

Đánh giá định lượng kết quả nghiên cứu khoa học

Hồ Tú Bảo
Trường Khoa học Tri thức
Viện Khoa học và Công nghệ Tiên tiến Nhật Bản

1

Lý do làm báo cáo này

- Chia sẻ một số thông tin tôi tìm hiểu thời gian qua về câu chuyện đánh giá kết quả nghiên cứu khoa học.
- Cùng trao đổi một số suy nghĩ về việc dùng đánh giá định lượng một cách thích hợp trong hoàn cảnh cụ thể của ta.

2

Nội dung

- Mở đầu
- Về các độ đo
 - Chỉ số trích dẫn
 - Hệ số ảnh hưởng của tạp chí
 - Chỉ số H
- Một vài nhận xét và ý kiến



3

Ba loại hình nghiên cứu

Nghiên cứu cơ bản

- Tìm tri thức khoa học nền tảng về thiên nhiên và xã hội.
- Chứng minh ước đoán Poincaré trong toán học hay việc xác định các gen gây ra bệnh tật.

Nghiên cứu ứng dụng

- Tìm tri thức khoa học mới và cần cho nhu cầu thực tiễn.
- Cách chẩn đoán cấp độ bệnh viêm gan dựa trên xét nghiệm máu nhưng không làm sinh thiết.

Nghiên cứu phát triển

- Tìm tri thức mới để làm ra các sản phẩm cụ thể (thường của doanh nghiệp).
- Hãng Microsoft nghiên cứu làm hệ điều hành máy tính Windows 7.

Sản phẩm cơ bản (bài báo, sách)



Sản phẩm ứng dụng và các ấn phẩm khoa học (bài báo, báo cáo kỹ thuật)



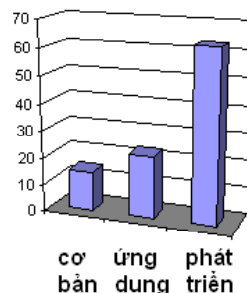
Sản phẩm chủ yếu (sản phẩm thương mại, công nghệ mới)



4

Kinh phí cho các loại hình nghiên cứu ở Nhật

	Nghiên cứu cơ bản	Nghiên cứu ứng dụng	Nghiên cứu phát triển
Đại học và cao học	55,1	35,8	9,1
Viện phi lợi nhuận	20,3	35,8	43,9
Viện nghiên cứu công	24,4	29,6	46,0
Doanh nghiệp	6,3	19,6	74,1
Tỷ lệ chung	14,3	22,8	62,9



5

Đánh giá ấn phẩm khoa học

- Đánh giá số lượng
- Đánh giá chất lượng ấn phẩm
 - Đánh giá chủ quan (peer review, đánh giá định tính)
 - Đánh giá khách quan (metrics-based, đánh giá khách quan)
- Độ đo định lượng tiêu biểu:
 - Chỉ số trích dẫn (citation index)
 - Hệ số ảnh hưởng (impact factor)
 - Chỉ số H (h-index).



6

Đo gì và cho ai?

- Tổng số bài báo (total paper)
- Tổng số trích dẫn (total citations)
- Trích dẫn của từng bài báo (cites per paper)
- Tỷ lệ bài được trích dẫn (percent cited paper)
- Ảnh hưởng tương đối trong ngành (impact relative to field)
- Xếp hạng trong ngành (percentile rank in field)
- Chỉ số hợp tác (collaboration indicators)
- Số trích dẫn trung bình (expected citation count)
- Tỷ lệ trích dẫn trên trích dẫn trung bình (ratio of citations to expected citation count)
- Chỉ số H (H index)
- Phân bố tần suất trích dẫn (citation frequency distribution)
- etc.

- Tác giả Authors
- Đơn vị Institutions
- Quốc gia Nations
- Chủ đề Topics
- Lĩnh vực Fields



7

Cơ sở dữ liệu thông tin khoa học

- Cơ sở dữ liệu ISI (Institute for Scientific Information, Thomson-Reuteur 4.2008)
- Cơ sở dữ liệu SCOPUS (Elsevier)
- Cơ sở dữ liệu Google Scholar (Google)
- Hơn 100 cơ sở dữ liệu và công cụ khác cho phép tìm kiếm chỉ số trích dẫn, như arXiv, CiteSeer, ScienceDirect, SciFinder Scholar, MEDLINE, ...



