ISSUE 8 NOVEMBER 2008

researchtrends

Welcome to the eighth edition of Research Trends. This issue takes a broad look at a number of approaches to citation analysis. We speak to Richard Klavans and Kevin Boyack, who have created a map of science, and to Jevin West, who is part of the Eigenfactor project team, which is using bibliometric data to tell some fascinating stories, including ranking journals, mapping science and revealing trends. We also ask two researchers why they believe their work is highly cited.

Hiroyuki Tomizawa at the OECD tells us about their recent decision to use Scopus data for research, analysis and benchmarking, and we analyze the publication output of various regions, focusing on the researcher to output ratio and ask why some countries appear to be more efficient than others.

We welcome your **feedback** to any of the topics covered.

Kind regards,
The Research Trends Editorial Board

DID YOU KNOW?

Book review: *The Impact Factor of Scientific and Scholarly Journals: Its Use and Misuse in Research Evaluation*

Compiled and edited by Braun, T. (2007) Scientometrics Guidebook Series, Volume 2. Hungary: Akadémiai Kiadó. ISBN: 978-963-05-8528-6. Smithsonian Institution Libraries.

Like the first volume in the Scientometrics Guidebook Series, this second volume draws content from past issues of the journal Scientometrics. The focus of this volume is the merits of the journal Impact Factor (IF) since its introduction as a statistical measure several decades ago. The IF is a standard measurement historically associated with the Institute for Scientific Information and is now being challenged by several statistical alternatives. These alternatives are recognized by many of the book's contributors (including the father of the IF himself, Eugene Garfield), who warn of the danger of using statistical data out of context and highlight the fact that the IF is only one of several measures that can be used to evaluate scholarship.

This book comes at an important time in the development of research metrics and serves as a historical look at a particular (and once nearly universal) measurement. The emergence of data-mining technologies will undoubtedly show new relationships between scholars and their publications and will alter many assumptions about the IF, which has achieved near iconic status in evaluating research output. However, despite imminent technological change in bibliometric methods and the fact that the book's contents have been previously published, this compilation is a worthy addition to the collection of any serious scientometrician.

Powered by SCOPUS™

ISSUE 8
NOVEMBER 2008

Custom Data fuels OECD's Innovation Strategy

Hiroyuki Tomizawa

The Organisation for Economic Co-Operation and Development (OECD) provides a forum for the governments of 30 like-minded market democracies to compare policy experiences, share best practices and seek answers to common economic, social and governance challenges.

Established in 1948 to lead the Marshall Plan for rebuilding Europe after the Second World War, the OECD has been collecting and analyzing statistical, economic and social data at the request of its members since 1961. These data are used to generate collective policy discussions, leading to decision-making and implementation. For instance, its Science, Technology and Industry Scoreboard, which comes out every two years, explores the interaction between knowledge and globalization at the heart of the ongoing transformation of OECD economies.

Impact of globalization on research

In recent years, the OECD has expanded its focus on its 30 member countries to offer analytical expertise and experience to over 100 developing and emerging market economies. This has to a certain extent been driven by globalization, which has made it virtually impossible to study specific areas in isolation. This has seen the scope of the OECD's work shift from the examination of individual policy areas within each member country to the analysis of how various policy areas interact with each other and with other countries, including those outside the OECD group.

The OECD has an ambitious publishing program, releasing a large amount of its research and accompanying data in 250 new titles every year. Through this output, the OECD aims to help governments foster sustainable economic growth, financial stability, trade, investment and innovation, while at the same time striving for environmental preservation, social equity and poverty reduction.

It has also come to realize that data alone are not enough; to truly help governments foster innovation, strategies are needed. To this end, the OECD is developing an Innovation Strategy, which will provide mutually reinforcing policies and recommendations to boost innovation performance, pointing to general and country-specific practices and, where appropriate, developing guidelines. This work will culminate in a report to ministers in 2010, but some patterns are already clear.

For instance, are governments doing enough to foster collaboration between universities and businesses, and not just within their borders? Many key inventions, such as the World Wide Web, have come from public research. Are governments doing enough to strengthen this bedrock of innovation?

