

TOPICS OF PROMINENCE, AN INSIGHT INTO IDENTIFYING NEW, EMERGING RESEARCH TRENDS, SETTING THE SCENE FOR TECHNOLOGIES OF THE FUTURE

Abstract

Our methodology aims at identifying the most dynamic research areas worldwide, understanding research trends thanks to a bottom-up approach, and driving to enhanced foresight, providing policy-makers with a well-informed base regarding research trends.

'Topics of Prominence' in Science can help answer the following questions:

- Which are the research fields with high momentum?
- Which stakeholders are leading the way in those Topics? Countries, institutions, researchers, sources?
- What are the related Topics adjacent to ones with high momentum?

Underlying methodology

The creation of Topic Prominence in Science is based upon years of research by Kevin Boyack and Richard Klavans and extensive institutional feedback and testing of SciVal product, User Experience and development teams.

Due to recent advances and improvements in quality of the underlying data, the development of better algorithms to cluster scientific papers and greater computing technology, we have now created a highly accurate model able to analyze the whole of the Scopus dataset at once.

What is a Topic?

Our technology takes into consideration 95% of the articles in ScopusTM and clusters them into 96,000 global, unique research topics based on citation patterns. Scopus publications are clustered into Topics based on a direct citation analysis. A Topic can be defined as a collection of documents with a common intellectual interest. Topics are dynamic and many Topics are multidisciplinary.

What is Prominence?

Calculating a Topic's Prominence combines 3 metrics to indicate the momentum of the Topic:

- Citation Count in year n to papers published in n and $n-1$.
- Scopus Views Count in year n to papers published in n and $n-1$, the number of clicks on the abstracts of the Scopus publications.
- Average CiteScore for year n . The CiteScore metrics calculate the citations from all documents in year one to all documents published in the prior three years for any serial indexed in Scopus.

Robustness of the model

We have identified a correlation between the prominence (momentum) of Topics and the amount of funding per author within those Topics. On average, the higher the momentum, the more money has been made available per author for research on that Topic. To prove this, SciTech Strategies assigned 314,000

grants worth \$203 billion from the STAR METRICS database, a large project-level funding database accounting for 24 percent of US federal funding, to all 97,000 Topics through textual similarity.

The Topics of Prominence allow policy-makers to analyze the emerging Research and Innovation domains which require support through policies. New research trends in fields such as fracking, GMOs may have been identified earlier, hence benefitting from adjusted strategies. Our methodology makes also possible to identify the emerging fields in which Europe is ahead and for which funding is needed in order to remain on the cutting edge. And when lagging behind in strategic fields, this helps make decisions speeding up Research developments or developing accurate collaborations.

Keywords: research trends, momentum, multidisciplinary

Introduction

Until now, research performance analytics tools have focused mainly on the past, showing what has been done and comparing that to the performance of peers. Increasingly, however, researchers and policy makers need to know where to focus their attention for the future. Those who do this are better positioned to monitor current Research trends, organize accurate funding programs, identify the key players and perform the most up-to-date analyses.

The purpose is to turn traditional research analysis around to face the future and build on the solid citation connections that are the hallmark of international and multidisciplinary science. Making it available in the research performance benchmarking platform [SciVal](#), offering new tools for monitoring research and identifying key players: Topic Prominence in Science.

[SciVal](#) offers quick, easy access to the research performance of 8,500 research institutions and 220 countries worldwide. It enables users to visualize research performance, benchmark relative to peers, develop collaborative partnerships and analyze research trends. As SciVal uses advanced data analytics, users can instantly configure and process enormous amounts of data, and generate on-demand data visualizations relevant to specific challenges. SciVal runs on Scopus data, which covers more than 69 million records from 22,000 journals and 5,000 publishers worldwide.

Dr. Richard Klavans, chairman of SciTech Strategies, the company that Elsevier partnered with to develop the new capability, explained why looking forward is so important: “Researchers and decision makers have limited time and resources. Knowing where the focus is make the process more efficient. Knowing which topics have the highest momentum can help policy makers take decisions upon the next programs and identify the key players and the position of a territory in a specific area of research.”

