Bibliometrics, Research Performance Evaluation, and Beyond: Towards Actionable Intelligence for Science Administrators, Policymakers, and Funders

INTELLECTUAL PROPERTY & SCIENCE

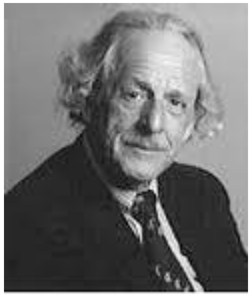
David A. Pendlebury, Consultant, Bibliometric Analysis
September 3, 2012

SMU
SINGAPORE MANAGEMENT UNIVERSITY

THOMSON REUTERS

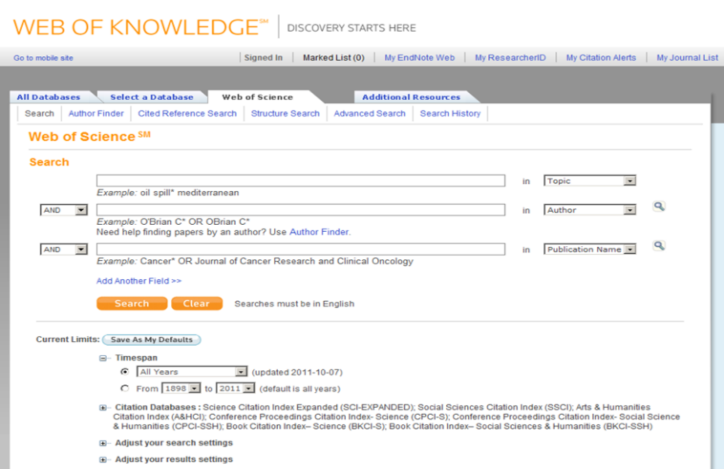
I. From Rejection to Acceptance: Some History

- 1955 Eugene Garfield's paper in *Science* on "Citation Indexes for Science"
- 1963 First Science Citation Index (ISI >Thomson >Thomson Reuters)
- 1972 U.S. National Science Foundation initiates *Science Indicators* (later *Science and Engineering Indicators*), including publication and citation data
- 1980s Rapid uptake of science indicators throughout Europe by governments (and research by SPRU, CWTS, Hungarian Academy of Sciences, as well as ISI)
- 1993 Mosaic introduced, sparking a revolution and aiding in more intuitive understanding of the nature of citation indexes – a turning point
- 1997 *Science Citation Index* and other databases move to web format, now under *Web of Knowledge* platform
- 2004 Elsevier's Scopus and Google Scholar are launched
- 2005 Hirsch introduces h-index
- 2000s Rankings proliferate; Experiments in visualization
- 2010s Other measures, services introduced and evaluated



What a Citation Index Offers: Search and Analysis

- Search
- Analysis
    - Structure, Dynamics of research
    - Research performance



- Single measures, rankings
- Multiple measures, contexts
    - Analytical tools
    - Visualization tools

THOMSON REUTERS

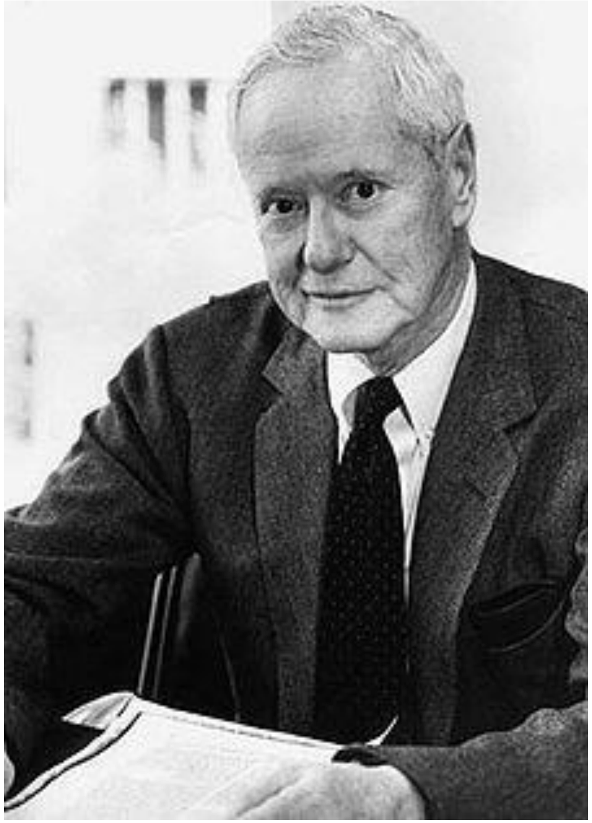
Bibliometrics or Scientometrics: Counting Publications and Citations

