

CHỈ TIÊU THỐNG KÊ VỀ CÔNG BỐ KHOA HỌC VÀ CÔNG NGHỆ

TS Lê Xuân Định
Cục Thông tin KH&CN Quốc gia

Giới thiệu một số phương pháp xác định chỉ tiêu thống kê về công bố KH&CN như trắc lượng thư mục, hệ số tác động, chỉ số trích dẫn. Mô tả nội dung, phương pháp tính và cách thể hiện của ba chỉ tiêu thống kê về công bố KH&CN của Việt Nam: số bài báo khoa học của Việt Nam công bố trong nước và ngoài nước, số bài báo khoa học của Việt Nam công bố ở nước ngoài được trích dẫn và số giải thưởng KH&CN trong nước và quốc tế được trao tặng, số người được trao tặng giải thưởng KH&CN trong nước và giải thưởng quốc tế.

1. Đặt vấn đề

Chỉ tiêu thống kê về công bố KH&CN (gọi tắt là chỉ tiêu) là công cụ quan trọng để đánh giá tính hiệu quả của các chương trình nghiên cứu, ghi nhận và đánh giá hiệu quả nghiên cứu của từng nhà khoa học, cơ quan nghiên cứu và lĩnh vực nghiên cứu, đáp ứng nhu cầu quản lý nhà nước và hoạch định chính sách về KH&CN. Chỉ tiêu luôn được coi là cần thiết cho nhà nghiên cứu, tổ chức nghiên cứu, cơ quan quản lý nhà nước về KH&CN và các tổ chức cấp kinh phí cho

nghiên cứu khoa học và phát triển công nghệ (NC&PT). Chỉ tiêu thường được nhìn nhận như một công cụ đánh giá định lượng các kết quả nghiên cứu, hỗ trợ trực tiếp cho công tác quản lý KH&CN.

Tầm quan trọng của việc thống kê về công bố KH&CN đã được thể hiện rõ trong Thông tư số 05/2009/TT-BKHCN về “Hệ thống chỉ tiêu thống kê ngành KH&CN” do Bộ KH&CN ban hành ngày 30/3/2009. Thông tư quy định rõ, nhóm chỉ tiêu thống kê về “Công bố KH&CN” bao gồm ba chỉ tiêu:

MÃ SỐ	NHÓM, TÊN CHỈ TIÊU	PHÂN TỔ CHỦ YẾU	KỲ CÔNG BỐ
	6. CÔNG BỐ KHOA HỌC VÀ CÔNG NGHỆ		
0601	Số bài báo KH&CN của Việt Nam công bố trong nước và ngoài nước	Lĩnh vực nghiên cứu Nơi công bố (trong nước, ngoài nước)	Năm
0602	Số bài báo KH&CN của Việt Nam công bố ở nước ngoài được trích dẫn	Lĩnh vực nghiên cứu Số lần trích dẫn	Năm
0603	Số giải thưởng KH&CN trong nước và quốc tế được trao tặng; Số người được trao tặng giải thưởng KH&CN trong nước và giải thưởng quốc tế	Loại giải thưởng Lĩnh vực nghiên cứu	Năm

Bài viết này nhằm cung cấp cho bạn đọc thông tin chi tiết và luận cứ khoa học liên quan trực tiếp đến việc xác định chỉ tiêu thống kê về công bố KH&CN nói trên.

2. Phương pháp xác định chỉ tiêu thống kê về công bố KH&CN: trắc lượng thư mục, hệ số tác động, chỉ số trích dẫn

2.1. Trắc lượng thư mục

Đánh giá năng lực khoa học của một nước là việc làm không đơn giản. Sản phẩm của nghiên cứu khoa học - nhất là nghiên cứu khoa học cơ bản - thường mang tính trừu tượng, có độ trễ và ít khi gắn trực tiếp với người tiêu dùng. Trong nhiều trường hợp, kết quả nghiên cứu khoa học được áp dụng vào cuộc sống sau khi công trình nghiên cứu hoàn tất đã vài ba chục năm.

