

NGHIÊN CỨU XÂY DỰNG QUY TRÌNH BIÊN SOẠN HỆ THỐNG THUẬT NGỮ GIẢI THÍCH CHO CHUYÊN NGÀNH HỆ THỐNG THÔNG TIN QUẢN LÝ

TS. Trần Thị Song Minh

Đại học Kinh tế Quốc dân

Trong các hoạt động đào tạo, nghiên cứu, tư vấn, ứng dụng và chuyển giao công nghệ rất cần đến các hệ thống thuật ngữ giải thích, giúp người dùng hiểu được các thuật ngữ mới hoặc ít phổ biến cũng như các thuật ngữ chuyên ngành một cách chính xác và cô đọng nhất. Bài viết sau đây nghiên cứu phương pháp luận và quy trình biên soạn hệ thống thuật ngữ giải thích trong môi trường ngôn ngữ tiếng Việt cho chuyên ngành Hệ thống thông tin quản lý với phương châm khai thác và tận dụng tối đa các công nghệ hiện đại và các nguồn thuật ngữ giải thích có sẵn.

Từ khóa: dịch tự động, hệ thống thông tin quản lý, hệ thống thuật ngữ, quy trình phát triển, thuật ngữ, từ điển giải thích

1. Sự cần thiết của hệ thống thuật ngữ giải thích chuyên ngành Hệ thống thông tin quản lý

Hệ thống thuật ngữ giải thích (glossary of terms) hay còn gọi là từ điển giải thích (dictionary of terms) là một hệ thống các thuật ngữ liên quan đến một phạm trù kiến thức, một lĩnh vực chuyên môn xác định, cùng với một định nghĩa hay giải thích ngắn gọn về từng thuật ngữ.

Khác với từ điển phổ thông, thuật ngữ giải thích có những đặc điểm riêng của nó: (1) thuật ngữ được giải thích theo ngữ cảnh, nghĩa là nội dung mô tả của thuật ngữ gắn với ngữ cảnh cụ thể của chuyên ngành và thường khác nhiều so với ngữ cảnh chung; (2) nội dung của một thuật ngữ có thể được giải thích theo các cách nhìn khác nhau, tùy thuộc mục đích của chủ thể; (3) thuật ngữ được giải thích theo phạm vi sử dụng, khi được dùng cho một phạm vi rộng rãi cần có nội dung đủ khái quát cho phạm vi đó, nhưng nếu nó chỉ dùng cho một phạm vi hẹp thì nội dung cần được giải thích cụ thể, chi tiết hơn.

Hệ thống thuật ngữ giải thích có thể ở dạng đơn ngữ, hoặc đa ngữ. Trong hệ thống thuật ngữ giải thích đơn ngữ, đi kèm với mỗi thuật ngữ là một giải thích ngắn gọn, súc tích cho thuật ngữ đó trong cùng một ngôn ngữ. Trong hệ thống thuật ngữ giải thích đa ngữ, thuật ngữ thường được liệt kê bằng một hoặc nhiều ngôn ngữ và giải thích kèm theo cũng có

thể được mô tả bằng một hay nhiều ngôn ngữ.

Hệ thống thuật ngữ giải thích có thể tồn tại ở dạng tài liệu tra cứu độc lập hoặc như một phần không thể tách rời của các tài liệu nghiệp vụ. Hệ thống thuật ngữ giải thích độc lập thường phục vụ nhu cầu tra cứu chung trong một lĩnh vực xác định, trong khi đó, phần từ vựng giải thích trong các tài liệu nghiệp vụ chỉ chứa các khái niệm, thuật ngữ mới được giới thiệu, hoặc ít thông dụng hoặc có tính chất chuyên ngành và giải thích ngữ nghĩa ngắn gọn giúp người đọc hiểu được nội dung trình bày trong các tài liệu đó.

Bằng việc cung cấp một định nghĩa, hay một giải thích ngắn gọn về các thuật ngữ chuyên môn, hệ thống thuật ngữ giải thích đóng vai trò quan trọng trong việc đảm bảo tính nhất quán khi thực hiện các hoạt động chuyên môn nghiệp vụ, quản lý tổ chức cũng như các hoạt động học tập, nghiên cứu hay ứng dụng và chuyển giao công nghệ.