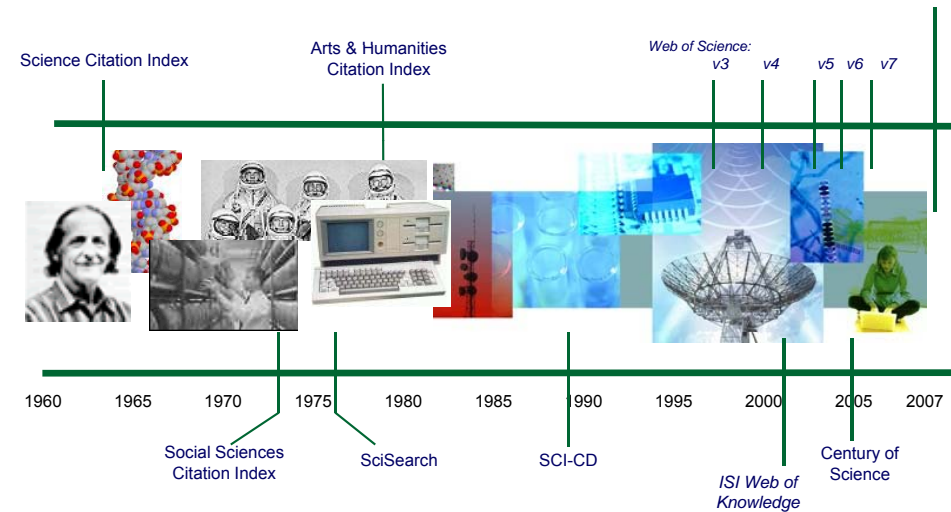
8

Cơ sở dữ liệu ISI

- **SCI** (Science Citation Index): từ 1964, 3773 tạp chí của 100 ngành; và **SCIE** (SCI mở rộng, Science Citation Index Expanded) với 8207 tạp chí của 150 ngành.
- **SSCI** (Social Sciences Citation Index): từ 1973, hiện có 2697 tạp chí và 3500 công trình của 50 ngành;
- **A&HCI** (Arts & Humanities Citation Index): từ 1978, hiện có 1470 tạp chí và 6000 công trình khác.
- **CPCI** (Conference Proceedings Citation Index): 110,000 tuyển tập 256 ngành khoa học tự nhiên, xã hội nhân văn
- Chuyển lên mạng (1997) dưới tên **Web of Science** (<http://isiwebofknowledge.com>).

9

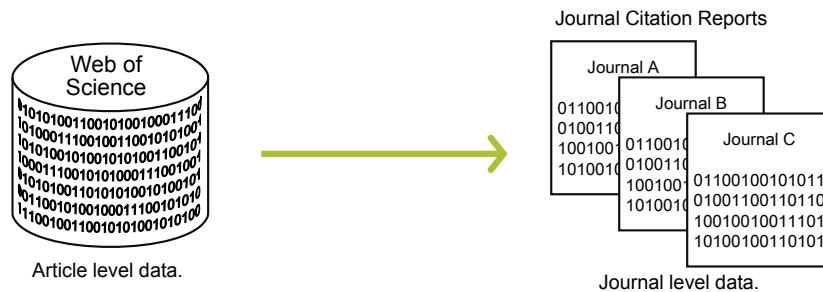
Lịch sử Web of Science



(from David Horky, Thomson-Reuteurs)

10

Journal citation reports



Web of Science chứa thông tin thư mục và trích dẫn các bài báo trong các tạp chí, có thể được xử lý để tạo ra nhiều độ đo đánh giá các tạp chí, như đo mức ảnh hưởng, xác định xu thế, phân và đặc trưng các tạp chí.

