

THƯ VIỆN SỐ XÃ HỘI

ThS Nguyễn Lê Phương Hoài
Viện Thông tin Khoa học Xã hội

Tóm tắt: Bài viết trình bày về Thư viện số hiện đại, Xã hội thông tin, Thư viện số xã hội, Mạng xã hội của các thư viện số xã hội. Đồng thời mô tả giải pháp của một số nhiệm vụ sử dụng bộ lý thuyết tập mờ và phương pháp đo lường thư mục cho kiểm soát quản lý thư viện bằng cách phân tích các mối quan hệ giữa người sử dụng và các nguồn tài nguyên của thư viện số.

Từ khóa: Thư viện số; xã hội thông tin; thư viện số xã hội; nguồn lực thông tin; người dùng tin.

Social Digital Library

Abstract: This paper provides overview on modern digital library, information society, social digital library, social network of social digital libraries. It also introduces the application of the fuzzy set theory and bibliometric methods for library management control by analyzing the relationships between users and digital library resources.

Keywords: digital library; information society; digital library; information resources; information user.

1. Thư viện số và xã hội thông tin

Hiện nay, thuật ngữ *Thư viện số* được sử dụng rộng rãi và thường xuyên. Dựa trên các cách tiếp cận khác nhau, có nhiều quan niệm khác nhau về Thư viện số. Thư viện số bao gồm một loạt các hệ thống làm việc và nguyên mẫu nghiên cứu, các bộ sưu tập các tài nguyên thông tin, và công nghệ [Ioannidis et al, 2005]. Đối với người dùng tin, thư viện số là một cơ sở dữ liệu lớn. Đối với người làm thông tin, thư viện số là một ứng dụng của Web. Theo quan điểm của ngành khoa học thư viện, thư viện số là bước tiếp theo trong tự động hóa thư viện... Hiệp hội Thư viện số Hoa Kỳ (Digital Library Federation) đã định nghĩa: “*Thư viện số là các tổ chức cung cấp tài nguyên, gồm các nhân*

viên chuyên biệt giúp lựa chọn, tổ chức, cung cấp khả năng truy xuất thông minh, chỉ dẫn, phân phối, bảo quản tính toàn vẹn và sự thống nhất của các bộ sưu tập theo thời gian để đảm bảo sao cho chúng luôn sẵn có để truy xuất một cách dễ dàng và kinh tế nhất đối với một cộng đồng người dùng hoặc một nhóm cộng đồng người dùng” [theo Water, D.J, 1998].

Xã hội thông tin là một xã hội dựa trên những ứng dụng phổ biến của công nghệ thông tin và truyền thông mới, tức là công nghệ máy tính điện tử và kỹ thuật số, mà đỉnh cao là công nghệ Internet [Nguyễn Thị Đông]. Trong xã hội thông tin, thông tin được sử dụng như một nguồn lực kinh tế, việc phát triển một ngành thông tin ở ngay trong nền kinh tế. Số

người sử dụng thông tin ngày càng tăng với các mục đích khác nhau. Nguồn thông tin trong các ngành nghề, lĩnh vực công tác của người dùng tin để đem lại hiệu quả lớn hơn, đòi hỏi các cơ quan thông tin phải tạo ra được các sản phẩm thông tin đa dạng, thiết thực hơn. Từ đó hình thành và phát triển một ngành công nghiệp sản sinh ra chính thông tin và các dòng tin được truyền tải trên các mạng: đó là các nhà cung cấp nội dung thông tin.

Việc ứng dụng các tính năng của công nghệ truyền thông, Internet, khái niệm mạng, thiết bị di động... đã làm thay đổi việc tra cứu, tiếp cận tài nguyên thông tin của người dùng tin theo hướng cởi mở, tiêu chuẩn, quy trình bản quyền, hệ thống quản lý thông tin... Các thư viện số hiện đại quan sát hành vi của người dùng tin để xem xét các phương pháp tiếp cận khác nhau trong nhu cầu tin của người dùng tin. Thư viện số hiện đại, không chỉ là một hệ thống thông tin, còn bao gồm việc chia sẻ, trao đổi, hợp tác...

