

VÀI NÉT ĐỐI SÁNH VỀ TỪ KHÓA TRONG HỆ THỐNG THÔNG TIN-THƯ VIỆN VÀ NGUỒN THÔNG TIN TRÊN INTERNET

TS Đoàn Thị Thu

Khoa Thư viện - Thông tin học, Trường Đại học KHXH&NV Tp. Hồ Chí Minh

Tóm tắt: Bài viết so sánh để làm rõ sự khác biệt về từ khóa giữa hệ thống thông tin thư viện với nguồn tin trên internet về các mặt: khái niệm, mục đích, cách thức và công cụ tạo lập từ khóa, tính năng và cơ chế tìm kiếm thông tin bằng từ khóa, phân loại và yêu cầu đối với từ khóa.

Từ khóa: Từ khóa; hệ thống thông tin; thư viện; internet.

KEYWORDS COMPARISON IN THE LIBRARY-INFORMATION SYSTEM AND INFORMATION SOURCES ON THE INTERNET

Abstract: The article compares to clarify the difference in keywords between library information systems and internet sources in terms of: concepts, purposes, methods and tools to generate keywords, features and information search mechanism by keywords, classification and requirements for keywords.

Keywords: Information system; library; internet.

ĐẶT VẤN ĐỀ

Ngày nay, từ khóa được xem là một trong các công cụ tìm kiếm thông tin phổ biến trong lĩnh vực thông tin-thư viện (TT-TV) cũng như trên internet. Trong lĩnh vực TT-TV, nhiều thư viện đã xây dựng được cơ sở dữ liệu với tính năng tìm kiếm thông tin bằng từ khóa. Với nguồn thông tin trên internet, người sử dụng cũng đã hình thành thói quen tìm kiếm thông tin bằng cách nhập từ khóa vào ô tìm kiếm. Tuy đều được người sử dụng yêu thích bởi ưu điểm tìm kiếm nhanh, linh hoạt và thân thiện nhưng việc tạo lập và sử dụng từ khóa vẫn có nhiều điểm khác biệt giữa hệ thống TT-TV với nguồn tin trên internet. Bài viết này tập trung làm rõ sự khác biệt của từ khóa về một số nội dung như: khái niệm, ý nghĩa, phân loại, cơ chế tạo lập cũng như mục đích sử dụng từ khóa giữa hệ thống TT-TV và nguồn tin trên internet.

1. KHÁI NIỆM TỪ KHÓA

Trong hệ thống TT-TV, khái niệm từ khóa được đề cập trong một số tài liệu:

Theo TCVN 5354:2009 (ISO 5127:2001) từ khóa là “từ có nghĩa được rút ra từ nhan đề của văn bản hoặc văn bản tài liệu để thể hiện nội dung” [Tiêu chuẩn Quốc gia, 2009].

Theo ALA -Từ điển giải nghĩa thư viện học, từ khóa là “một từ ngữ quan trọng (có ý nghĩa) trong một toát yếu, một nhan đề hay trong văn bản của một tác phẩm được dùng làm từ mô tả” [ALA, 1996].

Theo Phan Huy Quế, 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 thông tin tư liệu” [Phan Huy Quế, 2001].

Như vậy, mặc dù có sự khác biệt nhỏ về yêu cầu đối với từ khóa (TCVN 5354:2009 và từ điển ALA không đề cập tới yêu cầu đối với từ khóa, còn Phan Huy Quế lại chỉ rõ từ khóa phải là từ/cụm từ ổn định, đơn nghĩa) nhưng nhìn chung các khái niệm trong hệ thống TT-TV đều xem từ khóa là từ hoặc cụm từ thể hiện nội dung tài liệu.

Nguồn tin trên internet rất đa dạng, bao gồm các bài báo trên internet, các website cũng như các cơ sở dữ liệu lớn với các máy tìm tin (Search Engines). Với các nguồn thông tin trên internet, từ khóa thường đặt trong bối cảnh SEO (Search Engine Optimization - tối ưu hóa công cụ tìm kiếm). Người dùng internet thường sử dụng từ khóa trong việc tìm kiếm thông tin, trong khi

các doanh nghiệp lại quan tâm tới từ khóa như một công cụ để kết nối với khách hàng. Hiện có nhiều khái niệm khác nhau về từ khóa SEO. Tuy nhiên, tổng hợp các khái niệm về từ khóa SEO, tác giả nhận thấy có hai hướng tiếp cận:

- Hướng tiếp cận thứ nhất, xem từ khóa là những từ hoặc cụm từ được người dùng internet nhập vào ô tìm kiếm để tìm những thông tin mình mong muốn. Ở hướng tiếp cận này, có một số khái niệm cụ thể như sau:

Theo Ahref - một trong các công cụ phân tích và tạo lập từ khóa trong lĩnh vực SEO định nghĩa: “từ khóa là bất cứ từ hoặc cụm từ nào mà người tìm tin nhập vào ô tìm kiếm”.

Theo Apdia.vn (2020), từ khóa là “từ hay cụm từ có nghĩa mà người tìm kiếm sử dụng trên các công cụ tìm kiếm như Facebook, Google,...”

Từ khóa “là một từ hoặc cụm từ xác định một chủ đề, một đối tượng hoặc một khái niệm. Khi mọi người tìm kiếm thông tin về một trang web trên công cụ tìm kiếm, họ cần điền một số từ hoặc cụm từ trên công cụ tìm kiếm để được trả về danh sách các trang web liên quan đến từ hoặc cụm từ họ đang tìm kiếm” [ATP Media, 2020].

- Hướng tiếp cận thứ hai, xem từ khóa là những từ hoặc cụm từ được tác giả hay chủ các website sử dụng để mô tả nội dung của bài viết nhằm mục đích giúp người dùng internet tìm ra, cải thiện vị trí xếp hạng của từ khóa trong trang kết quả tìm kiếm. Cụ thể:

Theo Google Ads, từ khóa là những từ hoặc cụm từ dùng để so khớp quảng cáo của bạn với những cụm từ mà mọi người đang tìm kiếm.

Theo Lê Nam (2020), Từ khóa SEO là thuật ngữ để nói về những từ khóa được tối ưu trong một bài viết với mục đích cải thiện vị trí xếp hạng của thuật ngữ đó trên SERP (Search Engine Results Pages- trang kết quả của công cụ tìm kiếm).