(from Thomson Reuters: Journal Citation Reports® and the Impact Factor)

11

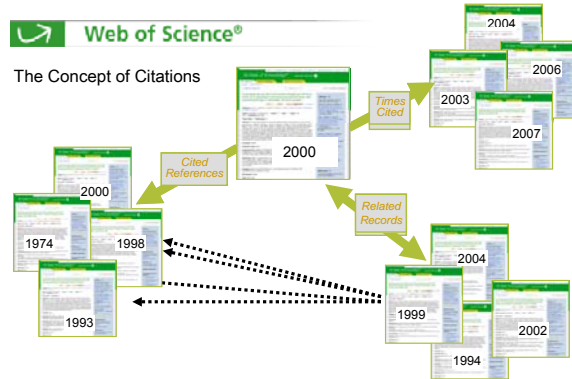
Scopus and Google Scholar

- **SCOPUS** (<http://info.scopus.com>, từ 2004)
 - 16.500 tạp chí,
 - 600 ấn phẩm nghề nghiệp,
 - 350 loạt sách chuyên khảo, ~3,6 triệu bài báo từ hội nghị.
- **Google Scholar** (<http://scholar.google.com>, từ 2005)
 - hầu hết các tạp chí của nhà xuất bản lớn Âu-Mỹ,
 - báo cáo kỹ thuật, luận văn, sách và nhiều loại tài liệu khác.

12

Chỉ số trích dẫn

- Số lần được trích dẫn, tham khảo trong các ấn phẩm khác (Garfield, 1955)



(Eugene Garfield, "Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas", *Science*, 122, 3159 (July 15, 1955):108)

13

Chỉ số trích dẫn

Một số đặc điểm

- Chỉ có ý nghĩa so sánh trong từng ngành khoa học.
- Thường khác nhau và có sai số khi tính từ các nguồn khác nhau
- Số lượng bài báo cũng như số citations của các ngành "to-nhỏ" là rất khác nhau (commutative algebra vs. machine learning).
- Không dễ xác định và phân biệt self-citation.
- Một số ngành như arts, humanities hay math, citations thường đến muộn, chậm nhưng trong thời gian dài.
- Ở ngành khác như life sciences, citations thường nhanh, đạt đỉnh trong vòng 2-3 năm ... rồi thôi.



14

Chỉ số trích dẫn

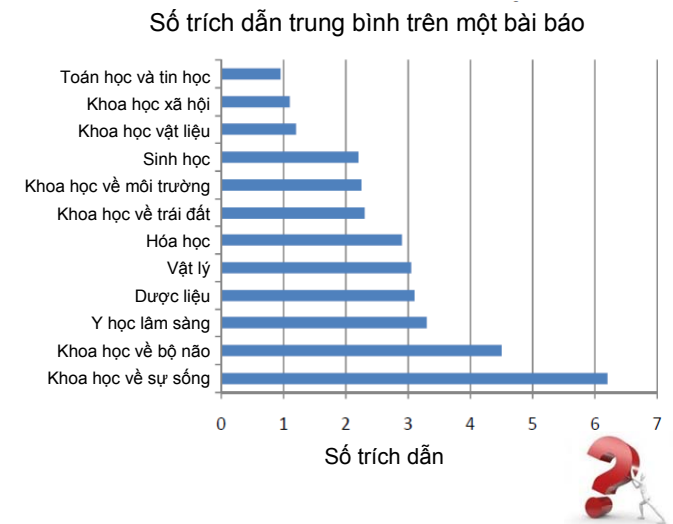
- Chỉ số trích dẫn chỉ có ý nghĩa so sánh trong từng ngành khoa học.
- Chỉ số trích dẫn tính từ các nguồn khác nhau thường khác nhau và có sai số.



15

Chỉ số trích dẫn

- Khác biệt về "văn hóa ngành".
- Phụ thuộc vào bản chất, cách làm và cách công bố.
- Khác nhau giữa nghiên cứu lý thuyết và thực nghiệm.



16

(http://www.elsevier.com/framework_editors/pdfs/Perspectives1.pdf)

Chỉ số trích dẫn

- Chỉ số trích dẫn tính được từ các nguồn khác nhau thường khác nhau và có lỗi.