Công nghệ hiện đại, có khả năng tạo ra liên kết giữa các nguồn thông tin, hệ thống, các nhóm, các cá nhân, đem lại hiệu quả trao đổi thông tin trong xã hội thông tin hiện đại. Các thư viện số hiện đại trên thế giới sử dụng các công cụ truyền thông xã hội như blog, Twitter, Facebook và Youtube cho phép cán bộ thư viện tham gia vào các cuộc hội thoại trực tuyến, các hoạt động chia sẻ thông tin, quảng bá và khuyến khích người dùng tin tra cứu, sử dụng các bộ sưu tập trực tuyến của thư viện. Các thư viện số hiện đại tập trung vào các khía cạnh tiềm năng chính của cấu trúc xã hội thông tin mới bao gồm: *Yếu tố con người* (một bên là tài nguyên trí tuệ và một bên là người tiêu dùng); *Nội dung* - sản phẩm cuối cùng của trí tuệ và môi trường để đáp ứng nhu cầu thông tin của các cá nhân hoặc nhóm người tiêu dùng;

Công nghệ thông tin và truyền thông (ICT) và phương pháp tiếp cận hệ tư tưởng hiện đại. Việc ứng dụng công nghệ thông tin và truyền thông vào hệ thống thông tin-thư viện giúp môi trường này mang tính công nghệ cao, tính toàn cầu, kết nối mạng mà không bị hạn chế về khoảng cách địa lý, phân biệt chủng tộc, tuổi tác và thân thiện hơn. Phương pháp tiếp cận hệ tư tưởng hiện đại như Quản lý hệ khách hàng (CRM) tạo ra những cách thức mới thiết lập cho môi trường thông tin - thư viện.

2. Thư viện số xã hội

Bước sang thế kỷ 21, công nghệ thông tin và truyền thông và các công cụ xã hội trực tuyến được khám phá. Nhưng từ cuối thế kỷ 19, khái niệm “mạng xã hội”, “nhóm xã hội” đã được ESmile Durkheim và Ferdinand Tunnies nghiên cứu [theo Wich Peter, 2014]. Mạng xã hội tạo ra một hệ thống kết nối các thành viên cùng sở thích với nhiều mục đích khác nhau, không phân biệt không gian và thời gian với những tính năng, như: kết bạn, chat, e-mail, phim ảnh, voice chat, chia sẻ file, blog và xã luận dựa trên nền Internet. Ngày nay, các khái niệm mạng xã hội (công cụ truyền thông, mạng và các vấn đề liên quan) là một phần của công nghệ phát triển nhanh và hiện đại, liên quan với hai khái niệm khác là công nghệ “Đám mây” và “Web 2.0”.

Công nghệ “Đám mây” (hay điện toán đám mây) được hiểu một cách tổng quát là phương thức lưu trữ dữ liệu khổng lồ trên máy chủ ảo. Tài nguyên, dữ liệu, phần mềm và các thông tin liên quan đều được chứa trên các server (chính là các “đám mây”). Mọi khả năng liên quan đến công nghệ thông tin đều được cung cấp dưới dạng các “dịch vụ”, cho phép người sử dụng truy cập các dịch vụ công nghệ từ một nhà cung cấp nào đó “trong đám mây” mà không cần phải có các kiến thức, kinh nghiệm

về công nghệ đó, cũng như cơ sở hạ tầng phục vụ công nghệ đó. Ứng dụng đám mây chính là những ứng dụng trực tuyến trên Internet. Trình duyệt là nơi ứng dụng hiện hữu và vận hành còn dữ liệu được lưu trữ và xử lý ở máy chủ của nhà cung cấp ứng dụng.

Web 2.0 là một cách tiếp cận mới để sử dụng web như là một nền tảng nơi mà người dùng cùng nhau tham gia vào việc tạo ra, chỉnh sửa và xuất bản thông tin thông qua những công cụ hợp tác sáng tạo nội dung trên nền web. Web 2.0 liên quan tới người dùng không chỉ ở chỗ người dùng tạo ra nội dung mà người dùng giúp thu thập, tổ chức, mô tả, cập nhật, chia sẻ, truyền bá, sắp xếp lại, bình luận, hiệu đính, và đóng gói nội dung [Nguyễn Công Hà, 2011]. Web 2.0 tạo cơ hội cho người dùng không còn là người tiếp nhận thông tin thụ động mà là người tham gia tạo nên nội dung của thông tin.