- Publications as indicators of output
- Citations as indicators of influence
- Citations per paper as indicators of impact (weighted influence)
- Various derivative measures such as relative indicators, for example, citations per paper relative to average citations per paper for field (normalized)
- Impact Factor
- Others, such as h-index and three dozen variants of the h-index

Theories of Citation and the Normative School

- Robert K. Merton, (1910-2003), sociologist of science, Columbia University. Normative theory.
- Citations as currency used to repay intellectual debts. Those with many citations have gained "credits" from their peers.
- The formal nature of publication and the moral imperative to cite.
- Other theories, including citations as rhetorical devices, constructivist theories.

Known for coining the concepts and phrases: "self-fulfilling prophecy," "role model," "focus group," "unanticipated consequences"



Research Evaluation, Qualitative vs. Quantitative: Two (Complementary) Types of Peer Review

Peer Review: Qualitative

- Small-scale, ground-up view
- Absolute counts, size colors perceptions and judgments
- Affected by work done long ago

Citation Analysis: Quantitative

- Global, top-down view
- Weighted and relative measures
- Can reveal recent contributions

The Impact Factor: Recommended Uses

- Designed to evaluate journals, especially in the context of acquisition decisions by librarians
- Formula: Citations in year 3 to journal articles in years 1 and 2, divided by the number of citable items in years 1 and 2 (citable items are regular discovery accounts and review articles). Thus, a short-term measure of average (mean) per paper performance for a journal
- Journal impact factor scores vary by field and are themselves skewed within a field (the 80:20 rule, pervasive at all levels)
- Thomson Reuters *discourages* the use of impact factors to evaluate individual articles or authors ("a mortal sin" – Ton Van Raan). Unfortunately, a very common 'quick and dirty' practice!

Special issue of *Scientometrics* devoted to discussion of impact factors: Vol. 92, No. 2, August 2012

The h-index: A Measure of Productivity and Influence

Jorge E. Hirsch, “An index to quantify an individual's scientific research output” *PNAS*, 102(46): 16569-16572, 2005.

- Formula: A researcher with an index of *h* has published *h* papers each of which has been cited at least *h* times.
- Represents an attempt to combine measures of productivity and influence. Like other measures, it is field dependent.
- Strengths: simple to calculate, combines output and impact, depicts “durable” performance and not single achievements, correlates with other measures of significance.
- Weaknesses: discriminates against young researchers, will not capture small but high-quality output, may not depict recent performance, h will never decline so one can “rest on one’s laurels,” AND correlates with other measures of significance.

Citation Analysis and Research Evaluation: National and Institutional to Individual

Some General Principles of Good Practice

- Basic better than applied sciences
- Large better than small datasets (macro and meso vs. micro analysis)
- Long better than short period
- Relative (normalized) better than absolute measures
- Multiple better than single measures ("The use of a single index ***crashes*** the ***multidimensional*** space of bibliometrics into one single dimension" – Wolfgang Glänzel)
- Top end of distribution better than middle and bottom to obtain strong, unambiguous signals

Above all, compare like with like, not "apples with oranges"

From Rejection to Acceptance…
But Has the Pendulum Swung Too Far?

- Much naïve, uninformed use of publication and citation data – not even accurately collected
- Formulaic use of data for evaluation, especially employing single measures, such as the impact factor, the h-index, and others
- The spread of “Impactitis” – Padmanabhan Balaram
- Perverse incentives in the form of ill-advised financial rewards to achieve specific outcomes
- “Horse before the cart” and “Tail wagging the dog”

II. Unintended Negative Consequences

- Law of Unintended Negative Consequences: Negative effects contrary to what was intended. Can stem from perverse incentives, and an emphasis of short- over long-term goals.
- Goodhart's Law (1975): "Once a social or economic indicator or other surrogate measure is made a target for the purpose of conducting social or economic policy and control, then it will lose the information content that would qualify it to play such a role."
- In both cases, setting a simple or crude measure of performance changes behavior as subjects attempt to optimize their performance – not only does this disturb behavior, it also destroys the utility of the measure.
- **Goal in science is not citations and not prizes: the goal is excellence in research.** Citations and prizes will follow.