The importance of a reliable data source

In October 2008, the OECD announced it had decided to use Scopus Custom Data in its research, analysis and benchmarking work. Hiroyuki Tomizawa, Principal Administrator in the Economic Analysis and Statistics Division of the Directorate for Science, Technology and Industry at the OECD, explains: "The three key factors behind this decision were the product's broad (international) coverage, clean, flexible data and advanced features, such as the ability to link between authors and institutions."

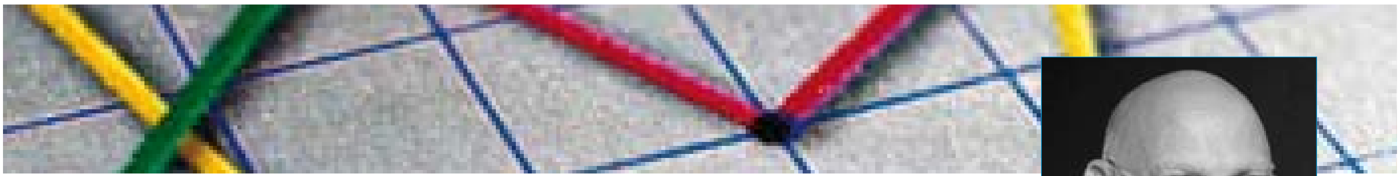
He adds that the OECD anticipates using the Scopus data for three main purposes:

- to analyze global trends and identify subject areas that are experiencing intense activity;
- to understand research activities at the country level in order to be able to make comparative analyses between countries;
- to understand co-authorship and collaboration across borders. In a competitive knowledge society, countries are deploying policies to attract the best talent, but it is not always easy for them to assess whether they were successful or not.

Three possible groups can benefit from the resulting OECD reports: policymakers, funding agencies, and governments and commercial research organizations. In this way, Scopus data will contribute to the OECD achieving its goals and will help to determine the direction of future economic decision-making.

More information on the OECD, including the full list of members and a wide range of publicly available reports, is available **here**.

Research Trends



Mapping unknown regions

Richard Klavans (pictured) and Kevin W. Boyack

A map of science is a diagram showing how different areas of science are related. The earliest maps tended to be hierarchical, starting with mathematics, then physics, chemistry and biology. Applied sciences would be like branches off this tree – electrical engineering branching off physics, chemical engineering branching off chemistry, and medicine and agricultural science branching off biology (with some chemistry).

But our analysis does not support the hypothesis that science is actually structured this way. We analyzed 20 maps of science. Two of these maps were made by experts, 17 were drawn through analysis of the citation patterns of millions of articles in thousands of peer-reviewed journals and one was based on course requirements at a university.

We found that science looks more like a circle than a hierarchy. Starting (arbitrarily) at mathematics, one can proceed through the areas mentioned above (physics, chemistry, engineering, earth sciences, biology, biotechnology, infectious diseases, medicine) and continue around the circle through health services, brain research, humanities, social sciences, computer science and back to mathematics (see figures 1 and 2) There isn't agreement about the order suggested here (some might put computer science next to biotechnology, others might put chemistry closer to medicine), but there is consensus that these are the most common connections for all of the maps we examined.

What is a map of science?

A map of science consists of a set of elements and the relationships between the elements. These elements can be any unit that represents a partition of science. Maps must have partitions, where science is separated into different parts, and these partitions must be linked, either explicitly (such as a line drawn between two partitions), or through proximate location (or physical adjacency) that explicitly denotes linkage.

Why use a circle?

Mapping science as a (non-hierarchical) circle is a useful aid for career counseling. A hierarchical map of science implies that one's path should always be aimed towards "central" areas (and correspondingly avoiding "peripheral" areas). A circle has the unique characteristic that there is no "center" (or, to say it more accurately, each point is the center). A circle illustrates what we need to communicate to a student – that many paths are equally valid.

Mapping science as a circle is also useful for understanding science policy. Governments support investments in science, just like one places weights on the edge of a wheel. Balancing the wheel of science reflects fundamental tradeoffs between supporting the arts, providing an understanding about how people behave, providing health and well-being to society, pursuing techno-economic goals, and supporting basic research which may have no immediate economic or social impact. Maps, presented as weights on a wheel of science, can play an important role in communicating the national orientation towards these different objectives.