Methodological approach

At SciTech Strategies, Dr. Klavans works with company president Dr. Kevin Boyack to develop cutting-edge technologies to understand and model the structure and dynamics of science and technology at a highly detailed level. Their goal has been to analyze topics across, rather than

within, a range of subjects and industries to help researchers and decision makers to see which topics have momentum today and predict which will likely have momentum tomorrow.

Topic Prominence in Science divides the full spectrum of published research – 95 percent of the Scopus corpus of data – into 97,000 distinct topics based on citation paths. They clustered all Scopus publications from 1996 onwards – about 35 million – into 97,000 different “Topics” using citation paths – the first time this level of analytics has been done across the entire Scopus corpus. This provides a much better reflection of reality than citations alone, showing who is working with whom on papers, for example. They defined a Topic as a collection of documents with a common intellectual interest. Topics can be large or small, old or new, growing or declining. They are dynamic and can evolve but can also be dormant. Researchers are not limited to one Topic – they can contribute to many, shaping the connections and the research map as they publish their work.

These topics are ranked using a new indicator called “Prominence,” which is calculated by weighing together the current impact of papers clustered in a topic, how much attention these papers receive and the prestige of the journals in which they are published ([CiteScore](#)). The prominence rank is indicative of the current momentum of a Topic.

The indicator is calculated by weighing 3 metrics for papers clustered in a Topic:

- Citation Count in year n to papers published in n and $n-1$
- Scopus Views Count in year n to papers published in n and $n-1$
- Average CiteScore for year n

Read more about the formula behind the indicator Prominence and its correlating with future funding: “[Research Portfolio Analysis and Topic Prominence](#)” by Richard Klavans and Kevin W. Boyak.

Results, discussion and implications

The predictive aspect of Prominence has been analyzed is its correlation with future funding. SciTech Strategies assigned 314,000 grants worth \$203 billion from the [STAR METRICS](#) database, a large project-level funding database that accounts for 24 percent of US federal funding, to all 97,000 Topics through textual similarity. The grant data were split into two time periods for each Topic and the correlation analyzed. The model showed that Prominence accounted for 37 percent of the variance of future funding.

