

Giá trị của việc ứng dụng công nghệ số trong lĩnh vực từ điển và bách khoa thư

Nguyễn Huy Bình^(*)

Tóm tắt: Hiện nay, việc ứng dụng công nghệ số ngày càng phổ biến đã mở ra cơ hội lớn cho sự phát triển của nhiều lĩnh vực, ngành nghề, trong đó có lĩnh vực từ điển và bách khoa thư. Với việc ứng dụng công nghệ số, chủ thể biên soạn sẽ xây dựng các công trình một cách tiện ích, hiệu quả; việc tiếp cận thông tin tri thức cũng trở nên dễ dàng ở mọi lúc, mọi nơi. Bài viết phân tích giá trị của việc ứng dụng công nghệ số trong biên soạn và vận hành từ điển và bách khoa thư.

Từ khóa: Ứng dụng công nghệ, Công nghệ số, Từ điển, Bách khoa thư

Abstract: The current popular application of digital technology has opened up great opportunities for the development of several fields and professions, including dictionaries and encyclopedias. Digital technology facilitates efficient and convenient compilation and provides anytime-anywhere access to academic knowledge and information. The paper analyzes the value of applying digital technology in compiling and running dictionaries and encyclopedias.

Keywords: Application of Technology, Digital Technology, Dictionary, Encyclopedia

1. Lời mở

Trong những năm đầu thế kỷ XXI và đặc biệt là khoảng 10 năm trở lại đây, vấn đề ứng dụng công nghệ số ở Việt Nam ngày càng được quan tâm. Năm 2020 là năm đánh dấu mốc quan trọng với “Chương trình chuyển đổi số quốc gia đến năm 2025, định hướng đến năm 2030” đã được Thủ tướng Chính phủ phê duyệt, điều này góp phần thúc đẩy mạnh mẽ quá trình chuyển đổi số tại Việt Nam. Có thể hiểu, ứng dụng công nghệ số là quá trình thay

đổi từ phương thức thủ công truyền thống sang áp dụng công nghệ với trung tâm là dữ liệu lớn, điện toán đám mây, trí tuệ nhân tạo. Ứng dụng công nghệ số làm thay đổi phương thức hoạt động, quy trình làm việc. Thực tế, việc ứng dụng công nghệ số sẽ tối giản hóa và giảm chi phí quản lý, vận hành; tạo ra nhiều sản phẩm, dịch vụ mới, nâng cao hiệu quả. Công nghệ số cũng mở ra triển vọng mới cho lĩnh vực biên soạn từ điển và bách khoa thư. Việc ứng dụng công nghệ số sẽ giúp chủ thể biên soạn và đối tượng tiếp nhận vượt qua những giới hạn của công nghệ truyền thống, góp phần tối ưu hóa trong cả khâu biên soạn, vận hành từ điển và bách khoa thư.

^(*) TS., Viện Từ điển học và Bách khoa thư Việt Nam, Viện Hàn lâm Khoa học xã hội Việt Nam; Email: huybinhvvh@gmail.com

2. Giá trị của việc ứng dụng công nghệ số trong biên soạn từ điển và bách khoa thư

Việc ứng dụng công nghệ số trong biên soạn từ điển và bách khoa thư có thể triển khai trên hai bình diện: số hóa tài liệu phục vụ cho việc biên soạn; tiến hành biên soạn mục từ trên nền tảng ứng dụng công nghệ số.

Đối với phương thức lưu trữ tài liệu truyền thống như lưu trữ theo kho, theo kệ, theo hộp, theo hồ sơ, theo văn bản,..., chủ thể biên soạn thực sự phải đối mặt với nhiều rủi ro, đó là sự tác động của các điều kiện khách quan như thời tiết, côn trùng và bản thân giấy sẽ tự hủy theo thời gian. Mặt khác, việc lưu trữ tài liệu trong kho khiến việc tìm kiếm một tài liệu cụ thể để sử dụng trong biên soạn từ điển và bách khoa thư sẽ khó khăn.