Tương tự như vậy, Vicent Do (2021) tổng hợp và đưa ra khái niệm từ khóa SEO là “những từ và cụm từ trong trang web có chức năng mô tả nội dung của bài viết/trang web chứa chúng. Từ khóa SEO được xem là một mắt xích vô cùng quan trọng giúp cho các con bot của bộ máy tìm kiếm và người dùng tìm thấy và hiểu nội dung trang web, qua đó cải thiện thứ hạng trang web trên trang kết quả tìm kiếm”.

Như vậy, sự khác biệt đầu tiên về từ khóa giữa hệ thống TT-TV và nguồn tin trên internet nằm ở khái niệm về từ khóa. Trong hệ thống TT-TV, nhìn chung các khái niệm đều xem từ khóa là những từ hoặc cụm từ thể hiện nội dung tài liệu. Ở cách tiếp cận này, có thể hiểu từ khóa là do những người làm công tác xử lý nội dung tài liệu là người tạo ra từ khóa và đảm bảo yêu cầu từ khóa phải phản ánh được nội dung tài liệu. Trong khi đó, cách tiếp cận từ khóa trong SEO lại theo hai hướng: hướng thứ nhất tiếp cận ở phía người dùng internet cho rằng từ khóa là những từ hoặc cụm từ họ sử dụng để tìm kiếm. Ở góc độ tiếp cận này, ta thấy người dùng internet chính là người tạo ra từ khóa. Hướng thứ hai trong SEO tiếp cận từ khóa ở góc độ người tạo lập ra thông tin (trên website) và người tạo lập công cụ tìm kiếm (các máy tìm) xem từ khóa như một sản phẩm được tạo lập giúp đạt mục đích đề ra như giúp người dùng tìm kiếm thông tin hay cải thiện thứ hạng của website. Ở góc độ tiếp cận này, có thể nhận thấy, người tạo lập thông tin cũng như người tạo lập công cụ tìm kiếm đang xem xét tạo lập các từ khóa phù hợp với các từ khóa mà người dùng internet cũng như máy tìm ưu tiên sử dụng để “khớp” với từ khóa được chủ website tạo lập. Sự khác biệt này về khái niệm từ khóa giữa hệ thống TT-TV và nguồn tin trên internet vừa phản ánh các nhìn nhận từ khóa ở hai hệ thống nhưng đồng thời là kết quả phản ánh sự khác biệt về mục đích, cách thức cũng như cơ chế tìm kiếm thông tin bằng từ khóa giữa hai hệ thống trên.

2. MỤC ĐÍCH, CÁCH THỨC VÀ CÔNG CỤ TẠO LẬP TỪ KHÓA

2.1. Về mục đích tạo lập từ khóa

Trong hệ thống TT-TV, việc các thư viện tạo lập từ khóa nhằm tạo các điểm truy cập, giúp người sử dụng tìm kiếm được tài liệu phù hợp. Mục đích này được quy định bởi thư viện được thành lập phi lợi nhuận với nhiệm vụ của thư viện là xây dựng hệ thống thông tin và bộ máy tra cứu phục vụ cho người sử dụng. Tuy nhiên, việc xây dựng công cụ tìm kiếm trên internet lại có mục đích hoàn toàn khác. Các doanh nghiệp xây dựng công cụ tra cứu cũng như các doanh nghiệp tạo lập tin trên internet đều có chung mục đích là quảng bá thương hiệu của doanh nghiệp tới đông đảo người sử dụng nhằm thu lợi nhuận từ việc kinh doanh hoặc thu phí từ các dịch vụ quảng cáo.

Đối với các doanh nghiệp tạo tin trên internet, có thể nhận thấy ngày càng nhiều website có tạo lập từ khóa cho các bài viết. Tuy nhiên, khi so sánh tiêu đề, nội dung bài viết với các từ khóa được tạo lập cho bài viết đó, tác giả nhận thấy có một số từ khóa không xuất hiện (hoặc được đề cập sơ) trong nội dung bài viết. Nhìn vào Hình 1 có thể thấy tên bài viết và các từ khóa được tạo lập cho bài viết chưa thực sự phù hợp. Nhan đề bài viết đề cập tới các đô thị vùng đệm (vùng xung quanh) sân bay Long Thành tại Đồng Nai, tuy nhiên các từ khóa của bài viết đều đề cập tới dự án bất động sản của tập đoàn Đất Xanh. Như vậy, có thể thấy, mục đích của các từ khóa được tạo lập cho bài viết này nhằm hướng sự quan tâm của người dùng tới các sản phẩm, dịch vụ của doanh nghiệp chi trả chi phí quảng cáo (mục đích của bài viết) chứ không đơn thuần là phản ánh nội dung bài viết.



Hình 1. Định từ khóa cho bài viết trên báo Thanh niên

Nhắc tới các doanh nghiệp tạo lập công cụ tìm kiếm trên internet, người dùng sẽ nhắc tới công cụ tìm kiếm Google. Hiện nay, việc sử dụng công cụ tìm kiếm Google đã trở nên quá quen thuộc đối với hầu hết người dùng internet. Người dùng sử dụng Google như một công cụ tìm kiếm thông tin khi có nhu cầu. Do vậy, rất nhiều người lầm tưởng Google được tạo lập với mục đích giúp người dùng internet tìm kiếm thông tin miễn phí (giống như mục đích của các hệ thống TT-TV). Không thể phủ nhận nhiều người dùng cá nhân đang sử dụng Google để

tìm kiếm thông tin miễn phí. Tuy nhiên, tìm hiểu sâu về cơ chế tìm kiếm của Google có thể thấy mục đích của Google khi xây dựng công cụ tìm kiếm không đơn thuần là cung cấp miễn phí cho người dùng internet mà hướng tới: (1) thu thập các thông tin từ người dùng internet và (2) giúp quảng cáo các thông tin/sản phẩm/dịch vụ của doanh nghiệp - có thu phí. Cụ thể, hiện nay Google đang thực hiện 6 loại hình quảng cáo như: quảng cáo tìm kiếm (Google Search), quảng cáo mạng hiển thị (Google Display Network), quảng cáo mua sắm (Google

Shopping Ads), quảng cáo video (Video Ads), quảng cáo ứng dụng toàn cầu (Apps), chiến dịch quảng cáo thông minh [Nguyễn Ngọc Hưng, 2021; Google Ads trợ giúp].