Bài viết xem xét khái niệm thư viện số xã hội theo ý tưởng thư viện số thế hệ kế tiếp. Các tính năng mạng xã hội đưa ra các tùy chọn cho các thư viện số như cá nhân hóa, đánh giá và phân loại người dùng tin, điều hướng, gợi ý các nguồn tài nguyên...

2.1. Cấu trúc lý thuyết của thư viện số xã hội

Tài nguyên thông tin trên mạng Internet có thể được khai thác dễ dàng, nhưng người dùng tin dễ bị lạc lối, đặc biệt là những người dùng tin mới tham gia vào thế giới nghiên cứu. Tất cả các hoạt động học thuật và các quy trình liên quan đều bắt đầu từ thư viện và trong cấu trúc và kiến thức cổ điển, có rất nhiều tài liệu tham khảo, hướng dẫn thư mục, mục lục... sẽ điều hướng người dùng tin đến các thông tin cần thiết. Ý tưởng và triết lý này xuất phát từ hoạt động quản lý thư viện truyền thống, nhưng mục tiêu của mô hình hiện đại này là

tổ chức thư viện số đi kèm với các công cụ truyền thông, quản lý và quy định thông tin, cơ sở hạ tầng kỹ thuật, hạ tầng công nghệ thông tin và các kỹ năng nghiên cứu, tìm kiếm. Nói cách khác, thông qua thư viện số xã hội, mạng xã hội có thể giúp người cán bộ thư viện và người sử dụng không chỉ tương tác, mà còn xử lý, tổ chức, hệ thống hóa, thêm ngữ nghĩa liên quan và chia sẻ nguồn lực thông tin trong môi trường điện tử [Worrall, 2014].

Cấu trúc lý thuyết của thư viện số xã hội, được xây dựng trên khoa học thông tin-thư viện và các tiêu chuẩn mạng xã hội, gồm 4 điểm truy cập chính và cơ sở dữ liệu có liên hệ với nhau là: Tác giả, Tài nguyên, Người dùng tin và Đề mục chủ đề.

Tác giả: Điểm truy cập này trong giai đoạn đầu tiên mở ra mạng lưới những người tạo ra các nguồn lực bên trong hoặc bên ngoài Thư viện số xã hội. Mỗi tác giả có thông tin cá nhân và mở được đăng ký dữ liệu mà về mặt ngữ nghĩa là được kết nối với các tác giả liên quan. Tập hợp dữ liệu chính có thể là: thông tin cá nhân, thông tin nghề nghiệp, kỹ năng liên quan... Bộ dữ liệu này có thể bao gồm thông tin cá nhân, gia đình hoặc tập đoàn...

Tài nguyên (sách điện tử, tài liệu âm thanh - hình ảnh, và hình ảnh): Điểm truy cập tính đến tất cả các nguồn thông tin được thu thập trong thư viện số xã hội. Những tài nguyên này trước hết được hệ thống hóa theo loại hình và đặc điểm như: trực tuyến, ngoại tuyến, văn bản, âm thanh-hình ảnh, hình ảnh... Sau đó được tạo siêu dữ liệu cá nhân cho mỗi tài nguyên với các mục chính và bổ sung, thông tin thư mục và xác nhận có thể theo các chuẩn khác nhau như: MARC21, MARCXML, Dublin Core, FRBR, FRAD,... Việc tạo ra các dữ liệu có quan hệ ngữ nghĩa giữa các tài liệu giúp tích hợp các "tài nguyên" và ba thực thể

điểm truy cập khác. Bộ sưu tập siêu dữ liệu này sẽ là phần mở rộng của các thư mục thư viện thể hệ mới. Các tính năng chính là: giao diện trực quan; tăng cường khả năng tìm kiếm bao gồm cả các thể thông tin do người dùng tạo ra, đánh giá và bình luận; kết quả xếp theo mức độ liên quan một cách trí tuệ; hỗ trợ tìm kiếm, nội dung do người dùng tạo ra, công cụ xã hội đặc biệt trong các hạng mục như gắn thẻ, đánh dấu trang, đánh giá, bình luận và khả năng liên quan khác. Ngoài ra dữ liệu thư mục sẽ có thêm thông tin về bài đọc thêm...