Sở dĩ đánh giá chất lượng nghiên cứu rất khó thực hiện vì cộng đồng khoa học vẫn chưa nhất trí về một chuẩn mực thống nhất cho tất cả các lĩnh vực nghiên cứu. Tuy nhiên, một số chỉ tiêu xếp hạng khả năng khoa học giữa các nước cũng đã được đa số chấp nhận. Trong đó, chỉ tiêu quan trọng nhất là số lượng bài báo khoa học và chất lượng nghiên cứu. Số lượng bài báo khoa học phản ánh “sản lượng” của một nền khoa học và mức độ đóng góp vào kho tri thức toàn cầu của một nước. Hiện nay, trên thế giới có hơn 100.000 tạp chí khoa học với chất lượng rất khác nhau, do đó, chỉ một phần trong số đó được cộng đồng khoa học quốc tế công nhận là nghiêm chỉnh và đáng tin cậy, và chúng thường được ghi trong danh mục các CSDL của Web of Science

và Scopus. Ví dụ, năm 2003, trong tổng số 4497 tạp chí khoa học của Trung Quốc, chỉ có 67 tạp chí được công nhận và liệt kê trong danh mục của Viện Thông tin Khoa học (ISI), Philadelphia, Hoa Kỳ. Hằng năm, tập đoàn Thomson thu thập tất cả các bài báo trên thế giới, kể cả các chi tiết như tên và địa chỉ của tác giả, nước xuất phát, đề tài, lĩnh vực nghiên cứu, năm công bố, số lần trích dẫn.

Về cơ bản, các chỉ tiêu về công bố KH&CN được xác định bằng phương pháp trắc lượng thư mục. Kết quả nghiên cứu khoa học được đánh giá khách quan và rộng rãi thông qua số lượng bài báo và số lần trích dẫn của một hay một nhóm tác giả (theo nhóm nghiên cứu, viện nghiên cứu, trường đại học, công ty) trong một khoảng thời gian nhất định. Sự kết nối trong nghiên cứu giữa các đơn vị nghiên cứu khoa học được đánh giá thông qua phân tích đồng tác giả và trích dẫn. Trắc lượng thư mục là sự ứng dụng các phương pháp toán và thống kê đối với các công bố khoa học với giả thiết rằng, hầu hết những phát kiến khoa học và kết quả nghiên cứu cuối cùng sẽ được công bố trên các tạp chí khoa học uy tín, nơi mà các kết quả này sẽ được đọc và trích dẫn trong các công trình nghiên cứu khác.

Trắc lượng thư mục có thể đánh giá mức độ ảnh hưởng về khoa học của một cá nhân hoặc đánh giá quan hệ hợp tác nghiên cứu giữa các nhà khoa học hay các đơn vị nghiên cứu (ví dụ: viện nghiên cứu – trường đại học – công ty). Trắc lượng thư mục không chỉ giới hạn ở việc

thu thập số liệu về số bài báo khoa học và các xuất bản phẩm khoa học, phân loại theo tổ chức, theo lĩnh vực nghiên cứu, nơi xuất bản để xác định năng suất NC&PT mà còn bao hàm cả số lần trích dẫn, tác giả trong nước, tác giả nước ngoài để đánh giá chất lượng nghiên cứu khoa học.

Hiện nay, hầu hết các số liệu về trích dẫn khoa học đều do các công ty dịch vụ thông tin hoặc các hiệp hội nghề nghiệp thu thập, xử lý và cung cấp. Nguồn cấp chủ yếu của số liệu này là những CSDL chỉ dẫn trích dẫn của ISI (như CSDL SCI – Science Citation Index, SSCI – Social Science Citation Index, và AHCI – Arts and Humanities Citation Index) của Viện Thông tin Khoa học (ISI), Philadelphia, Hoa Kỳ hay CSDL Scopus của Nhà xuất bản Elsevier. Các CSDL của ISI trên Web of Science là công cụ cho phép tìm kiếm và đánh giá chất lượng các công trình khoa học trên cơ sở các trích dẫn khoa học từ hơn 8.600 tạp chí nổi tiếng thế giới, trong đó, 6.100 tạp chí thuộc lĩnh vực khoa học tự nhiên, khoa học công nghệ và 1.790 tạp chí thuộc lĩnh vực khoa học xã hội và nhân văn. Web of Science cho phép thực hiện các lệnh tìm đặc biệt về chỉ dẫn trích dẫn (citation index); xác định những bài báo và lĩnh vực được trích dẫn nhiều nhất; hỗ trợ việc đánh giá tác động của tạp chí hoặc bài báo; xác định vị trí của từng tổ chức nghiên cứu và của từng quốc gia trong lĩnh vực KH&CN (các CSDL này thường chỉ bao gồm tài liệu tiếng Anh).