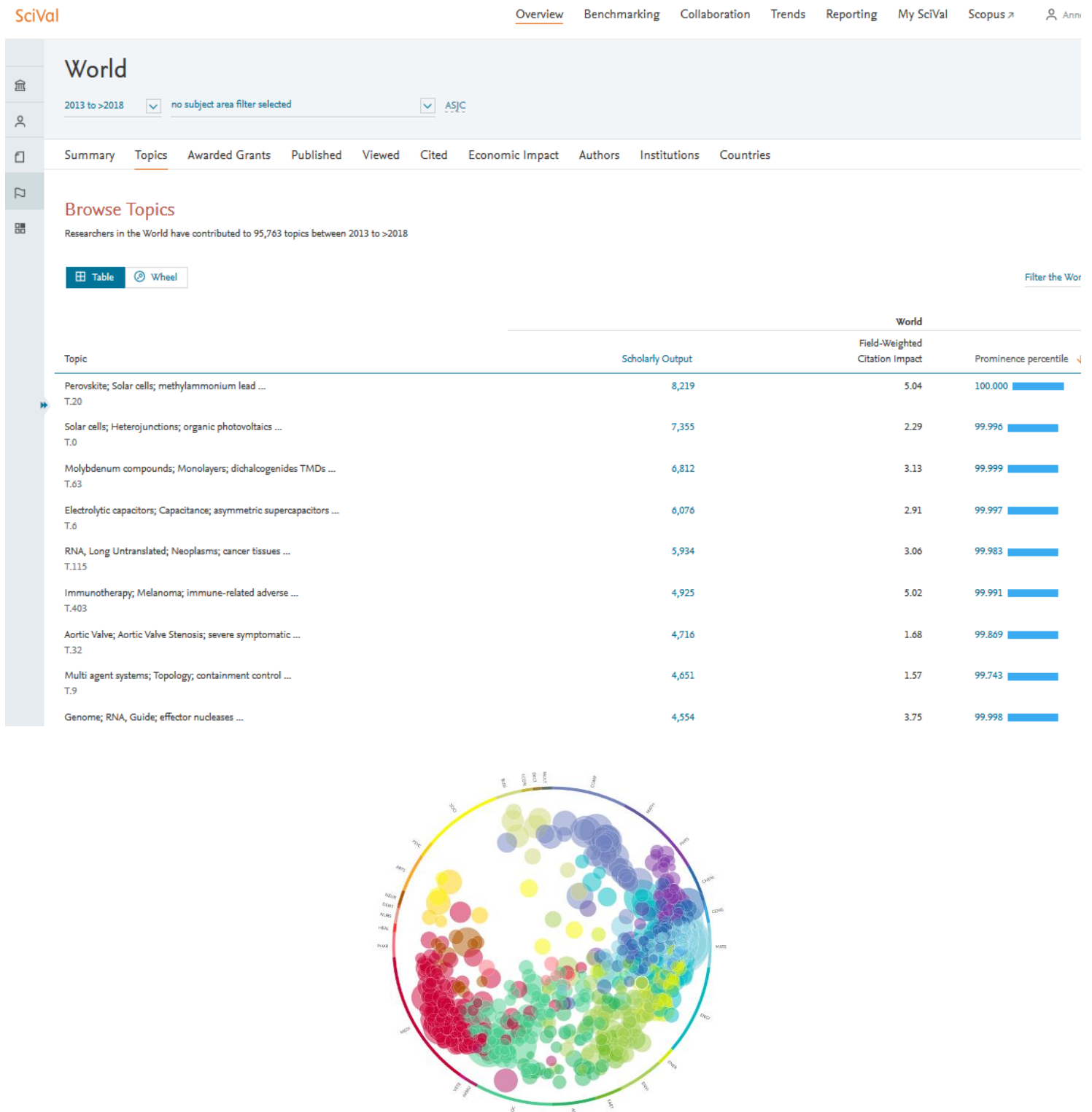
“If we split the grants into two 3-year periods, we find that Prominence combined with early funding explains 71 percent of the natural variability in funding in the final three years, which is extremely good,” Dr. Klavans explained. “From the perspective of SciVal, this adds a layer of usability that is a gold mine for researchers and research managers: looking at what has high Prominence can help them discover what is likely to get funded.”

Topic Prominence responds to the evolving information needs of researchers and policy makers. Dr. Klavans explained: “Every researcher needs to justify that what they’ve done has impact. But they also need to look at where they’re going in the future. For example, if you are

an academic, there might be five Topics related to the work you do. The Topic where you publish the most is very familiar to you – you know the leaders, their biases and where they apply for funding. That's your home community; although it might not get much funding, you care about it, and it's the source of your ideas. But look down the list and you might come across a different, related Topic that is rising in Prominence and therefore likely to be a way to generate funding for your work. Intelligence gathering is an important activity, and scientists are doing it all the time." As Dr. Klavans said, "the problem isn't knowing your own area, it's knowing your neighbor's neighbor."

Research managers and policy makers will look at different things even if it is worth adopting the researcher perspective as well. Which aspects of a Research field should we support and fund? How do disciplines combine among each other? How do you justify assigning funding to one research group over another? What evidence can you gather and use to advocate for the establishment of a new grant? Topic Prominence becomes a source of evidence and a talking point, both internally and externally. There's also a showcasing role, as Dr. Klavans explained: "If you want to demonstrate you're at the forefront, leading the way in a field, Prominence offers an alternative way of doing that, adding to the collection of indicators already available that demonstrate research impact. It's a good way to decide which areas of research to put in the spotlight."