Như vậy, việc tạo lập từ khóa giúp người dùng internet tìm kiếm thông tin cũng là một trong các cách thức để hoạt động của doanh nghiệp, còn việc giúp người sử dụng tìm kiếm thông tin chỉ là lợi ích đi kèm. Chính sự khác biệt về mục đích tạo lập từ khóa này sẽ quyết định tới cách thức cũng như công cụ tạo lập từ khóa giữa hệ thống thông tin thư viện và các công cụ tra cứu trên internet.

2.2. Về công cụ tạo lập từ khóa

Trong hệ thống TT-TV, việc tạo lập từ khóa cho tài liệu được dựa trên nội dung của tài liệu. Theo đó, để định từ khóa cho tài liệu, người định từ khóa cần thực hiện theo quy trình: phân tích nội dung tài liệu, xác định các khái niệm tiêu biểu đặc trưng cho nội dung tài liệu (từ/cụm từ chỉ đối tượng nghiên cứu và phương diện nghiên cứu), chuyển đổi từ khóa và nhập từ khóa vào hệ thống. Tùy theo lĩnh vực chuyên môn của tài liệu mà mỗi hệ thống TT-TV lựa chọn định từ khóa tự do hay có kiểm soát. Tuy nhiên, hiện nay để chuẩn hóa nghiệp vụ, các thư viện đều hướng tới sử dụng các công cụ trong tạo lập từ khóa như: Bộ từ khóa của Thư viện Quốc gia Việt Nam, bộ từ khóa Khoa học Xã hội và Nhân văn, Từ điển Từ khóa Khoa học và Công nghệ,...

Đối với nguồn tin trên internet, mặc dù việc sử dụng từ khóa để tìm kiếm thông tin đã dần trở nên phổ biến với người sử dụng, tuy nhiên việc tạo lập từ khóa chưa thực sự được chú trọng. Nhìn chung các công cụ tìm kiếm trên internet đều sử dụng cơ chế tìm kiếm bằng cách so sánh từng ký tự trong từ khóa với ký tự trong kho dữ liệu. Trong thời gian gần đây, một số trang báo điện tử đã bắt đầu tiến hành định từ khóa các bài viết. Cụ thể, dưới mỗi bài viết sẽ được gắn các từ khóa hoặc ký hiệu hashtag (Hình 1).

Nhìn vào Hình 1 có thể thấy, mặc dù mỗi

bài viết đều có các từ khóa đi kèm. Về lý thuyết, các từ khóa này sẽ giúp người đọc hiểu hơn về nội dung bài viết. Tuy nhiên, có thể nhận thấy nội dung bài viết và từ khóa chưa thực sự khớp nhau, thậm chí từ khóa chưa phản ánh được nội dung bài viết. Điều này có thể được lý giải rằng bài viết và từ khóa được tạo lập nhằm mục đích quảng cáo sản phẩm cho các doanh nghiệp. Bên cạnh đó, khi tìm kiếm bằng từ khóa, một số tài liệu hiển thị trong kết quả tìm kiếm nhưng nội dung không phù hợp là do các từ khóa được tạo lập cho tài liệu tuy khớp với từ khóa tìm kiếm nhưng lại không phản ánh đúng nội dung tài liệu.

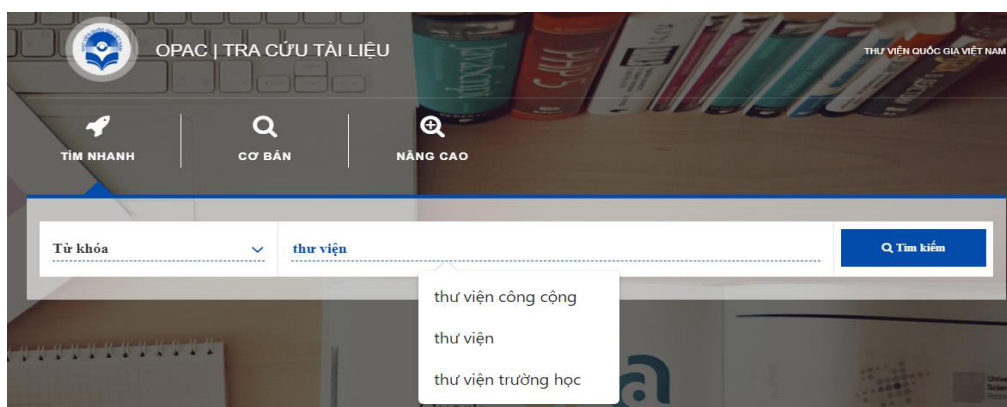
Khác với lĩnh vực TT-TV, các từ khóa được tạo lập cho các nguồn tin trên internet đều là các từ khóa tự do (người định từ khóa tự tạo từ khóa mà không cần công cụ kiểm soát từ khóa nào). Điều này là do việc tạo lập từ khóa cho các nguồn tin trên internet mới được chú trọng gần đây (trong khi lĩnh vực thư viện đã có thời gian dài nghiên cứu và tạo lập), hơn nữa do nội dung thông tin trên internet quá đa dạng và nguồn tin khó được kiểm soát.