Trắc lượng thư mục là một phương

pháp đánh giá dòng chảy tri thức quốc gia, là công cụ định lượng kết quả/sản phẩm của các chương trình, đề tài nghiên cứu khoa học, đang được sử dụng rộng rãi trên thế giới. Hướng tới mục tiêu này, một số quốc gia cũng đã tiến hành xây dựng các CSDL trích dẫn khoa học của riêng mình như Nhật Bản (Citation Database for Japanese Papers – CJP từ năm 1995), Trung Quốc (Chinese Science Citation Database; China Scientific and Technical Papers and Citations Database; Chinese Social Science Citation Index). Ở nước ta, Cục Thông tin KH&CN Quốc gia (NASATI) cũng tiến hành thu thập và xử lý thư mục một số lượng khá lớn và đầy đủ nhất Việt Nam các bài báo từ tạp chí KH&CN trong nước trong nhiều năm qua, một phần trong số đó đã được số hóa trong thư viện điện tử của Cục.

2.2. Hệ số tác động

Để đánh giá chất lượng một công trình nghiên cứu khoa học, một tiêu chí quan trọng là công trình đó được công bố trên tạp chí nào? Chất lượng và độ tin cậy của tạp chí đó được đánh giá như thế nào? Nói cách khác, hệ số tác động (Impact Factor – IF) của tạp chí là bao nhiêu?

Hệ số tác động do E. Garfield đề xuất năm 1955 và là cơ sở để xây dựng SCI ở Hoa Kỳ vào đầu những năm 1960. Hệ số tác động của một tạp chí được tính toán bằng cách chia số lần được trích dẫn của các bài báo trong năm hiện tại và số bài báo được đăng trên tạp chí đó trong vòng 2 năm trước đó. Ví dụ, trong hai năm 2003 và 2004, tạp chí y học

“Lancet” đã công bố 450 bài báo khoa học, trong năm 2005 đã có 10.500 bài báo khác trích dẫn 450 bài báo đó, như vậy Lancet có hệ số $IF = 10.500/450 = 23,3$. Những công trình nghiên cứu được công bố trên các tạp chí có hệ số tác động cao đồng nghĩa với chất lượng cao của công trình nghiên cứu. Tuy nhiên, hệ số tác động chỉ phản ánh uy tín của tạp chí nhưng không phản ánh chất lượng của một bài báo cụ thể. Ngoài ra, hệ số tác động của tạp chí cũng phụ thuộc vào bộ môn khoa học (chẳng hạn, các tạp chí thuộc khoa học thực nghiệm thường có hệ số tác động cao hơn các tạp chí ngành khoa học tự nhiên và toán học).

2.3. Chỉ số trích dẫn khoa học

Thành quả khoa học phải được lượng hóa, cũng như chất lượng của công trình nghiên cứu phải được xem xét. Vấn đề trở nên nan giải khi tìm một công thức quân bình giữa hai yếu tố số lượng và chất lượng. Trong quá khứ, thành quả của một nhà khoa học thường được đánh giá theo: số lượng bài báo khoa học công bố; tổng số lần trích dẫn các bài báo; và hệ số tác động của tạp chí khoa học. Số lượng bài báo công bố trên các tạp chí quốc tế thường được xem là một thước đo về tính tích cực và năng suất nghiên cứu của một nhà khoa học. Tuy nhiên, số lượng bài báo chỉ phản ánh phần lượng vì một nhà nghiên cứu có thể công bố nhiều bài báo khoa học, nhưng chất lượng nghiên cứu có thể không cao. Vậy lấy gì để đánh giá “chất lượng” nghiên cứu của một nhà khoa học? Trong khi chưa có một thước

