

THƯ VIỆN SỐ VÀ VĂN HÓA XÃ HỘI

TS. Đỗ Quang Vinh

Khoa Công nghệ Thông tin, Trường Đại học Văn hóa Hà Nội

Tóm tắt: Thư viện số là một trong những hướng nghiên cứu chính về ứng dụng công nghệ thông tin và truyền thông trong lĩnh vực thông tin – thư viện hiện nay. Bài viết trình bày khái niệm thư viện số và một số vấn đề liên quan đến thư viện số như vấn đề kinh tế, xã hội, luật pháp...

Từ khóa: Thư viện số, Văn hóa xã hội

1. Mở đầu

Ngày nay, cách tìm kiếm và thu thập thông tin thay đổi đáng kể. Người dùng tin không cần phải ra khỏi nhà mà vẫn có thể truy cập được lượng thông tin lớn sẵn có trực tuyến, được cung cấp bởi nhiều nhà cung cấp thông tin như: các thư viện số (DL – Digital Library), NXB điện tử, doanh nghiệp, tổ chức, cá nhân. Việc thu thập thông tin không còn bị giới hạn trong các cuốn sách hay tạp chí sẵn có ở thư viện gần nhất, mà có thể từ các kho chứa thông tin phân tán toàn cầu.

Thông tin không chỉ là văn bản và dữ liệu số mà còn bao gồm các dạng khác như hình ảnh, âm thanh/ tiếng nói, video, audio, multimedia. Điều này làm tăng tính đa dạng của thông tin, cho phép người sử dụng thực hiện các cuộc du lịch ảo qua các viện bảo tàng, các di tích lịch sử và các kỳ quan thiên nhiên, tham dự những buổi hòa nhạc và biểu diễn sân khấu ảo, xem phim, nghe giảng, đọc sách báo... thông qua thư viện số.

2. Định nghĩa Thư viện số

Dưới đây là một số định nghĩa về thư viện số:

Định nghĩa 1: Theo Arms W.Y. (2003), thư viện số là một kho thông tin có tổ chức với các dịch vụ liên kết, trong đó thông tin được lưu trữ ở dạng số và có thể truy cập qua một mạng.

Ý chính của định nghĩa này là thông tin có tổ chức. DL chứa các loại kho thông tin khác nhau được dùng bởi nhiều NSD khác nhau. DL có quy mô từ nhỏ đến rất lớn. DL có thể sử dụng bất kỳ loại thiết bị tính toán nào và bất kỳ phần mềm phù hợp. Chủ đề thống nhất là thông tin được tổ chức trên máy tính và có sẵn trên mạng với các thủ tục lựa chọn tài liệu trong các kho để tổ chức, làm cho sẵn có với người sử dụng (NSD) và lưu trữ.

Định nghĩa 2: Reddy R., Wladawsky-Berger I. (2001) cho rằng, thư viện số là các kho dữ liệu mạng về tài liệu văn bản số, ảnh, âm thanh, dữ liệu khoa học và phần mềm là lõi của Internet hiện nay và các kho dữ liệu số có thể truy cập phổ biến về tất cả tri thức của loài người trong tương lai.

Định nghĩa 3: Liên đoàn Thư viện số - The Digital Library Federation đã định nghĩa thư viện số là những tổ chức cung cấp các nguồn lực gồm cả cán bộ chuyên môn để lựa chọn, xây dựng, truy cập tri thức, giải thích, phân phát, bảo tồn tính toàn vẹn và đảm bảo

tính bền vững vượt thời gian của các kho tài liệu số, do đó chúng luôn sẵn sàng đáp ứng nhu cầu sử dụng của một cộng đồng cụ thể hoặc của một nhóm cộng đồng.

Tóm lại, thư viện số là một kho thông tin số khổng lồ có tổ chức với các dịch vụ liên kết qua mạng.

3. Các vấn đề văn hóa xã hội của Thư viện số

Thư viện số cho phép con người tương tác với nhau và thông tin theo các cách mới và đem lại nhiều lợi ích cho người sử dụng. Tuy nhiên, việc phát triển thư viện số cũng gặp một số vấn đề về kinh tế, xã hội và luật pháp.

3.1 Kinh tế

Khi thông tin điện tử được sản xuất, lưu trữ, sửa đổi và phân phối rẻ, xác định giá thị trường và chi phí cho nó là khó hơn nhiều so với cho một vật thể vật lý. Cho đến nay, không có một mô hình kinh tế đạt được sự chấp nhận chung có thể xác định chính xác và rành mạch giá trị và chi phí đối với các dịch vụ thư viện số.