Here there be dragons

In the 13th and 14th centuries, maps of the world showed the known world floating on a sea of uncertainty with unexplored regions marked "here there be dragons". This metaphor is still important today. Science education should be about communicating that there are many more areas yet to be discovered, students can take part in this process, and society, as a whole, can benefit from this discovery process.

One can communicate this same sense of excitement by placing what we "know" on the edge of a circle and what is "unknown" as the white space inside the circle. We should communicate, to both students and the public at large, that there still need to be explorations into the heart of the unknown. More and more of this exploration is interdisciplinary, which means it's further from the known edges of the circle of science. Deep inside the circle are the dragons that the next generation must face and conquer.

Useful links:
Places & Spaces
All the original maps and their codings

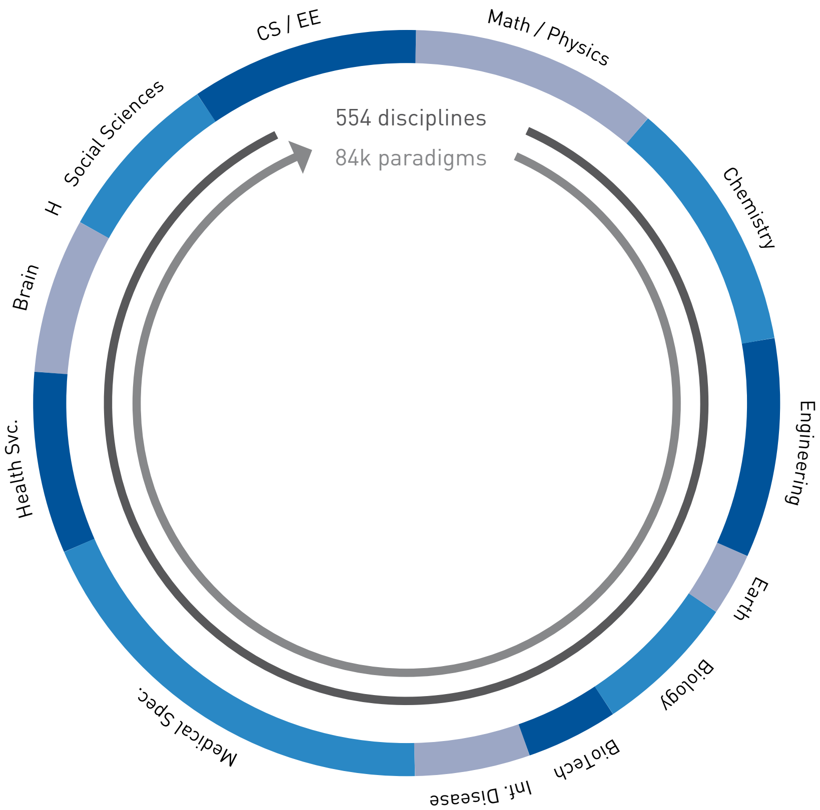


Figure 1- In this map, fields are arranged around a circle based on the meta-analysis of 20 maps of science. The order of the 554 disciplines (journal categories) is based on multiple factor analyses and the 84,000 paradigms (co-citation clusters) are ordered around the circle by discipline.

