

NGHIÊN CỨU THỰC TRẠNG CÔNG TÁC ĐỊNH TỪ KHÓA TẠI THƯ VIỆN TRƯỜNG ĐẠI HỌC HOA LƯ

Nguyễn Thị Thảo

Trường Đại học Hoa Lư

Tóm tắt: Bài báo trình bày quy trình định từ khóa, trên cơ sở đó đánh giá, phân tích được thực trạng quy trình của việc mô tả nội dung tài liệu bằng từ khóa, chất lượng mô tả nội dung tài liệu bằng từ khóa tại Thư viện Trường Đại học Hoa Lư.

Từ khóa: Định từ khóa; thư viện; xử lý nội dung tài liệu; thực trạng; Đại học Hoa Lư.

CURRENT STATUS OF KEYWORD IDENTIFICATION AT HOA LU UNIVERSITY LIBRARY

Abstract: This article presents a keyword identification process and, based on this process, evaluates and analyzes the current status of keyword-based content description of documents, the quality of keyword-based content description of documents at the Hoa Lu University Library.

Keywords: Keyword identification; library; document content processing; current status; Hoa Lu University.

ĐẶT VẤN ĐỀ

Trường Đại học Hoa Lư được thành lập theo quyết định số 407/QĐ-TTg ngày 09 tháng 4 năm 2007 của Thủ tướng Chính phủ trên cơ sở nâng cấp trường Cao đẳng Sư phạm Ninh Bình. Trong gần 20 năm qua, Trường Đại học Hoa Lư đã hoàn thành tốt nhiệm vụ chính trị của mình, đào tạo hàng nghìn học viên đạt trình độ cử nhân sư phạm, cử nhân kinh tế,... và trình độ cao đẳng sư phạm. Trường đã góp phần đào tạo, bồi dưỡng đội ngũ cán bộ và người lao động chất lượng cao cho tỉnh Ninh Bình và đất nước.

Thư viện trường thuộc Trung tâm Thư viện-Thiết bị. Với sự quan tâm, chỉ đạo sát sao của lãnh đạo trường, đến nay Thư viện đã có một sở hạ tầng khang trang, rộng rãi, cơ sở vật chất được cải thiện. Với số lượng người dùng tin ngày càng tăng, chất lượng thông tin ngày càng cao, đòi hỏi Thư viện phải có một chiến lược phát triển hợp

lý, nhằm bảo đảm nguồn thông tin chính xác, kịp thời. Chính vì vậy, công tác mô tả nội dung tài liệu nói chung tại Thư viện và công tác định từ khóa nói riêng phải được đặt lên hàng đầu.

Nghiên cứu công tác định từ khóa tại Thư viện được thực hiện nhằm đánh giá, phân tích thực trạng quy trình của việc mô tả nội dung tài liệu bằng từ khóa, chất lượng mô tả nội dung tài liệu bằng từ khóa tại Thư viện, góp phần nâng cao chất lượng hoạt động của Thư viện phục vụ cho công tác đào tạo và nghiên cứu khoa học của Trường Đại học Hoa Lư.

1. CƠ SỞ LÝ LUẬN VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Định từ khóa là một trong các khâu xử lý nghiệp vụ về nội dung tài liệu đang được áp dụng rộng rãi ở các thư viện Việt Nam. Để giúp cho người đọc và người dùng tin có

thể tra cứu được tài liệu một cách dễ dàng, thuận lợi các thư viện trên thế giới nói chung và các thư viện ở Việt Nam nói riêng đã luôn dành sự quan tâm đặc biệt cho khâu xử lý nội dung tài liệu này.

Từ khóa là một từ hoặc một cụm từ được rút ra từ tên hoặc chính văn của tài liệu nhằm biểu đạt nội dung chính yếu của tài liệu. Tài liệu có thể có một hoặc một tập hợp từ khóa. Đó chính là những dấu hiệu để tạo nên các biểu thức tìm tài liệu trong cơ sở dữ liệu.

Điều quan trọng nhất của quá trình định từ khóa là phải xác định được những khái niệm cơ bản được phản ánh nội dung của tài liệu và thể hiện chúng bằng các từ khóa đã được kiểm soát [3].

Mục đích của công tác định từ khóa là thiết lập điểm truy nhập nội dung tài liệu bằng từ ngữ. Từ điểm truy nhập này, người dùng tin có thể tiếp cận và khai thác tài liệu có nội dung thể hiện bằng tập hợp các từ ngữ phù hợp với yêu cầu của họ.

Trong nghiệp vụ xử lý nội dung tài liệu của các cơ quan TT-TV có hai phương pháp định từ khóa là: Phương pháp định từ khóa tự do và Phương pháp định từ khóa có kiểm soát.