đo hoàn chỉnh, người ta thường dựa vào hệ số tác động (IF) của tạp chí mà nhà khoa học công bố. Tuy nhiên, vì IF chỉ phản ánh chất lượng của tạp chí, nên phải tìm một chỉ số khác phản ánh chất lượng nghiên cứu của cá nhân nhà khoa học. Một trong những chỉ số hấp dẫn là chỉ số trích dẫn trung bình (average citation) của cá nhân nhà khoa học. Chỉ số trích dẫn trung bình được tính bằng cách lấy tổng số lần trích dẫn chia cho số lượng bài báo khoa học của một tác giả. Việc diễn giải chỉ số trích dẫn trung bình đôi khi cũng gặp khó khăn. Theo phân tích của ISI, trong tất cả các bài báo khoa học công bố trên thế giới, có khoảng 55% các bài không bao giờ được ai (kể cả chính tác giả) trích dẫn sau 5 năm công bố.

2.4. Chỉ số H – đánh giá năng lực nghiên cứu của cá nhân nhà khoa học

Chỉ số H (chỉ số Hirsch, hay H index) được J. Hirsch (Đại học California San Diego, Hoa Kỳ) đề xuất năm 2005 như một chỉ tiêu khác của Hệ số tác động. Chỉ số H là thước đo về năng suất làm việc, ảnh hưởng của công trình nghiên cứu đến cộng đồng khoa học quốc tế và sự ghi nhận của đồng nghiệp trong ngành của một nhà khoa học.

Chỉ số H bổ sung cho cách tính tổng số lần trích dẫn đơn thuần và giúp phân biệt được những nhà khoa học có ảnh hưởng lớn với những người chỉ đơn thuần viết nhiều bài. Chỉ số H được nhiều trường đại học trên thế giới tham khảo, sử dụng khi xét phong tặng chức danh hàn lâm, xét cấp kinh phí nghiên cứu, thậm chí cả dự báo giải thưởng Nobel.

Theo J. Hirsch, chỉ số H được tính toán dựa vào số công trình công bố và số lần trích dẫn. Mục tiêu của chỉ số H là đo mức độ ảnh hưởng của một nhà khoa học. Chỉ số H của một nhà khoa học là H công trình trong số N công trình của nhà khoa học đó được trích dẫn ít nhất là H lần, và $(N - H)$ được trích dẫn dưới H lần. Ví dụ, nếu một nhà khoa học có chỉ số $H = 20$ có nghĩa là nhà khoa học này có 20 công trình nghiên cứu với mỗi công trình được trích dẫn ít nhất là 20 lần. Chỉ số H phản ánh thành quả tích lũy của một nhà khoa học. Một nhà nghiên cứu có thể công bố hàng trăm bài báo khoa học, nhưng nếu chỉ một số ít trong số này được trích dẫn thì chỉ số H vẫn không cao. Có thể nói, lợi thế lớn nhất của chỉ số H là chẳng những bao gồm hai yếu tố lượng và chất, mà còn quân bình được hai yếu tố này.

Tuy nhiên, do chỉ số H luôn tăng theo thời gian nên phụ thuộc vào độ tuổi của nhà nghiên cứu và thời gian làm nghiên cứu. Chẳng hạn, chỉ số H của một người đã làm nghiên cứu 35 năm thường cao hơn so với người mới làm khoa học. Chỉ số H cũng phụ thuộc vào ngành khoa học. Nói chung, các ngành khoa học tự nhiên và thực nghiệm (như vật lý, y sinh học) có xu hướng công bố nhiều công trình nghiên cứu và thường hay trích dẫn nhau hơn các ngành khoa học như toán học hay xã hội học. Do đó, một số hệ số điều chỉnh đã được nêu ra nhằm so sánh tương đối giữa các ngành khoa học với nhau. Tuy vẫn còn khiếm khuyết nhưng chỉ số H là một thước đo thành quả khoa học khách quan nhất so với các chỉ số

hiện nay. Việc ứng dụng chỉ số H cũng là một cách đưa hoạt động khoa học nước ta từng bước hội nhập quốc tế.