- Sách "Quantum Computation and Quantum Information" (2000, Cambridge Univ. Press) có trích dẫn
 - 2800 citations từ Web of Science
 - 3150 từ Scopus,
 - 4310 từ Google (gấp 1,53 so với WoS)
- 283 bài báo từ 3 tạp chí y học có trích dẫn
 - 68088 từ Web of Science
 - 82076 từ Scopus
 - 83538 từ Google Scholar 1,226 lần cao hơn WoS)

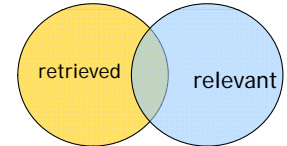
(Kulkarni, A.V., et al., "Comparisons of Citations in Web of Science, Scopus, and Google Scholar for Articles Published in General Medical Journals", JAMA, 302(10), 2009.)

17

Chỉ số trích dẫn

- Precision (độ chính xác) và recall (khả năng tìm hết)
- Lỗi tính toán thường do tên người trùng hoặc giống nhau.

- Nguyễn Anh Tuấn có 80 bài báo được trích dẫn
- Nguyễn Anh Tuấn (khác) có 5
- Nguyễn Anh Tuấn có 35 bài
- Google tìm được 100 bài được trích dẫn với "N.A. Tuan", trong số này có 60 bài của Nguyễn Anh Tuấn
- Precision = $60/100 = 0.6$
- Recall = $60/80 = 0.75$



$$\text{precision} = \frac{\text{retrieved} \cap \text{relevant}}{\text{retrieved}}$$

$$\text{recall} = \frac{\text{retrieved} \cap \text{relevant}}{\text{relevant}}$$

retrieved = {retrieved documents}
relevant = {relevant documents}

18

Hệ số ảnh hưởng của tạp chí (IF)

- IF của tạp chí T trong năm N tính bằng $IF = A/B$,
 - A = số trích dẫn năm N đến bài của T trong hai năm ngay trước N
 - B = tổng số các bài đăng trên T trong hai năm này.

$$IF_{2009} = \frac{\text{so trích dẫn trong 2009 đến bài của tạp chí đang trong 2007 và 2008}}{\text{so bài của tạp chí đang trong năm 2007 và 2008}}$$

- IF 2009 là 7,180: trung bình mỗi bài đăng năm 2007 và 2008 được trích dẫn 7,180 lần trong 2009
- IF có thể thay đổi nhiều theo thời gian (Bioinformatics)

2008	2007	2004	2003	2002	2001
4,328	4,894	5,742	6,701	4,615	3,421

19

Hệ số ảnh hưởng khác nhau giữa các ngành

Y học		Toán học	
(1) A Cancer Journal for Clinicians	74,575 [1]	(1) Communications on Pure and Applied Mathematics	3,806 [711]
(2) The New England Journal of Medicine	50,017 [2]	(2) Bulletin of the American Mathematical Society	3,5 [851]
...		...	
(100) Epilepsia	3,733 [739]	(100) Monatshefte Fur Mathematik	0,584 [5248]

Số trong (x) và [y] là thứ tự của IF trong chuyên ngành và toàn bộ các ngành.

- Thứ tự IF không luôn đồng nhất với uy tín của tạp chí trong chuyên ngành (Annals of Mathematics uy tín nhất trong Toán)
- Có thể sai gì khi so sánh thô sơ IF giữa các ngành?

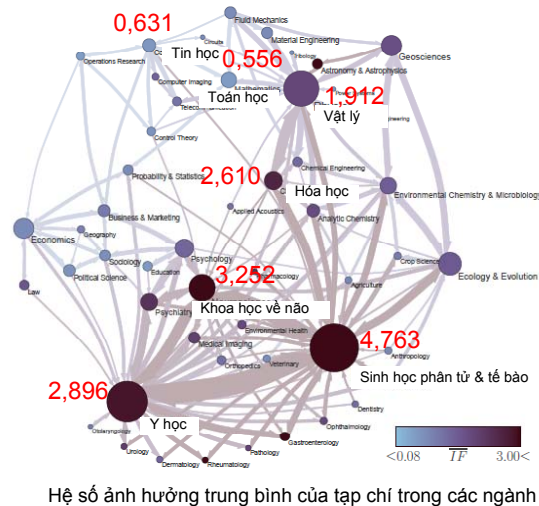
(Journal Citation Reports của Web of Science, 2008)



20

Hệ số ảnh hưởng trung bình theo ngành

“Giả sử IF tăng tuyến tính trong các ngành, thì có thể nói nôm na rằng IF 9,5 ngành sinh học phân tử và tế bào có ảnh hưởng ở ngành này quăng như ảnh hưởng của một tạp chí có IF 6 trong ngành y, tạp chí IF 4 trong vật lý hay một tạp chí có IF 1,2 trong toán và tin học.”