3. TÍNH NĂNG CỦA TỪ KHÓA

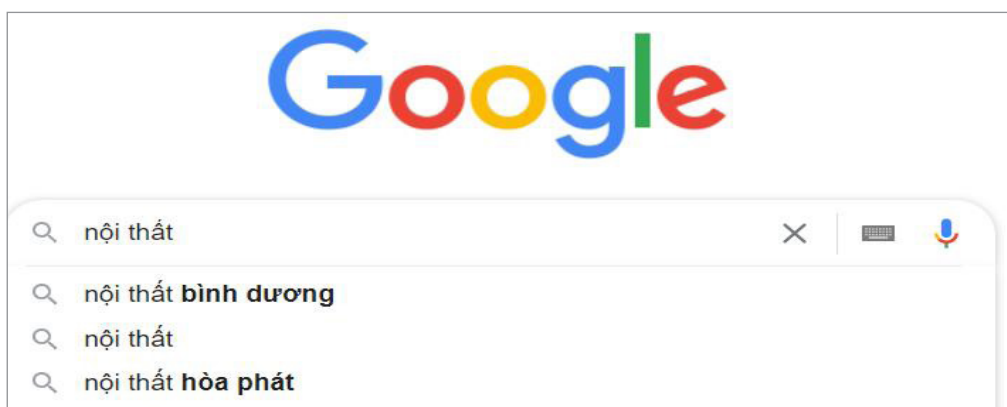
Có thể nói, vai trò lớn nhất của từ khóa là giúp người sử dụng tìm kiếm thông tin. Sở dĩ từ khóa được sử dụng phổ biến trong các hệ thống thông tin cũng như trên internet bởi nó có ưu điểm là tìm kiếm thông tin nhanh, thân thiện và linh hoạt. Cụ thể là:

- Khả năng tìm tin nhanh của từ khóa được thể hiện ở chỗ sau khi nhập từ khóa vào ô tìm kiếm, tùy từng hệ thống thông tin hay công cụ tìm kiếm nhưng trung bình người sử dụng chỉ cần chờ vài giây, thậm chí một phần của giây đã được hệ thống hiển thị kết quả tìm kiếm.

- Tính thân thiện của từ khóa được thể hiện ở việc khi nhập từ khóa vào ô tìm kiếm, cả hệ thống TT-TV lẫn công cụ tìm kiếm trên internet đều có các từ khóa gợi ý cho người sử dụng (Hình 2 và 3)



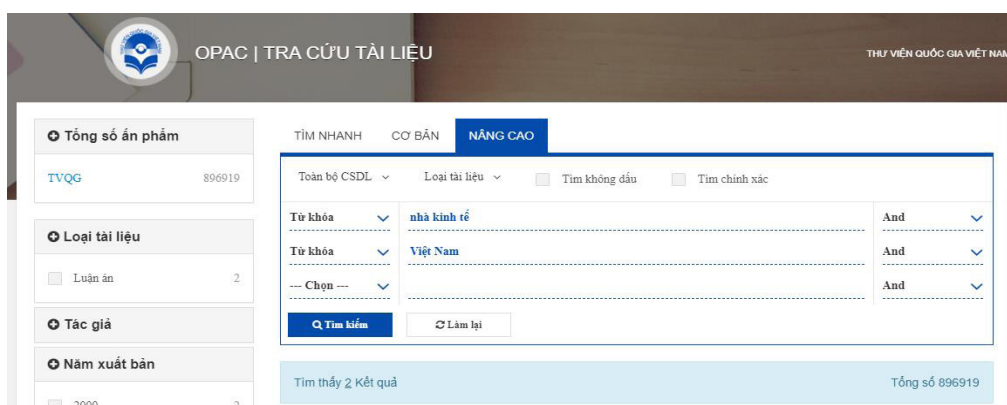
Hình 2. Tính năng gợi ý khi tìm tin bằng từ khóa trên OPAC của Thư viện Quốc gia Việt Nam



Hình 3. Tính năng gợi ý khi tìm tin bằng từ khóa trên Google

- Tính linh hoạt khi tìm kiếm bằng từ khóa được thể hiện ở việc hệ thống sử dụng các toán tử trong việc thu hẹp hoặc mở rộng kết quả tìm kiếm. Cụ thể, trong hệ thống

thông tin thư viện, có thể mở rộng hoặc thu hẹp kết quả tìm kiếm dựa vào việc tìm kiếm nâng cao (sử dụng toán tử Boolean - And, Or, Not) như Hình 4.



Hình 4. Tính năng tìm kiếm nâng cao trong OPAC của Thư viện Quốc gia Việt Nam

Đối với Google, việc tìm kiếm bằng từ khóa thường hiện thị nhiều kết quả tìm kiếm do việc so sánh từng từ trong từ khóa tìm kiếm với từng từ trong tài liệu. Tuy nhiên, Google cũng hướng dẫn người sử dụng thu hẹp kết quả tìm kiếm bằng các toán tử như: bỏ khoảng trống giữa các từ khóa, sử dụng các biểu tượng (@ khi tìm tin trên mạng xã hội, \$ khi muốn tìm giá tiền, # hashtag, đặt dấu “-” trước từ khóa nếu muốn loại trừ kết quả tìm kiếm),... [Google tìm kiếm trợ giúp].

Vai trò thứ hai của từ khóa là giúp thông báo nội dung tài liệu [Phan Huy Quế, 2001]. Cụ thể: trong hệ thống TT-TV, từ khóa được sử dụng với vai trò thông báo nội dung tài liệu, giúp người sử dụng nắm được thông tin tài liệu. Hình 5 minh họa cho việc từ khóa giúp người sử dụng phán đoán được nội dung tài liệu dựa vào từ khóa, đặc biệt với những tài liệu có nhan đề không thể hiện rõ nội dung.



Hình 5. Vai trò của từ khóa trong việc thông báo nội dung tài liệu

Đối với nguồn tin trên internet, hiện nay nhiều trang báo điện tử đã tiến hành định từ khóa cho các báo như báo Thanh niên, Tuổi

trẻ... Cuối mỗi bài báo đều có các từ khóa đi kèm (Hình 6).



Hình 6. Từ khóa sau mỗi bài viết trên báo Tuổi trẻ

Tuy nhiên, ngoài vai trò thông báo nội dung bài viết, mỗi từ khóa này còn đóng vai trò như một chủ đề bởi lẽ khi đọc giả click

vào từ khóa sẽ được dẫn sang một chủ đề với các tài liệu có liên quan (Hình 7).