Mỗi ngành, lĩnh vực đều có ngôn ngữ riêng của nó và Hệ thống thông tin quản lý cũng không phải là một ngoại lệ. Việc xây dựng một tài liệu tra cứu mang tính học thuật và thống nhất về các khái niệm đặc biệt là các thuật ngữ mới liên quan đến công nghệ mới, ứng dụng mới trong chuyên ngành Hệ thống thông tin quản lý là hết sức cần thiết.

Thứ nhất, hệ thống thông tin quản lý là một ngành đa lĩnh vực. Khi thực hiện các hoạt động

nghiên cứu, đào tạo, tư vấn hay ứng dụng hệ thống thông tin quản lý, đều rất cần đến những kiến thức đa lĩnh vực. Hệ thống thuật ngữ giải thích chuyên ngành sẽ giúp người sử dụng trong thời gian ngắn nhất hiểu được những thuật ngữ cơ bản trong lĩnh vực liên quan một cách ngắn gọn và súc tích nhất;

Thứ hai, phần lớn các tài liệu chuyên môn hệ thống thông tin quản lý là bằng tiếng Anh. Khi tiếp cận với các tài liệu này, rào cản lớn nhất đối với người sử dụng là vấn đề ngôn ngữ. Trong môi trường xã hội sử dụng tiếng Việt, tất cả những khái niệm và nội dung của tài liệu bằng tiếng Anh cần được thể hiện bằng ngôn ngữ tiếng Việt. Hệ thống thuật ngữ giải thích chuyên ngành là một công cụ tốt nhất, hiệu quả nhất để trợ giúp hoạt động này.

Thứ ba, trên thế giới đã có hệ thống thuật ngữ giải thích cho chuyên ngành này ở những dạng thức khác nhau và phần lớn bằng tiếng Anh. Ở Việt Nam chưa có một hệ thống thuật ngữ giải thích nào về hệ thống thông tin quản lý bằng tiếng Việt, nếu có cũng chỉ dừng ở mức từ điển, tức là chỉ liệt kê tên gọi các khái niệm mà không có giải thích ý nghĩa chuyên ngành đi kèm.

Thứ tư, trong hai thập kỷ gần đây, công nghệ thông tin phát triển như vũ bão. Nhiều công nghệ mới, nhiều ứng dụng mới hình thành và được đưa vào triển khai mau chóng trong các nền kinh tế phát triển. Hệ thống thuật ngữ giải thích chuyên ngành *Hệ thống thông tin quản lý* sẽ trợ giúp rất nhiều trong việc tiếp cận và chuyển giao công nghệ tiên tiến từ các nước phát triển về lĩnh vực này.

2. Đề xuất quy trình biên soạn hệ thống thuật ngữ giải thích

Do nhu cầu thực tế, không ít những hệ thống thuật ngữ giải thích đã được nghiên cứu, biên tập và xuất bản. Tuy nhiên chúng nằm rải rác ở nhiều nơi, dưới nhiều hình thức lưu trữ với cách tổ chức và thể hiện khác nhau (bản cứng, bản mềm, trong các tập sách hay trên mạng). Đặc biệt, các tài liệu về thuật ngữ giải thích chuyên ngành lại chủ yếu bằng tiếng nước ngoài (chủ yếu là tiếng Anh). Về mặt kỹ thuật, các mục từ trong hệ thống thuật ngữ giải thích phải được phân thành từng vắn, mỗi vắn sắp xếp theo một trình tự xác định. Những công việc này đòi hỏi một khối lượng xử lý rất lớn.

