



CONFINED TO GROW?

PUBLICATION DYNAMICS AND THE PROLIFERATION OF SCIENTIFIC JOURNALS

■ **Istvan Daruka** – Institute of Semiconductor and Solid State Physics – Johannes Kepler University
Altenbergerstr. 69, 4040 Linz, Austria – Istvan.Daruka@jku.at – DOI: 10.1051/ePN/2014102

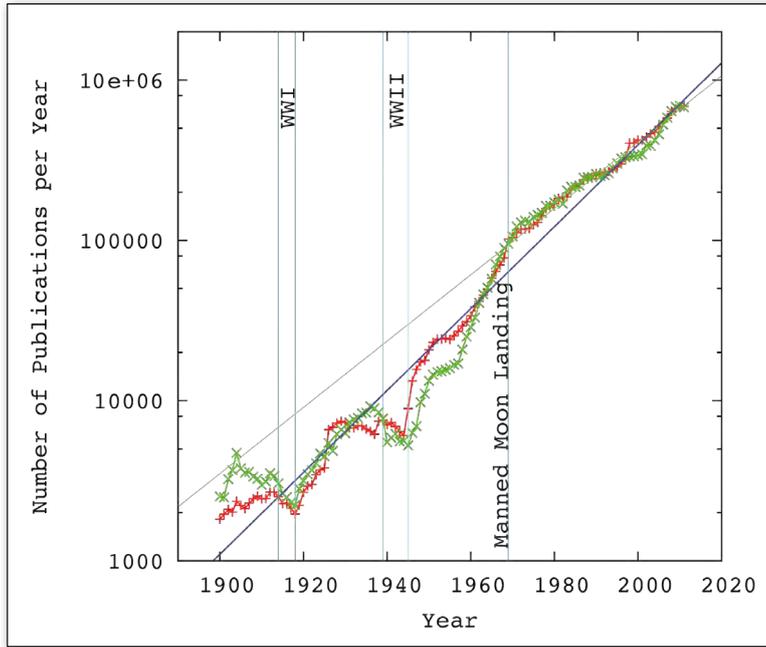
The number of research papers published yearly shows a staggering exponential growth. One can in fact witness a century-long publication avalanche (Fig.1). According to the two major databases INSPEC and Thomson Reuters, there is a 300-fold increase in the number of published items since the year 1900. The world of science has recently reached such a level of proliferating complexity that its structural evolution itself poses challenging scientific problems.

In Fig. 1 we plotted the number of published papers using the INSPEC database that represents the fields of physics, engineering, and information technology. As a control plot, we used the all-inclusive yearly publication data from Thomson Reuters Web of Science Database. The latter dataset was divided by a factor of 11 to overlap the data for better visualization. The two dips on the curves correspond to the First and Second World Wars. After WWII a rather rapid, cold war era burst can be observed lasting until 1970. After 1970 until today a somewhat smaller average slope sets in. Furthermore, it is intriguing to observe that in spite of the 11-fold difference in the average number of yearly publications,

both datasets show strongly correlated publication-record dynamics. This suggests a more generic growth pattern spanning through different research fields. Such an exponential increase over more than a century can be aligned with the empirical Moore's law describing the exponential growth of microprocessors transistor counts over the past four decades. We note that similar growth trends were established by the paper of Larsen *et al.* [1].

A few decades ago the number of scientific papers and the number of independent citations seemed to be adequate to assess a researcher's activity. Due to the unparalleled expansion of the scientific activity, this has

▲ © iStockPhoto



▲ FIG. 1: Total number of publications as a function of years using the INSPEC (green curve) and the Thomson Reuters Web of Knowledge databases (red curve). The latter curve was divided by a factor of 11 to obtain a data collapse. Two exponential functions are fitted to guide the eye.

fundamentally changed by, leading to the proliferation of scientometry itself [2]. After the introduction of the Hirsch-index in 2005, many other indices followed rapidly, such as the related Wu-index, the g-index, and the i10-index, just to name a few frequently used ones. Furthermore, these scientometric indices can also be used to assess the output of different research units (groups, laboratories, universities, etc.). During a rather short period of time, such refined indicators have become vitally important and play a more decisive role for research funding and even for determining carrier paths.