Tại Thư viện Trường Đại học Hoa Lư (TV ĐHHL) phương pháp định từ khóa được cán bộ thực hiện là phương pháp định từ khóa tự do (lựa chọn từ khóa theo quan điểm của người định từ khóa khi mô tả các khái niệm mà tài liệu đề cập đến). Đây là quá trình xử lý tài liệu mang tính chủ quan và sáng tạo, phụ thuộc vào kỹ năng và trình độ của người định từ khóa [3].

Để nghiên cứu công tác định từ khóa của TV ĐHHL, tác giả áp dụng phương pháp nghiên cứu định tính, tổng quan, thu thập dữ liệu. Đồng thời, áp dụng kết hợp phương pháp phỏng vấn trực tiếp, rút mẫu nghiên cứu, so sánh và hỏi ý kiến của chuyên gia. Với mẫu nghiên cứu được xác định theo phương pháp chọn mẫu ngẫu nhiên. Tác giả đã rút ngẫu nhiên 30 biểu ghi thuộc nhiều

lĩnh vực trong cơ sở dữ liệu sách và cơ sở dữ liệu tài liệu nội sinh của thư viện. Trên cơ sở dữ liệu được rút ra, tiến hành phân tích đánh giá ở hai khía cạnh: thực trạng quy trình của việc mô tả nội dung tài liệu bằng từ khóa và chất lượng mô tả nội dung tài liệu bằng từ khóa.

2. QUY TRÌNH ĐỊNH TỪ KHÓA

“Định từ khóa tài liệu là quá trình xử lý tài liệu, sử dụng các từ khóa, từ chuẩn để mô tả nội dung, đối tượng, phương diện nghiên cứu cũng như hình thức của tài liệu” [3]. Định từ khóa tài liệu có liên quan chặt chẽ đến nhiều khâu công tác khác nhau trong các cơ quan TT-TV như: xây dựng ngôn ngữ tìm tin theo từ khóa, tổ chức các phương tiện tra cứu, xây dựng các điểm tiếp cận để tìm kiếm trong các cơ sở dữ liệu và biên soạn các bản thư mục,... Do đó, đòi hỏi các thư viện và cơ quan thông tin phải tiến hành việc định từ khóa cho tài liệu. Hiện tại, TV ĐHHL có tổng số bốn cán bộ thư viện. Trong đó, công tác xử lý tài liệu nói chung và công tác định từ khóa nói riêng của thư viện được giao cho một cán bộ xử lý, không có công tác kiểm duyệt sau khi xử lý. Do đó, công tác định từ khóa tài liệu của thư viện hoàn toàn mang tính chủ quan và sáng tạo của cán bộ làm công tác này. Tuy nhiên, công tác định từ khóa tài liệu vẫn phải tuân theo một quy trình nhất định và tác giả đã tiến hành nghiên cứu thực tế việc áp dụng quy trình này tại TV ĐHHL, bao gồm các công đoạn:

- Phân tích nội dung tài liệu.
- Diễn đạt từ khóa.
- Hoàn chỉnh quá trình định từ khóa.

Để nắm bắt được cách thức thực hiện các công đoạn đó, trên cơ sở 30 biểu ghi được rút ngẫu nhiên, tác giả tiến hành phân tích quy trình định từ khóa của cán bộ thư viện.

Danh sách các tài liệu được dùng để khảo sát chất lượng định từ khóa được thể hiện ở Bảng 1.

Bảng 1. Danh sách biểu ghi được rút ngẫu nhiên trong các CSDL của TV ĐHHL

STT	Biểu ghi	Tên tài liệu	Từ khóa
1	121	Truyện Kiều	Nguyễn Du, văn học cận đại, truyện thơ
2	122	Những danh tướng của Trung Hoa thời xưa	Nhân vật lịch sử, phong kiến, Trung Hoa, danh tướng, quân sự
3	161	Sự sống ngoài trái đất	Hành tinh, sự sống, sách thường thức, thiên văn học, vũ trụ
4	166	Minh Tâm bảo giám	Nho giáo, đạo đức, Trung Quốc
5	351	Nhật Kí trong tù	Thơ, Hồ Chí Minh, Việt Nam, văn học hiện đại
6	368	Đôxtôpexki cuộc đời và sự nghiệp	Văn học, cuộc đời, sự nghiệp, Nga
7	384	Nghiệp vụ buôn bán quốc tế	Kinh tế, thương mại, hợp đồng buôn bán, quốc tế
8	401	Cẩm nang tuổi mới lớn	Sinh lí, tuổi dậy thì, tâm lí, khoa học thường thức
9	430	Ngụ ngôn La Phông-ten	Truyện ngụ ngôn, La Phông-ten, văn học hiện đại, Pháp
10	452	Tiếng Việt - giáo trình đào tạo giáo viên mầm non	Giáo trình, Tiếng Việt, mầm non
11	474	Đại số đại cương	Đại số, bài tập, giáo trình
12	1001	Kĩ thuật điện tử	Kỹ thuật điện tử, điện tử, giáo trình
13	1017	Bài tập cơ học lý thuyết	Cơ học lý thuyết, bài tập, giáo trình, tĩnh học, động học
14	1030	Văn hóa các dân tộc thiểu số ở Việt Nam	Văn hóa các dân tộc, bản sắc văn hóa, văn hóa Việt Nam
15	5762	Hướng dẫn ôn tập môn Lịch sử Đảng Cộng sản Việt Nam	Đảng Cộng sản Việt Nam, lịch sử

