin một cách dễ dàng và thuận tiện nhất.

### **3. Kết quả đạt được**

*Các nguồn thông tin trở nên hấp dẫn, sinh động và lôi cuốn người dùng tin hơn:*  
Trong hệ thống mục lục công cộng trực tuyến OPAC sẽ thể hiện đầy đủ dữ liệu được tạo lập từ việc mô tả thông tin, lựa chọn và tạo lập các điểm truy cập, thiết lập các đường liên kết hữu ích ... Giúp cho người dùng dễ dàng tìm, định vị, thu nhận những thông tin họ cần. Ví dụ, một tài liệu được thể hiện ở các dạng khác nhau như bản in, bản số, dạng vi phiếu, băng âm thanh, video ... và có thể được liên kết với các cơ sở dữ liệu khác, được nhóm lại và thể hiện cùng một lúc, cho phép người dùng tin xem, nhận biết mối quan hệ giữa các dạng của tài liệu một cách dễ dàng và tiện ích hơn; từ đó đưa ra yêu cầu tin chính xác cũng như giúp cho người dùng tin yêu thích sử dụng loại hình tài liệu nào, để tiếp thu nội dung thông tin chứa đựng trong tài liệu đó thì có thể lựa chọn loại hình mình yêu thích hơn. Ví dụ: giữa tài liệu ở dạng in là một cuốn sách, ở dạng điện tử trực tuyến là bài giảng hay là một bản tài liệu số hóa thì khả năng thu nhận thông tin cũng như sự ưa thích và có sự lựa chọn nguồn tin cũng sẽ khác nhau.

*Các nguồn thông tin trở nên linh hoạt, không tĩnh, luôn có sự tương tác, phản hồi đa chiều:* thông qua các đường link liên kết được tạo lập trong và giữa các tài liệu điện tử/số, người dùng tin có thể tìm thấy thông tin liên quan tới tác giả hoặc cá nhân có

trách nhiệm liên quan tới nguồn tin để có thể trao đổi thông tin, chia sẻ thông tin. Với các đường link liên kết người dùng không chỉ tiếp cận được với nguồn tài liệu mình cần mà còn được chỉ dẫn tới các thông tin khác có liên quan tới chủ đề đó hoặc một tập hợp các tài liệu của chính tác giả đó, các tài liệu viết về tác phẩm và tác giả đó, các tài liệu của các tác giả khác có cùng chủ đề, các nguồn tham khảo hay các tác giả khác đã trích dẫn trong công trình; liên kết tới các nguồn tin khác ngoài nguồn tin có trong thư viện. Như vậy, mỗi thông tin trong tài liệu số không chỉ là một thông tin đơn lẻ mà nó là thông tin chứa đựng các liên kết, các siêu liên kết, mang tính chỉ dẫn, định hướng thông tin cho người sử dụng. Nó ẩn chứa nhiều thông tin tiềm năng khác và cũng là một trong những cơ sở giúp cho người dùng tin sáng tạo, tạo lập các thông tin mới có giá trị gia tăng. Giúp họ mất ít thời gian nhất nhưng lại tìm được nhiều thông tin nhất.

Ví dụ minh họa về đường liên kết giữa các tài liệu trong CSDL của Trung tâm: Từ công tìm kiếm tập trung, tài liệu được liên kết đến CSDL thư mục, đến Thư viện Quốc hội Mỹ và Thư viện số Trung tâm

### Cổng tìm kiếm tập trung (URL)

The screenshot shows the VNU LIC search portal. The search bar contains the text "Computer aided design, engineering, and manufacturing : systems techniques and applications. Vol. 5. The design of manufacturing systems". The search results list two items:

1. **Computer aided design, engineering, and manufacturing : systems techniques and applications. Vol. 5. The design of manufacturing systems**  
Boca Raton, FL : CRC Press, 2001-- (TS155.6 .C6455 2001)  
Kiểm tra tình trạng sẵn có của tài liệu  
Đặt mượn | Vị trí tài liệu | Chi tiết | Liên kết mở rộng
2. **Computer integrated design systems in concurrent engineering**  
Jeppsson, Peter; Luleå tekniska universitet, Institutionen för teknikvetenskap och matematik, Produkt- och produktionsutveckling

### Liên kết đến CSDL Thư mục của Thư viện

The screenshot shows the catalog page for the book. The title is "Computer aided design, engineering, and manufacturing : systems techniques and applications. editor, Cornelius Leondes." by Leondes, Cornelius T. The page includes a search bar, a "Toàn bộ biểu ghi" (Full Record) section, and a "Liên kết tài liệu số" (Digital Item View Page) section.