Những đặc thù và khó khăn như trên cho thấy rằng, công việc biên soạn hệ thống thuật ngữ giải thích là rất khó khăn và phức tạp. Đây là công việc

tốn nhiều thời gian và công sức, cần sự tham gia của nhiều người có kiến thức và kinh nghiệm chuyên môn sâu về lĩnh vực liên quan. Vấn đề là làm thế nào để có thể xây dựng các hệ thống thuật ngữ giải thích một cách chuyên nghiệp và hiệu quả? Một quy trình nghiệp vụ chuẩn mực có tính khoa học cao và tận dụng các lợi thế về công nghệ và chuyên môn sẽ là câu trả lời cho vấn đề này. Cụ thể:

Cần phải tận dụng các công nghệ hiện đại như nhận dạng và quét văn bản, chuyển định dạng văn bản và dịch văn bản tự động trong quá trình biên soạn thuật ngữ giải thích để giảm thiểu công việc nhập liệu và đẩy nhanh tiến độ biên soạn. Không có ứng dụng công nghệ thông tin, việc biên soạn hệ thống thuật ngữ giải thích sẽ vô cùng khó khăn và vất vả;

Cần khai thác và tận dụng các thuật ngữ có sẵn từ nhiều nguồn khác nhau, ở nhiều dạng thức khác nhau để phục vụ cho công việc biên soạn thuật ngữ trong môi trường tiếng Việt. Một trong những nguồn thuật ngữ lớn nhất chính là nguồn từ Internet. Nếu không bắt đầu từ những nguồn có sẵn thì công việc biên soạn sẽ rất khó khả thi.

Khi xây dựng hệ thống thuật ngữ giải thích cho một chuyên ngành hay lĩnh vực xác định, cần giải quyết các vấn đề chính sau đây:

Xác định phạm vi và đối tượng như thế nào để bao quát được nội dung mong muốn?

Tìm các nguồn tư liệu từ đâu và thế nào là đầy đủ?

Làm thế nào để tư liệu có thể xử lý tự động trên máy tính?

Khai thác khả năng xử lý tự động trên máy tính phục vụ quá trình biên dịch đến đâu?

Biên dịch thuật ngữ sang tiếng Việt như thế nào để đảm bảo tính chính xác và phù hợp về mặt chuyên môn và ngôn ngữ, đặc biệt là các thuật ngữ mới?

Biên dịch nội dung giải thích như thế nào đảm bảo tính chính xác, đầy đủ và phù hợp về mặt chuyên môn và ngôn ngữ?

Tổ chức và xuất bản tài liệu như thế nào cho thích hợp và hiệu quả nhất?

Về mặt nguyên tắc, quá trình xây dựng hệ thống thuật ngữ giải thích phải bao gồm các hoạt động chính sau đây:

Nghiên cứu tổng quan về chuyên ngành hoặc lĩnh

vực liên quan và xác định yêu cầu;

Xác định và hệ thống hóa các nguồn thuật ngữ giải thích hiện tượng có liên quan đến chuyên ngành hoặc lĩnh vực liên quan;

Thu thập các nguồn thuật ngữ giải thích có sẵn;

Đánh giá mức độ đáp ứng của các thuật ngữ giải thích thu thập được và tìm kiếm bổ sung các thuật ngữ;

Xử lý, tổng hợp các thuật ngữ giải thích;

Biên dịch các thuật ngữ và biên tập nội dung các giải thích;

Thẩm định, đánh giá và hoàn thiện hệ thống thuật ngữ giải thích;

Tổ chức sắp xếp, biên tập để xuất bản hệ thống thuật ngữ giải thích cho nhu cầu sử dụng.

Một trong những phương pháp phát triển hệ thống thuật ngữ giải thích hiện đại là phát triển trực tuyến, dần dần thông qua tương tác trên web. Việc thực hiện các cuộc họp ảo trên mạng cho phép tập hợp các ý kiến phản hồi rộng rãi của cộng đồng mạng, phục vụ mục tiêu bổ sung, cập nhật hệ thống thuật ngữ giải thích. Công cụ IBM InfoSphere Business Glossary là một ví dụ minh họa. Công cụ này hỗ trợ tất cả các hoạt động: từ lập kế hoạch, đến thiết kế và xuất bản từ điển giải thích nghiệp vụ trong kinh doanh.

Nếu không sử dụng phương pháp phát triển hệ thống thuật ngữ chuyên dụng trực tuyến như trên, mà xây dựng hệ thống thuật ngữ giải thích dưới dạng một tài liệu, trong một thời gian nhất định có thể sử dụng quy trình 15 bước được mô tả ở hình 1, thực hiện các hoạt động chính của quá trình xây dựng hệ thống thuật ngữ giải thích nêu trên.