21

Vài đặc điểm của hệ số ảnh hưởng

- Giá trị trung bình về ảnh hưởng của các bài trong tạp chí
 - Theo thứ tự trích dẫn: 15% bài đầu góp 50% số trích dẫn, 50% bài đầu góp 90% số trích dẫn, các bài ở nhóm 50% đứng đầu có trích dẫn khoảng 10 lần hơn các bài ở nhóm 50% cuối.
- Tính IF trong 2 năm sau công bố là ngắn và sớm đối với một số ngành → JCR thêm IF tính trong 5 năm.
- Chưa tính bài được dùng nhưng không được trích dẫn
- Biến động đáng kể từ năm này qua năm khác
- Quăng 10-15 nghìn trên 110 nghìn tạp chí tham gia vào việc tính hệ số ảnh hưởng. Tạp chí không dùng tiếng Anh và các ngành “nhỏ” ít được tính đến IF.

22

Ý kiến chuyên gia

(Philip Campbell, Editor-in-chief of *Nature*)

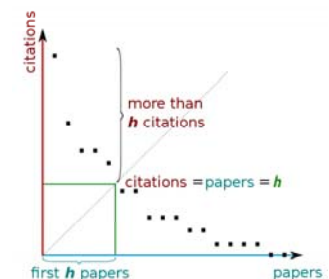
- “Nature’s IF for 2004 was 32,2 ... 89% of our impact factor was generated by just 25% of our papers.”
- “I am concerned by the tendency within academic administrations to focus on a journal’s impact factor when judging the worth of scientific contributions by researchers, affecting promotions, recruitment and, in some countries, financial bonuses for each paper.”
- “Although the current system may be effective at measuring merit on national and institutional scales, the most effective and fair analysis of a person’s contribution derives from a direct assessment of individual papers, regardless of where they were published.”

(“Escape from the Impact Factor”, *Ethics in Science & Env. Politics*, 8, 5-7, 2008)

23

H-index

- Cân bằng số lượng và chất lượng
- Một người có chỉ số H là N nếu có N bài báo với chỉ số trích dẫn ít nhất là N, và các bài còn lại có chỉ số trích dẫn nhiều nhất là N.
 - Một người công bố 40 bài báo, 9 bài có nhiều hơn 10 trích dẫn, 5 bài có 10 trích dẫn và 26 bài còn lại có ít hơn 10 trích dẫn, thì chỉ số H của người này là 10.
- Thử ở <http://delab.csd.auth.gr/~lakritid/index.php?lan=1&s=2>



Hirsch, J.E., “An index to quantify an individual’s scientific research output”, *PNAS* **102**, 2005.

24

Đặc điểm của chỉ số H

- Chỉ số H có ý nghĩa hơn với những người thuộc cùng một ngành, không luôn thích hợp khi so sánh khác ngành.
- Không phân biệt được sự đóng góp khác nhau của các tác giả trong một công trình. Bốn kiểu thứ tự tác giả:
 - Thứ tự đóng góp do các tác giả xác định (định tính);
 - Thứ tự chữ cái của tên khi các tác giả xem đóng góp mọi người như nhau;
 - Thứ tự “đầu-cuối”, nhấn mạnh tác giả đầu và cuối là quan trọng nhất, và những người ở giữa theo đóng góp;
 - Thứ tự đóng góp do các tác giả xác định một cách định lượng,

Tscharntke et al. “Author Sequence and Credit for Contributions in Multiauthored Publications”, *PLoS Biology*, 5(1), 2007.