STT	Biểu ghi	Tên tài liệu	Từ khóa
16	6700	Hàn Mặc Tử tác phẩm chọn lọc	Phê bình văn học, văn học hiện đại, Việt Nam, kịch thơ, thơ
17	6722	Con đường tơ lụa: quá khứ và tương lai	Quan hệ quốc tế, giao thông, lịch sử, văn hóa, thương mại, thế giới, Trung Quốc
18	6755	Sư phạm học tiểu học	Giáo dục, tiểu học, giáo trình
19	7711	New English file - Elementary student's book and workbook	Tiếng Anh, giáo trình
20	7695	Photoshop thật là đơn giản	Phần mềm Photoshop, đồ họa, ảnh kỹ thuật số
21	472	Bài tập giải tích toán học	Toán giải tích, bài tập, giáo trình
22	542	Giáo trình phần mềm kế toán máy	Kế toán, máy vi tính, ứng dụng, phần mềm, SAS INOVA
23	599	Toán học trong hoạt động thư viện - thông tin	Lí thuyết tập hợp, thư viện, toán logic, toán thống kê, toán ứng dụng, tìm tin, giáo trình
24	966	Phương pháp toán trong địa lí	Toán học, giáo trình, địa lí
25	7464	Giáo trình mô hình toán kinh tế	Mô hình, toán kinh tế, kinh tế, giáo trình
26	7927	Hệ thống câu hỏi và bài tập môn học kế toán công ty	Kế toán, công ty, câu hỏi, bài tập, kế toán doanh nghiệp, (giáo trình)
27	7965	Luật kế toán và kiểm toán	Luật kế toán, luật kiểm toán, văn bản pháp luật, pháp luật
28	8362	Lý luận phê bình văn học Việt Nam đầu thế kỉ XX	Lí luận văn học, phê bình văn học, văn học hiện đại, Việt Nam, chuyên khảo (Nghiên cứu văn học, thế kỉ 20)
29	8405	Giáo trình cảm thụ văn học	Cảm thụ, văn học, giáo trình
30	8517	Đảng bộ tỉnh Ninh Bình lãnh đạo xây dựng và phát triển đời sống văn hóa từ năm 1998 đến năm 2010	Lược văn, tài liệu nội bộ, văn hóa, Đảng bộ, Ninh Bình, Đảng Cộng sản

2.1. Phân tích nội dung tài liệu

Phân tích nội dung tài liệu là nghiên cứu các yếu tố chứa đựng thông tin về nội dung của tài liệu. Mục đích của việc phân tích nội dung tài liệu để định từ khóa là xác định rõ tài liệu đề cập tới vấn đề gì và đối tượng này được đề cập như thế nào. Đây chính là để xác định rõ đối tượng bậc 1, đối tượng bậc 2, các phương diện nghiên cứu của đối tượng và phương pháp nghiên cứu đặc thù được đề cập trong nội dung tài liệu.

Đối tượng nghiên cứu của tài liệu là một phần của thực tại khách quan được đưa vào nghiên cứu và phản ánh nội dung tài liệu- đây là đặc trưng quan trọng nhất của tài liệu. Trong một tài liệu có thể có một hoặc một số đối tượng nghiên cứu. Đối tượng nghiên cứu có thể là đối tượng bậc 1, bậc 2.

Phương diện nghiên cứu của đối tượng là thông tin trả lời câu hỏi: đối tượng được nghiên cứu ở góc độ nào, trên quan điểm nào, ở đâu, thời gian nào, hình thức nào,... Phương diện cũng được phân loại thành 2 bậc: phương diện bậc 1 và phương diện bậc 2 căn cứ vào mối quan hệ đối với đối tượng nghiên cứu.