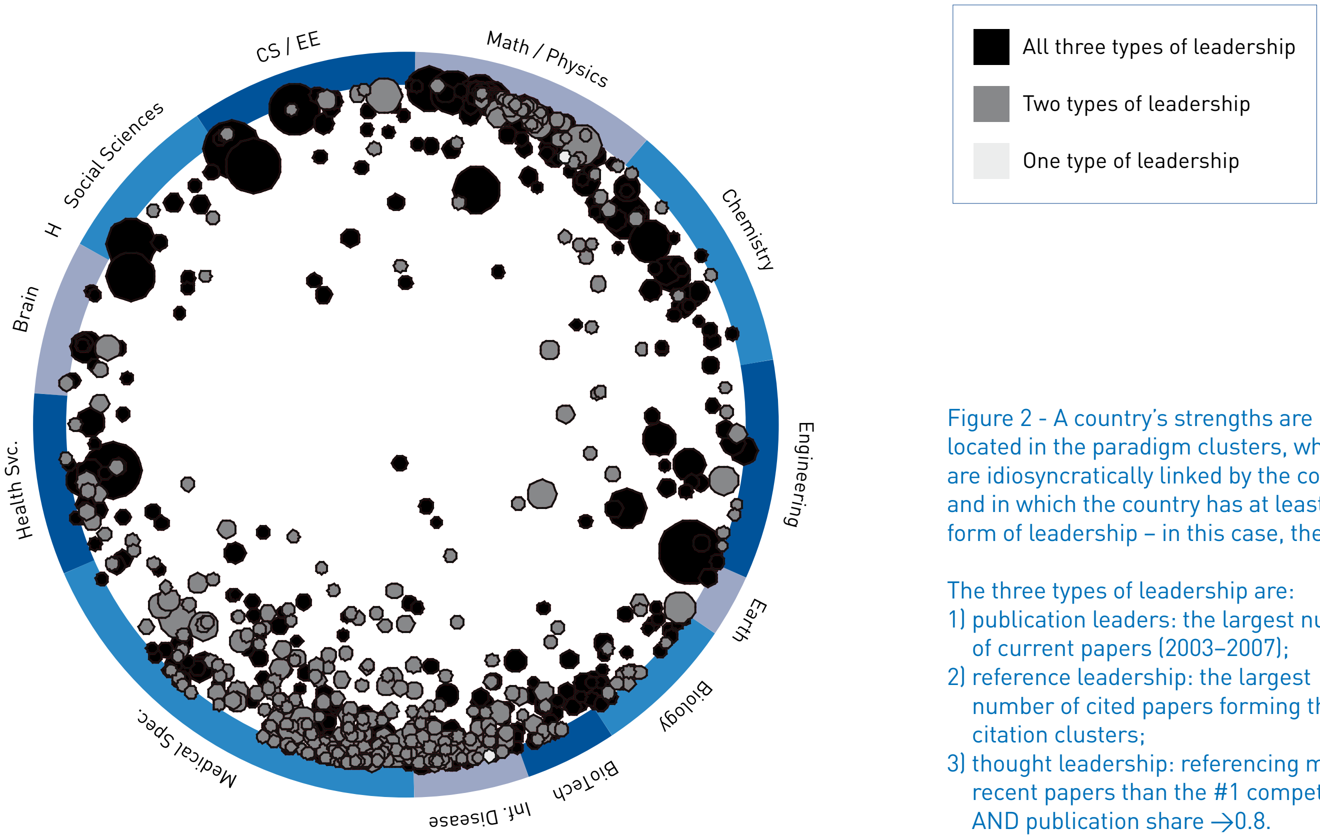
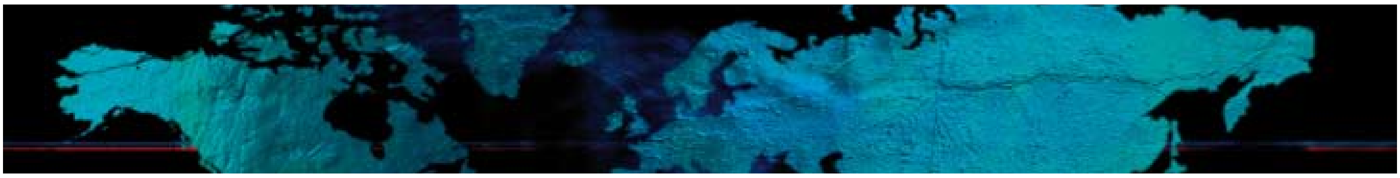


Figure 2 - A country's strengths are located in the paradigm clusters, which are idiosyncratically linked by the country and in which the country has at least one form of leadership – in this case, the USA.

The three types of leadership are:

1) publication leaders: the largest number of current papers (2003–2007);
2) reference leadership: the largest number of cited papers forming the co-citation clusters;
3) thought leadership: referencing more recent papers than the #1 competitor AND publication share →0.8.



Geographical trends of research output

The publication of journal articles worldwide follows a consistent pattern associated with the number of researchers based in a particular country. Unsurprisingly, the share of world articles is dominated by those countries with the most researchers, with countries such as the United States, Japan, the United Kingdom and Germany ranked highest. The geographical distribution of citations shows a similar pattern, with the same four countries appearing in the top four places according to citations received, albeit in a slightly different order. The growth in Chinese researcher numbers and research output has been previously discussed in Research Trends.

