



A Collapsing Academy, Part III: Scientometrics and Metric Mania

Hal Berghel, University of Nevada, Las Vegas

In the golden era of higher education, an institution advertised itself by Latin inscriptions like fiat lux. Now the operative watch phrase is fiat numerus.

When it comes to higher education in the United States, the 20th century will be known as the century of retrodorsal, reactive, regression. For all of the progress made in science and technology in the past century, we have precious little to show for an understanding about the nature and causes of our success. Don't misunderstand me. Science and technology are doing fine by themselves—it's our post facto assessment of how we accomplished what we did that falls short. Nowhere is this more obvious than in the penchant for the absurd quantification of research quality by citation metrics.

Digital Object Identifier 10.1109/MC.2022.3142542
Date of current version: 11 March 2022

To repurpose one of Edsger Dijkstra's observations, this has been a mistake carried through to perfection.

METRIC MANIA

Socrates is reported to have said that an unexamined life is not worth living. In the academy, this amounts to continually challenging ourselves with questions. Are we accomplishing anything of enduring value? Is what we do worth doing? Instead, we all too often fall into herd-like behavior and substitute relatively

meaningless goals like, for example, maximizing citation scores. Unfortunately, this has found acceptance in the academy. At any given moment in time, it seems like the academy is one bad idea, strategic plan, or strength, weakness, opportunities and threat (SWOT) analysis away from total breakdown. When it comes to the assessment of scholarship, we've wandered away from legitimate appraisal to bean counting. To further compound the problem, we're even counting the wrong beans.

Trying to quantify the quality of scholarship and research is akin to measuring the intensity of emotion: we may recognize qualitative difference when we see it, but it's impossible to score it objectively, fairly, and without bias. But in the world of a self-absorbed, short-term, upwardly mobile,

narrow-focused, criticism-averse modern university administrator, bureaucrat, politician, or corporate leader, bean counting all too often becomes a go-to tool for institutional assessment. A textbook case is the rank-and-yank personnel evaluation program that contributed so significantly to Enron's ascendancy as the corporate paradigm of excellence it enjoys today.

To be fair, such administrative shortcomings are in no small measure a response to myriad external pressures that come with the job. A central administrator has to simultaneously satisfy key stakeholders, like trustees and regents, politicians, potential and benefactors, legislators, and business leaders—any one of whom can throw the proverbial guano into the institutional punchbowl. It goes without saying that faculty and students are above reproach (I say this with tongue firmly placed in cheek). This falls under the rubric of a kind of administrative anorexia wherein administrators discover that, while their ability to actually improve their academy is largely beyond their control, they are in total control of the flow of memos, reports, budget requests, meeting schedules, off-campus retreats, strategic plans, SWOT analyses, and evaluation cycles. Therefore, these latter activities occupy much of their waking moments because—wait for it—they can control them. The watchword is: if you can't figure out a way to achieve institutional greatness, issue a memo, demand a report, or schedule an off-campus retreat. At least that way you can document that you've done something for your salary.

STAGES OF SOPHISTRY

The use of metrics in the evaluation of scholarship appears to have begun in the 1950s when Eugene Garfield created the field of scientometrics—the quantitative measurement of scholarly work—for his Institute for Scientific Information.¹ This evolved into the cottage industry of citation indexing of scholarly journals that

we know today. To be clear, I take no particular issue with either Garfield's scholarly work or citation indexing in general. Both fall within the sphere of legitimate curiosity-inspired research. Rather, it is the subsequent use to which it has been put that is problematic. What has happened is that the relatively obscure and moderately interesting study of scientometrics has been adopted as a legitimate indicator of quality, relevance, and importance in the evaluations and assessments of faculty scholarship—in some cases, on equal footing with peer review. Some faculty now wear their citation index scores as a badge of honor without any critical assessment of the accuracy, precision, or value of the processes involved. Regrettably, scientometrics is too often being considered on a par with more relevant, subjective, and impossible-to-quantify criteria in decision making when it comes to the evaluation of scholarship. The de facto mantra of this metric mania should be "cost savings before insight." There ought to be a placard to that effect on the wall of every provost and dean so that prospective faculty realize what they're up against when it's time for promotion and tenure.