6th International Conference on Future-Oriented Technology Analysis (FTA) – Future in the Making
Brussels, 4-5 June 2018



*A sample of the world's dashboard with the "Wheel of Science," which gives an overview of the Prominent Topics the researchers are active in, here worldwide.
(Source: SciVal's Topic Prominence in Science; Scopus data)*

This powerful predictive analysis is already providing valuable insights for institutions around the world. [Simoné “Sisi” Nutt](#), User Experience Lead for SciVal, has tested Topic Prominence with SciVal users, and the feedback has been overwhelmingly positive: “We haven’t seen anything like this before in terms of meeting users’ needs. Researchers relate to the Topics, they strongly resonate. The Topics really reflect the research areas they are working in, who they’re working with, the topics they’re covering and what their peers are doing. Researchers are excited by the possibilities Prominence gives them, and they are already getting a better understanding of what they can do to make their grant applications more efficient and their work more impactful.”

Also, having the ability to cut across the traditional journal subject areas, and zoom into the granularity of analysis the citation clustering brings, more powerfully reflects what is actually happening in science, who is leading and who is emerging, what competitors are doing – and on which Topic *exactly*.

Table

Wheel

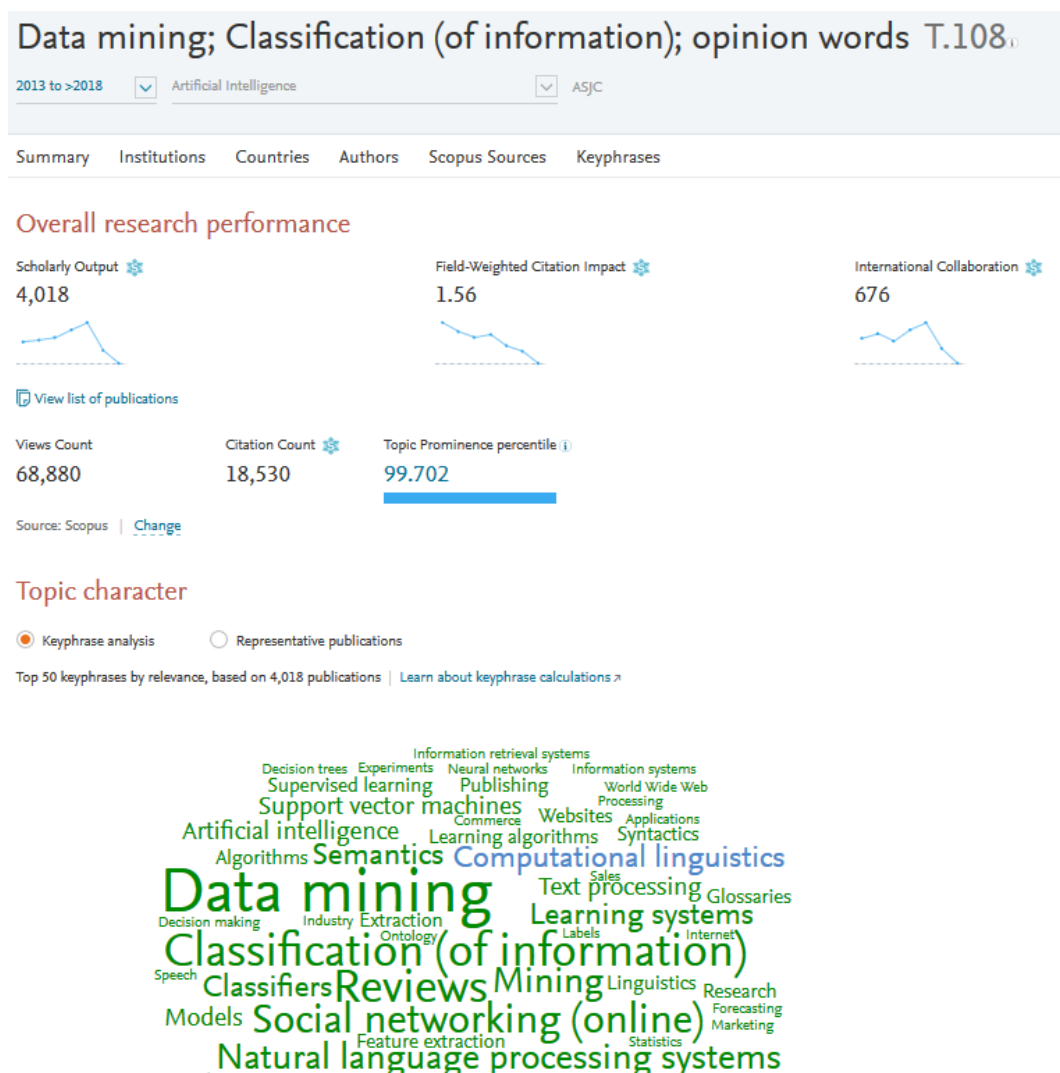
Search this Institution's Topics

Topic	At this Institution			Worldwide
	Scholarly Output	Publication Share	Field-Weighted Citation Impact	Prominence percentile
planet; planets; planet candidates ... T.131 Analyze Topic at Institution Analyze Topic worldwide	348	16.69%	3.17	99.733
HIV-1; Antibodies, Neutralizing; external region ... T.404	245	17.64%	5.42	99.812
Genome; RNA, Guide; effector nucleases ... T.456	237	8.09%	15.26	99.997
disks; protoplanetary disks; outer disk ... T.179	217	18.45%	2.04	99.186
Melanoma; Proto-Oncogene Proteins B-raf; MAPK pathway ... T.636	197	8.51%	8.39	99.831
General Surgery; Income; surgical disease ... T.5367	168	16.14%	3.25	97.850
masers; star formation; high-mass star-forming ... T.264	165	14.16%	1.33	98.028
Brain; Transcranial Magnetic Stimulation; anodal transcranial ... T.2202	164	9.84%	3.39	99.865

The table view displays the Topics an Institution is active in, which serves as a starting point for further analysis. Those results can be sorted by the Institution’s Scholarly output or by the Global Prominence of the Topic. (Source: SciVal’s Topic Prominence in Science; Scopus data)