Hiện tại, ít nhất có hai mô hình định giá cơ bản cho thông tin trên Internet:

- Cho phép truy cập tự do nhưng tính giá cho nội dung, nghĩa là, truy cập tự do bằng chỉ dẫn và bằng nội dung, nhưng tính giá cho bất kỳ thứ nào khác,
- Tính giá cho truy cập nhưng cho phép nghiên cứu và sử dụng tự do nội dung.

Hai mô hình này là không loại trừ lẫn nhau và cùng tồn tại trên Internet.

Một vài mô hình tài trợ thư viện số khác nhau được đề xuất, nhưng các mô hình cơ bản là dựa vào thời gian, dựa vào yêu cầu. Một số mô hình được đề xuất bao gồm:

- Tiền bao cấp cơ quan (chung và riêng) - mô hình hiện tại đối với hầu hết các thư viện số;
- Các dịch vụ chung “miễn phí” và tính giá đối với các dịch vụ không thông thường, đặc biệt đòi hỏi sự can thiệp của con người;
- Tính giá cho mọi thứ, bao gồm phí cho thời gian kết nối, cách sử dụng CPU, phí cho mỗi lần tìm kiếm, phí cho mỗi lần truy tìm và phục hồi và phí tải xuống;
- Bao cấp các dịch vụ thông qua quảng cáo, điển hình là các tạp chí, tivi và Web;
- Các cơ chế bao cấp khác, chẳng hạn, kêu gọi quyên góp công khai tương tự với truyền hình và phát thanh công cộng;
- Thuế hoặc các nguồn tài trợ công cộng khác;
- Tiền đóng trả trước cho một thời hạn/ tiền đăng ký;
- Các hội tương tự với câu lạc bộ bán hàng, trong đó các khách hàng riêng lẻ góp các tài nguyên của họ cho phép truy cập thông tin;
- Tính giá cho các tác giả một phí đơn vị đối với “quyền” có thông tin của họ và các dịch vụ truy cập được, sau đó tính giá cho những người dùng đối với giá truy cập thông tin tăng không đáng kể;
- Chi phí hợp lý: chi phí hợp lý được xác định bởi quan hệ giữa cung và cầu đối với một tài nguyên cho trước, sao cho chi phí hợp lý của một tài nguyên không dùng đến là gần bằng 0, nhưng của một tài nguyên sử dụng nhiều là cao đến mức là không thể có;

- Sử dụng một thuật toán tính giá từng byte chi tiết.

Các mô hình chi phí và công cụ tài chính hiện thời đã dùng trong sản xuất và sử dụng thông tin truyền thống không đáp ứng thỏa đáng yêu cầu của thư viện số. Các mô hình chi phí cố định là không nhạy cảm với những thay đổi ở nội dung và chi phí. Thông tin điện tử với tính đa dạng có những chi phí sản xuất và phân phối liên kết khác nhau. Các mô hình chi phí linh động và thích nghi được đòi hỏi để điều khiển tính đa dạng và tính phức tạp. Các mô hình kinh tế đối với thư viện số đòi hỏi một loạt giải thuật định giá và ước lượng chuyên dụng có thể xác định chi phí và giá thông tin hoặc dịch vụ và sửa đổi mô hình với nhiều yếu tố môi trường.

3.2 Luật pháp

Các vấn đề luật pháp liên quan thư viện số bao gồm:

- Vấn đề chủ sở hữu: khi một thư viện làm chủ một bản sao chép vật lý, những quyết định về thu thập và lưu trữ là tương đối dễ. Nếu một thư viện số chỉ làm chủ một liên kết tới thông tin, các loại bài toán chủ sở hữu nhất định tăng lên. Chẳng hạn, nếu một thư viện số quyết định hủy bỏ việc đặt mua dài hạn cho thông tin xuất bản bình thường, sự truy cập sẽ được kiểm soát như thế nào? Rõ ràng, sự truy cập tới các số phát hành tương lai sẽ không được cho phép nhưng quyền truy cập các số phát hành quá khứ hoàn toàn được thỏa thuận. Phải giữ gìn động loại thông tin này dẫn đến làm phức tạp các chính sách, thủ tục và quá trình điều khiển truy cập và giữ gìn bản ghi đối với thư viện số. Điều gì nên làm khi một nhà cung cấp thông tin “vượt khỏi kinh doanh” hoặc một mục thông tin “vượt khỏi in ấn”. Trong cả hai trường hợp nhà cung cấp thông tin có thể không còn đủ khả năng trợ giúp lưu trữ vật lý. Các quyền của chủ sở hữu về các liên kết tới thông tin đó được bảo vệ như thế nào?
- Truy cập không hợp pháp: thông tin điện tử dường như dễ bị tấn công theo hướng truy cập không hợp pháp, lấy trộm và gian lận hơn các bản sao chép vật lý vì những xâm nhập như thế khó phát hiện hơn. Nhiều kỹ thuật đang được nghiên cứu để trợ giúp bảo vệ thông tin điện tử, bao gồm tường lửa, chữ ký điện tử, mật mã, phần mềm và phần cứng biến đổi đặc biệt và các hình mờ điện tử.
- Trách nhiệm pháp lý: luật truyền thống phân biệt tác giả và nhà xuất bản là người có trách nhiệm đối với thông tin họ sản xuất và các nhà phân phối (bưu điện, thư viện và hiệu sách) là không có trách nhiệm. Thư viện số có thể phân phối cũng như sản xuất thông tin và làm nảy sinh các vấn đề về trách nhiệm của họ đối với thông tin đã xuất bản, trình bày hoặc phân phối từ các điểm của họ. Trong các trường hợp thông tin điện tử có nhiều tác giả và nhiều ấn bản, ý kiến chuyên môn có thể được xác định và gán trách nhiệm như thế nào?
- Sự vi phạm nhãn hiệu đăng ký: biểu tượng đăng ký thương mại, ví dụ như một con dấu trường đại học hoặc một bức tranh thương mại, có thể bị sao chép, hoặc quét, hoặc sử dụng như là giấy dán tường hoặc hình ảnh trong thông tin điện tử. Nhiều

tổ chức yêu cầu thông báo và/hoặc trả tiền cho việc sử dụng các nhãn hiệu đăng ký. Những quyền này được bảo vệ như thế nào?

- Bản quyền và quyền sở hữu trí tuệ: các vấn đề bản quyền, trong sự vi phạm bản quyền nói riêng và các vấn đề quyền sở hữu trí tuệ có liên quan, là các vấn đề luật cơ bản của thư viện số.

Hầu như bất kỳ thứ gì được giữ bản quyền cũng có thể được số hóa. Một lần số hóa, bất kỳ người nào với một máy tính có thể sao chép nó, sửa đổi nó và phân phối nó tới bất kỳ người khác có truy cập tới một mạng. Thông tin điện tử là dễ sao chép và phân phối lại, nhưng khó phân biệt một bản sao chép hợp pháp với một bản không hợp pháp. Những điều chỉnh hiện có, như là không tải xuống toàn bộ; không lưu trữ điện tử; không sao chép và phân phối, ngay cả nội bộ; không sao chép và phân phối tới thành phần thứ ba và những hạn chế sử dụng riêng biệt khác nhau, phần lớn bị phớt lờ. Các nhà cung cấp thông tin hoàn toàn chứng thực hành vi này bằng cách “nhìn hướng khác” trong nhiều trường hợp. Những trách nhiệm của thư viện số là gì trong việc tuân theo các luật bản quyền áp dụng cho thông tin và các dịch vụ nó cung cấp?

Các luật bản quyền mới và thực thi, tối thiểu đối với thông tin điện tử, cần được tạo ra bởi vì tốc độ của các thành tựu công nghệ bỏ lại xa phía sau các hệ thống luật. Khi các luật mới có hiệu lực và được thực thi, chúng phải dựa vào công nghệ mới để trợ giúp bảo vệ tài liệu có bản quyền khỏi truy cập không hợp pháp, tái sản xuất, thao tác, phân phối và trình bày. Công nghệ mới cũng có thể trợ giúp phát hiện những vi phạm bản quyền thông qua các phương pháp mới của xác thực, quản lý tài liệu được bảo vệ bản quyền và các kỹ thuật cấp giấy phép.

Một số phương pháp bảo vệ các quyền sở hữu trí tuệ và bản quyền của thông tin điện tử hiện tại đang được nghiên cứu, bao gồm :