In an earlier academy (pre-1980), the assessment of scholarly prowess relied primarily on local peers—colleagues and academic administrators (as opposed to career administrators)—complemented with external peer review from alleged experts. In most subject areas, the grist for this mill was primarily scholarly and published work presented in some form appropriate to the discipline. In that bygone era, it was not unusual for all responsible parties to actually read publications and make an informed judgment of quality. This method was not without shortcomings, but it was light-years ahead of bean counting.

But those halcyon days are largely past. A by-product of our current post-modern, digital age is this enthusiastic acceptance of putatively objective measures like metrics to "assist" in

scholarly assessments. Many varieties have been developed for such purposes, and several online resources utilize them. Google Scholar and Microsoft Academic Search are well-known websites that build metrics into their indexing services. Computer professionals who are long enough in the tooth will recall that the idea of web resources that would integrate document search with some sort of metadata analysis dates back at least to the 1990s, when CiteSeer was developed at the NEC Research Institute.² While CiteSeer's goal focused primarily on the addition of citations to the index of bibliographic entries in a central database or digital library, later web tools sought the expansion to all online resources. The slide from listing citations to calculating metrics was transformative, for the latter can be used like a rapier by the unenlightened and self-serving administrator—not to mention narcissistic, competitive faculty. As the saying goes, figures don't lie, but liars figure.

What must be remembered is that metrics are disentangled from the underlying scholarship and can be manipulated into pretty much any narrative one chooses because they are decontextualized. This flexibility makes the reliance on metrics inherently subject to abuse. There is a close parallel between metrics and scholarship, and movie reviews and motion pictures. But with movies, it is understood by all that reviews may be ad hoc and arbitrary and are not always reliable indicators of quality. However, with scholarship, metrics are assumed to be reliable because they are quantitative in nature, objectively determined, and based on publicly accessible data derived from the Internet. What could possibly be wrong with that?

Clearly, indexing hyperlinks of scholarly resources, analyzing and comparing their document metadata, indexing and analyzing nontextual data such as imagery and graphs, performing text analysis, and using a host of advanced file management techniques that

support efficient and effective comprehension, and so on are legitimate activities in service to scholarship. These ideas can all be traced back to the pioneering work of Vannevar Bush,³ Theodor Nelson,⁴ and Douglas Engelbart,⁵ dating back more than a half century. (Nelson's account of much of this history has been captured in a video interview by Devon Zuegel; see <https://www.notion.so/blog/ted-nelson>.) But the intention of this triumvirate was to facilitate the acquisition of knowledge, not to evaluate and rank it. The difference between comprehension and critique is profound and irreconcilable.

METRICITY DECONSTRUCTED

So what metrics might be useful to measure scholarly quality? Most academics would agree that any metric worthy of the name should measure impact. That should be our starting point. But how do we do that? Worthy assessment of a scholarly paper, for example, might include a spectrum of reasonable measures that consider a multitude of factors. For example:

1. Did anyone read it?
2. Did anyone react to it (for example, with a review or a letter to the editor)?
3. Did any teachers outside the author's circle of influence assign it to their students? Is there any reason to suspect that such students reacted positively?
4. Was it included as a course reading assignment by teachers?
5. Was it referenced in M.S. theses or Ph.D. dissertations?
6. Did anyone support or challenge it in a review?
7. Was it republished or reprinted?
8. If it was included in a digital library, how many full-page downloads did it produce?
9. Do any standard reference sources (encyclopedias or texts) refer to it?

10. Was it cited? If so, by whom?
 - a. How many citations were self-citations?
 - b. How many citations were produced by coauthors and students?
 - c. How many coauthors were involved?
 - d. How many citations were venue specific?
 - e. How many citations were external to the knowledge domain?
 - f. How many citations were produced by domain-knowledge experts?
 - g. What percentage of citations was produced by archival (journal) publications?

Note that citation counting is but one measure of many. And, like the other measures, it has its own unique set of problems. For one, using citation metrics in the measurement of impact must necessarily include an estimate of the credibility and importance of the citation sources. Obviously, in the special case where the list of coauthors is coextensive with the sources of the citations, we might question whether the citations mean much. The phrase "write-only publication" has been used to describe such cases. At the other extreme, citations could come from leading scholars in the field who have little or no connection to the authors. In this case, the citations would seem to be a much more accurate measure of impact. Unfortunately, any situation between these two extremes requires interpretation, and any reasonable interpretation would require contextualization, which in turn would require reading the source material and having familiarity with the literature. So for the majority of cases, the interpretation of the metrics takes as much effort as the old-timey method during the halcyon days.