2.5. Chỉ số trích dẫn tương đối – đánh giá năng lực quốc gia

Để so sánh ở quy mô lớn hơn có thể dựa vào **Chỉ số trích dẫn tương đối (relative citation index - RCI)** là chỉ số phản ánh mức độ ảnh hưởng trung bình của một nước hay một khu vực. Chỉ số trích dẫn tương đối là tần suất trích dẫn của một quốc gia hoặc khu vực và được tính bằng cách lấy trung bình số lần trích dẫn của một bài báo của quốc gia đó rồi chia cho số lần trích dẫn trung bình toàn thế giới của các bài báo khoa học. Chỉ số RCI được tính như sau:

n_i : số bài báo xuất phát từ nước i ,

c_i : số lần trích dẫn từ những bài báo đó,

N : tổng số bài báo của tất cả các nước trên thế giới,

C : tổng số trích dẫn của tất cả các bài báo trên thế giới.

$p_i = n_i / N$: tỷ số bài báo cho từng nước p_i (so với tổng số trên thế giới),

$q_i = c_i / C$: tỷ số lần trích dẫn cho từng nước q_i

Chỉ số trích dẫn tương đối được ước tính sẽ là:

$$RCI_i = p_i / q_i.$$

Như vậy, nếu chỉ số $RCI_i = 1$ có nghĩa là chất lượng và ảnh hưởng của nghiên cứu khoa học của nước đó tương đương với chất lượng trung bình trên thế giới; nếu $RCI_i > 1$ hay $RCI_i < 1$ có nghĩa là chất lượng và ảnh hưởng của

Nghiên cứu - Trao đổi

nghiên cứu khoa học nước đó cao hơn hay thấp hơn trung bình thế giới. Một nước có thể công bố nhiều bài báo khoa học (như Trung Quốc chẳng hạn), nhưng ít ai trích dẫn hay quan tâm, và do đó chỉ số RCI rất thấp (0,27).

3. Các chỉ tiêu thống kê về công bố KH&CN ở Việt Nam: nội dung, phương pháp tính và cách thể hiện

3.1. Số bài báo khoa học của Việt Nam công bố trong nước và ngoài nước

Nội dung: phản ánh tổng số bài báo khoa học của tác giả người Việt Nam đã được đăng trên các tạp chí khoa học trong và ngoài nước. Bài báo khoa học (scientific paper) phải hàm chứa kết quả nghiên cứu của công trình khoa học và được công bố trên một tạp chí khoa học chuyên ngành (scientific journal) đã qua hệ thống bình duyệt (peer-review) của tạp chí.

Bài báo đăng trong các tạp chí quốc tế thì tạp chí phải thuộc loại có uy tín (có tên trong danh mục các tạp chí của ISI hoặc Scopus).

Bài báo đăng trong tạp chí trong nước thì tạp chí phải đáp ứng ba điều kiện: được Bộ Thông tin và Truyền thông cho phép xuất bản, hội đồng biên tập gồm các nhà khoa học thuộc chuyên ngành khoa học của tạp chí, bài báo được bình duyệt trước khi đăng.

Phương pháp tính và cách thể hiện nội dung:

a. Phạm vi thu thập số liệu: toàn bộ các bài báo khoa học của tác giả người Việt Nam được đăng trong các tạp chí

khoa học trong và ngoài nước trong kỳ báo cáo hoặc khoảng thời gian điều tra.

b. Thời kỳ thu thập số liệu: hết ngày 31/12 hàng năm đối với báo cáo thống kê hoặc theo thời kỳ điều tra của các cuộc điều tra KH&CN.

c. Cách thể hiện:

Đơn vị: số lượng bài báo.

Chỉ tiêu này được chia theo các phân tổ chính là:

- Lĩnh vực nghiên cứu: ghi rõ đến cấp III theo Bảng phân loại lĩnh vực nghiên cứu KH&CN do Bộ KH&CN ban hành kèm theo quyết định số 12/2008/QĐ-BKHCN ngày 04/09/2008 về ban hành một số bảng phân loại thống kê KH&CN.

- Nơi công bố: ghi rõ nơi công bố là trong hay ngoài nước.

3.2. Số bài báo khoa học của Việt Nam công bố ở nước ngoài được trích dẫn

Nội dung: phản ánh số bài báo khoa học của tác giả người Việt Nam công bố ở nước ngoài được trích dẫn và số lần được trích dẫn. Bài báo phải được đăng trong các tạp chí có uy tín (có tên trong danh mục các tạp chí của ISI hoặc Scopus).