Quantitative scientometric comparison and ranking have become key factors in the competition for research funding. Therefore, it is of outmost importance to establish scientometric schemes allowing the best possible comparisons among different fields. In fact, different fields have significant differences in their citation frequency. For example, medicine and molecular biology receive much more citations than physics [3]. On the other hand, there are other research fields such as mathematics or sociology, which get along with even

less citations per paper [3]. There are promising developments regarding a field-based normalization of citation numbers that should enable a better comparison between fields [4], even though

the use of such measures remains limited [2].

The other prevailing problem – due to the rather diversified forms of scientific activities – is how to compare the contributions of large collaborations (for example in high-energy physics) to those of small groups or

individuals. Researchers at Harvard University offered an effective way to alleviate this problem by introducing an automated citation normalization service [5]. Their normalization scheme simply divides the received number of citations by the number of co-authors of a paper. According to this methodology, the reduction in the number of citations is probably the most remarkable in high-energy physics where papers with a thousand co-authors are not uncommon.

The Society of Scientific Journals

Citation dynamics can also be used to trace the developmental activity of scientific fields and to determine their functional interrelation. A very insightful dynamical network, a map of science (Fig. 2) was created this way by Eigenfactor.org [6]. The evolution and the interaction of different fields has some analogy with the star formation in an early galaxy: new fields emerge, grow, while some others shrink in an interactive, dynamic ambiance.

The enormous multitude of papers is published in a wide range of an estimated 25.000 scientific journals. The number of scientific journals also exhibits a rapid growth at a rate of about 3% each year [7]. Journals can be ranked by their impact factors. In Fig. 3 the impact factors of 8300 scientific journals is plotted versus their ranking for two different years, as derived from the Thomson Reuters Journal Citation Report Database. A power-law with an approximate slope of -0.5 nicely describes these distributions. Interestingly, the same exponent was obtained for the citation rank distribution of scientific publications [8].

Power-law for ranking distributions might indicate hierarchical correlations among the constituents involved. One can think of the Zipf plots (*i.e.*, log of the variable being analyzed vs. log of the rank) also for word occurrence distributions in different languages, or city size distributions. Such a hierarchical arrangement could explain the related growth of the number of scientific journals. A new discovery induces a flow of workforce and leads to a related nucleation of new journals in a hierarchical way [7,9].

According to Fig. 3, one can intuitively distinguish three classes of journals: top journals (trend setters), middle-class competitors (power-law scaling regime), and follower-class journals. Such a hierarchical structure of the middle-class journals can be compared with the case of turbulence. A large eddy, *i.e.*, a pioneering scientific idea or a new methodology, creates smaller ones in a cascading way, leading to a self-similar, power-law scaling behavior [9].

The parallel citation and impact factor competitions induce a dual, interconnected expansion. Recently, Physical Review Letters and Applied Physics Letters opened up the usage of an extra page for citations only. Such a change will probably further inflate their own impact factors, but at the same time will contribute to an overall inflation in the use of references for bibliometric purposes

“ The number of yearly published research papers shows a staggering exponential growth, one can in fact witness a century-long publication avalanche ”