Phương pháp nghiên cứu đặc thù trả lời câu hỏi: việc nghiên cứu đối tượng có sử dụng phương pháp gì đặc biệt.

“Cán bộ định từ khóa sau khi nắm bắt được những vấn đề được trình bày trong tài liệu sẽ phải nghiên cứu xem tại sao tài liệu về vấn đề đó lại được lựa chọn bổ sung vào vốn tài liệu của thư viện và bạn đọc sẽ thích thú ở điểm gì,... để từ đó định từ khóa được chuẩn xác nhất” [5].

Bản chất của việc định từ khóa là sử dụng từ vựng khoa học để mô tả nội dung tài liệu. Do vậy, đặc điểm nội dung tài liệu

có ảnh hưởng nhiều tới việc lựa chọn từ khóa. Việc lựa chọn từ khóa chủ yếu căn cứ vào các yêu cầu chung về nội dung và hình thức của từ khóa nhưng việc áp dụng chúng tại các thư viện đặc biệt là các thư viện chuyên ngành có những đặc điểm khác nhau. TV ĐHHL là một thư viện đa ngành nhưng khối tài liệu của một số chuyên ngành chiếm phần lớn trong vốn tài liệu là: tài liệu phục vụ các ngành sư phạm (Văn, Toán, Hóa, Lý, Sinh, Địa, Sử, Giáo dục mầm non, Tiểu học,...) và tài liệu của ngành kế toán, quản trị kinh doanh, du lịch, văn hóa học. Đặc điểm của nhóm này là nội dung đề cập tới những đối tượng hữu hình, cụ thể: tài liệu thường có các đối tượng bậc 1, bậc 2 và phương diện bậc 1, bậc 2. Đặc điểm này của tài liệu sẽ dẫn đến việc xác định các đối tượng và phương diện đối tượng dễ dàng khi phân tích nội dung tài liệu. Đặc điểm nữa chi phối cách thức phân tích nội dung tài liệu, đó là các dạng tài liệu khác nhau trong hệ thống. Với dạng tài liệu là luận văn, luận án, đề tài nghiên cứu khoa học cần đặc biệt quan tâm tới việc mô tả phương diện, phương pháp và kết quả nghiên cứu của tài liệu. Đây là những đặc điểm đòi hỏi người cán bộ thư viện cần phải nắm vững trong quá trình phân tích nội dung tài liệu.

Tác giả đã tiến hành phỏng vấn về cách thức phân tích nội dung tài liệu trong quá trình định từ khóa tài liệu của cán bộ TV ĐHHL.

Kết quả cán bộ định từ khóa cho rằng:

- Khi phân tích nội dung tài liệu phải đọc các yếu tố: thư mục, mục lục, lời giới thiệu, lời nói đầu,...

- Khi tiến hành phân tích nội dung tài liệu cần phải chú ý tới lĩnh vực nghiên cứu của tác giả và cần tìm ra các khái niệm đặc

trung cho tài liệu gồm: đối tượng nghiên cứu, phương diện và phương pháp nghiên cứu.

- Với những loại hình tài liệu khác nhau cần chú ý tới các đặc trưng của chúng để việc phân tích nội dung tài liệu đạt được hiệu quả cao nhất.

Qua phân tích biểu ghi được rút ngẫu nhiên trong cơ sở dữ liệu của TV ĐHHL, sau đó tác giả đã tiến hành tham khảo ý

kiến chuyên gia, kết quả cho thấy:

- Số lượng biểu ghi phân tích đúng các đặc trưng nội dung tài liệu là: 28/30 tài liệu (chiếm 93,3%).

- Số lượng biểu ghi phân tích thiếu đặc trưng nội dung tài liệu là: 2/30 tài liệu (chiếm 6,7%).

Kết quả điều tra về cách thức phân tích nội dung tài liệu và chọn lọc đặc trưng tài liệu được trình bày ở Bảng 2.

Bảng 2. Đánh giá cách thức phân tích nội dung tài liệu của cán bộ

Chất lượng Tổng số biểu ghi	Chính xác	Sai một phần	Sai hoàn toàn
30	28	2	0
100%	93,3%	6,7%	0%

2.2. Diễn đạt từ khóa

Từ khóa là từ hoặc cụm từ ổn định, đơn nghĩa được sử dụng để mô tả nội dung chính của tài liệu và để tìm tin trong hệ thống tìm tin tư liệu.

Mục đích của việc diễn đạt từ khóa là để chuyển đổi tượng nghiên cứu, phương diện nghiên cứu của đối tượng và phương pháp nghiên cứu được thể hiện bằng ngôn ngữ tự nhiên của tài liệu sang dạng ngôn ngữ tư liệu - thể hiện bằng từ khóa.