At first glance, having so many Topics seems to complicate the picture, where traditionally 27 subject areas with 334 subcategories would have been used. But repeatedly over the last 25

years, it was noticed that researchers identify much more closely with fine-grained Topics that define the structure of science at the problem level. The granularity also makes analysis easier, opening the door to this new indicator.



The table view displays the Trends in one given Topic of Prominence, including the Field weighted citation impact, the most frequent keyphrases, and access within various tabs to the players being the most active.
(Source: SciVal's Topic Prominence in Science; Scopus data)

"We will learn more over time to what extent research managers and policy makers would need a more aggregated level of Topics," said Dr. Klavans. "If that is the case, it's something we can implement. From what we've seen, the fine-grained Topics are welcomed by researchers. Overall we're excited to offer our audience this window onto their future planning."

Conclusions

The Topics of Prominence can support in an innovative way the policy-makers allowing analyzes of the emerging Research and Innovation domains which require support through policies. New research trends in fields such as fracking, GMOs may have been identified earlier, hence benefitting from adjusted strategies. Our methodology makes also possible to identify the emerging fields in which Europe is ahead and for which funding is needed in order to remain on the cutting edge. And when lagging behind in strategic fields, this helps make decisions speeding up Research developments or developing accurate collaborations.

References

Richard Klavans, Kevin W. Boyack, (2017), Research portfolio analysis and topic prominence, *Journal of Infometrics*, 11, 1158-1178

Those references below are the ones of the above paper:

Börner, K., Klavans, R., Patek, M., Zoss, A. M., Biberstine, J. R., Light, R. P., et al. (2012). Design and update of a classification system: the UCSD map of science. *PLoS ONE*, 7(7), e39464.

Boyack, K. W., & Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology*, 61(12), 2389–2404.

Boyack, K. W., & Klavans, R. (2014a). Creation of a highly detailed, dynamic, global model and map of science. *Journal of the Association for Information Science and Technology*, 65(4), 670–685.