Trong quá trình biên soạn hệ thống thuật ngữ giải thích cần lưu ý một số vấn đề sau đây:

Các khái niệm, thuật ngữ phải được liệt kê theo vần alphabet;

Giải thích, định nghĩa các khái niệm, thuật ngữ cần ngắn gọn;

Khi có một số cách giải thích cho một thuật ngữ thì chọn cách tốt nhất nếu cùng quan điểm, hoặc chọn cả nếu theo quan điểm khác nhau;

Khi cần định nghĩa một động từ, nên chuyển động từ thành danh từ, vì định nghĩa một danh từ dễ hơn định nghĩa một động từ;

Đối với những từ khóa bắt đầu bằng số thì cần

căn cứ vào vần khi viết số đó ở dạng chữ để xếp vần;

Đối với các chữ viết tắt, không đưa ra định nghĩa cho chữ viết tắt mà chỉ viết rõ, đầy đủ cụm từ liên quan, sau đó đưa ra tham chiếu đến cụm từ đầy đủ và giải thích của cụm từ đó.

3. Xây dựng hệ thống thuật ngữ giải thích chuyên ngành Hệ thống thông tin quản lý

Vận dụng quy trình được đề xuất ở mục 2, quá trình biên soạn hệ thống từ vựng giải thích cho chuyên ngành *Hệ thống thông tin quản lý* đã được thực hiện đúng quy trình và thành công, với kết quả là một hệ thống thuật ngữ giải thích gồm hơn 500 thuật ngữ. Sau đây là tóm tắt kết quả của quá trình vận dụng này.

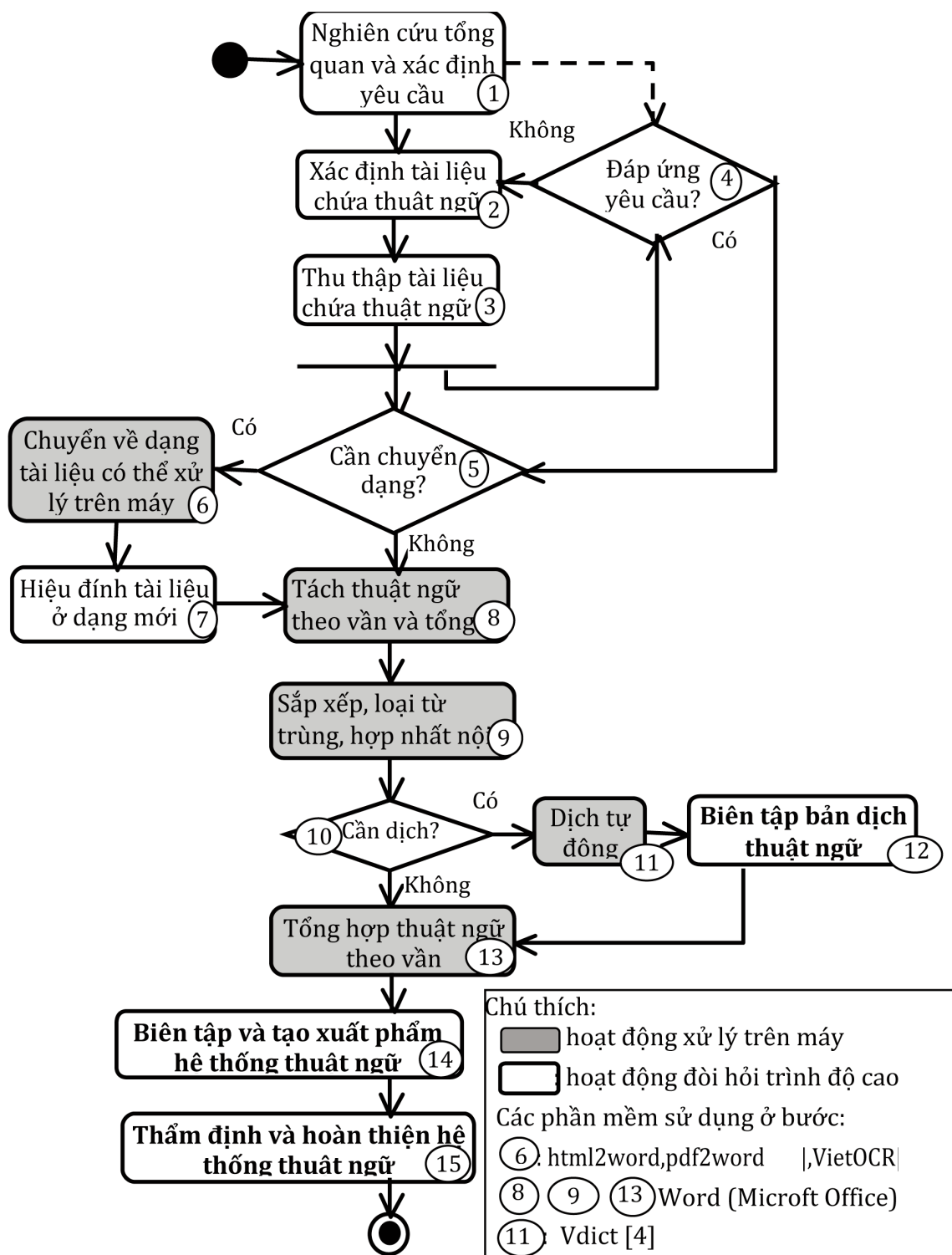
Ở bước nghiên cứu tổng quan và xác định yêu cầu, đã xác định được các phạm trù chuyên môn chính của chuyên ngành *Hệ thống thông tin quản lý* bao gồm: (1) Hệ thống thông tin với vai trò chiến lược trong các tổ chức; (2) Các công nghệ thông tin cơ bản ứng dụng trong hệ thống thông tin; (3) Các loại hình hệ thống thông tin ứng dụng trong quản lý và quản trị kinh doanh; (4) Các phương pháp triển khai hệ thống thông tin quản lý; (5) Phát triển và quản trị các hệ thống thông tin quản lý.