Khác với phương thức lưu trữ truyền thống, hiện nay, việc thu thập tài liệu qua ứng dụng công nghệ số hóa tạo ra các văn bản điện tử đã mang lại những giá trị tối ưu trong hoạt động tìm kiếm thông tin tri thức phục vụ quá trình biên soạn bách khoa thư. Việc số hóa tài liệu giúp bảo quản, duy trì “tuổi thọ” của tài liệu lâu hơn. Công việc số hóa tài liệu sẽ dễ dàng mở rộng phạm vi cộng đồng người sử dụng nguồn thông tin, thư viện, giúp việc truy xuất, tìm kiếm, chia sẻ thông tin nhanh chóng, dễ dàng, ở bất kỳ đâu, vào bất cứ thời điểm nào.

Số hóa tài liệu cũng giúp cho việc lưu trữ thông tin trở nên thuận tiện hơn, giảm tối đa chi phí quản lý và không gian lưu trữ, tăng cường khả năng bảo mật thông tin. Ngoài ra, số hóa dữ liệu giúp chủ thể biên soạn từ điển và bách khoa thư có thể chỉnh sửa và tái sử dụng tài liệu, linh hoạt trong việc chuyển đổi sang các dạng tài liệu số khác nhau.

Trong lĩnh vực biên soạn từ điển và bách khoa thư, việc ứng dụng công nghệ

số được thực hiện trong xây dựng bảng từ, lựa chọn mục từ, tổ chức viết mục từ (định nghĩa mục từ, các thông tin tri thức, ví dụ và hình ảnh minh họa,...), biên tập mục từ, lập bảng tra cứu, giải quyết mối quan hệ giữa các mục từ,... Cụ thể, việc tiến hành xây dựng kho ngữ liệu, xử lý các ngữ liệu, khai thác ngữ liệu phục vụ biên soạn đến nay đều được thực hiện trên máy tính, tích hợp và gửi thông tin qua không gian mạng.

Việc xử lý các ngữ liệu phục vụ biên soạn còn được thực hiện qua gán nhãn tự động bằng trí tuệ nhân tạo. Ứng dụng công nghệ số giúp chuyển ngữ liệu từ điển ở dạng cấu trúc tri thức trừu tượng thành ngữ liệu được đánh dấu hay gọi là gán nhãn, bảo đảm được lượng thông tin ban đầu. Nguồn ngữ liệu từ điển được mã hóa sẽ cung cấp các ứng dụng công nghệ trong nhiều lĩnh vực với một số ưu điểm như sau: *Thứ nhất*, cung cấp các trường nhập thông tin theo cấu trúc vi mô đầy đủ của một mục từ; giúp người biên soạn có một phương tiện làm việc hệ thống và nhất quán; giúp tránh bỏ sót và tránh nhầm lẫn khi nhập thông tin. *Thứ hai*, mục từ khi biên soạn được tự động sắp xếp theo trật tự bảng chữ cái... và có thể quan sát ngay trong danh sách mục từ. *Thứ ba*, nếu mục từ biên soạn bị trùng lặp thì chương trình sẽ gửi nhắc nhở đến người biên soạn. *Thứ tư*, ngữ liệu từ điển được chuyển sang định dạng sách, từ điển điện tử (dạng cài trên máy tính hoặc máy từ điển, và dạng tra cứu trực tuyến qua mạng Internet) một cách tự động và nhanh chóng (Theo: Viện Từ điển học và Bách khoa thư Việt Nam, 2020: 344-345). Trên thực tế, việc mã hóa có thể thực hiện đối với cả công trình từ điển và bách khoa thư, giúp việc thực hành biên soạn cũng như vận hành các công trình được linh hoạt và hiệu quả hơn.

Việc ứng dụng công nghệ số hóa trong tiến hành biên soạn từ điển và bách khoa

thư đã mang lại hiệu quả cao trong từng khâu đoạn. Trước hết là vấn đề huy động tài liệu điện tử trên máy tính, chủ thể biên soạn dễ dàng tìm kiếm, sử dụng thông tin lưu trữ phục vụ cho việc viết mục từ. Tiếp đó, mỗi mục từ của từ điển và bách khoa thư sau khi hoàn thành được lưu trữ trên hệ thống phần mềm hoặc trên không gian mạng sẽ rất tiện ích cho việc sắp xếp theo bảng mục từ, lập bảng tra cứu, chỉnh sửa lỗi kỹ thuật hay nội dung khoa học.