Khi diễn đạt từ khóa, các cán bộ thư viện đã chú trọng tới đặc điểm đối tượng dùng tin. Do người dùng tin ở TV ĐHHL chủ yếu là cán bộ nghiên cứu, giảng dạy và sinh viên mà nhu cầu tin của họ không chỉ về một vấn đề cụ thể nào đó, các khía cạnh của vấn đề mà còn về những vấn đề có liên quan nên cần đặc biệt chú trọng tới việc mô

tả các đối tượng bậc 2 và phương diện của đối tượng bậc 1.

Qua khảo sát các biểu ghi trong các cơ sở dữ liệu tại TV ĐHHL cho thấy:

- Cán bộ thư viện đã mô tả nội dung chính của tài liệu bằng cách định từ khóa. Đối với các thuật ngữ trong tài liệu được chọn làm từ khóa đã đáp ứng các yêu cầu về nội dung và hình thức của từ khóa thì được giữ nguyên, nếu chưa đáp ứng được các yêu cầu của từ khóa thì đã được chỉnh sửa để trở thành từ khóa. Đặc biệt, có trường hợp các thuật ngữ trong tài liệu được cho là chuẩn xác thì cán bộ thư viện không phải chỉnh sửa gì mà lấy nguyên vẹn ra để làm từ khóa. Qua phân tích 30 tài liệu rút ngẫu nhiên trong cơ sở dữ liệu cho thấy lượng tài liệu được định từ khóa dựa chủ yếu vào các thuật ngữ trong tài liệu mà không cần chỉnh sửa hoặc chỉ bổ sung thêm từ khóa thể hiện hình thức của tài liệu (2/30 tài liệu).

- Từ khóa đã đảm bảo độ chính xác, ngắn gọn, súc tích, thông dụng, khách quan.

- Cán bộ đã chú trọng tới các từ khóa mô tả các đối tượng bậc 2 và phương diện của đối tượng bậc 1 bởi đặc thù nghiên cứu giảng dạy của đối tượng người dùng tin.

2.3. Hoàn chỉnh quá trình định từ khóa

Sau khi xác định được các từ khóa của tài liệu, công việc tiếp theo là hoàn chỉnh quá trình định từ khóa. Mục đích của công đoạn này là đảm bảo độ đầy đủ, chính xác, hiệu suất đáp ứng yêu cầu tin của các từ khóa đã chọn. Việc hoàn chỉnh quá trình định từ khóa bao gồm hoàn chỉnh về mặt định lượng và hoàn chỉnh về mặt định tính.

Công đoạn hoàn chỉnh về mặt định lượng từ khóa tại TV ĐHHL là việc cán bộ tiến hành kiểm tra xem các từ khóa đã phản ánh đầy đủ các yếu tố cần thiết của nội dung tài liệu hay chưa bằng cách đối chiếu các từ khóa đã chọn với kết quả phân tích nội dung tài liệu, nếu chưa thấy đầy đủ phải bổ sung các từ khóa còn thiếu.

Công đoạn hoàn chỉnh về mặt định tính từ khóa tại TV ĐHHL là việc cán bộ tiến hành bổ sung thêm từ khóa trong một số trường hợp:

- Từ vựng mang khái niệm tương đương với từ vựng mô tả đối tượng bậc 1.

- Từ khóa theo quan hệ liên đới với từ vựng mô tả đối tượng bậc 1.

Trong trường hợp đối tượng, phương diện của đối tượng mới xuất hiện, cần xem xét bổ sung thêm từ khóa mô tả đối tượng, phương diện đối tượng bậc cao hơn gần nhất.

Trong quá trình hoàn chỉnh quá trình định từ khóa các cán bộ đã chú ý tới đặc điểm của người dùng tin để bổ sung những từ khóa mới theo quan hệ liên đới và quan hệ tương đương, từ đồng nghĩa. Tuy vậy, có một số trường hợp sử dụng quá nhiều từ khóa cho một đối tượng dẫn đến hiện tượng nhiễu tin, làm đầy bộ nhớ của phần mềm thư viện điện tử.

Kết quả phỏng vấn cán bộ thư viện cho thấy:

- Cán bộ thư viện đã diễn đạt các đặc trưng của tài liệu bằng các từ khóa sau đó tiến hành kiểm tra lại để bổ sung những từ

khóa thiếu và loại bỏ những từ khóa thừa.

- Cán bộ thư viện cho rằng trong một số trường hợp, cần bổ sung thêm từ khóa mới theo quan hệ tương đương; mặt khác, cần bổ sung từ khóa ở cấp cao hơn khi mô tả đối tượng và càng bổ sung được nhiều từ khóa ở cấp cao hơn càng tốt.