25

Đặc điểm của chỉ số H

- Precision và recall thường không cao khi tính tự động do chưa phân biệt rõ tên người
- Không đánh giá được công trình ảnh hưởng rất lớn
- Không phân biệt được lý do trích dẫn một bài báo:
 - (a) Ghi nhận công trạng của tác giả
 - (b) Thể hiện sự kính trọng tác giả
 - (c) Nói về phương pháp liên quan
 - (d) Cung cấp thông tin nền có ích
 - (e) Trích dẫn để phê phán
 - (f) Trích dẫn để làm cơ sở cho bài báo
- Chỉ số G: đề cao những bài báo có rất nhiều trích dẫn

Lawani, “Citation Analysis and the Quality of Scientific Productivity”, *BioScience*, 27, 1977

26

Một số nhận xét và trao đổi

1. Cần hiểu rõ đặc điểm và dùng đúng (hợp lý) các độ đo đánh giá định lượng
 - Tránh dùng sai khi đánh giá các cá thể
2. Nên dùng kết hợp nhiều độ đo đánh giá định lượng
3. Cần có thêm thẩm định của chuyên gia khi cần
 - Chưa phân biệt tự trích dẫn
 - Trích dẫn để phê phán
 - Đóng góp của tác giả trong bài báo



27

Một số nhận xét và trao đổi

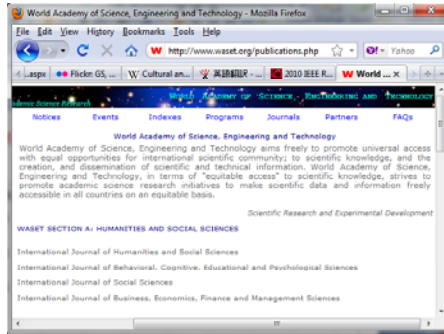
4. Cần chú ý đến ảnh hưởng của khác biệt văn hóa ngành lên các độ đo và số lượng ấn phẩm
 - (a) Khác nhau về độ dài (NB Châu 197 trang)
 - (b) Khác nhau về số tác giả → khác về số lượng
 - (c) Khác nhau về vai trò của hội nghị khoa học, thí dụ nhiều hội nghị đầu ngành CNTT được đánh giá cao như (hơn) nhiều tạp chí.
- Còn khó trong đánh giá hội nghị
- ISI Conference Proceedings Citation Index
- Có thể (nên) tham khảo các danh sách “conference ranking”
- <http://core.edu.au/index.php/categories/conference%20rankings>

28

Một số nhận xét và trao đổi

5. Những “sân chơi” khác nhau và những “sân chơi” mới

- ❑ 90% bài báo trên tạp chí có trích dẫn = 0, 50% không bao giờ được ai khác đọc ngoài chính tác giả và những người phản biện.
- ❑ WASET (World Academy of Science, Engineering and Technology, <http://www.waset.org>)
- ❑ WSEAS (World Scientific and Engineering Academy and Society, <http://www.worldses.org/wseas.htm>).
- ❑ Hai khía cạnh? (Biết và Cần)



Kumar, M.J., “Evaluating Scientists: Citations, Impact Factor, h-Index, Online Hits and What Else?”, *IETI Technical Review*, 26(3), 165-168, 2009

29

Một số nhận xét và trao đổi

6. Không ngừng cải tiến các độ đo

- ❑ H index → G index, số lần download
- ❑ Hệ số riêng (eigenfactor) học từ Pagerank của Google.
- Cần luôn chú ý đến các cải tiến, các độ đo tốt hơn.

7. Thông tin khoa học ngày càng phong phú và minh bạch

- ❑ Dễ xác định đạo văn, nộp nhiều nơi, ...
- ❑ “... bia web vẫn còn trơ trơ ...”
- ❑ Đạo đức khoa học
- ❑ Sự nhanh nhạy của tuổi trẻ

30

Một số nhận xét và trao đổi

8. Về một số đánh giá tình hình nghiên cứu của ta

- ❑ Các nghiên cứu so sánh → có giá trị tham khảo
- ❑ SCImago Journal & Country Rank <http://www.scimagojr.com/index.php>
- ❑ Nghĩ và đánh giá chúng hợp lý?