In addition, special consideration must be given to self-citations, which measure only the esteem in which

the author(s) regard their own work. While self-citations are useful for reconstructing research lineage, they are not valuable for much else. We may generalize one step beyond self-citation to citations produced by "circles of influence," including the authors and subsequent generations of their students and postgraduates. These citation threads work much as natural selection does in evolution: long-term success cannot be achieved in the absence of diversity. And, as with evolution, the long-term effects of inbreeding in scholarship are likely to be undesirable and unsustainable. Such was the case with Lysenkoism, geocentric astronomy, cold fusion, and the endless outpouring of cargo cult science⁶ that has been enjoying a renaissance for the past 50 years. It is worth noting that this very phenomenon caused Einstein to criticize quantum physics. It wasn't that he thought quantum physics was wrong; he considered it to be an incomplete product of a herd mentality. Predictive capacity wasn't enough for Einstein. Explanatory capacity was also required. Quantum entanglement without the latter was but spooky action at a distance—akin to vaporware in computing.

You can see where this is going: the use of citations as a measure of quality is necessarily sketchy. Citations are like recommendations. They are only as good as the citer/recommender. This leads to an infinite regress as the burden of credibility passes down step by step through succeeding generations of recommendations. What is more, naked metrics do not distinguish between critical citations and supportive ones. So we need another metric for the strength of each endorsement, and so forth. In short, citation metrics are worthless apart from interpretation, and meaningful interpretation requires the same level of domain knowledge and understanding as peer review.

And so far, we've only dealt with parameter 10. What about 1–9? Are they equal in value to 10? They are certainly

relevant in some circumstances, so why aren't they also quantified? The plain fact is that only parameter 10—citations—is easy to count and simple to automate. Citations are the lowest-hanging fruit on the tree of assessment, not because they provide any more meaningful assessments than the alternatives, but rather because they are inexpensive to generate. Let's label citation analysis for what it is: scholarly evaluation on the cheap.

The point is that questions like 1-10 are all collectively necessary for any thorough evaluation of scholarship, but no subset is, in itself, sufficient. All require an interpretation cycle that is as challenging as the understanding of the scholarly work that they purport to represent. Failure to appreciate this is the first of two primary causes of the evaluative nonsense facing university faculty today. I'll drive home this point with two extreme examples: the 15 May 2015 issue of *Physical Review Letters*, announcing the discovery of the Higgs boson mass by the CERN Large Hadron Collider, which listed 5,154 coauthors,⁷ and a 1996 article by Alan Sokol, apropos of nothing in particular, in a social science journal.⁸

In the case of the nine-page CERN article, the authors/page ratio is 573:1, and the authors/word ratio is 1:1. These are accurate, objective, and quantitative measures, to be sure. But no reasonable person would suggest that they accurately reflect the importance of the article. One wonders whether the coauthor list is due to gatecrashing (coauthors are listed who are largely unfamiliar with the content), proforma recognition of level of effort, horse trading between coauthors, misuse of authority or power, or the inflation of author lists through group, lab, institutional, or political associations. There is no way to interpret these ratios (metrics) without context. About all that we can conclude from the metrics is that the optics are wrong because they involve the reader in the politics behind the publication.⁹

While the CERN article example illustrates the difficulty in apportioning scholarly credit, the Sokol article highlights the fragility of the peer-review system.¹⁰ Sokol submitted a paper that was obtuse, unintelligible nonsense. He suspected that inattentive editors might be so biased toward a submission that appeared to be rigorous and formal that they would fail to verify the relevance of the references. He was correct. Sokol snowed the editors with footwork that purported to integrate weighty topics like relativity, quantum physics, Minkowskian space-time, and Einstein's field equations with philosophical topics such as phenomenology, semiotics, deconstructivism, and hermeneutics. The prospect that quantum physics and relativity could be unified with the social sciences was a temptation too great to be ignored, so the editors published it and, in so doing, received the 1996 Ig Nobel Prize "for eagerly publishing research that they could not understand, that the author said was meaningless, and which claimed that reality does not exist."¹¹ (The induction ceremony is available online on C-SPAN2 at <https://www.c-span.org/video/?75657-1/1996-ig-nobel-awards>.) In its own way, the Sokol hoax confirms our point that there are no shortcuts in the evaluation of scholarship. As the good folks at literacy central have been saying for 70 years, reading is fundamental (<https://www.rif.org/>).