Như vậy, thông qua hệ thống từ khóa trong các cơ sở dữ liệu của TV ĐHHL đã phản ánh khá rõ nét về trình độ, kỹ năng của cán bộ. Trước và sau khi định từ khóa cho tài liệu của thư viện, cán bộ định từ khóa của thư viện có tham khảo kết quả định từ khóa của tài liệu được phản ánh trong các cơ sở dữ liệu của Thư viện Quốc gia Việt Nam. Tuy nhiên, cán bộ định từ khóa cần lưu ý hơn trong việc sử dụng từ khóa thể hiện hình thức của tài liệu (nhất là đối với loại giáo trình) để khắc phục hiện tượng thiếu từ khóa hình thức.

2.4. Trình bày từ khóa trong biểu ghi

Mục đích của trình bày từ khóa trong biểu ghi là định ra từ khóa chính và từ khóa phụ nhằm tăng cường vai trò thông báo về nội dung tài liệu của các từ khóa, cũng như định hướng cho người dùng tin về những vấn đề cơ bản của nội dung tài liệu. Bởi đối với người dùng tin, đặc biệt là các nhà nghiên cứu, giảng dạy thường có nhu cầu tin về các vấn đề cụ thể, nếu các từ khóa của một tài liệu được sắp xếp một cách có chủ ý để định hướng cho họ về những vấn đề quan trọng của tài liệu thì sẽ giúp họ có được những nhận dạng cơ bản ban đầu về nội dung tài liệu mà chưa cần tham khảo các yếu tố khác của biểu ghi.

Từ khóa chính gồm: các từ khóa về đối tượng bậc 1, từ khóa về đối tượng bậc 2 chiếm nhiều thông tin trong tài liệu.

Từ khóa phụ gồm: các từ khóa về đối tượng bậc 2 chiếm ít thông tin trong tài liệu, từ khóa về phương diện, phương pháp nghiên cứu của tài liệu. Hiện tại thư viện không thực hiện việc sắp xếp các từ khóa thành từ khóa chính và từ khóa phụ nhưng vẫn có những quan điểm khác nhau về vấn đề này. Qua phỏng vấn cho thấy, các cán bộ cho rằng không cần phải sắp xếp các từ khóa thành các từ khóa chính và từ khóa

phụ vì các từ khóa đều có vị trí quan trọng như nhau trong việc tra tìm tài liệu.

3. CHẤT LƯỢNG CÔNG TÁC ĐỊNH TỪ KHÓA

Trên cơ sở 30 biểu ghi của Bảng 1, để khảo sát chất lượng xử lý nội dung tài liệu bằng từ khóa tác giả đã xem xét ở 2 phương diện:

- Chất lượng phản ánh nội dung tài liệu.
- Chất lượng từ khóa độc lập.

3.1. Chất lượng phản ánh nội dung tài liệu

Để đánh giá chất lượng phản ánh nội dung tài liệu, tác giả đánh giá từng biểu ghi trong 30 biểu ghi đã được lựa chọn ngẫu nhiên theo các tiêu chí: độ chính xác và độ đầy đủ.

Độ chính xác

Độ chính xác của từng biểu ghi được xác

định bằng tỷ số giữa số lượng các từ khóa được lựa chọn phản ánh đúng đối tượng, phương diện và phương pháp với tổng số từ khóa được lựa chọn trong biểu ghi.

Ví dụ:

Biểu ghi số 7965 trong Bảng 1, có 5 từ khóa: *Luật kế toán, Luật kiểm toán, Văn bản pháp luật, Việt Nam, Pháp luật* thì từ khóa *Pháp luật* là phản ánh không chính xác nội dung tài liệu (từ khóa quá rộng). Từ khóa *Luật kế toán, Luật kiểm toán, Văn bản pháp luật, Việt Nam* là phản ánh đúng nội dung tài liệu. Như vậy, độ chính xác của biểu ghi này đạt 4/5.

Kết quả phân tích các biểu ghi được rút ngẫu nhiên trong các cơ sở dữ liệu cho thấy:

- Lượng biểu ghi có từ khóa quá rộng có thể gây nhiễu tin khi người dùng tin tra cứu là: 5/30.