Boyack, K. W., & Klavans, R. (2014b). Including non-source items in a large-scale map of science: What difference does it make? *Journal of Informetrics*, 8, 569–580.

Boyack, K. W., Newman, D., Duhon, R. J., Klavans, R., Patek, M., Biberstine, J. R., et al. (2011). Clustering more than two million biomedical publications: Comparing the accuracies of nine text-based similarity approaches. *PLoS ONE*, 6(3), e18029.

Boyack, K. W., Klavans, R., Small, H., & Ungar, L. (2014). Characterizing the emergence of two nanotechnology topics using a contemporaneous global micro-model of science. *Journal of Engineering and Technology Management*, 32, 147–159.

Boyack, K. W. (2017). Investigating the effect of global data on topic detection. *Scientometrics*, 111, 999–1015.

Chen, C., Ibekwe-SanJuan, F., & Hou, J. (2010). The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis. *Journal of the American Society for Information Science and Technology*, 61(7), 1386–1409.

Clarivate. (2016). Research fronts 2016.

Crane, D. (1972). *Invisible colleges. Diffusion of knowledge in scientific communities*. Chicago: University of Chicago Press.

Dalrymple, D. G. (2006). Setting the agenda for science and technology in the public sector: The case of international agricultural research. *Science and Public Policy*, 33(4), 277–290.

Emmons, S., Kobourov, S., Gallant, M., & Börner, K. (2016). Analysis of network clustering algorithms and cluster quality metrics at scale. *PLoS ONE*, 11(7), e0159161.

Evans, J. A., Shim, J.-M., & Ioannidis, J. P. A. (2014). Attention to local health burden and the global disparity of health research. *PLoS ONE*, 9(4), e90147.

Fisher, R. L. (2005). *The research productivity of scientists: How gender, organizational culture, and the problem choice process influence the productivity of scientists*. Lanham: University Press of America.

Foster, J. G., Rzhetsky, A., & Evans, J. A. (2015). Tradition and innovation in scientists' research strategies. *American Sociological Review*, 80(5), 875–908.