Hình 7. Tính năng tập hợp các bài viết theo từ khóa của chủ đề trên báo Tuổi trẻ

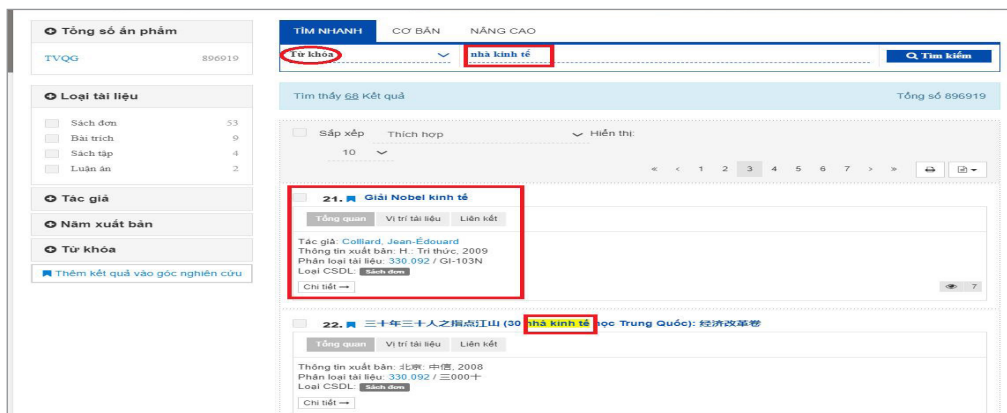
Như vậy, hệ thống TT-TV và công cụ tìm kiếm trên internet đều khai thác được các tính năng cơ bản của từ khóa là tìm kiếm tài liệu có nội dung phù hợp đồng thời giúp thông báo nội dung của tài liệu đến độc giả. Tuy nhiên, vẫn có điểm khác biệt nhỏ ở chỗ các nguồn tin trên internet đang lồng ghép vai trò của từ khóa và đề mục chủ đề trong việc tập hợp các tài liệu có liên quan.

4. CƠ CHẾ TÌM KIẾM THÔNG TIN BẰNG TỪ KHÓA

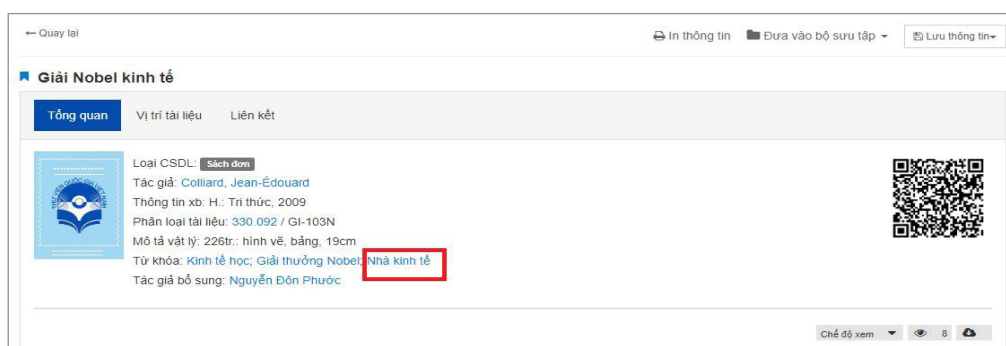
Trong hệ thống TT-TV, để người sử dụng có thể tìm kiếm thông tin bằng từ khóa, thư viện cần: (1) xây dựng kho dữ liệu với các tài liệu đã được định từ khóa (dựa trên nội dung tài liệu), (2) xây dựng bộ máy tra cứu với tính năng giúp người sử dụng nhập từ khóa vào ô tìm kiếm (dựa trên nội dung yêu cầu tin). Khi có yêu cầu tin (người sử

dụng nhập từ khóa vào ô tìm kiếm), máy tính sẽ sử dụng từ khóa đó để quét trong kho dữ liệu nhằm tìm ra các tài liệu có từ khóa (được cán bộ thư viện tạo lập trong quá trình xử lý) trùng khớp hoàn toàn với từ khóa trong yêu cầu tin.

Để mô phỏng cơ chế tìm kiếm thông tin bằng từ khóa trong hệ thống thông tin, tác giả truy cập vào OPAC của Thư viện Quốc gia Việt Nam và tìm kiếm với từ khóa “nhà kinh tế”. Kết quả tìm kiếm cho thấy 68 tài liệu phù hợp, trong đó có tài liệu có từ/ cụm từ trong nhan đề tài liệu trùng hợp với từ khóa tìm kiếm (tài liệu số 22 - Hình 8) và tài liệu tuy nhan đề tài liệu không có từ trùng khớp với từ khóa (tài liệu số 21 - Hình 8) nhưng nội dung lại trùng khớp về từ khóa (Hình 9).



Hình 8. Kết quả tìm kiếm bằng từ khóa “nhà kinh tế” trên OPAC của Thư viện Quốc gia Việt Nam

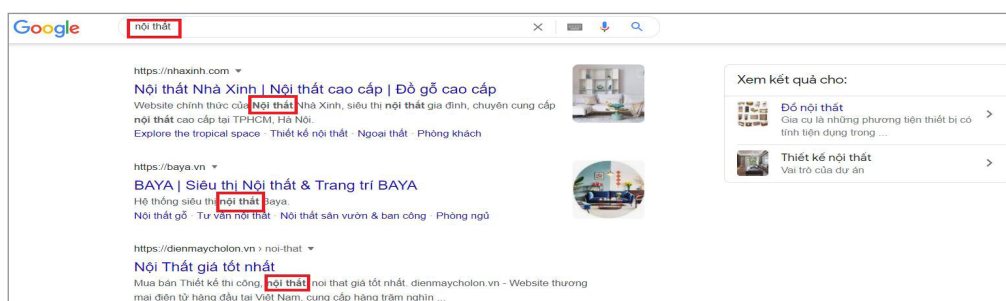


Hình 9. Thông tin chi tiết tài liệu tìm được khi tìm kiếm bằng từ khóa “nhà kinh tế” trên OPAC của Thư viện Quốc gia Việt Nam

Cơ chế tìm kiếm thông tin bằng từ khóa trên internet thường được gọi là công cụ tìm kiếm (Search Engine) với mục đích tra cứu, tìm hiểu thông tin trên internet. Hiểu một cách đơn giản thì đây chính là một trang web mà tại đó người dùng nhập các từ khóa vào thanh công cụ tìm kiếm và được trả về các kết quả hiển thị có chứa từ khóa đó hay một phần của từ khóa đó và được sắp xếp

theo một thứ tự xếp hạng cụ thể.