Người dùng tin: Điểm truy cập cho sự kết nối bước đầu tiên với những người dùng khác. Đây là một mức độ khởi đầu mà mỗi người dùng riêng lẻ là một phần của bất kỳ nhóm nào trong hệ thống. Cấu trúc tương tự như cơ sở dữ liệu “Tác giả” nhưng khác nhau về chức năng và mối quan hệ. Cơ sở dữ liệu “Người dùng tin” và cấu trúc điểm truy cập sẽ được xây dựng trên mô hình mạng xã hội cổ điển và có cùng chức năng và các chức năng bổ sung. Các tính năng chính là những tài khoản người dùng có thể tùy biến; cơ hội chia sẻ đăng ký, danh mục dữ liệu: bản ghi thư mục, bản ghi đăng ký cá nhân với các mạng xã hội khác; khả năng tương thích với các thiết bị di động làm cho các mạng xã hội khác; khả năng tương thích với các thiết bị di động làm cho các hoạt động xã hội này rộng hơn trong phạm vi của xã hội; khả năng thực hiện nghiên cứu không chỉ trong một cơ sở dữ liệu, mà còn để kết nối với các trang web thư viện, cơ sở dữ liệu và các nguồn khác nhau có liên quan.

Để mục chủ đề: Điểm truy cập đề mục chủ đề và cơ sở dữ liệu là sự phân loại kiến thức và trong tình huống hiện tại liên kết với các cơ sở dữ liệu khác, hệ thống hóa và tổ chức giá trị thông tin lớn giữa các cơ sở dữ liệu “Tác giả”, “Người dùng tin” và “Tài nguyên”.

2.2. Mạng xã hội của các thư viện số xã hội

Việc tổ chức bốn cơ sở dữ liệu chính và dữ liệu liên kết giữa các thông tin này sẽ tạo thành mạng xã hội (social network). Điểm mấu chốt của Thư viện số xã hội bao gồm mối quan hệ giữa bốn cơ sở dữ liệu kể trên. Mỗi dữ liệu đăng ký có liên kết ngữ nghĩa với thông tin cơ sở dữ liệu khác, ví dụ nếu người dùng tin nghiên cứu bất kỳ một chuyên khảo cụ thể nào, họ sẽ tiếp cận thông tin tác giả, các bình luận, đánh giá, đánh dấu được thực hiện trong tài liệu hiện hành, và các thông tin về nhóm người dùng tin đang quan tâm tới tài liệu này, theo đó họ sẽ truy cập vào cơ sở dữ liệu đầy đủ của các chủ đề liên quan. Tiếp theo, người dùng tin có thể truy cập sâu vào mạng xã hội của Thư viện số xã hội. Công cụ mạng xã hội và cơ chế hoạt động tuân theo bốn cơ sở dữ liệu chính.

Liên kết giữa các nguồn thông tin: hệ thống hóa, phân loại và liên kết các nguồn thông tin cần có trong Thư viện số xã hội. Các tiêu chí cho quy trình này là nguồn thông tin, cách sử dụng, đối tượng và các vấn đề liên quan khác.

Liên kết giữ người dùng tin: thư viện viên có thể sử dụng công cụ này để kết nối với người dùng tin, hoặc người dùng tin có thể tự mình kết bạn, tổ chức các cuộc thảo luận trực tuyến và chia sẻ tài nguyên. Các cá nhân có thể tạo blog để phổ biến thông tin, trò chuyện trực tuyến hoặc dịch vụ tham khảo ảo bởi các chuyên gia thông tin-thư viện. Mỗi người dùng tin có thể tạo ra không gian điện toán đám mây cá nhân (thư viện ảo), các lớp học, và hội thảo, khuyến khích các nhóm người dùng tin liên quan và đang hoạt động chia sẻ các bài thuyết trình và ý tưởng của mình cho các cộng đồng lớn hơn.

Dữ liệu được hệ thống hóa, kết nối và kiểm soát về người sáng tạo (các tác giả, biên tập viên...) là dữ liệu đăng ký được kết nối riêng biệt về cộng đồng trí thức, các tác giả khác

nhau có thể lập các nhóm lợi ích đặc biệt, hội thảo khoa học và các hoạt động có giá trị khác.