- Glänzel, W., & Heeffer, S. (2014). Cross-national preferences and similarities in downloads and citations of scientific articles: a pilot study. In *Science and technology indicators conference 2014* (pp. 207–215).
- Gläser, J., Glänzel, W., & Scharnhorst, A. (2017). Same data – different results? Towards a comparative approach to the identification of thematic structures in science. *Scientometrics*, 111, 981–998.
- Grandjean, P., Eriksen, M. L., Ellegaard, O., & Wallin, J. A. (2011). The Matthew effect in environmental science publication: A bibliometric analysis of chemical substances in journal articles. *Environmental Health*, 11, 96.
- Herbert, D. L., Barnett, A. G., Clarke, P., & Graves, N. (2013). On the time spent preparing grant proposals: An observational study of Australian researchers. *BMJ Open*, 3, e002800.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520, 429–431.
- Hicks, D. (2004). The four literatures of the social science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research: The use of publication and patent statistics in studies of S&T systems* (pp. 473–495). Dordrecht: Springer.
- Klavans, R., & Boyack, K. W. (2008). Thought leadership: A new indicator for national and institutional comparison. *Scientometrics*, 75(2), 239–250.
- Klavans, R., & Boyack, K. W. (2009). Toward a consensus map of science. *Journal of the American Society for Information Science and Technology*, 60(3), 455–476.
- Klavans, R., & Boyack, K. W. (2011). Using global mapping to create more accurate document-level maps of research fields. *Journal of the American Society for Information Science and Technology*, 62(1), 1–18.
- Klavans, R., & Boyack, K. W. (2017a). The research focus of nations: Economic vs. altruistic motivations. *PLoS ONE*, 12, e169383.
- Klavans, R., & Boyack, K. W. (2017b). Which type of citation analysis generates the most accurate taxonomy of scientific and technical knowledge? *Journal of the Association for Information Science and Technology*, 68(4), 984–998.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.
- Kurtz, M. J., & Bollen, J. (2010). Usage bibliometrics. *Annual Review of Information Science and Technology*, 44, 3–64.
- Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact. *Journal of the Association for Information Science and Technology*, 66(7), 1323–1332.
- Lewis, M. L. (2003). *Moneyball: The art of winning an unfair game*. W. W Norton & Company.
- Martin, S., Brown, W. M., Klavans, R., & Boyack, K. W. (2011). OpenOrd: An open-source toolbox for large graph layout. *Proceedings of SPIE – The International Society for Optical Engineering*, 7868, 786806.
- Moed, H. (2005). *Citation analysis in research evaluation*. Dordrecht: Springer.
- Moed, H. F., & Halevi, G. (2016). On full text download and citation distributions in scientific-scholarly journals. *Journal of the Association for Information Science and Technology*, 67(2), 412–431.
- Mullins, N. C. (1973). *Theories and theory groups in contemporary American sociology*. New York: Harper & Row.
- National Science Board. (2016). *Science and Engineering Indicators 2016*. Arlington, VA: National Science Foundation (NSB 2016-1).
- Price, D. J. D., & Gürsey, S. (1975). Studies in scientometrics I: Transience and continuance in scientific authorship. *Ci. Inf. Rio de Janeiro*, 4(1), 27–40.
- Price, D. J. D. (1963). *Little science, big science*. New York: Columbia University Press.
- Rafols, I., Ciarli, T., & Chavarro, D. (2015). Under-reporting research relevant to local needs in the global south. Database biases in the representation of knowledge on rice. In *15th international conference of the international society on scientometrics and informetrics*.
- Ruiz-Castillo, J., & Waltman, L. (2015). Field-normalized citation impact indicators using algorithmically constructed classification systems of science. *Journal of Informetrics*, 9, 102–117.
- Sarewitz, D., & Pielke, R. A., Jr. (2007). The neglected heart of science policy: Reconciling supply of and demand for science. *Environmental Science & Policy*, 10, 5–16.
- Schulz, C., Mazloumian, A., Petersen, A. M., Penner, O., & Helbing, D. (2014). Exploiting citation networks for large-scale author name disambiguation. *EPJ Data Science*, 1, 11.
- Small, H., Boyack, K. W., & Klavans, R. (2014). Identifying emerging topics in science and technology. *Research Policy*, 43, 1450–1467.
- Small, H. (1976). Structural dynamics of scientific literature. *International Classification*, 3(2), 67–74.
- Small, H. (2009). Critical threshold for co-citation clusters and emergence of the giant component. *Journal of Informetrics*, 3, 332–340.

- Sparck Jones, K., Walker, S., & Robertson, S. E. (2000). A probabilistic model of information retrieval: development and comparative experiments. Part 1. *Information Processing & Management*, 36(6), 779–808.
- Von Hippel, T., & Von Hippel, C. (2015). To apply or not to apply: A survey analysis of grant writing costs and benefits. *PLoS ONE*, 10(3), e0118494.
- Wallace, M. L., & Rafols, I. (2015). Research portfolio analysis in science policy: Moving from financial returns to societal benefits. *Minerva*, 53(89–115).
- Waltman, L., & van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, 63(12), 2378–2392.
- Waltman, L., & van Eck, N. J. (2013). A smart local moving algorithm for large-scale modularity-based community detection. *European Physical Journal B*, 86, 471.
- Waltman, L., Boyack, K. W., Colavizza, G., & Van Eck, N. J. (2017). A principled approach for comparing relatedness measures for clustering publications. In 16th international conference of the international society on scientometrics and informetrics.
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10, 365–391.
- Wang, J., & Shapira, P. (2015). Is there a relationship between research sponsorship and publication impact? An analysis of funding acknowledgments in nanotechnology papers. *PLoS ONE*, 10(2), e0117727.
- Zuckerman, H. (1978). Theory choice and problem choice in science. *Sociological Inquiry*, 48(3–4), 65–95.