Table 1 illustrates the rank of countries according to their share of world articles and indicates the equivalent rank for each country according to citations received.

Rank by articles	Rank by citations	Country	Articles	Cites	Researchers	% docs	% cites
1	1	United States	3,437,213	43,436,526	7,442,000	25.9%	37.6%
2	4	Japan	983,020	7,167,200	896,211	7.4%	6.2%
3	2	United Kingdom	962,640	9,895,817	313,848	7.3%	8.6%
4	3	Germany	888,287	8,377,298	470,729	6.7%	7.2%
5	13	China, Peoples' Republic of	758,042	1,629,993	1,152,617	5.7%	1.4%
6	5	France	640,163	5,795,531	348,714	4.8%	5.0%
7	6	Canada	473,763	4,728,874	199,060	3.6%	4.1%
8	7	Italy	461,292	3,821,440	164,026	3.5%	3.3%
9	11	Spain	330,399	2,350,185	161,932	2.5%	2.0%
10	17	Russian Federation	330,020	1,064,077	951,569	2.5%	0.9%
11	9	Australia	295,977	2,566,649	118,145	2.2%	2.2%
12	19	India	286,109	994,561	N/A	2.2%	0.9%
13	8	Netherlands	264,565	3,012,291	915,65	2.0%	2.6%
14	18	Korea, Republic of	217,879	1,018,532	194,055	1.6%	0.9%
15	12	Sweden	194,921	2,188,026	72,459	1.5%	1.9%
16	10	Switzerland	188,134	2,384,981	52,250	1.4%	2.1%
17	22	Taiwan, Province of China	164,823	769,206	138,604	1.2%	0.7%
18	23	Brazil	163,550	752,658	N/A	1.2%	0.7%
19	24	Poland	159,536	682,354	78,362	1.2%	0.6%
20	14	Belgium	141,737	1,347,624	52,252	1.1%	1.2%

Table 1 – Geographical distribution of world articles 2004–2007 – top 20 countries. Source: Scopus. Researcher data taken from OECD Main Science & Technology Indicators, 2008 edition; data is for 2004 FTE researchers. US Researcher Data taken from Science & Engineering Indicators 2008, Table 3.1.

Country Trends

Table 2 illustrates that if the 2004 output of articles in Scopus is compared to researcher numbers in 2004 for these countries, some interesting trends develop. For instance, the number of researchers per research article published varies remarkably. It is important to note that this is different to authors per published article; in this case, we are calculating the ratio of total researchers in a country to the publication output of the country. In many cases there are researchers who never appear on articles as authors, and this is an important distinction to consider.

In Russia, there are 30 researchers for each research article published, while in the US there are 23 researchers. Switzerland has the lowest number of researchers per article at 2.5, followed by the UK at 3.2.

Country	**Number of researchers (2004)**	**Number of articles (2004)**	**Ratio of researchers per article**
Russian Federation	951,569	31,134	30.6
United States	7,442,000	315,161	23.6
China, Peoples' Republic of	1,152,617	101,685	11.3
Japan	896,211	97,579	9.2
Taiwan, Province of China	138,604	20,054	6.9
Korea, Republic of	194,055	28,943	6.7
France	348,714	64,909	5.4
Germany	470,729	91,881	5.1
Spain	161,932	36,849	4.4
Poland	78,362	18,524	4.2
Canada	199,060	50,904	3.9
Australia	118,145	32,837	3.6
Sweden	72,459	20,057	3.6
Belgium	52,252	15,451	3.4
Italy	164,026	49,592	3.3
United Kingdom	313,848	97,671	3.2
Netherlands	91,565	28,309	3.2
Switzerland	52,250	20,623	2.5

Table 2 – Researcher numbers for 2004 (source: OECD, US Data from NSF Science & Engineering Indicators 2008, table 3.1) and articles published in 2004 (source: Scopus).

The question follows, why do these countries have such differences in the researcher per article ratio?

Of course this is a difficult question to answer and has many dimensions, all of which will contribute in different amounts in different countries.