Bên cạnh đó, việc ứng dụng công nghệ số trong biên soạn từ điển và bách khoa thư sẽ góp phần tối ưu hóa hoạt động của chủ thể biên soạn, giúp giảm tải những công việc thủ công, rút ngắn các khâu được thực hiện lặp đi lặp lại nhờ hệ thống máy móc tự động. Chủ thể biên soạn có thể giảm tải khối lượng công việc, dành thời gian tập trung vào chuyên môn. Việc áp dụng công nghệ số vào vận hành kho dữ liệu giúp nâng cao khả năng sử dụng tư liệu trong quá trình thực hành biên soạn từ điển và bách khoa thư.

Với những lợi thế trên, việc ứng dụng công nghệ số sẽ là một khâu quan trọng nhằm nâng cao hiệu quả biên soạn từ điển và bách khoa thư, vấn đề huy động các dữ liệu của từng mục từ được kích hoạt nhanh chóng, độ chính xác cao.

3. Giá trị của việc ứng dụng công nghệ số trong vận hành từ điển và bách khoa thư

Từ điển và bách khoa thư khi được vận hành trên nền tảng công nghệ số, sử dụng qua phương tiện Internet để truyền tải thông tin tri thức đến người sử dụng sẽ mang lại giá trị to lớn trên nhiều phương diện.

Trên thế giới, việc ứng dụng công nghệ số trong vận hành các công trình từ điển và bách khoa thư đã trở nên phổ biến. Hiện đã có một số tổ chức biên soạn từ điển và bách khoa thư cho xây dựng và vận hành các công trình từ điển và bách khoa thư bằng

hình thức trực tuyến (online)¹. Các công trình đã mang lại giá trị to lớn cho cả chủ thể biên soạn và đối tượng tiếp nhận.

Xét từ phương thức vận hành, việc ứng dụng công nghệ số sẽ tạo ra một số loại hình từ điển và bách khoa thư tồn tại ở dạng chuyên biệt, đó là sự kết hợp giữa kênh chữ và kênh hình, thậm chí có thể kết hợp cả dạng âm thanh một cách dễ dàng và tiện ích. So với các công trình truyền thống, từ điển và bách khoa thư trực tuyến đã làm thay đổi cơ bản cách thức quản lý, giúp giảm thiểu thời gian tra cứu thông tin, góp phần gìn giữ tri thức và phổ biến tri thức trên phạm vi toàn cầu một cách tiện ích và năng động hơn nhiều so với loại hình truyền thống.

Thực tế, việc xây dựng và vận hành các công trình từ điển và bách khoa thư trực tuyến hiện nay đã được thực hiện trên cơ sở ứng dụng công nghệ số. *Thứ nhất*, các nhà biên soạn từ điển và bách khoa thư xác lập một cách cụ thể nội dung thông tin, quy cách của công trình rồi tiến hành biên soạn nội dung trên phần mềm được thiết kế một cách phù hợp với tiêu chí chung của từng công trình cụ thể. *Thứ hai*, để vận hành từ điển và bách khoa thư trực tuyến, các nhà làm từ điển và bách khoa thư cần đưa ra một tập bản thảo được định sẵn, hoàn chỉnh về mặt cấu trúc và nội dung trên văn bản word; sau đó đưa vào phần mềm để vận hành trực tuyến trên hệ thống Internet (Nguyễn Huy Bình, Ngô Huy Biên, 2018).

Từ thực tiễn ứng dụng kỹ thuật số trong vận hành các công trình từ điển và bách khoa thư trực tuyến cho thấy, điểm mạnh nổi bật nhất của loại hình này là dễ sử dụng, tính năng động cao, dễ dàng thay đổi, nâng cấp và có tính phổ cập rộng.

¹ Tiêu biểu như: *leerit.com*; *envisafety.com*; *tran thanhxuan.name.vn*; *bostonenglish.edu.vn*; *wikipedia*,...