Ở giai đoạn thí điểm, hai nguồn thuật ngữ chủ yếu đã được lựa chọn để hệ thống hóa các từ vựng có sẵn là sách chuyên ngành hệ thống thông tin quản lý và mạng Internet. Các tài liệu chứa thuật ngữ giải thích *Hệ thống thông tin quản lý* thu thập được tồn tại ở các định dạng rất khác nhau: tài liệu điện tử dạng WORD, dạng HTML, dạng PDF, thậm chí ở dạng bản in cứng. Trong số các loại tài liệu này chỉ có tài liệu điện tử dạng WORD là có thể xử lý ngay được trên máy tính bằng phần mềm soạn thảo văn bản WORD, nên những tài liệu dạng khác đã được chuyển về dạng WORD trước khi được xử lý tiếp trên máy tính.

Việc chuyển định dạng cho các tài liệu được thực hiện với sự trợ giúp của các phần mềm chuyên dụng: phần mềm *pdf2word* dùng để chuyển từ định dạng PDF sang WORD, phần mềm *html2word* dùng để chuyển từ định dạng HTML sang WORD, phần mềm *VietOCR* dùng để chuyển tài liệu văn bản dạng ảnh (được quét từ bản cứng) chuyển sang WORD.

Các tài liệu sau khi chuyển định dạng được rà soát và đối chiếu với tài liệu gốc, phát hiện lỗi và tiến hành hiệu đính, vì quá trình chuyển định dạng

Hình 1: Quy trình xây dựng hệ thống thuật ngữ giải thích



Chú thích: **Đường nét đậm** chỉ trình tự thực hiện công việc (có thể bao gồm cả thông tin), đường đứt nét chỉ luồng thông tin (theo chuẩn của UML).

bằng các phần mềm chuyên dụng đã có một số lỗi nhỏ, ví dụ “*its final customer*” “sau khi quét thành ảnh và nhận dạng trở thành “*;tz final customer*”.