Dữ liệu được hệ thống hóa, kết nối và kiểm soát về đề mục chủ đề liên quan với các nguồn thông tin gọi là các bộ từ điển chuyên đề, từ vựng được kiểm soát mà chức năng chính là thu thập, tổ chức và phổ biến các tài liệu. Các công cụ đề mục chủ đề phổ biến có thể là Đề mục chủ đề Thư viện Quốc hội (LCSH), đề mục chủ đề Canada, Đề mục chủ đề y học-MeSH hoặc các đề mục chủ đề khác.

Liên kết giữa các nguồn thông tin và nguồn lực con người: phân phối thông tin là cách tốt nhất để marketing tài liệu mới, thư viện số xã hội có thể chia sẻ sách điện tử, hình ảnh mới... Sau các quy trình kỹ thuật nội bộ người dùng có thể giúp thu thập thông tin còn thiếu về tài liệu được xử lý và các tài liệu liên quan khác.

3. Phân tích các mối liên hệ giữa người dùng tin và các nguồn lực trong các thư viện số

Một ý tưởng khác cho các cộng đồng thư viện số xã hội hiện đại là kết nối người dùng tin có các lợi ích tương tự hoặc các nguồn lực có các đặc tính tương tự với nhau để phục vụ công tác quản lý thư viện số [Bramoullé, Djebbari, Fortin, 2009]. Phân tích mối liên hệ giữa người dùng tin của thư viện hoặc các nguồn lực sử dụng các chỉ số khác nhau giúp thư viện xác định các nguồn lực phù hợp nhất hoặc những người sử dụng tích cực nhất. Thông tin phân tích về mối liên hệ này rất cần thiết trong quản lý thư viện.

Để phân loại các nguồn lực được đưa ra và để đo lường mức độ phù hợp của các nguồn lực, các nhà thư viện học trên thế giới sử dụng mô hình tập mờ. Kể từ khi khởi xướng vào năm 1965, lý thuyết tập mờ đã phát triển theo rất nhiều cách thức và nhiều nguyên tắc. Sau năm 1992, lý thuyết tập mờ, lý thuyết về mạng trung lập và khu vực lập trình tiến hóa đã trở nên

nổi tiếng với tên gọi “trí thông minh điện toán” hoặc “điện toán mềm” [Zimmerman, 2010]. Lý thuyết tập mờ đã được ứng dụng trong đo lường thư mục và đánh giá nghiên cứu, ví dụ như phương pháp xếp loại mờ, các mối quan hệ giống nhau mờ để xếp hạng tạp chí, phương pháp tập hợp nhóm mờ để xác định thuật ngữ tự động trong việc lập bản đồ đo lường thư mục.

Trong bài viết *Khai thác mạng xã hội trong môi trường thư viện số hiện đại* [1], các tác giả đề xuất sử dụng tập mờ để phân loại các nguồn lực được đưa ra và giải quyết nhằm định nghĩa về mức độ phù hợp của các nguồn lực. Gọi B là tập hợp tài nguyên trong thư viện. Thư viện có N tài nguyên (sách). Ta có công thức và các đặc điểm như sau:

$$B = \{b_1, b_2 \dots b_N\}$$

Gọi số lượng người dùng đối với tài nguyên thứ k là u_k

Gọi số truy vấn cho tài nguyên thứ k là q_k

Xác định thuật ngữ sử dụng cho u_k là: “vài”, “nhiều”, “rất nhiều”.

Thuật ngữ sử dụng cho q_k là: “một ít”, “vài”, “nhiều” và “rất nhiều”.

Thuật ngữ chỉ sự phù hợp của các tài nguyên là: “không phù hợp”, “phù hợp ít”, “phù hợp”, “phù hợp nhiều”.

Quy tắc cơ bản xây dựng như sau:

Nếu u_k = “vài” và q_k = “một ít” thì b_k = “không phù hợp”

Nếu u_k = “vài” và q_k = “vài” thì b_k = “phù hợp ít”

Nếu u_k = “nhiều” và q_k = “nhiều” thì b_k = “phù hợp”

Nếu u_k = “rất nhiều” và q_k = “rất nhiều” thì b_k = “phù hợp hơn”.