Một triển vọng lớn của việc vận hành từ điển và bách khoa thư trực tuyến là khả năng xuất hiện những công trình với nhiều phương thức truyền đạt ở các dạng thức khác nhau như kênh chữ, kênh hình, kênh tiếng. Bên cạnh đó, các công trình còn ở dạng tích hợp, bao gồm cả dạng kênh chữ, kênh hình, video, âm thanh,... Có thể nói, ở phương diện vận hành trực tuyến, loại hình từ điển và bách khoa thư nào cũng có những tính năng ưu việt, vượt trội so với phương thức xuất bản bằng giấy.

Khi các công trình này đã được đồng bộ hóa có thể tạo ra một chuỗi liên kết qua đường link, giúp đối tượng tra cứu dễ dàng tiếp cận được hệ thống thông tin tri thức của nhiều mục có từ khóa giống nhau. Mặt khác, các tài liệu tham khảo đã được đưa lên mạng Internet cũng có thể được đưa vào đường link, giúp người sử dụng có thể tiếp cận văn bản gốc.

Với việc đồng bộ hóa các phần mềm từ điển và bách khoa thư trực tuyến, đối tượng tiếp nhận chỉ cần đưa ra lệnh tìm kiếm đối với một từ khóa là từ đầu mục thì hệ thống Internet sẽ tự động cung cấp tất cả thông tin về từ khóa đó. Đây là điều kiện lý tưởng để người tiếp nhận có được một cái nhìn tổng thể về các thông tin tri thức ở mỗi công trình.

So với phương thức truyền thống là bản in giấy, chi phí vận hành từ điển và bách khoa thư theo phương thức trực tuyến thấp hơn đáng kể. Bên cạnh đó, loại hình từ điển và bách khoa thư trực tuyến còn tiện ích cho việc bổ sung và chỉnh sửa thông tin tri thức. Đối với các công trình từ điển và bách khoa thư xuất bản trên giấy, nếu muốn bổ sung và chỉnh sửa thông tin tri thức hoặc kỹ thuật thì phải tiến hành tái bản, bổ sung; còn đối với loại hình trực tuyến, chủ thể biên soạn có thể bổ sung, chỉnh sửa ngay trên phần mềm đang được vận hành. Thông qua hình thức vận hành trực tuyến, các công trình từ điển

và bách khoa thư có thể mở kênh thông tin để lấy ý kiến của cộng đồng bằng việc tạo ra mục chuyên biệt để tiếp nhận phản hồi.

Nhìn từ góc độ của người sử dụng, việc ứng dụng công nghệ số trong vận hành từ điển và bách khoa thư trực tuyến đã đáp ứng được những yêu cầu đang đặt ra trong xã hội hiện đại mà các loại hình vận hành trước đó không đáp ứng được. Trong xu hướng phát triển mạnh mẽ của hệ thống công nghệ số ở Việt Nam hiện nay, việc vận hành công trình từ điển và bách khoa thư trực tuyến không chỉ đáp ứng được nhu cầu của chủ thể biên soạn mà còn đáp ứng được nhu cầu của cộng đồng trong việc tìm kiếm thông tin, nó sẽ trở thành một công cụ hữu ích, được đối tượng tiếp nhận sử dụng thường xuyên.

Ở Việt Nam, trước đây việc vận hành từ điển và bách khoa thư vẫn chủ yếu bằng hình thức xuất bản trên giấy, việc ứng dụng công nghệ số trong công tác xây dựng từ điển và bách khoa thư vẫn chưa có quy mô lớn, đa phần các công trình từ điển và bách khoa thư trực tuyến có quy mô nhỏ, hẹp. Hiện nay, việc sử dụng Internet cho hoạt động tra cứu thông tin trên từ điển và bách khoa thư đã trở nên phổ biến, thay thế dần cho việc tra cứu thông tin trên những ấn phẩm truyền thống. Thực tế, một số ngành khoa học đã ứng dụng công nghệ số để xây dựng và vận hành từ điển chuyên ngành, đã có một số công trình từ điển tra cứu trực tuyến trên Internet, bước đầu nhận được sự hưởng ứng của cộng đồng¹. Loại hình từ điển trực tuyến này được dùng tra cứu phổ biến trên máy tính, điện thoại, máy tính bảng.