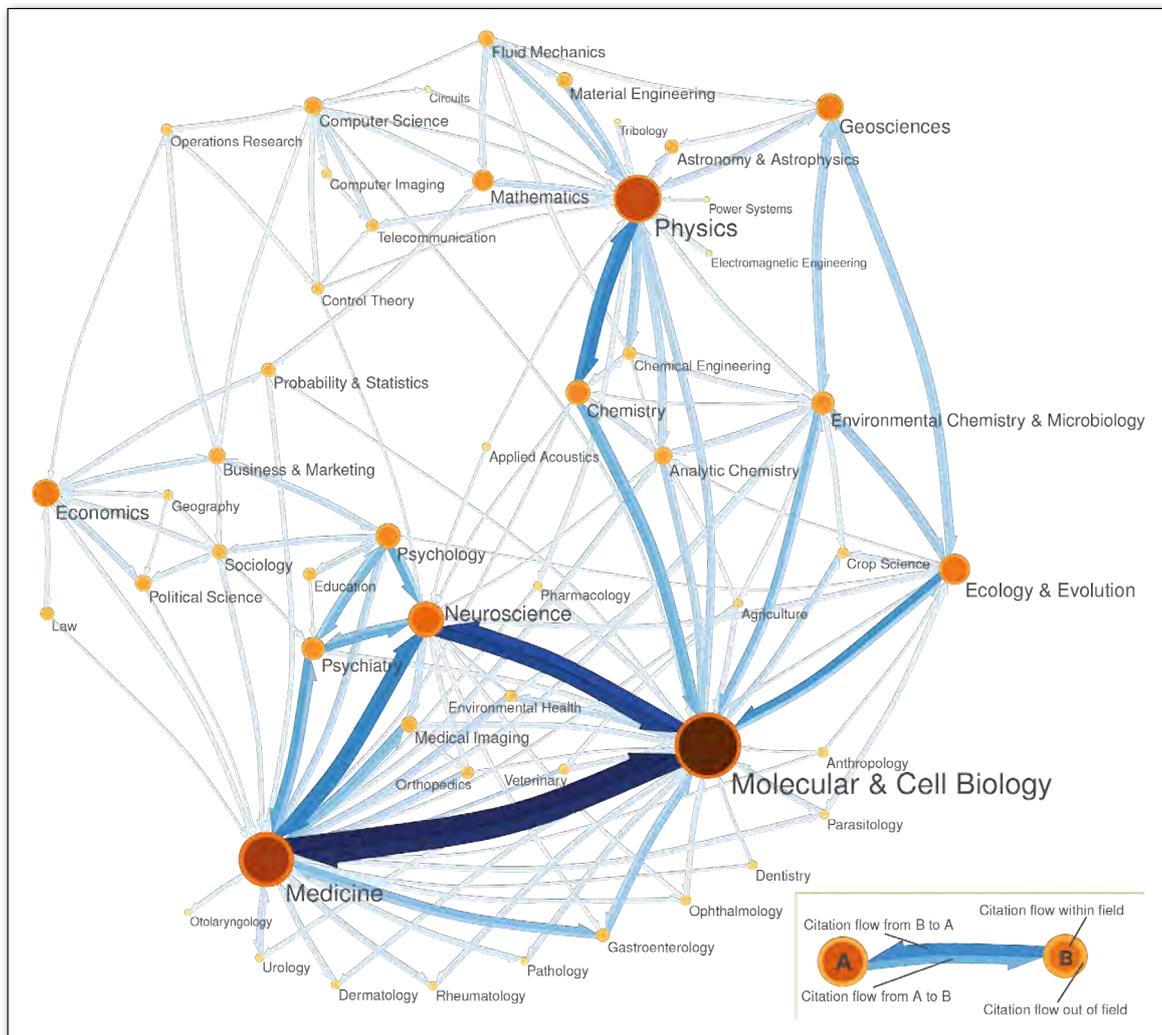
[3]. Another influent parameter will be the generalization of the Open Access models, which opens new avenues to publishers by providing a barrier-free access to their publications [10]. However, the long-term effects of such novelties are not easy to foresee.

Table 1 shows the top ten research journals sorted by their impact factors for years 2012 and 2005. The impact factor vs. rank of Fig. 3 and Table 1 demonstrates the aforementioned expansion process: one can in fact witness a 20-30 percent overall inflation of the journal impact factors over the considered short period of time. This corresponds to an annual impact factor inflation rate of 3%, in agreement with the earlier findings of [3]. For the leading impact factor, one can even see a startling tripling. This hyper-expansional trend could probably be attributed to an enrichment/segregation process to publish in the most prestigious journals.