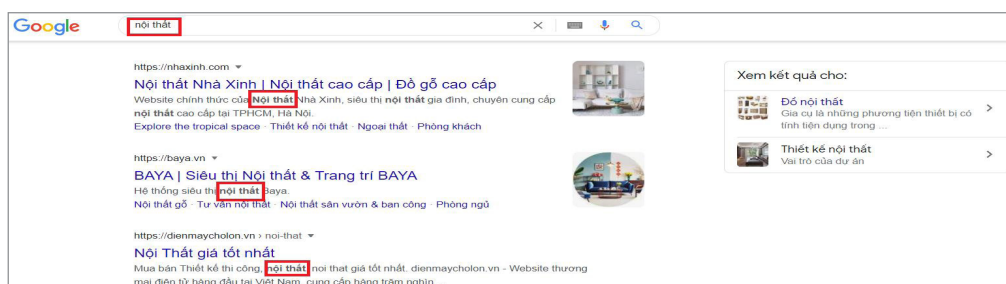
Để mô phỏng cơ chế tìm kiếm thông tin bằng từ khóa trên internet, tác giả sử dụng từ khóa “nội thất” trên trang chủ Google. Kết quả tìm kiếm trả về có 137.000.000 tài liệu phù hợp. Nhìn vào từng kết quả tìm kiếm, có thể nhận thấy tất cả kết quả trả về đều có chứa từ khóa giống với từ khóa được nhập vào ô tìm kiếm (Hình 10).



Hình 10. Kết quả tìm kiếm bằng từ khóa “nội thất” trên Google

Tuy nhiên, hiển thị trên đầu danh sách kết quả tìm kiếm là các thông tin có gắn chữ “Quảng cáo”. Đồng thời nhìn kỹ kết quả có

thể thấy một số thông tin trả về không có chứa từ khóa tìm kiếm hoặc chỉ chứa một phần từ khóa tìm kiếm (Hình 11).



Hình 11. Thông tin được gắn “quảng cáo” trong kết quả tìm kiếm từ khóa “nội thất” trên Google

Sở dĩ việc tìm kiếm từ khóa trên Google có tốc độ nhanh và số lượng kết quả phù hợp nhiều như vậy là bởi Google có cơ chế thu thập dữ liệu bằng cách sử dụng các phần mềm có tên “trình thu thập dữ liệu web” để khám phá các trang web công khai, lưu các tín hiệu này và lập chỉ mục (định từ khóa) cho mỗi từ trên trang web [Trung tâm tìm kiếm Google]. Do vậy, có thể hiểu đơn giản là cơ chế tìm kiếm thông tin bằng từ khóa trên google được thực hiện bằng cách quét từng từ của từ khóa với các dữ liệu được lưu trữ trong kho và đưa ra kết quả phù hợp. Tuy nhiên, thứ tự xếp hạng các kết quả tìm kiếm không phải ngẫu nhiên mà được sắp xếp theo các tiêu chí như: ưu tiên các website có chạy quảng cáo tìm kiếm (Google search) (trả phí quảng cáo cho Google), chỉ số ranking của website (số lượng người truy cập), cá nhân hóa kết quả tìm kiếm dựa trên địa điểm (nơi tài khoản cá nhân đăng nhập, ngôn ngữ và thiết bị của người dùng), tính cập nhật của thông tin hoặc phân tích ngữ nghĩa của từ khóa web [Trung tâm tìm kiếm Google].

Như vậy, có sự khác biệt trong cơ chế tìm kiếm thông tin bằng từ khóa giữa hệ thống TT-TV và các công cụ tìm kiếm trên internet. Trong hệ thống TT-TV, kết quả tìm kiếm dựa trên việc so sánh và trùng khớp hoàn toàn giữa từ khóa tìm kiếm với từ khóa được định sẵn cho tài liệu trong quá trình xử lý hoặc từ/cụm từ xuất hiện trong nhan đề tài liệu. Đối với việc tìm kiếm trên internet, kết quả tìm kiếm tuy cũng được căn cứ vào độ phù hợp giữa từ khóa (nhập vào ô tìm kiếm) với các từ/cụm từ có trong văn bản nhưng tùy vào từng website mà có sự ưu tiên thứ hạng hiện thị trong kết quả tìm kiếm dựa trên các yếu tố như: website có chi trả chi phí quảng cáo, chỉ số ranking của website, thông tin cá nhân của người dùng,...

5. PHÂN LOẠI VÀ YÊU CẦU ĐỐI VỚI TỪ KHÓA

5.1. Xét về thành phần từ vựng

Trong lĩnh vực TT-TV, có các từ khóa như danh từ chung, danh từ riêng, động từ được danh từ hóa, cụm danh từ kết hợp [Phan Huy Quế, 2001]. Tuy có nhiều loại từ khóa khác nhau, song các từ khóa trong lĩnh

vực TT-TV đều giống nhau ở điểm là phản ánh thông tin của tài liệu được xử lý. Trong lĩnh vực SEO (nguồn tin trên internet), có ba loại từ khóa: từ khóa cung cấp thông tin (Information keywords), từ khóa giao dịch (Transactional keyword) và từ khóa điều hướng (Navigational keywords). Cụ thể: từ khóa cung cấp thông tin được hiểu là những từ người dùng muốn tìm kiếm, từ khóa giao dịch là loại từ khóa được người sử dụng gõ vào ô tìm kiếm khi mua sản phẩm/dịch vụ; từ khóa điều hướng là loại từ khóa cụ thể giúp người sử dụng đến trang cụ thể nhưng do họ không nhớ đầy đủ link hay không muốn gõ ra. Trong ba loại từ khóa trên, từ khóa giao dịch được đánh giá là từ khóa được các doanh nghiệp yêu thích nhất, đầu tư xây dựng nhất nhưng cũng cạnh tranh nhất vì cơ hội đem lại lợi nhuận trực tiếp so với hai loại từ khóa còn lại.