Trong xu hướng chung của thế giới, sự ra đời của các công trình từ điển và bách khoa thư trực tuyến là tất yếu, nó sẽ góp

¹ Đơn cử như Bộ *Từ điển bách khoa Việt Nam* (4 tập) gồm khoảng 40.000 mục từ với 40 chuyên ngành khác nhau.

phần làm thay đổi phương thức tồn tại truyền thống theo chiều hướng tích cực, giúp việc tra cứu thông tin tri thức được thuận lợi và dễ dàng hơn. Tuy nhiên, từ điển và bách khoa thư trực tuyến không vì thế mà mất đi tính chất, đặc trưng của từ điển và bách khoa thư, nó vẫn đảm bảo tính khoa học, khách quan của tư liệu. Có thể nói, trong bối cảnh ứng dụng công nghệ số trở thành mục tiêu của nhiều ngành nghề thì lĩnh vực từ điển và bách khoa thư cũng không thể đứng ngoài việc xây dựng và vận hành các công trình trực tuyến. Đây có thể coi là một loại hình đang mở ra cơ hội và triển vọng lớn cho lĩnh vực từ điển và bách khoa thư ở Việt Nam.

4. Kết luận và kiến nghị

Thời đại công nghệ số 4.0 đã và đang diễn ra với quy mô lớn, có tác động đến tất cả các ngành nghề, lĩnh vực trong đời sống xã hội. Việc ứng dụng công nghệ số hóa hay vấn đề sử dụng một tập hợp bao gồm Internet, hệ thống vật lý không gian mạng, hệ thống trí tuệ nhân tạo,... đã tạo ra cơ hội lớn trong lĩnh vực biên soạn từ điển và bách khoa thư. Việc ứng dụng công nghệ số trong xây dựng từ điển và bách khoa thư sẽ góp phần giảm thiểu công sức lao động, cũng như thời gian xây dựng. Mặt khác, việc ứng dụng công nghệ số trong vận hành từ điển và bách khoa thư trực tuyến còn làm tăng khả năng cung cấp dịch vụ thông tin, đáp ứng nhu cầu của người sử dụng thông tin, phổ biến tri thức và gìn giữ tri thức cho xã hội.

Hiện nay, Việt Nam đang trong bối cảnh phát triển công nghệ số và hội nhập

quốc tế, nhu cầu của người sử dụng từ điển và bách khoa thư trực tuyến ngày càng gia tăng, việc xây dựng các bộ công cụ tra cứu này được đặt ra như một đòi hỏi cấp bách. Để thực hiện được những mục tiêu và nhiệm vụ xây dựng từ điển và bách khoa thư dưới dạng số hóa, cần đặt ra một số vấn đề cơ bản như: *Thứ nhất*, cần có sự nhận thức đúng đắn về giá trị của các tổ chức trong việc ứng dụng công nghệ số để biên soạn, vận hành từ điển và bách khoa thư. *Thứ hai*, việc ứng dụng công nghệ số trong biên soạn, vận hành từ điển và bách khoa thư cần có sự nhập cuộc và kết hợp của nhiều ngành, lĩnh vực để có thể phối hợp một cách đồng bộ. *Thứ ba*, cần xây dựng đội ngũ tham gia biên soạn, vận hành từ điển và bách khoa thư có trình độ kỹ thuật tốt, trong đó có đội ngũ những người làm về công nghệ thông tin được đào tạo bài bản và thực hành tốt các kỹ năng nghiệp vụ. *Thứ tư*, cần có sự chuẩn bị tốt về cơ sở vật chất, các thiết bị máy móc hiện đại phục vụ cho việc biên soạn, vận hành từ điển và bách khoa thư trực tuyến □

Tài liệu tham khảo

1. Nguyễn Huy Bình, Ngô Huy Biên (2018), “Loại hình từ điển online ở Việt Nam: Thực trạng và triển vọng”, Tạp chí *Từ điển học và bách khoa thư*, số 7.
2. Viện Từ điển học và Bách khoa thư Việt Nam (2020), *Từ điển học và Bách khoa thư học Việt Nam: Lý luận và thực tiễn* (Kỷ yếu hội thảo khoa học 2020), Nxb. Dân trí, Hà Nội.