Bảng 3. Danh sách biểu ghi có từ khóa quá rộng

STT	Biểu ghi	Tên tài liệu	Các từ khóa	Từ khóa quá rộng	Sửa lại
1	122	Những danh tướng của Trung Hoa thời xưa	Nhân vật lịch sử, lịch sử, phong kiến, Trung Hoa, Trung Quốc, danh tướng, quân sự	Lịch sử	Lịch sử cổ đại
2	368	Đôxtôpexki cuộc đời và sự nghiệp	Văn học, cuộc đời, sự nghiệp, Nga	Văn học	Văn học hiện đại, nghiên cứu văn học, lí luận văn học, phê bình văn học
3	384	Nghiệp vụ buôn bán quốc tế	Kinh tế, thương mại, hợp đồng buôn bán, quốc tế	Kinh tế	Bỏ từ khóa thừa
4	7965	Luật kế toán và kiểm toán	Luật kế toán, luật kiểm toán, văn bản pháp luật, Việt Nam, pháp luật	Pháp luật	Bỏ từ khóa thừa
5	966	Phương pháp toán trong địa lí	Toán học, giáo trình, địa lí	Toán học	Toán ứng dụng

- Lượng biểu ghi thừa từ khóa có thể gây nhiễu tin cho người dùng khi tra cứu tài liệu là: 1/30 tài liệu

Bảng 4. Danh sách biểu ghi có từ khóa bị thừa

STT	Biểu ghi	Tên tài liệu	Các từ khóa	Từ khóa quá rộng	Sửa lại
1	122	Con đường tơ lụa quá khứ và tương lai	Quan hệ quốc tế, giao thông, lịch sử, văn hóa, thương mại, thế giới, Trung Quốc	Thế giới	Bỏ từ khóa thừa

Độ đầy đủ

Độ đầy đủ của từng biểu ghi được xác định bằng tỷ số giữa số lượng các đối tượng, phương diện, phương pháp được mô tả và tổng số đối tượng, phương diện,

phương pháp được đề cập đến trong nội dung tài liệu.

- Lượng biểu ghi thiếu từ khóa có thể gây khó khăn trong việc tra tìm tài liệu của người dùng tin là: 2/30 tài liệu

Bảng 5. Danh sách biểu ghi có từ khóa bị thiếu

STT	Biểu ghi	Tên tài liệu	Các từ khóa	Từ khóa bị thiếu cần bổ sung thêm
1	8362	Lí luận phê bình văn học Việt Nam đầu thế kỉ XX	Lí luận văn học, phê bình văn học, văn học hiện đại, Việt Nam, chuyên khảo	Thế kỉ 20
2	7927	Hệ thống câu hỏi và bài tập môn học kế toán công ty	Kế toán, công ty, câu hỏi, bài tập, kế toán doanh nghiệp	Giáo trình

- Lượng biểu ghi có các từ khóa mang ý nghĩa tương đương giúp người dùng tin thuận lợi hơn khi tra cứu tài liệu là:

2/30 tài liệu. Những trường hợp thiếu từ khóa tương đương cần phải được bổ sung.

Bảng 6. Danh sách biểu ghi cần bổ sung từ khóa có nghĩa tương đương

STT	Biểu ghi	Tên tài liệu	Các từ khóa	Từ khóa bị thiếu cần bổ sung thêm
1	122	Những danh tướng của Trung Hoa thời xưa	Nhân vật lịch sử, lịch sử, phong kiến, Trung Hoa, danh tướng, quân sự	Trung Quốc
2	8362	Lí luận phê bình văn học Việt Nam đầu thế kỉ XX	Lí luận văn học, phê bình văn học, văn học hiện đại, Việt Nam, chuyên khảo	Nghiên cứu văn học

3.2. Chất lượng từ khóa độc lập

Để đánh giá chất lượng từ khóa độc lập, tác giả xem xét về nội dung và hình thức của từ khóa dựa trên những chuẩn yêu cầu nghiệp vụ của công tác định từ khóa tại các cơ quan TT-TV.

Tại TV ĐHHL công tác định từ khóa đã đảm bảo được các yêu cầu về nội dung từ khóa đó là:

- Thông dụng, đúng dẫn theo thuật ngữ khoa học: Hệ thống từ khóa là những từ khoa học, thông dụng trong lĩnh vực mà nội dung tài liệu đề cập, không sử dụng thuật ngữ và các từ nghĩa bóng,...

- Súc tích: Từ khóa thể hiện nội dung thông tin dưới hình thức ngắn gọn nhất; cán bộ thư viện đã lựa chọn được những từ khóa thực sự mang thông tin phản ánh nội dung tài liệu và loại bỏ những từ không có ích trong việc tra cứu.

- Ngắn gọn: Cán bộ định từ khóa đã tách các khái niệm phức tạp thành các khái niệm đơn giản nhất có thể nhằm tạo điều kiện thuận lợi cho việc sử dụng các toán tử tìm tin của người dùng trong quá trình tìm kiếm tài liệu trên các cơ sở dữ liệu.