Three Cautionary Examples

1 The Australian government in the 1990s used publication output as a measure of research performance evaluation. The result: Australian scientists published more, but in lower impact journals.

2 A nation's universities offered financial incentives to researchers to publish in ISI-indexed journals, and rewards were specifically geared to journal rank. The result: a few individuals published in such great quantity in low impact titles – easier to publish in and claim an award – that the entire nation's impact in chemistry declined.

3 Analysts have detected an increase in error and fraud in nations that have attempted to build research capacity quickly through use of crude metrics (single measures), formulaic assessments, and 'pay for paper' financial rewards. "Piece rates for professors"

The issue is not that performance measures suffer from formulaic evaluation and perverse incentives, but rather that scientists have been diverted from their main work and that precious resources have been wasted. Or worse…as illustrated in the last example.

“What Does it Take to Get a Nobel Prize?” Advice from Nobel Laureate Ahmed Zewail

- First and foremost, the priority should be on education in science, technology, mathematics, and engineering. Capacity in R&D requires the best young minds. Large buildings and massive funds will not produce much without the right people.
- Second, nurture an atmosphere of intellectual exchange. To distract faculty with the writing of extensive and numerous proposals or to turn them into managers is the beginning of the end.
- Third, without resources little can be achieved, no matter how creative the mind. Obviously, investment in science is needed ….Countries and institutions that provide the requisite infrastructure and the funding for ideas will be the homes of discoveries. But such support should follow the vision of creative researchers, not be built merely to lure money or to force people into fashionable research areas.

paraphrased from: Ahmed Zewail, “Curiouser and curiouser: Managing discovery making,” *Nature*, 468: 347, 18 November 2010.

III. Meeting the Needs of Research Policymakers, Managers, Funders

- Support for science from citizens requires policymakers and administrators to ensure effectiveness and efficiency. Scientists are accountable for the support they receive.
- Citation analysis combined with peer review can often add substantially to research assessment and improve decisions made by administrators and policymakers.
- But using metrics in simple ways to control outcomes can change behavior and actually institutionalize uniformity or even mediocrity in research.
- This may dampen creativity and derail "revolutionary science" (Thomas Kuhn), the type recognized as excellent and of "Nobel-class."
- The ideal is informed, thoughtful, and wise assessment coupled with directed support related to national and institutional goals. And this takes work! The results?

Berkshire Hathaway vs. S&P 500: Cumulative Results of Informed Decisions

Policy and Funding Decisions Aligned with Bibliometric Distributions: Equity and Excellence

"The tension between equity and excellence is fundamental in science policy. This tension might appear to be resolved through the use of merit-based evaluation as a criterion for research funding. This is not the case.

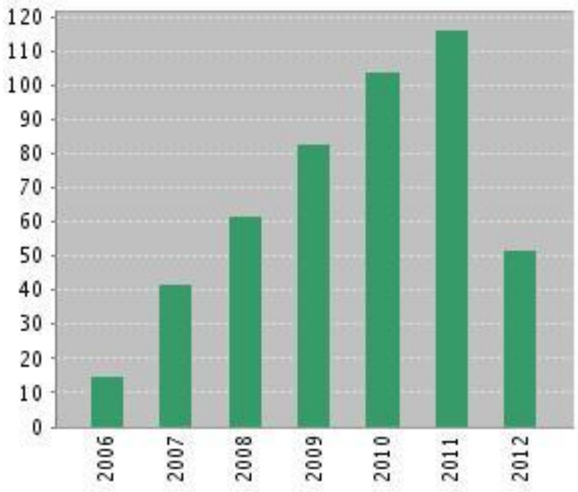
Merit-based decision making alone is insufficient because of inequality aversion, a fundamental tendency of people to avoid extremely unequal distributions. The distribution of performance in science is extremely unequal, and no decision maker with the power to establish a distribution of public money would dare to match the level of inequality in research performance. We argue that decision makers who increase concentration of resources because they accept that research resources should be distributed according to merit probably implement less inequality than would be justified by differences in research performance. Here we show that the consequences are likely to be suppression of incentives for the very best scientists.