Ví dụ: để cho u_1 - là việc sử dụng nguồn lực đầu tiên bằng 32 và q_1 - số lượng truy vấn

bằng 56, trong trường hợp này mức độ phù hợp của nguồn lực thứ nhất sẽ bằng 0.88, Kết quả nhận được cho thấy sự phù hợp cao của nguồn lực thứ nhất.

Như vậy, thư viện có thể xác định mức độ hoạt động của người dùng và xác định các quá trình hành động tiếp theo để tăng cường hoạt động của họ bằng cách phân tích khu vực quan tâm hoặc cách sống của người dùng. Các biến số đầu vào để định nghĩa về hoạt động của người dùng có thể được xác định bằng nhiều cách khác nhau như tính toán số giờ truy cập trên trang nhất định, số lượng nguồn lực được sử dụng, số lượng câu hỏi... Đồng thời, thư viện cũng có thể gộp các nguồn lực theo các nguyên tắc phân tích chuyên ngành của người dùng hoặc các nguồn lực khác khi được sử dụng cùng nhau.

Kết luận

Các công nghệ web, kỹ thuật số đưa ra phương pháp mới để tổ chức các nguồn lực thông tin; tạo lập xã hội số, mạng lưới; dữ liệu cá nhân...; quan hệ ngữ nghĩa, đo lường thư mục, các phương pháp phân tích trí tuệ, các đặc điểm và chức năng. Thư viện số xã hội là mô hình phù hợp hơn với một xã hội kết nối mạng hiện nay và tiếp tục đóng vai trò là cầu nối kiến thức cho người dùng tin thư viện. Phân tích các cơ sở dữ liệu cốt lõi trong môi trường thư viện số xã hội - cơ sở dữ liệu của người sáng tạo và người dùng tin, cơ sở dữ liệu các nguồn lực. Sau đó, tìm hiểu hành động của cơ sở dữ liệu này theo bề ngoài của mối quan hệ liên kết giữa người dùng, giữa các nguồn lực, giữa người dùng tin với các nguồn lực với nhau để cho ra nhiều cơ hội nhóm các nguồn lực lại hơn theo các nguyên tắc của chúng, sự phù hợp, và tập hợp nhóm những người dùng tin theo hoạt động, mối quan tâm, chuyên ngành và sở thích và đặc điểm của họ. Thông tin nhận được là kết quả của sự phân tích này rất quan trọng cho việc quản lý thư viện số, thông

tin được hệ thống hóa và tập hợp nhóm, chính sách và quy trình hiện đại.

TÀI LIỆU THAM KHẢO

1. Alguliyev R., Jafarov J., Mammadov E., Ismayilova N., Mammadova R. *Extraction of social networks in modern digital library environment*, Economics and Sociology, Vol.8, Số 1, 2015, tr. 308-317.

2. Bramoullé, Djebbari, Fortin (2009), Identification of peer effects through social networks, *Journal of Econometrics*, 50: 41-50.

3. Ioannidis et al (2005), Digital Library Information-Technology Infrastructures, *International Journal on Digital Libraries* 5(4): 1462-1480.

4. Nguyễn Công Hà (2011), *Web 2.0 và Thư viện cộng đồng*, Bản tin Thư viện - Công nghệ thông tin, tháng 11.

5. Nguyễn Thị Đông, *Nhân lực thông tin - thư viện trong xã hội thông tin và xã hội tri thức*

<http://thuviennesoi.vu.dreamlib.vn:8080/phamquangquyen/bitstream/123456789/681/1/3.pdf>

6. Waters, D. J. (1998), *What are digital libraries?* CLIR Issues, <https://www.clir.org/pubs/issues/issues04.html>.

7. Winch, Peter (2014), *Community and Community Interventions*, Health Behavior Change at the Individual, Household and Community Levels 224.689. US, Baltimor.

8. Worrall (2014, Social Digital Libraries: Their Roles Within and Across Social Worlds, Information Worlds, and Communities. *Bulletin of the IEEE Technical Committee on Digital Libraries* 9 (2).

9. Zimmerman (2010), Fuzzy Set Theory, *Wiley interdisciplinary Reviews: Computational Statistics* 2, tr.317-332.

(Ngày Tòa soạn nhận được bài: 18-4-2017; Ngày phản biện đánh giá: 5-7-2017; Ngày chấp nhận đăng: 20-8-2017).