- 1) Truy cập phân đoạn: chỉ áp dụng cho các nguồn tin rất lớn và giá trị của chúng nằm trong lượng thông tin và tri thức có thể được lượm lặt từ phân tích toàn bộ tập hợp. Không có lợi thế kinh tế nào để sao chép những phần nhỏ dữ liệu và sao chép không hợp pháp toàn bộ nguồn dữ liệu là tương đối dễ phát hiện.
- 2) Điều khiển giao diện: đòi hỏi một giao diện giữ độc quyền để truy cập thông tin, ngụ ý rằng truy cập phổ quát là không thể.
- 3) Khóa phần cứng: truy cập thông tin bị hạn chế bởi phần cứng truy cập độc quyền.
- 4) Kho chứa thông tin: các bản sao chép hợp pháp chỉ có sẵn từ một kho chứa lớn; bất kỳ bản sao chép khác là không hợp pháp.
- 5) Hình mờ số: nhúng các thông báo ẩn vào trong thông tin. Mỗi một bản sao chép hợp pháp được gán nhãn với một số định danh khác nhau cho phép các bản sao chép không hợp pháp để lại dấu vết đằng sau người mua gốc, nghĩa là hình mờ số. Các bài toán chính là thông báo dễ di chuyển, khó chèn và khi mà phương pháp có vẻ như hợp với ảnh phức tạp, lại không hợp với văn bản đơn giản và không thể áp dụng ngay cả cho dữ liệu audio.

- 6) Mật mã: thông tin được mã hóa và không thể dịch được khi không có một khóa mật mã.
- 7) Các cách tiếp cận kinh tế: định rõ các cách làm cho sao chép không hợp pháp thông tin điện tử trở nên không kinh tế. Các ý tưởng bao gồm: các trang nhà cung cấp tính giá để làm giảm giá mỗi lần sao chép, điểm đăng ký để làm giảm điểm lừa đảo và nhà quảng cáo trợ giúp các xuất bản phẩm.
- 8) Lập lờ: công nghệ thông tin cho phép một khách hàng xem nhưng không lưu trữ thông tin.

3.3 Vấn đề chất lượng và an toàn

Chất lượng thông tin trực tuyến rất khó xác định. Hiện nay không có một công cụ tìm kiếm nào có một cách đánh giá chất lượng thông tin điện tử và vì thế không có cách nào về sắp xếp hoặc lọc thông tin dựa trên chất lượng.

Sự thiếu thông tin về chất lượng có xu hướng hạn chế tìm kiếm tới các chuyên gia nổi tiếng. Một số công cụ tìm kiếm thực sự sử dụng một profile của chuyên gia riêng lẻ hoặc một profile nhóm để hỏi thông tin. Tuy nhiên, toàn vẹn thông tin vẫn là một vấn đề bởi vì thông tin điện tử không dễ sửa đổi như vậy. Thông tin điện tử hiếm khi được bảo đảm là được thực sự tạo ra hoặc chứng thực bởi chuyên gia. Hiện tại, xuất bản khoa học điện tử là một chủ đề tranh luận nóng. Nhiều khái niệm tập trung quanh các câu hỏi về thông tin và phẩm chất trí tuệ.

Các vấn đề an toàn và điều khiển liên quan tới các vấn đề chất lượng. Thư viện số cần xem xét an toàn ít nhất ở bốn khía cạnh:

- 1) Độ tin cậy: bảo vệ truy cập tới nội dung thông tin, đặc biệt là thông tin nhạy cảm, như là thông tin cá nhân, tài chính hoặc sức khỏe và thông tin kinh doanh chiến lược hoặc quốc gia, khỏi truy cập và phân phối không hợp pháp.
- 2) Tính xác thực: quy thông tin cho tác giả chính xác và xác nhận nó là gốc, chính xác và gán đúng. Điều này có thể là khó khăn đặc biệt trong một môi trường nhiều tác giả và nhiều ấn bản.
- 3) Tính toàn vẹn: bảo vệ nội dung thông tin khỏi sửa đổi không hợp pháp. Loại an toàn này đòi hỏi một sự cân bằng giữa cho việc phép cập nhật hợp pháp dễ dàng và ngăn cản cập nhật không hợp pháp. Sự chứng thực sửa đổi phải được xác nhận.
- 4) Tính chính xác: bảo vệ truy cập thông tin và mẫu sử dụng khỏi sự truy cập và bán lại không hợp pháp.

Một thách thức quan trọng đối với sự cài đặt của bất kỳ kỹ thuật an toàn là sự cân bằng nhu cầu đối với an toàn với nhu cầu đối với thực hiện. Truy cập và sửa đổi hợp pháp không phải là khó đến nỗi nó không bao giờ được thử. Tương tự, khi các kỹ thuật xác nhận phải chính xác cũng như khả thi về mặt kỹ thuật, chúng không thể đòi hỏi nhiều thời gian và nguồn sao cho tính truy cập được và tính kịp thời của thông tin được thỏa hiệp. Thông tin không có giá trị nếu nó không truy cập được đúng lúc và có ích.