Fundamentally, overall population density, economic factors such as GDP and per capita expenditure and infrastructure will be significant factors in the ability to support research, but we are quick to point out that the countries that have the lower ratios, such as the UK and Switzerland, have some of the highest economic capabilities and strongest infrastructures in the world – this illustrates the issues in trying to understand these differences. Certainly, research funding from both governmental and private sources will affect the maintenance of research institutions and the ability to recruit research personnel. In the US, many research institutions have huge programs that require a substantial amount of staff members, which will increase the researcher numbers in our ratio.

In addition, the ability of a country to actually encourage students to follow a research path can often be problematic – in recent times in the UK there has been commentary on the problems of filling university places in subjects such as chemistry and physics and a study by Olivieri & Rowlands (2006) indicated that acquiring research staff was one of biggest barriers to research performance, which could be a significant factor to understand these interesting ratios between articles and researchers.

Expert Opinion

Eigenfactor: pulling the stories out of the data

Jevin West

Carl Bergstrom has been researching journal economics for over a decade. One fruit of those efforts, the Eigenfactor project, is drawing interest from editors, authors, researchers, policy-makers and evaluators seeking new measures of journal influence.

Jevin West, graduate student at the University of Washington in Bergstrom's research group, recalls: "It all started with Ted Bergstrom, Carl Bergstrom and I chatting about evaluation tools over a beer in December 2005. Carl was getting a lot of flack for using Impact Factors (IFs) in his work on journal economics, so we decided to come up with another way of evaluating the scholarly literature.

"I come from the theoretical side of biology and I'm interested in applying tools and concepts from network science and information theory to various problems, and that extends beyond biology to other fields, including bibliometrics. Fortunately, the nature of citation networks means that many of the models we use in biology are easily transferable."

Phenomenal interest

The Eigenfactor works like Google's PageRank, both of which are based on social network theory; where Google follows page links, Eigenfactor uses citations. They evaluate the importance of each journal (or Web page) based on the structure of the entire network.

The IF, in comparison, only looks one citation away and it ignores where they come from. "We take into account where the citation came from just as Google takes into account where a hyperlink comes from," says West.

When the Eigenfactor Web site was launched in January 2007, it attracted comment in numerous blogs, which raised its profile. "It has been far beyond anything we could have imagined. The interest has been phenomenal," says West.

"We also have our critics, and this is healthy," West adds. "I think all metrics should be criticized. Nothing beats reading an individual article in a journal to assess its value, and nothing ever will. But with time and budget constraints being what they are, there is a legitimate need for tools like this."

How the tools work

The Eigenfactor works by taking a random journal and following a random citation in that journal to another journal, then selecting another random citation from the second journal and following that to the next journal and so on. The Eigenfactor calculates the percentage of time you would spend at each journal. For instance, a search of all journals in 2006 gives *Nature* the highest Eigenfactor score. If you followed random citations infinitely, you would spend 1.9917% of your time at *Nature*.

The Article Influence score is calculated by dividing the Eigenfactor score for a particular journal by the number of articles published by that journal. All journals are normalized to 1. *The Annual Review of Immunology* comes out top, at 27.454 times normal.

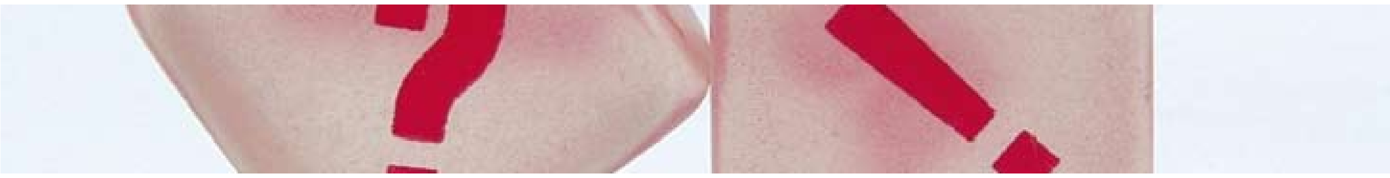
More and better tools

Journal ranking is just one of many stories the Eigenfactor team are pulling out of the data. They have created a cost-effectiveness score to help librarians manage their budgets efficiently, science maps and motion charts showing trends over time, which are particularly popular with editors and authors, and researchers interested in the history of science.