The consequences for the performance of a national research system may be substantial. Decision makers are unaware of the issue, as they operate with distributional assumptions of normality that guide our everyday intuitions."

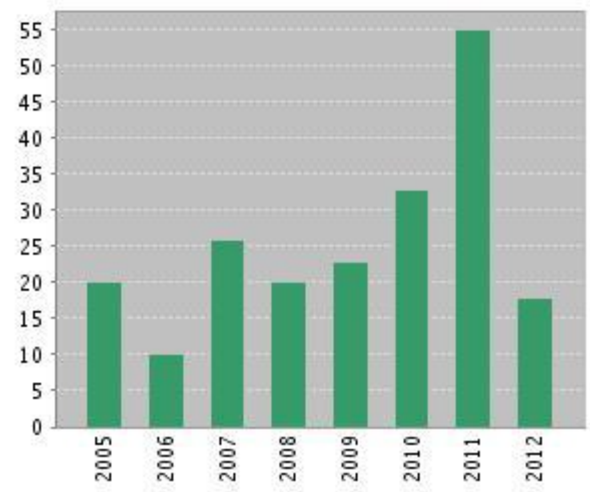
Diana Hicks and J. Sylvan Katz, "Equity and excellence in research funding," *Minerva*, 49 (2): 137-151, June 2011

Interests of Academic Bibliometricians vs. Needs of Research Policymakers, Managers, Funders

- Published items in each year: “H-index”



- Published items in each year: “Research Evaluation”



17th International Conference on Scientific and Technology Indicators, 5-8 September 2012, Montreal, Quebec, Canada: 71 papers, 34 posters

Only 14 of 71 papers (20%) and 6 of 34 posters (18%) addressed needs of research policymakers, managers, and funders – this based on a very liberal classification scheme

Topics Featured at 17th STI Conference: Some New and Emerging Trends

STI 2012 Montréal
17th International Conference on Science and Technology Indicators

Old:

- Validity of bibliometric indicators (recently especially university rankings)
- Determinants of productivity and impact (collaboration, migration, interdisciplinarity)
- Internationalization, globalization, status of emerging nations
- Research Fronts and detection of emerging, “hot” areas
- Patent citation analysis and the connections between fundamental research and applied research

New:

- New indicators based on: downloads from full-text databases and repositories, data derived from social media, consolidated data from multiple sources, and funding acknowledgements; also, recent interest in percentiles vs. means
- Open access and its characteristics, influence on citations impact
- Citation analysis for social sciences and humanities (*Thomson Reuters Book Citation Index*)
- Social and economic impact of basic research (Henk F. Moed: ‘not politically neutral’)

…and not featured at the conference:

- Visualization: spatial scientometrics

Funding Acknowledgement Analysis: Linking Inputs to Outputs and Impacts

REUTERS/Tony Gentile

FUNDING ACKNOWLEDGEMENTS IN THE JOURNAL LITERATURE:

USES AND BENEFITS FOR FUNDERS, RECIPIENTS AND ANALYSTS

MARCH 2012

KAREN A. GURNEY AND DAVID PENDLEBURY

THOMSON REUTERS

Analytical Tools for Research Evaluation
An Example: Thomson Reuters InCites

SUMMARY METRICS

Citation Metrics

Times Cited	12,294
Web of Science Documents	486
Cites per Document	25.30
h-index	58
Median Cites	10
2nd Generation Citations	295,320
2nd Generation Citations per Citing Document	43.04

Self Citation Metrics

Self Cites	1,257
% Self Cites	10.22%
Times Cited without Self Cites	11,037
Cites per Document without Self Cites	22.71
h-index without Self Cites	55