These two examples illustrate just how deceptive citations can be. It isn't at all clear what a citation to the CERN article would imply for any particular author, and, in the case of Sokol's paper, it isn't clear whether any of the included citations should be assigned any significance. The critical point is that decontextualized citations are vacuous. Even a statistical correlation between number of citations and quality or importance requires a thoughtful, measured assessment of quality. Correlations can be coincidental. Failure to recognize this by universities is one of the two primary causes of the

evaluative nonsense facing university faculty today. The other is the elephant in the room: external funding, which we'll defer to another forum!

METRIC MANIA REVISITED

Metrics may be thought of as quantitative metadata.¹² Currently fashionable metrics include the Hirsch index (h-index), hi-k index (the h index over k years), m-values, Carbon h-index, and so on, all of which are assumed to be "useful" measures of research and publication quality in some circles. The h-index, which is defined as the number N of publications that have been cited at least N times in a rank-ordered list, seems to predominate. There is now an online cottage industry associated with the development of research quality metrics called "reputation" sites. One use of reputation sites is the assessment of the alleged impact of publications for purposes of faculty evaluation. Unfortunately, faculty have been encouraged to support these for-profit sites by creating "profiles" that provide metric summaries of their work because they are quantitative, objective, and, of course, free. But this is misguided. Throwing papers down a staircase and assigning weights depending on the number of stairs they have traversed is also quantitative, objective, and free, but it hardly qualifies as a useful tool for meaningful assessment. Encouraging the use of metric ranking services introduces a moral hazard into the evaluation of scholarly work—it encourages the participants to chase the metrics rather than focus on creating durable scholarship. It also blurs the distinction between correlation and causation.

The overarching concern of the use of metrics in such evaluation is the open question: What do such measures measure? It must be remembered that correlations are no guarantee of causal connection. They are just estimates, and there is no standard method for determining confidence in them.¹³ Consider the following two statistical "facts": 1) total revenue generated by

arcades correlates with the number of computer science doctorates awarded in the United States (98.51%, $r = 0.985065$), and 2) spending on science, space, and technology correlates with suicides by hanging, strangulation, and suffocation at the level of 99.79% ($r = 0.99789126$).¹⁴ Even if we concede that these correlations are objective and unbiased, we are certainly not committed to accept them as important, relevant, or useful. That metrics and statistics are not infallible guides to optimal decision making, can be misleading, and are supplements to, and not surrogates for, careful reasoning, are well documented¹⁵⁻¹⁷ but seemingly ignored by consumers of ranking services.

This is not to say that evaluative metrics have no use in the academy. For example, metrics can be useful at identifying so-called *predatory* or *deceptive* publishing (such as open-access publications that lack rigorous peer review). But, as University of Colorado librarian Jeffrey Beall discovered, such identifications can lead to defamation litigation, threats, and persecution.^{18,19} So personal risk may be attendant with the use of metrics in this domain.

The use of metrics in the evaluation of scholarship has had its share of detractors.²⁰ As Hicks et al. observe, “research evaluations are now routine and reliant on metrics. The problem is that evaluation is led by data rather than judgment.”²¹ The so-called Leiden Manifesto for research metrics offers 10 principles as a “distillation of best practice in metrics-based research assessment” that are worthy of consideration. These principles bring into focus the challenges of using metrics for this purpose. A parallel effort has been undertaken by the San Francisco Declaration on Research Assessment (DORA), which offers equally valuable suggestions for best practices for organizations, publishers, researchers, and funding agencies in their Declaration on Research Assessment.²² An online document titled “5 common myths about the perceived value of

journal impact factor (JIF)” is particularly noteworthy.²³ A separate 2015 report commissioned by the U.K. Higher Education Funding Council for England elaborates on the uses and misuses of research metrics and indicators,²⁴ and the supplementary correlation analysis revealed that author-based metrics yield different results than the peer-review process.²⁵ The companion extensive literature review is noteworthy.²⁶

there is nothing to prohibit the demand for scholarly excellence according to standards befitting the discipline. That demand is spoofy enough to avoid legal scrutiny.