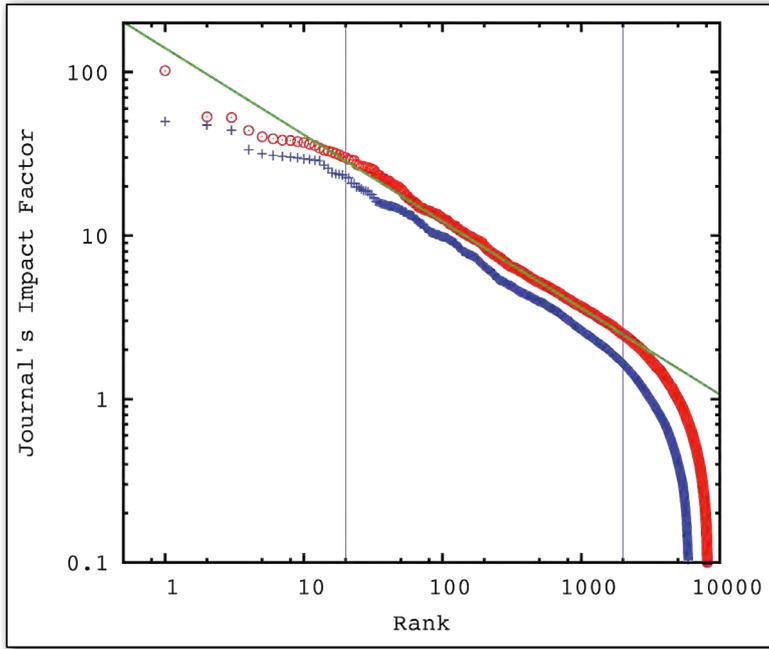
Perspectives on Scientific Research and Publication Dynamics

The discussed bursting trends, both in publication dynamics and in the growth of the society of scientific journals, support the intuitive picture that positive feedbacks are inherently encoded into the scientific activity.

The rich-gets-richer effect [9,11] creates segregation/broadening trends, which contributes to more competitive research environments but also to a rather broad and diverse carrier spectrum. At the top end, one can speak of star scientists, who are often skilled managers and businessmen at the same time. At the other end, one can find the existential scientists who constantly work for their survival and contribute mainly to an incremental research. Furthermore, the increasing pace at which the research environment changes seems to lead to an overall loss of job security [12]. The job spectrum seems to shift towards more temporal options requiring faster adaptation and higher mobility at the same time. In terms of the development of research institutes and collaborations, this novel dynamics might be somewhat similar to the process of Ostwald ripening: the larger/more prestigious units grow at the expense of the smaller/provincial ones in a research environment with limited resources. The concomitant dispersion of workforce might lead to a significant broadening of the pursued research quality as well. This trend, together with the hyper-competition induced compulsive publication habits pose severe problems that are in the focus of very recent discussions [13].



◀ FIG. 2: A map of science, based on the citation network within and among a vast range of scientific fields. [Reproduced with the kind permission of C.T. Bergstrom, M. Rosvall, and J. West at eigenfactor.org]



▲ FIG. 3: Rank distribution of scientific journals' impact factors for years 2011 (red curve) and 2005 (blue curve) using the Thomson Reuters Journal Citation Report Database. Low rank numbers mean high reputation.

With the ever accelerating pace of research it is natural to ask whether this steady state of growth has actually become the ultimate goal of scientific research itself. Indeed, mega/giga/tera-projects such as large scale particle accelerators, knowledge accelerators to be funded by EU Flagship calls [14], extensive scientific networks of super-centers around the globe, data/text mining robots, together with the unparalleled growth of computational and modeling capacity [15] for “big data” are likely to outline the beginning of a new emerging era. Is science actually confined to grow and accelerate further?