	Số bài báo có ĐQQT	Tác giả nội địa chủ trì	Phần trăm nội địa
Bốn đại học hàng đầu Việt Nam	160	87	54
Viện Khoa học và Công nghệ Việt Nam	156	74	47
Toàn bộ số bài báo của Việt Nam	806	307	37
Đại học Chulalongkorn	869	715	82
Đại học Mahidol	817	598	73

Số công bố quốc tế năm 2008 của các tổ chức nghiên cứu và phát triển hàng đầu của Việt Nam và Thái Lan

Pham Duy Hien, “A Comparative Study of Research Capacities of East Asian Countries and Implication for Vietnam”, *Higher Education*, Springer, Feb. 2010.

31

Một số nhận xét và trao đổi

9. Về Quỹ NAFOSTED, SCI và SCIE

- ❑ Có thể rất nhiều bài cuối SCIE

10. Liên hệ với tình hình của ta

- ❑ Dùng các độ đo định lượng hợp lý: hiểu và dùng chúng đúng đắn, và hợp với tình hình khoa học của ta (ngưỡng)
- ❑ Nâng cấp một số tạp chí trong nước.
- ❑ Dùng để thay cho một phần cho cách “tính điểm” trong xét GS-PGS?

32

Kết luận

- Dù còn những hạn chế, tính được tự động các độ đo là một bước tiến nhiều ý nghĩa và việc sử dụng chúng đang là xu thế trong phát triển khoa học.
- Hiểu rõ và dùng hợp lý các độ đo định lượng để đánh giá kết quả nghiên cứu có tác dụng thúc đẩy phát triển khoa học và công nghệ của chúng ta.

<http://www.jaist.ac.jp/~bao/writinginvietnamese>



33

Tài liệu tham khảo chính

- Adler, R., et al. "Citation Analysis", *Statistical Science*, 24(1), 1-14, 2009.
- Althouse, B.M., et al., "Differences in Impact Factor Across Fields and Over Time", *American Society for Information Science and Technology*, 60(1), 27-34, 2009.
- Bakkalbasi, N., et al., "Three Options for Citation Tracking: Google Scholar, Scopus and Web of Science", *Biomedical Digital Libraries*, 3(7), 1-8, 2006.
- Bergstrom, C.T., "Eigenfactor: Measuring the Value and Prestige of Scholarly Journals", *College & Research Library News*, 68(5), 2007.
- Hồ Tú Bảo, "Tổ chức và quản lý đề tài nghiên cứu khoa học ở Nhật", *Tạp chí Tia Sáng*, 8.2008.
- Campbell, P., "Escape from the Impact Factor", *Ethics in Science & Environmental Politics*, 8, 5-7, 2008.
- Eugene Garfield, "Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas", *Science*, 122(3159), 1955.
- Harzing, A. K, Wal, R., "Google Scholar as a New Source for Citation Analysis", *Ethics in Science & Environmental Politics*, Vol. 8, 61-73, 2008.

34

Tài liệu tham khảo chính

- Hirsch, J.E., "An Index to Quantify an Individual's Scientific Research Output", *Proceedings of the National Academy of Sciences*, 102 (46), 16569-16572, 2005.
- Iglesias, J.E., Pecharromen C., "Scaling the h-index for Different Scientific ISI Fields", *Scientometrics*, 73(3), 303-320, 2007.
- Kulkarni, A.V., et al., "Comparisons of Citations in Web of Science, Scopus, and Google Scholar for Articles Published in General Medical Journals", *The Journal of the American Medical Association*, 302(10), 1092-1096, 2009.
- Kumar, M.J., "Evaluating Scientists: Citations, Impact Factor, h-Index, Online Hits and What Else?", *IETI Technical Review*, 26(3), 165-168, 2009.
- Lawani S.M. "Citation Analysis and the Quality of Scientific Productivity", *BioScience*, 27, 26-31, 1977.
- Meho, L.I., "The Rise and Rise of Citation Analysis", *Physics World*, 20(1), 32-36, 2007.
- Pham Duy Hien, "A Comparative Study of Research Capacities of East Asian Countries and Implication for Vietnam", *Higher Education*, Springer, Feb. 2010.
- Tschamtko, T., Hochberg, M.E., Rand, T.A., Resh, V.H., Krauss, J., "Author Sequence and Credit for Contributions in Multiauthored Publications", *PLoS Biology*, 5(1), 0013-0014, 2007.

35