Như đã trình bày ở trên, năng lực khoa học của cá nhân nhà nghiên cứu hoặc tổ chức nghiên cứu được thể hiện gián tiếp qua số lần trích dẫn các bài báo đã công bố của cá nhân nhà khoa học/tổ chức nghiên cứu đó. Số liệu này có thể được rút ra từ CSDL “Chỉ dẫn Trích dẫn Khoa học” của Viện Thông tin Khoa học Hoa

Nghiên cứu - Trao đổi

Kỳ hay từ CSDL Scopus của Nhà xuất bản Elsevier.

Phương pháp tính và cách thể hiện nội dung

a. *Phạm vi thu thập số liệu:* toàn bộ các bài báo khoa học của tập thể/cá nhân người Việt Nam được đăng trong các tạp chí khoa học quốc tế trong kỳ báo cáo hoặc khoảng thời gian điều tra.

b. *Thời kỳ thu thập số liệu:* tính đến hết ngày 31/12 hằng năm.

c. *Cách thể hiện:*

Đơn vị: số lượng bài báo được trích dẫn, số lần trích dẫn của mỗi bài.

Chỉ tiêu này được chia theo các phân tổ chính là:

- Lĩnh vực nghiên cứu: ghi rõ đến cấp III theo Bảng phân loại lĩnh vực nghiên cứu KH&CN do Bộ KH&CN ban hành kèm theo quyết định số 12/2008/QĐ-BKHCN ngày 04/09/2008 về ban hành một số bảng phân loại thống kê KH&CN.

- Số lần trích dẫn: tổng số lần bài báo đã công bố của tập thể/cá nhân được trích dẫn bởi các bài báo khác trên các tạp chí khoa học quốc tế.

4. Kết luận

Chỉ tiêu về công bố khoa học và công nghệ là một công cụ quan trọng để đánh giá tính hiệu quả của các chương trình nghiên cứu, ghi nhận và đánh giá hiệu quả nghiên cứu của từng nhà khoa học cho đến cơ quan nghiên cứu và lĩnh vực nghiên cứu trong tiến trình phát triển và hội nhập. Xác định được chỉ tiêu công bố khoa học và công nghệ là xây dựng được một công cụ đánh giá có tính định lượng đối với kết quả nghiên cứu của từng nhà khoa học, từng tổ chức khoa học và từng chuyên ngành khoa học, giúp ích cho công tác quản lý khoa học và công nghệ sát sao và hiệu quả hơn, đồng thời đáp ứng nhu cầu thông tin thống kê KH&CN của xã hội.

Tài liệu tham khảo

1. Garfield E., 2002. The impact factor [internet] Current Contents 1994 20; 3-7
2. Garfield E., 2006. The History and Meaning of the Journal Impact Factor. Journal of the American Medical Association (JAMA), (293): 90-93.
3. Hirsch J. E., 2005. An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA 2005;102(46):16569-16572
4. Nguyễn Văn Tuấn, 2005. Thế nào là một "bài báo khoa học"? , Tia Sáng số tháng 12/2005.
5. Nguyễn Văn Tuấn, 2007. Đánh giá nghiên cứu khoa học: không nên chỉ căn cứ vào số lượng ấn phẩm! Vietsciences-04/07/2007
6. Nguyễn Văn Tuấn, Nguyễn Đình Nguyên, 2008. Chất lượng nghiên cứu khoa học ở Việt Nam qua chỉ số trích dẫn. Vietsciences-10/2008
7. Phạm Duy Hiển, 2006. Nhìn lại 10 năm công bố ấn phẩm khoa học Việt Nam: các nhà khoa học nghĩ gì qua những bài báo đăng trên tạp chí quốc tế?, Tạp chí Hoạt động Khoa học, số tháng 1 năm 2006.
8. Phạm Duy Hiển, 2008. Đánh giá ảnh hưởng trong nghiên cứu khoa học qua chỉ số H. Tia Sáng, 03/06/2008.
9. Phelan TJ, 1999. A compendium of issues for citation analysis. Scientometrics 1999; 45:117-36.
10. Yoshiko Okubo, 1997. Bibliometric indicators and analysis of research systems: Methods and examples STI working papers 1997, OECD.