Evaluative metrics should always be taken for what they are: anecdote, no more, no less. Consider how the greatest scholars in history would have fared were their work assessed on the basis of current metrics. One may get an idea of the absurdity of

Encouraging the use of metric ranking services introduces a moral hazard into the evaluation of scholarly work—it encourages the participants to chase the metrics rather than focus on creating durable scholarship.

So it is not that the employment of metrics in the evaluation of scholarship is without criticism that accounts for its increasing use, but rather that the criticism is largely ignored by evaluators. This is not an oversight, but an expedient! To deal with the political and ethical issues that accompany the misuse of metrics, one would have to challenge some basic tenets of academic capitalism, not the least of which are cost savings and reducing the administrative burden.

The moral hazards and perverse incentives that accompany the use of metrics in the academy are acceptable within the framework of the need for cost-effective, objective, and uncontentious evaluations. The evaluation of scholarship with metrics may be thought of as a parallel construction to another hidden metric-based evaluation based on external funding, where the standards are even less prescriptive but, in some disciplines, even more important. In a sense, the former metric can serve as a façade to the latter—required to give legitimacy to the formal process and administrative code. While it would be *déclassé* to demand of faculty that they become institutional profit centers in writing,

the current fascination with metrics in some circles by considering how Copernicus, Galileo, Newton, Einstein, Pasteur, and Salk would have reacted to the use of quantitative metrics in the evaluation of their work. The fact that Einstein’s work on relativity was passed over by the Nobel Committee should not be overlooked. We should look at h-indices, JIFs, and the like, just as we do Facebook likes, number of retweets, Amazon review scores, page views, number of downloads, and Google Scholar citation counts—along with collegiate football rankings, political and opinion polls, customer satisfaction surveys, and sundry performance rankings. They are all just artificial abstractions from data that are subject to wide interpretation and certainly neither necessary nor sufficient criteria for the evaluation of anything in particular. Even indicators as basic as statistical measures of central tendency (various means, medians, modes, and norms) must be interpreted and are not immune to misunderstanding and misuse. There is only one way to evaluate scholarship: invest the time to read and understand it.

The issue is not whether the use of metrics in the evaluation of scholarship has kept abreast of new technology, or whether it is sensitive to societal goals regarding diversity, equal opportunities, and objectivity. The emphasis of metrics is a de facto corruption of the reward system as it encourages counterproductive behavior. So why are we still calculating h-indexes and JIFs? The answer is that they are inexpensive and objective surrogates for legitimate peer review. While bibliographic databases [ScienceDirect (Elsevier), Scopus (Elsevier), Mendeley (Elsevier), Web of Science (Clarivate Analytics, formerly Thomson Reuters), Springer Link (Springer), and Westlaw (Thomson Reuters)] and academic social networking sites [ResearchGate (private), Google Scholar (Google), and Academia (private)] may be objective and unbiased sources of information, they suffer from the same deficiencies as throwing papers down the staircase or counting words—they circumvent the necessity of personal, measured, intellectually engaged assessment.

Anything that draws our attention away from the actual content of a scholarly work is misplaced, and the necessary perusal entails 1) familiarization with the literature on the subject, 2) intense study of the object under review, and 3) assessment and verification of positions taken. This process cannot be outsourced and automated. The fact that something has been referenced, downloaded, or liked is *prima facie* irrelevant for purposes of evaluation. It is not that quantitative metrics are not yet the equal of peer review in evaluation, but rather that they will never be the equal of peer review in just the same way that watching faux news will never be a legitimate substitute for a formal education.

In conclusion, we need to be cynical about the utility and value of metrics—especially in the evaluation of scholarship. And, to paraphrase Lily Tomlin, no amount of cynicism is

ever sufficient in this context (https://www.brainyquote.com/quotes/lily_tomlin_383212). **■**