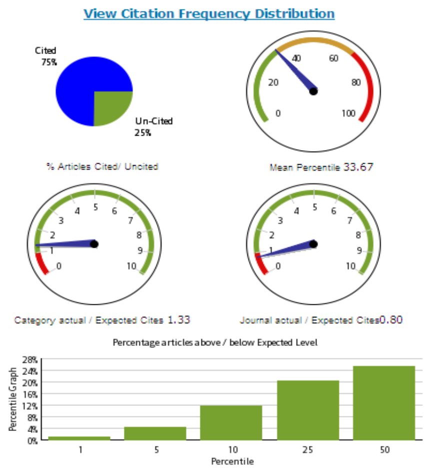
Disciplinarity Metrics

Disciplinarity index	0.18
Interdisciplinarity index	0.44

Collaboration Metrics

Unique Authors	985

View Citation Frequency Distribution



Percentile	1	5	10	25	50
Number of Documents	6	30	69	145	241
Percent of	1.88%	9.38%	21.56%	45.31%	75.31%

Summary Metrics communicate the "big picyure" for a dataset, Metrics are provided for citations, discipline, collaboration, and more.

Subject Areas 1 - 20 of 239 Sort By: Times Cited

Rank	Subject Area	Times Cited	Web of Science Documents	Average Cites per Document	h-index	Journal Actual/Expected Citations	Category Actual/Expected Citations	Average Percentile
1	BIOCHEMISTRY & MOLECULAR BIOLOGY	284,235	7,844	36.24	194	1.14	1.42	39.34
2	IMMUNOLOGY	211,464	6,236	33.91	182	1.19	1.57	38.41
3	CELL BIOLOGY	137,876	3,986	34.59	158	1.14	1.32	38.56
4	[illegible]	[illegible]	[illegible]	22.89	152	1.23	1.61	41.14
5	[illegible]	[illegible]	[illegible]	27.00	153	1.26	1.59	39.99
6	[illegible]	[illegible]	[illegible]	26.12	145	1.23	1.55	39.66
7	[illegible]	[illegible]	[illegible]	34.92	145	1.08	1.58	36.37
8	[illegible]	[illegible]	[illegible]	25.84	135	1.36	1.84	36.27
9	[illegible]	[illegible]	[illegible]	27.66	118	1.28	1.42	44.22
10	[illegible]	[illegible]	[illegible]	22.88	116	1.13	1.31	41.41
11	PHARMACOLOGY & PHARMACY	78,299	3,802	20.59	106	1.14	1.47	40.39
12	PHYSICS, MULTIDISCIPLINARY	78,286	2,528					
13	PERIPHERAL VASCULAR DISEASE	77,018	3,253					
14	CHEMISTRY, MULTIDISCIPLINARY	72,601	2,490					
15	ECOLOGY	72,244	1,941					
16	ASTRONOMY & ASTROPHYSICS	67,337	2,750					
17	SURGERY	59,550	3,588					
18	CLINICAL NEUROLOGY	55,791	2,312					
19	CHEMISTRY, PHYSICAL	51,163	2,319					
20	MICROBIOLOGY	51,064	1,824					

Subject Area ranking of your articles shows output and impact

SUBJECT AREA RANKING (CITING ARTICLE SET)