3.4 Các vấn đề văn hóa xã hội khác

Các vấn đề khác bao gồm:

- Học vấn: để sử dụng một thư viện số phải có một trình độ học vấn cơ bản nhất định về vận hành máy tính. Ai sẽ có trách nhiệm đào tạo các kỹ năng sử dụng máy tính cơ bản? Đào tạo thông qua các hệ thống giáo dục công cộng, hoặc là một phần trong những dịch vụ cung cấp bởi thư viện số? Sự tiếp cận các chương trình đào tạo cũng như tiếp cận thiết bị máy tính và các phương tiện thích hợp có chia đôi xã hội thành có và không có thông tin không?
- Ảnh hưởng đến văn hóa: lọc và tổ chức thông tin điện tử để hỗ trợ đối phó với sự quá tải thông tin là một dịch vụ hữu ích. Tuy nhiên, kết quả có thể ảnh hưởng đến văn hóa và các giá trị xã hội của nhà cung cấp dịch vụ. Ví dụ đơn giản nhất là ảnh hưởng đến ngôn ngữ. Thông tin nên được truy cập bằng ngôn ngữ gốc của nó hay nên dịch ra ngôn ngữ ưa thích của khách hàng không?
- Nên dịch tài liệu ở các dạng thức khác nhau như thế nào? Hơn nữa, thông tin nói chung có thể thích hợp đối với nhóm người này nhưng lại có thể là xúc phạm hoặc thậm chí không hợp pháp đối với nhóm khác. Một giải pháp có thể áp dụng là phát triển giao diện người dùng tùy ý để có thể thích nghi với một cá nhân hoặc văn hóa và sở thích ngôn ngữ của nhóm.
- Sự suy xét đạo đức: truy cập phổ quát kéo theo các vấn đề đạo đức liên quan tới công tác kiểm duyệt và ảnh hưởng đến văn hóa. Không phải tất cả thông tin đều thích hợp đối với mọi nhóm người. Thư viện số nên có hạn chế về loại thông tin có thể cung cấp truy cập tới? Hoặc nên có hạn chế về loại thông tin một cá nhân có thể nhận được?
- Sự bình đẳng: câu hỏi cần quan tâm là các cá nhân có cơ hội truy cập bình đẳng tới thông tin như nhau hay không? Khả năng bảo đảm truy cập bình đẳng của ấn bản điện tử cao hơn so với ấn bản in, vì thông tin được cung cấp đầy đủ và bất kỳ ai đều có thể truy cập nó. Không có khó khăn nào về cung cấp thông tin điện tử bình đẳng, nhưng các khách hàng có thể được khuyến khích truy cập một cách bình đẳng như thế nào?

4. Kết luận

Nói tóm lại, thư viện số ngày càng phát triển và đem lại nhiều lợi ích cho người sử dụng. Tuy nhiên, vẫn còn nhiều vấn đề liên quan quá trình xây dựng và phát triển thư viện số như vấn đề kinh tế, pháp luật, an toàn, chất lượng... Nhiều giải pháp đã được đưa ra để giải quyết các vấn đề này nhằm nâng cao vai trò của thư viện số trong bối cảnh thông tin gia tăng nhanh chóng và công nghệ liên tục đổi mới hiện nay.

Tài liệu tham khảo

1. Arms W.Y. (2003), *Digital Libraries*, MIT Press, Cambridge.
2. Đỗ Quang Vinh (2009), *Thư viện số - Chỉ mục và Tìm kiếm*, Đại học Quốc gia Hà Nội.
3. Lesk M. (2005), *Understanding Digital Libraries*, 2nd Edition, Morgan Kaufmann, San Francisco.
4. Marcum D.B. (2001), *Development of Digital Libraries*, Greenwood Press, Connecticut.
5. Pace A.K. (2003), *The Ultimate Digital Library*, American Library Association, Chicago.
6. Reddy R., Wladawsky-Berger I. (2001). *Digital Libraries: Universal Access To Human Knowledge*. A Report to the President. President's Information Technology Advisory Committee Panel on Digital Libraries.
Available at <https://www.nitrd.gov/pubs/pitac/pitac> [Accessed 9 October 2016]
7. Sun Microsystems (2002), *Digital Library Technology Trends*.
8. Theng Y.L., Foo S. (2005), *Design and Usability of Digital Libraries*, Information Science Publishing, London.
9. Witten I.H., Bainbridge D. (2003), *How to Build a Digital Library*, Morgan Kaufmann, San Francisco.
10. <https://www.diglib.org>