**Toàn bộ biểu ghi**  
Xem | MARC  
Tổng số yêu cầu 0 Số yêu cầu không hài lòng 0  
Bạn đã tìm kiếm CSDL VNU HN - Mọi nơi: "Computer aided design, engineering, and manufacturing : systems techniques and applications. editor, Cornelius Leondes."  
**Nhan đề**  
Computer aided design, engineering, and manufacturing : systems techniques and applications. editor, Cornelius Leondes.  
**Cá nhân liên quan**  
Leondes, Cornelius T.  
**Xuất bản**  
Boca Raton, FL : CRC Press, 2001-  
**Số trang**  
7 v. :  
**Chủ đề/Từ khóa**  
CAD/CAM systems.  
**Chủ đề/Từ khóa**  
Hệ thống CAD/CAM  
**Liên kết tài liệu số**  
<http://www.loc.gov/catdir/enhancements/fy0646/00046856-d.html>  
**Liên kết tài liệu số**  
<http://db.lic.vnu.edu.vn:2129/iii/cpro/DigitalItemViewPage.external?lang=vie&sp=1064669&sp=T&sp=Pal%2CRu100001%40%2CQtns005158>

The image shows two side-by-side screenshots of library catalogs. The left screenshot is from the Library of Congress Online Catalog, displaying a search result for a book titled "Computer aided design, engineering, and manufacturing : systems techniques and applications". The right screenshot is from the Vietnam National Library and Information Center Digital Library, showing a search result for the same book. A blue arrow points to the search bar in the right screenshot.

#### 4. Kết luận

Sự phát triển của khoa học và công nghệ, đặc biệt là công nghệ thông tin đã tác động sâu sắc, làm biến đổi về chất trong hoạt động của các cơ quan thông tin thư viện. Thư viện truyền thống đã và đang chuyển sang mô hình thư viện số; để xây dựng thành công thư viện số tại Trung tâm, các cán bộ thư viện nói chung, cán bộ xử lý thông tin nói riêng đang không ngừng nỗ lực nghiên cứu, áp dụng các chuẩn trong xử lý thông tin, xây dựng các bộ sưu tập số, các cơ sở dữ liệu có giá trị thông tin cao đáp ứng tối đa nhu cầu tin cho người dùng. Nâng cao chất lượng xử lý tài liệu là bước đi đúng hướng của các cơ quan thông tin thư viện trong quá trình xây dựng nguồn tài liệu số hiện nay. Đồng thời cũng cần có sự định hướng, quan tâm chỉ đạo từ các cấp lãnh đạo nhằm thống nhất nhận thức cũng như hành động hướng tới xây dựng thư viện số thành công. Qua đó góp phần nâng cao chất lượng giáo dục, đào tạo, nghiên cứu khoa học của Đại học Quốc gia Hà Nội.

#### TÀI LIỆU THAM KHẢO

- 1.The Application of Semantics Web in Digital Library Knowledge Management / Du Liangxian,Qi Junxia,Guo Pengfei. Physics Procedia 24 (2012) 2180 – 2186
- 2.Cataloging and organizing digital resources: a how-to-do-it manual for librarians / Anne M. Mitchell and Brian E. Surratt. 2005. 139 p.
3. Changing roles of the librarians in the virtual/digital era / P. Sarasvathy, Nambratha G.R and D. Giddaiah. SRELS Journal of Information Management Vol. 49, No. 5, October 2012, Paper AX. p495-500.

4. Digital Library Standards: Digital Collections & Services: Access to print, pictorial and audio-visual collections and other digital services

**<https://www.loc.gov/library/digitalstandards.html>**

5. "The future of open access and library publishing", *New Library World*, Vol. 115 Iss 5/6 pp. 225 – 236. Permanent link to this document: <http://dx.doi.org/10.1108/NLW-05-2014-0049>

6. Knowledge and skills for the digital era academic library / J. Raju. *The Journal of Academic Librarianship* 40 (2014) 163–170

7. Using METS and MODS to Create an XML Standards-based Digital Library Application / Morgan Cundiff, Network Development and MARC Standards Office, Library of Congress. <https://www.loc.gov/standards/mods/presentations/mets-mods-morgan-ala07/>

8. Website Trung tâm Thông tin Thư viện, Đại học Quốc gia Hà Nội: <http://lic.vnu.edu.vn/>