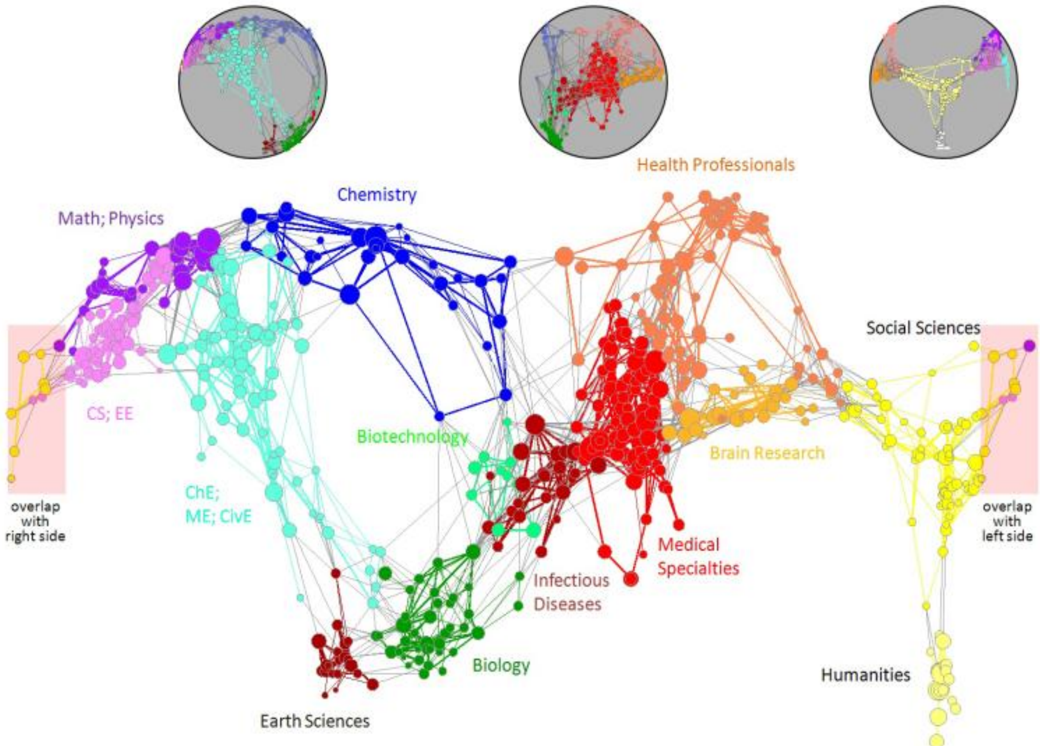
Report Limited To:
Dataset:
Report Name:
Time Period:
Additional Infor

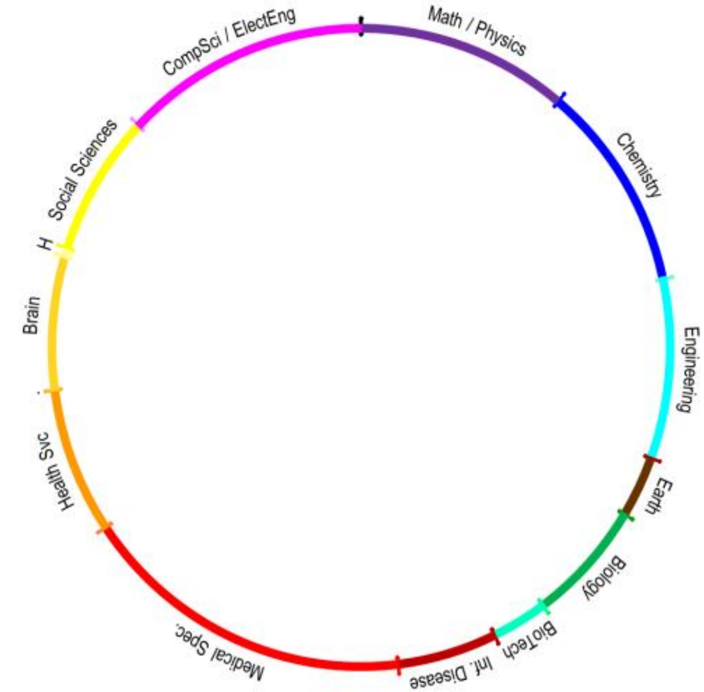
Subject Area ranking through citing articles shows fields you're impacting most

Subject Areas 1 - 20

Rank	Subject Area	Times Cited	Web of Science Documents	Average Cites per Document
1	BIOCHEMISTRY & MOLECULAR BIOLOGY	4,851,651	135,003	35.94
2	IMMUNOLOGY	3,169,055	95,874	33.05
3	NEUROSCIENCES	2,531,318	75,046	33.73
4	CELL BIOLOGY	2,528,996	69,539	36.37
5	ONCOLOGY	1,829,427	61,560	29.72
6	HEMATOLOGY	1,820,037	51,405	35.41
7	CARDIAC & CARDIOVASCULAR SYSTEMS	1,735,769	66,472	26.11
8	[illegible] & HEREDITY	1,613,632	47,720	33.81
9	PHAR[illegible] & PHARMACY	1,550,796	73,644	21.06
10	CHEMISTRY, [illegible]SCIPLINARY	1,335,782	39,599	33.73
11	ENDOCRINOLOG[illegible]ABOLISM	1,320,326	43,788	30.15
12	PERIPHERAL VASCUL[illegible]	[illegible]7,736	38,102	32.75
13	CLINICAL NEUROL[illegible]	[illegible]	38,975	23.63
14	MICROBIOLOGY	[illegible]	34,226	26.41
15	RESPIRATORY [illegible]	[illegible]	38,425	23.47
[illegible]	PHYSIOLOGY	[illegible]	31,493	27.14
[illegible]	[illegible]COLOGY	[illegible]	31,545	26.08
[illegible]	[illegible]TRONOMY & A[illegible]	[illegible]	35,681	22.17
[illegible]	[illegible]MISTRY, PHYSIC[illegible]	[illegible]283	37,455	20.75
[illegible]	[illegible]ERY	748,092	42,384	17.65