- Đơn nghĩa: Hệ thống từ khóa gồm những từ đơn nghĩa (chỉ mang một nghĩa duy nhất và một khái niệm/đối tượng cụ thể, chỉ được mô tả bằng một từ khóa duy nhất); các hiện tượng đa nghĩa và đồng nghĩa được sử dụng khi cần thiết nhằm tạo điều kiện thuận lợi cho người tra cứu thông tin.

- Chính xác và hiện đại: Hệ thống từ khóa đã phản ánh chính xác những khái niệm cơ bản của nội dung tài liệu, đồng thời cũng gồm các thuật ngữ được dùng phổ biến trong các chuyên ngành.

- Khách quan: Hệ thống từ khóa độc

lập với văn cảnh của nội dung tài liệu gốc, tức là có thể hiểu được từ khóa mà không cần đọc nội dung tài liệu gốc, đồng thời từ khóa không mang sắc thái phê phán, đánh giá của người cán bộ xử lý từ khóa.

KẾT LUẬN

Hiện nay, TV ĐHHL vẫn sử dụng phương pháp định từ khóa tự do trong công tác xử lý nội dung tài liệu. So với ngôn ngữ từ khóa có kiểm soát, ngôn ngữ từ khóa tự do có khối lượng từ phong phú và đảm bảo cao nhất về tính thời sự và cập nhật từ mới. Ngoài ra, nó còn giúp quá trình xử lý nội dung tài liệu nhận biết mẫu tìm dễ dàng, công việc xử lý tiến triển nhanh vì không phải thực hiện công đoạn kiểm soát từ. Tuy nhiên, ngôn ngữ từ khóa tự do có mức độ hình thức hóa thấp, thiếu thống nhất nên ảnh hưởng đến quá trình tìm tin. Vì sự thiếu thống nhất, thiếu chặt chẽ của ngôn ngữ từ khóa tự do nên trong cơ sở dữ liệu của TV ĐHHL còn tồn tại các từ khóa quá dài, đa nghĩa và đồng nghĩa. Việc sử dụng các từ khóa này có thể tăng độ chính xác khi tìm tin, nhưng cũng có thể gây mất hoặc sót tin. Điều này làm cho người tìm tin không tìm được đầy đủ các tài liệu cần thiết do hiện tượng nhiễu tin dù các tài liệu cần tìm có trong cơ sở dữ liệu. Hiện tượng này làm hạn chế khả năng khai thác cơ sở dữ liệu của người dùng tin tại TV ĐHHL. Chính vì vậy, ngoài việc định từ khóa tự do, cán bộ định từ khóa tài liệu cần sử dụng các công cụ tra cứu khác như bộ từ khóa do Thư viện Quốc gia biên soạn và "Bộ từ khóa đa ngành khoa học và công nghệ"

gồm 2 tập do Trung tâm Thông tin Khoa học và Công nghệ Quốc gia biên soạn để nâng cao tính chuẩn hóa và kiểm soát tính thống nhất trong cách định từ khóa tài liệu tại Thư viện. Đặc biệt, Thư viện có thể sử dụng kết quả định từ khóa tài liệu của Thư viện Quốc gia Việt Nam xử lý. Bởi lẽ, Thư viện Quốc gia Việt Nam là cơ quan có chuyên môn nghiệp vụ cao nhất về thư viện và kết quả xử lý của Thư viện là kết quả đáng tin cậy vì nó được kiểm soát chặt chẽ cả về số lượng và chất lượng.

TÀI LIỆU THAM KHẢO

1. Bộ từ khoá. - H.: Thư viện Quốc gia Việt Nam, 2005.
2. Bộ từ khoá khoa học xã hội và nhân văn. - H.: Viện Thông tin Khoa học Xã hội, 2005. - 2T.
3. Nguyễn Yến Vân, Vũ Dương Thúy Nga (2006). Thư viện học đại cương. - H.: Trường Đại học Văn hóa Hà Nội, 2006.
4. Nghiên cứu xây dựng bộ từ khoá cho các cơ sở dữ liệu đa ngành khoa học tự nhiên và kỹ thuật: báo cáo tổng kết đề tài nghiên cứu khoa học cấp Bộ.- H.: Trung tâm Thông tin Khoa học và Công nghệ Quốc gia, 1997.
5. Trần Thị Quý, Nguyễn Thị Đào (2007). Xử lý tài liệu trong hoạt động TtT-TV, Nxb Đại học Quốc gia Hà Nội, Hà Nội.

*(Ngày Tòa soạn nhận được bài: 15-9-2024;
Ngày phản biện đánh giá: 5-10-2024;
Ngày chấp nhận đăng: 15-11-2024).*