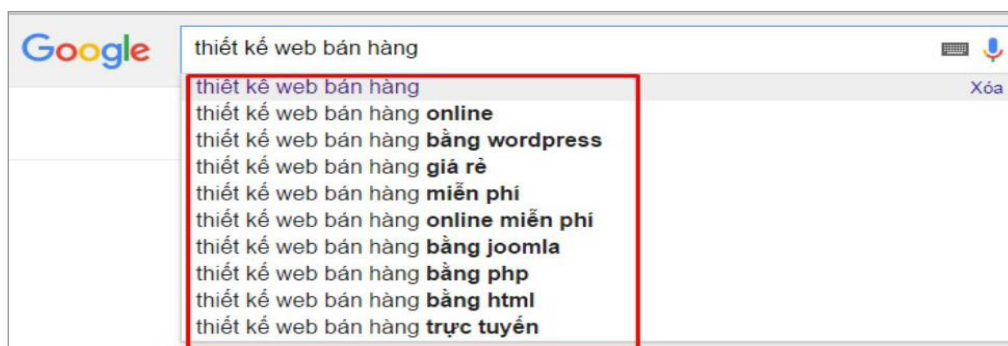
5.2. Xét về tỷ trọng của từ khóa trong việc thể hiện nội dung thông tin

Trong lĩnh vực TT-TV, có hai loại từ khóa: từ khóa chủ đề (từ khóa chính) và từ khóa phương diện (từ khóa phụ) [Phan Huy Quế, 2011]. Trong đó từ khóa chủ đề được hiểu là từ khóa phản ánh đối tượng nghiên cứu được đề cập trong tài liệu, còn từ khóa phương diện là từ khóa thể hiện một khía cạnh/nội dung của đối tượng nghiên cứu được đề cập trong tài liệu.

Ví dụ: tài liệu viết về thực trạng phát triển nền kinh tế Việt Nam sẽ có từ khóa chủ đề là kinh tế, từ khóa phương diện là thực trạng (chỉ nội dung) và Việt Nam (chỉ địa lý).

Việc phân loại từ khóa này trong hệ thống TT-TV có ý nghĩa trong việc giúp người sử dụng ưu tiên sử dụng các từ khóa chủ đề trong việc tìm kiếm cũng như kết hợp giữa từ khóa chủ đề và từ khóa phương diện để thu hẹp phép tìm.

Trong nguồn tin trên internet, cũng có sự phân chia theo tỷ trọng từ khóa bằng cách đặt từ khóa chính và từ khóa phụ. Cụ thể: từ khóa chính được xem là từ khóa ngắn nhất, thể hiện nội dung chính, từ khóa phụ là từ khóa dài hơn, mô tả sâu hơn và có nội dung liên quan tới từ khóa chính (Hình 12)

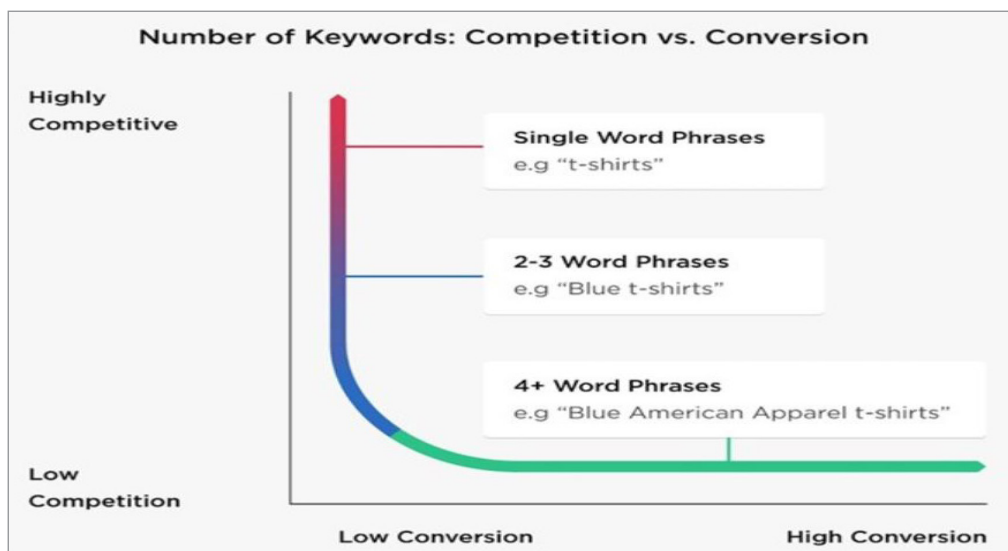


Hình 12. Từ khóa chính và từ khóa phụ trên internet

5.3. Về yêu cầu đối với từ khóa

Trong hệ thống TT-TV, các từ khóa được yêu cầu phải là những từ/ cụm từ ngắn gọn, súc tích và đơn nghĩa [Phan Huy Quế, 2001]. Trong khi đó, các nguồn tin trên internet lại chia từ khóa thành hai loại: từ khóa ngắn và từ khóa dài. Cụ thể, từ khóa ngắn (short-tail keywords) là các từ/cụm từ

ngắn (không quá 3 từ), từ khóa dài (long-tail keywords) là từ/ cụm từ dài (từ 3 từ trở lên) [Vicent Do, 2021]. Trong đó, nhiều ý kiến cho rằng từ khóa dài chiếm lợi thế hơn bởi nó giúp giảm tính cạnh tranh và tăng tính hiệu quả vì tìm đúng được khách hàng. Do vậy, nhiều doanh nghiệp hiện nay đang tập trung vào xây dựng từ khóa dài (Hình 13).



Hình 13. Sơ đồ mô tả tính cạnh tranh và khả năng chuyển đổi từ khóa trong SEO [Nguồn: Vicent Do, 2021]

Tuy nhiên, bên cạnh đó cũng có nhiều ý kiến cho rằng việc tạo lập các từ khóa dài không còn chiếm ưu thế bởi Google ngày càng cải thiện các thuật toán để nâng cao chất lượng tìm tin bằng từ khóa. Thuật toán Google Hummingbird hiện đã chú ý nhiều hơn đến ngữ nghĩa của từ khóa, phân biệt từ khóa chính - phụ cũng như các từ

khóa bổ sung chứ không phải đơn thuần là so sánh kỹ thuật các ký tự của từ khó [Nguyễn Công Anh, 2021]. Do đó, các doanh nghiệp cung cấp thông tin trên internet cũng cần chú trọng hơn nữa vào khâu xử lý nội dung, xây dựng các từ khóa chuyên sâu và tạo sự kết nối giữa các từ khóa (từ khóa đối tượng - từ khóa phương diện)