Experiments in Visualization: Recent Analysis of Structure of Research





Börner K, Klavans R, Patek M, Zoss AM, et al. (2012) Design and Update of a Classification System: The UCSD Map of Science. PLoS ONE 7(7): e39464. doi:10.1371/journal.pone.0039464
http://www.plosone.org/article/info:doi/10.1371/journal.pone.0039464

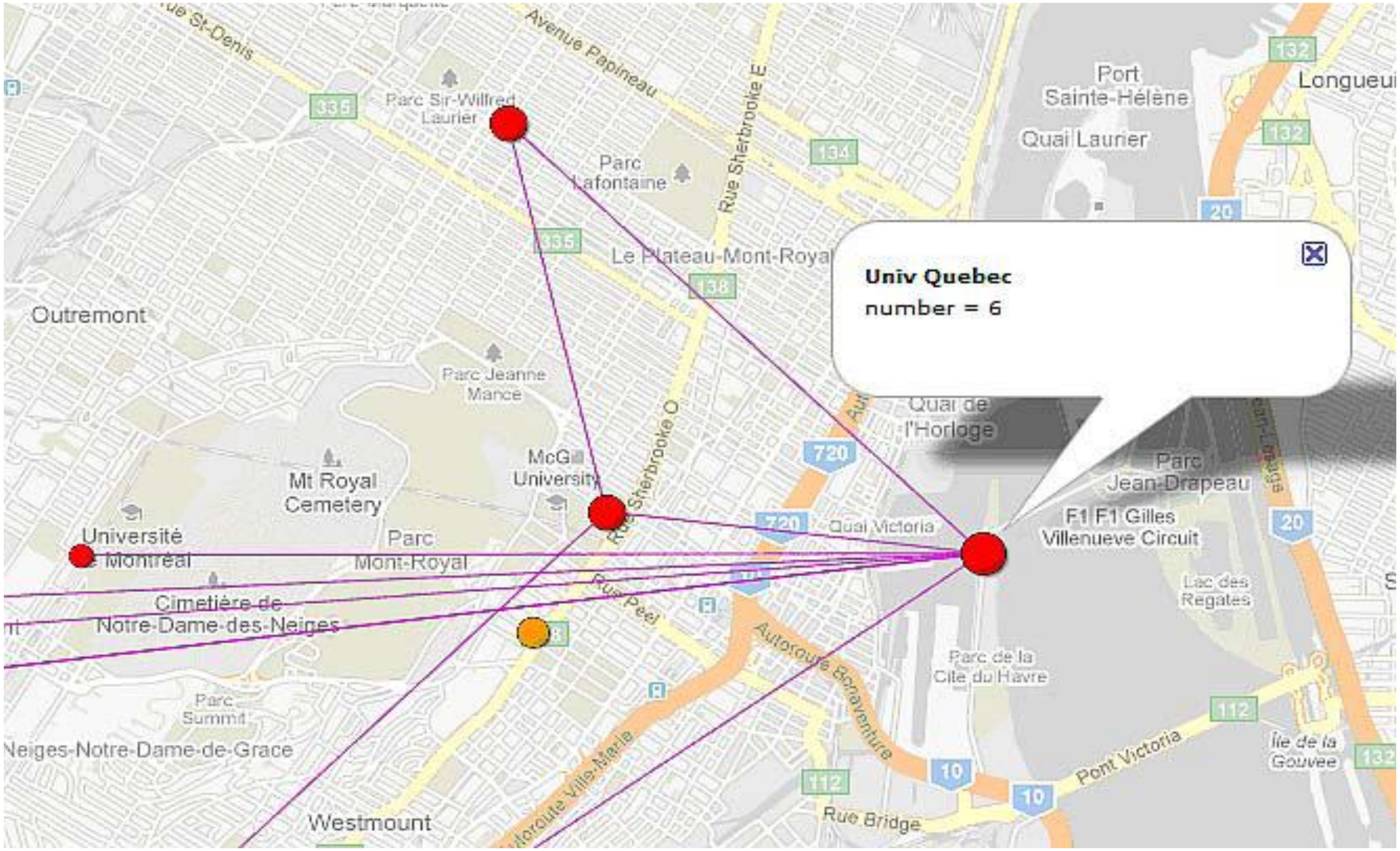
THOMSON REUTERS

PLoS one

A New Direction in Visualization: Spatial Scientometrics

Loet Leydesdorff and Olle Persson, "Mapping the geography of science: Distribution patterns and networks of relations among cities and institutes," *Journal of the American Society for Information Science & Technology*, 61 (8): 1622-1634, August 2010

Potential of Analytical Tools *Combined* with Visualization for Policy, Management, and Funding



Loet Leydesdorff and Olle Persson, “Mapping the geography of science: Distribution patterns and networks of relations among cities and institutes,” *Journal of the American Society for Information Science & Technology*, 61 (8): 1622-1634, August 2010

Citation Analysis and Research Evaluation: Select Bibliography

Books:

- **Henk F. Moed**, *Citation Analysis in Research Evaluation*, Springer, 2005
- **Nicola De Bellis**, *Bibliometrics and Citation Analysis: From the Science Citation Index to Cybermetrics*, Scarecrow Press, 2009
- **Katy Börner**, *Atlas of Science: Visualizing What We Know*, MIT Press, 2010

Journals:

- ***Scientometrics***, *(1978 – present)*
- ***Journal of Informetrics***, *(2007 – present)*
- ***Journal of the American Society for Information Science and Technology***, *(1950 – present)*
- ***Research Evaluation***, *(1992 – present)*

Conference (providing a review of contemporary research concerns)**:**

- **17th International Conference on Scientific and Technology Indicators, 5-8 September 2012, Montreal, Quebec, Canada**

  http://2012.sticonference.org/index.php?page=prog