Các thuật ngữ giải thích trong mỗi tệp tài liệu (dạng WORD) sau đó được tách theo vắn và dồn vào thành từng tệp WORD, mỗi tệp chỉ bao gồm các thuật ngữ thuộc một vắn xác định. Khi sắp xếp các thuật ngữ trong mỗi tệp theo vắn, đã phát hiện nhiều thuật ngữ bị trùng lặp do chúng được tập hợp từ nhiều nguồn tài liệu khác nhau. Đối với những thuật ngữ bị trùng lặp, chỉ một tên thuật ngữ với cách giải thích đầy đủ, chính xác nhất được giữ lại và bổ sung các ý giải thích cần thiết khác.

Các tài liệu chứa thuật ngữ giải thích chuyên ngành *Hệ thống thông tin quản lý* có sẵn phần lớn bằng tiếng Anh, vậy nên phần mềm *Vdict* đã được sử dụng để dịch tự động các thuật ngữ cùng với giải thích đi kèm (bằng tiếng Anh) sang tiếng Việt. Sử dụng phần mềm dịch tự động giúp rút ngắn đáng kể thời gian biên dịch các giải thích. Tuy nhiên, chất lượng tài liệu dịch từ tiếng Anh sang tiếng Việt (của mọi chương trình dịch hiện có) còn rất hạn chế, nên đã phải tiến hành hiệu đính bản dịch để có thể dùng được.

Sau khi được biên dịch thành công sang tiếng Việt, nguồn thuật ngữ (nguyên gốc bằng tiếng Anh) được “hợp nhất” với nguồn thuật ngữ giải thích có sẵn trong tiếng Việt, chuẩn bị cho việc tạo xuất hệ thống thuật ngữ giải thích. Tiếp theo, hệ thống thuật

ngữ giải thích hợp nhất được biên tập lại, theo đó các thuật ngữ của mỗi vắn được sắp xếp theo trình tự và mỗi thuật ngữ chỉ xuất hiện một lần với một giải thích đầy đủ, chính xác nhất đi kèm.

Để phục vụ nhu cầu tra cứu, các kết cấu sau đây đã được sử dụng để tổ chức hệ thống thuật ngữ giải thích chuyên ngành *Hệ thống thông tin quản lý* trong môi trường tiếng Việt: (1) Thuật ngữ tiếng Anh– Thuật ngữ tiếng Việt - Giải thích tiếng Việt; (2) Thuật ngữ tiếng Việt– Thuật ngữ tiếng Anh- Giải thích tiếng Việt. Theo đúng quy trình, hệ thống thuật ngữ giải thích được thẩm định và hoàn thiện với sự tham gia của các chuyên gia có uy tín và có kiến thức toàn diện, chuyên sâu trong lĩnh vực Hệ thống thông tin quản lý, trước khi xuất bản và chính thức đưa vào sử dụng.

4. Kết luận

Việc biên soạn hệ thống thuật ngữ giải thích chuyên ngành theo một quy trình chuẩn mực, có ứng dụng công nghệ thông tin, phục vụ nhu cầu tra cứu, tiếp cận và chuyển giao công nghệ tiên tiến từ các nước phát triển trong lĩnh vực liên quan là hết sức cần thiết. Với kết quả áp dụng thành công cho chuyên ngành *Hệ thống thông tin quản lý*, quy trình biên soạn hệ thống thuật ngữ giải thích đề cập trong bài có thể được xem xét để vận dụng cho nhiều chuyên ngành khác. □

Tài liệu tham khảo:

1. Balzan P. (2012), *Business Driven Information Systems*, third Edition, Mc Graw Hill,
2. James A. O'Brien, George M. Marakas. (2006), *Management Information Systems*, 7/E, McGraw-Hill.
3. Kenneth C. Laudon, Jane P. Laudons. (2006), *Management Information Systems: Managing the Digital Firm*, 9/E, Prentice Hall Publishing House.
4. Trần Thị Song Minh (Chủ nhiệm đề tài), Nguyễn Văn Vy, Trần Thị Thu Hà, Nguyễn Văn Thoan. (2011), *Đề tài NCKH cấp cơ sở mã số T.2011.31*.
5. Trần Thị Song Minh, Trương Văn Tú, Cao Đình Thi (2011), *Giáo trình Hệ thống thông tin quản lý*, NXB Đại học Kinh tế quốc dân.