If such inflative publication trends further prevail, by the end of this century the highest journal impact factors would probably exceed 1000. In this way, we might have more than 100.000 peer-reviewed scientific journals available with a total mass of 10.000.000 yearly publications. Such an enormous growth potential will certainly raise several socio-economical concerns and stability issues, the analysis of which extends well beyond the scope of this paper.

▼ TABLE 1: Top ten research journals sorted by their 2012 impact factors, using the Thomson Reuters Journal Citation Report Database.

Rank (2012)	Journal	IF (2012)	IF (2005)
1	CA-A Cancer Journal for Clinicians	153.459	49.794
2	New England Journal of Medicine	51.658	44.016
3	Reviews of Modern Physics	44.982	30.254
4	Chemical Reviews	41.298	20.869
5	Nature Reviews Genetics	41.063	19.211
6	Lancet	39.060	23.878
7	Nature	38.597	29.273
8	Nature Reviews Molecular Cell Biology	37.162	29.852
9	Annual Review of Immunology	36.556	47.400
10	Nature Materials	35.749	15.941

In conclusion, it is believed that some of the potentially dangerous runaway trends in publication dynamics are triggered by underlying epistemological features of scientific research itself. In order to attain a deeper understanding of the involved branching processes and to cultivate a sustainable scientific development, the paper calls for an open, multilateral discussion on the mitigation of the multiplication trends in the publication landscape. ■

About the Author

Istvan Daruka graduated in physics from the University of Notre Dame (USA). After working at several research institutions worldwide, he has been affiliated to the Johannes Kepler University in Linz as a guest scientist since 2012. His research interest focuses on materials science and complex systems, including multidisciplinary questions on social dynamics.

Acknowledgements

The author acknowledges inspiring discussions with G. Brunthaler, A. Csoto, J. Messner, N. Mingo, F. Schäffler, and G. Springholz. Furthermore, the critical reading and valuable suggestions on the manuscript by C.P. Rossel are greatly appreciated.

References

- [1] P.O. Larsen and M. von Ins, *Scientometrics* **84**, 575 (2010).
- [2] R.V. Noorden, *Nature* **465**, 864 (2010).
- [3] B.M. Althouse, J.D. West, C.T. Bergstrom, and T. Bergstrom, *J. Am. Soc. Information Science and Technology* **60**, 27 (2009).
- [4] F. Radicchi, S. Fortunato, and C. Castellano, *PNAS* **105**, 17268 (2008).
- [5] The SAO/NASA Astrophysics Data System, <http://adsabs.harvard.edu/> (Sort by normalized citations.)
- [6] M. Rosvall and C.T. Bergstrom, *PNAS* **105**, 1118 (2008); see also www.eigenfactor.org, and www.MapEquation.org
- [7] M. Mabe and M. Amin, *Scientometrics* **51**, 147 (2001).
- [8] S. Redner, *Eur. Phys. J. B* **4**, 131 (1998).
- [9] A. Mazloumian, Y.-H. Eom, D. Helbing, S. Lozano, S. Fortunato, *PLoS ONE* **6**, e18975 (2011).
- [10] I. Sample, *The Guardian*, 24 April 2012. www.theguardian.com/science/2012/apr/24/harvard-university-journal-publishers-prices
- [11] A.M. Petersen, W.-S. Jung, J.-S. Yang, and H. E. Stanley, *PNAS* **108**, 18 (2011).
- [12] V. Doronina, The Life Scientist Bubble, http://sciencecareers.sciencemag.org/career_magazine/previous_issues/articles/2013_08_28/caredit.a1300184; and Permanent jobs scarce, *Nature* **493**, 711 (2013).
- [13] *Science* **342**, 6154 (2013), featuring “Communication in Science”; *Nature* **502**, 7471 (2013), featuring “Impact”; *The Economist*, “How Science Goes Wrong”, October 19th (2013).
- [14] K. Kupferschmidt, *Science* **339**, 28 (2013).
- [15] A. Szalay and J. Gray, *Nature* **440**, 413 (2006).