The team are planning to improve the tools they have and develop new ones, and they hope to bring in richer data. Also, over the longer-term, they want to apply these tools to other areas. "We're curious about how science has changed over time and we're interested in applying these tools to non-bibliometric areas as well," says West.

Even though this is just a side project, the team are enjoying themselves. "This has all come together at the right time. The data is available and some very sophisticated tools have been developed over recent years. We can now analyze data in some very exciting ways. We're having a blast!"

Useful links:
eigenFACTOR
Related papers



Why am I Cited...?

Citation is an essential part of science. It places a researcher's thinking within a continuum of thought, indicating sources of ideas and theories that the author agrees or disagrees with.

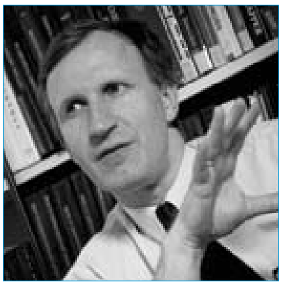
A highly cited paper is normally considered to be very relevant within its field, and increasingly across disciplines. It has, somehow, resonated throughout the scientific community.

Professor Jos H. Beijnen, pharmacist at Slotervaart Hospital and the Netherlands Cancer Institute in Amsterdam and appointed at Utrecht University, and Peter N. Nijkamp, based at the Department of Spatial Economics, Vrije Universiteit, Amsterdam, both believe that collaboration and relevant research have helped them become two of the most highly cited Dutch researchers.

Collaboration is inspirational

Beijnen is the most active Dutch author in Life Sciences and his most cited paper *(1)* has received over 950 citations. He attributes his remarkable output to efficient use of his time. He adds, "The selection of collaborators, in my case mostly young pharmacy students who want to do their Ph.D. in my group, is crucial. Their enthusiasm for research fuels me and gives me the energy to work seven days a week."

Nijkamp agrees: "The biggest challenge for a scientist is to find promising and bright young talents. I have been lucky to find so many interesting young people all over the world with whom I have worked and from whom I have learned a lot." Nijkamp is the most prolific of Dutch authors in Social Sciences. His most-cited document, with 59 citations, covers un-tolled congestion pricing *(2)*.



Jos H. Beijnen

Relevant research

Beijnen also believes research should be aimed at tackling issues that directly benefit society. "Our research is always based on a clinical research question. Our research should be beneficial to our patients, and that is what we always keep in mind."

Peter N. Nijkamp

Nijkamp feels the same: "Most of my research finds its inspiration in pressing societal problems, so it's no wonder that the information is then shared across a wide audience."

Cross-discipline, cross-border communication

Nijkamp adds that collaboration with his students and peers, as well as with researchers overseas and in different disciplines, has helped him maintain high output. "Modern quantitative economic research is a fascinating activity, where tools from various disciplines are extensively used. This leads to often surprising findings, with great scientific and policy-making value. Research in the social sciences is no longer a solitary activity. Increasingly, modern research in economics is based on collaboration with dozens of good people abroad. I have produced most of my publications together with many people outside the Netherlands."

The most relevant and highly cited papers are not produced in a vacuum; they offer insight into important questions in their field and their work has wide-ranging value thanks to the collaboration of students, peers in different geographical regions and, increasingly, from different fields. As Nijkamp concludes: "The knowledge society is indeed operating on a global market."

References:

(1) Schinkel, A.H.; Smit, J.J.M.; Van Tellingen, O.; Beijnen, J.H.; Wagenaar, E.; Van Deemter, L.; Mol, C.A.A.M.; Van der Valk, M.A.; Robanus-Maandag, E.C.; te Riele, H.P.J.; Berns, A.J.M.; Borst, P. (1994) "Disruption of the mouse *mdr1a* P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs", Cell, Vol. 77, pp. 491–502.

(2) *Verhoef*, E., Nijkamp, P., Rietveld, P. (1996) "Second-best congestion pricing: The case of an untolled alternative", *Journal of Urban Economics*, Vol. 40, No. 3, pp. 279–302.

researchtrends

EDITORIAL BOARD:

Iris Kisjes
Gert Jan Geraeds
Andrew Plume
David Tempest
Judith Kamalski
Myrto Arvaniti
Michelle Pirotta, The Write Company

You can find more information on www.researchtrends.com
or contact us at researchtrends@elsevier.com