KẾT LUẬN

Tuy đều được sử dụng như là công cụ hữu hiệu trong việc tìm kiếm thông tin phù hợp, song vẫn có nhiều sự khác biệt trong quá trình tạo lập và sử dụng từ khóa giữa hệ thống TT-TV và các công cụ tra cứu trên internet. Căn nguyên của sự khác biệt này nằm ở mục đích của việc tạo lập từ khóa. Đề cao tính phi lợi nhuận, các hệ thống TT-TV đang từng bước nỗ lực xây dựng kho dữ liệu và cung cấp thông tin chất lượng, khách quan tới người sử dụng. Trong khi đó, mục đích sống còn của các doanh nghiệp khi cung cấp nguồn tin cũng như bộ máy tra cứu trên internet là lợi nhuận, do vậy việc sử dụng từ khóa được xem như một công cụ để quảng bá thương hiệu hay quảng cáo và là mục tiêu được ưu tiên hàng đầu, việc giúp người sử dụng tìm kiếm thông tin phù hợp chỉ là thứ yếu. Chính điều này đã dẫn đến sự khác biệt trong cách tạo lập cũng như sử dụng từ khóa giữa hai hệ thống thông tin. Tuy nhiên, cũng có thể nhận thấy rằng dù Google đang được đông đảo người sử dụng nhưng cũng đang từng bước cải thiện chất lượng tìm kiếm thông tin bằng cách chú trọng xây dựng các thuật toán phân tích ngữ nghĩa của từ khóa cũng như thiết lập các chỉ số để xếp hạng kết quả tìm kiếm. Điều này đặt ra bài toán đối với các hệ thống TT-TV trong việc không ngừng tìm hiểu nhu cầu cũng như khó khăn của người sử dụng khi tìm kiếm thông tin trên hệ thống, từ đó cải thiện tính năng của hệ thống nhằm nâng cao trải nghiệm của người sử dụng.

TÀI LIỆU THAM KHẢO

1. ALA (1996). ALA: từ điển giải nghĩa thư viện học và tin học Anh - Việt: Glossary of Library and Information Science. Phạm Thị Lệ Hương, Lâm Thế Vinh, Nguyễn Thị Nga dịch.- Tucson : Galen press, 1996.
2. ATPMedia (2020). Từ khóa chính và từ khóa phụ là gì? Truy cập 21/6/2022 tại: <https://atpseo.com/blog/tu-khoa-chinh-va-tu-khoa-phu-la-gi/>
3. Joshua Hardwich (2020). What are keywords? How to use them for SEO? Truy cập 21/6/2022 tại: <https://ahrefs.com/blog/what-are-keywords/>
4. Google Ads. Giới thiệu về từ khóa trong các chiến dịch trên mạng tìm kiếm. Truy cập 05/7/2022 tại <https://support.google.com/google-ads/answer/1704371>.
5. Google Ads trợ giúp. Các hoạt động của chiến dịch thông minh. Truy cập 05/7/2022 tại: https://support.google.com/google-ads/answer/7652860?hl=vi&ref_topic=9024773
6. Google tìm kiếm trợ giúp. Thu hẹp kết quả tìm kiếm trên web. Truy cập 20/6/2022 tại: <https://support.google.com/websearch/answer/2466433>
7. Lê Đức Tuấn (2022). Từ khóa là gì? Từ khóa trong SEO quan trọng như thế nào? Truy cập 05/7/2022 tại: <https://dichvuseohot.com/tu-khoa-la-gi/>
8. Lê Nam (2020). Từ khóa SEO: 8 loại từ khóa bạn cần biết. Truy cập 05/7/2022 tại: <https://vietmoz.edu.vn/tu-khoa-la-gi/>
9. Nguyễn Công Anh (2021). Google Hummingbird là gì? Tầm quan trọng và ảnh hưởng đối với SEO. Truy cập 21/6/2022 tại: <https://vietmoz.edu.vn/hummingbird-la-gi/>
10. Nguyễn Ngọc Hưng (2021). Phân tích 6 loại quảng cáo trên Google hiện nay. Truy cập 05/7/2022 tại: <https://seongon.com/blog/google-ads/cac-loai-quang-cao-google.html>
11. Phan Huy Quế (2001). Mô tả nội dung tài liệu bằng từ khóa : Tài liệu hướng dẫn.-H.: Trung tâm TTKH&CNQG
12. Tiêu chuẩn Quốc gia (2009). TCVN 5354:2009 (ISO 5127:2001) Thông tin và Tư liệu - Từ vựng, 2009.
13. Tiên Kiều (2022). Từ khóa là gì? Cách xác định và ví dụ cụ thể. Truy cập 05/7/2022 tại: <https://atp.vn/tu-khoa-la-gi/>
14. Từ khóa là gì? Cách nghiên cứu từ khóa hiệu quả. Truy cập 05/7/2022 tại: <https://adpia.vn/tu-khoa-la-gi-cach-nghien-cuu-tu-khoa-hieu-qua-t2411.html>
15. Từ khóa là gì? Vai trò quan trọng của từ khóa trong SEO website. Truy cập 05/7/2022 tại: <https://mona.media/tu-khoa-la-gi/>
16. Từ khóa là gì? Cách tối ưu từ khóa khi SEO website. Truy cập 05/7/2022 tại: <https://carly.com.vn/blog/tu-khoa-la-gi/#>
17. Trung tâm tìm kiếm Google. Hướng dẫn chuyên sâu về cách thức hoạt động của Google tìm kiếm. Truy cập 20/6/2022 tại: <https://developers.google.com/search/docs/advanced/guidelines/how-search-works?hl=vi>
18. Vicent Do (2021). Từ khóa SEO là gì? Hướng dẫn tối ưu SEO website 2022. Truy cập 21/6/2022 tại: <https://gtvseo.com/tu-khoa-seo/>
19. Thư viện Quốc gia Việt Nam. Truy cập tại: <http://opac.nlv.gov.vn/pages/opac/wp-id-home.html>
20. <https://www.google.com/>

(Ngày Tòa soạn nhận được bài: 10-5-2022; Ngày phản biện đánh giá: 15-6-2022; Ngày chấp nhận đăng: 15-7-2022).