Citation Analysis and Research Evaluation: Select Bibliography

Articles:

- **Linda Butler**, "Modifying publication practices in response to funding formulas," *Research Evaluation*, 12 (1): 39-46, April 2003
- **Peter Weingart**, "Impact of bibliometrics upon the science system: Inadvertent consequences?" *Scientometrics*, 61 (1): 117-131, January 2005
- **Anthony F.J. Van Raan**, "Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods," *Scientometrics*, 61 (1): 133-143, January 2005
- **Henk F. Moed**, "UK research assessment exercises: Informed judgments on research quality or quantity?" *Scientometrics*, 74 (1): 153-161, January 2008
- **Koen Frenken, Sjoerd Hardeman, Jarmo Hoekman**, "Spatial scientometrics: a cumulative research program," *Journal of Informetrics*, 3 (3): 222-232, July 2009
- **Jonathan Adams**, "The use of bibliometrics to measure research quality in UK higher education institutions," *Archivum Immunologiae et Therapiae Experimentalis*, 57 (1): 19-32, February 2009
- **Loet Leydesdorff and Olle Persson**, "Mapping the geography of science: Distribution patterns and networks of relations among cities and institutes," *Journal of the American Society for Information Science & Technology*, 61 (8): 1622-1634, August 2010
- **Diana Hicks and J. Sylvan Katz**, "Equity and excellence in research funding," *Minerva*, 49 (2): 137-151, June 2011
- **Katy Börner, Richard Klavans, Michael Patek, Angela M. Zoss, Joseph R. Biberstine, Robert P. Light, Vincent Larivière, and Kevin W. Boyack**, "Design and update of a classification system: The UCSD map of science," *PLoS One*, 7 (7), article no. e39464, July 12, 2012

Contact Information for David A. Pendlebury, Consultant, Bibliometric Analysis, and for Shannen Dan, Sales Manager, Government, Thomson Reuters

Email: david.pendlebury@thomsonreuters.com

Email: shannen.dan@thomsonreuters.com