REFERENCES

1. E. Garfield, "Citation indexes for science: A new dimension in documentation through association of ideas," *Science*, vol. 122, no. 3159, pp. 108–111, Jul. 15, 1955, doi: 10.1126/science.122.3159.108.
2. S. Lawrence, C. Lee Giles, and K. Bollacker, "Digital libraries and autonomous citation indexing," *Computer*, vol. 32, no. 6, pp. 67–71, Jun. 1999, doi: 10.1109/2.769447.
3. V. Bush, "As we may think," *Atlantic*, Jul. 1945. [Online]. Available: <https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>
4. T. H. Nelson, "Complex information processing: A file structure for the complex, the changing and the indeterminate," in *Proc. ACM'65 20th Nat. Comput. Conf.*, 1965, pp. 84–100, doi: 10.1145/800197.806036
5. D. C. Engelbart, "Augmenting human intellect: A conceptual framework," Air Force Office of Scientific Research, Washington, DC, USA, SRI Summary Rep. AFOSR-3223, SRI Project No. 3578, Oct. 1962. [Online]. Available: <http://dougengelbart.org/pubs/augment-3906.html>
6. R. P. Feynman, "Some remarks on science, pseudoscience, and learning how to not fool yourself. Caltech's 1974 commencement address," *Cargo Cult Science*, Jun. 1974. <https://calteches.library.caltech.edu/51/2/CargoCult.pdf>
7. G. Aad et al., "Combined measurement of the Higgs Boson mass in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments," *Phys. Rev. Lett.*, vol. 114, no. 19, p. 191,803, 2015, doi: 10.1103/PhysRevLett.114.191803.
8. A. Sokol, "Transgressing the boundaries: Towards a transformative hermeneutics of quantum gravity," *Soc. Text*, vol. 46/47, pp. 217–252, Spring/Summer 1996. [Online]. Available: https://physics.nyu.edu/faculty/sokal/transgress_v2/transgress_v2_singlefile.html
9. P. Lawrence, "The politics of publication," *Nature*, vol. 422, pp. 259–261, Mar. 20, 2003, doi: 10.1038/422259a.
10. H. Berghel, "The Sokol Hoax: A 25-year retrospective," *Computer*, vol. 53, no. 3, pp. 67–72, Mar. 2020, doi: 10.1109/MC.2020.2964894.
11. "1996 Ig Nobel Prize Winners," *Ann. Improbable Res.*, Harvard Comput. Soc., 1996. [Online]. Available: <https://www.improbable.com/2021-ceremony/winners/#ig1996>
12. C. Carpenter, D. Cone, and C. Sarli, "Using publication metrics to highlight academic productivity and research impact," *Acad. Emerg. Med.*, vol. 21, no. 10, pp. 1160–1172, Oct. 2014, doi: 10.1111/acem.12482.
13. D. Deutsch, R. Dror, and D. Roth, "A statistical analysis of summarization evaluation metrics using resampling methods," *Trans. Assoc. Comput. Linguistics*, vol. 9, pp. 1132–1146, Oct. 2021, doi: 10.1162/tacl_a_00417.
14. T. Vigen, *Spurious Correlations*. New York, NY, USA: Hachette Books, 2015.
15. D. Huff, *How to Lie with Statistics*, reissue. New York, NY, USA: Norton, 1993.
16. H. Rosling, A. Rosling Ronnlund, and O. Rosling, *Factfulness: Ten Reasons We're Wrong About the World – And Why Things are Better Than You Think*, reprint. New York, NY, USA: Flatiron Books, 2020.
17. E. Tufte, *The Visual Display of Quantitative Information*, 2nd ed. Cheshire, CT, USA: Graphics Press, 2001.
18. P. Basken, "Why Beall's list died – And what it left unresolved about open access," *Chronicle of Higher Education*, Washington, DC, USA, Sep. 12, 2017. [Online]. Available: <https://www.chronicle.com/article/why-bealls-list-died-and-what-it-left-unresolved-about-open-access/>
19. C. Straumsheim, "Academic terrorist," *Inside Higher Ed*, Washington, DC, USA, Jun. 2, 2017. [Online]. Available: <https://www.insidehighered.com>

- com/news/2017/06/02/librarian-behind-list-predatory-publishers-still-faces-harassment-online
20. P. A. Lawrence, "The mismeasurement of science," *Current Biol.*, vol. 17, no. 15, pp. R583–R585, Aug. 7, 2007. [Online]. Available: [https://www.cell.com/current-biology/fulltext/S0960-9822\(07\)01516-3#secsection0020](https://www.cell.com/current-biology/fulltext/S0960-9822(07)01516-3#secsection0020), doi: 10.1016/j.cub.2007.06.014.
 21. D. Hicks, P. Wouters, L. Waltman, S. de Rijcke, and I. Rafols, "Bibliometrics: The Leiden Manifesto for research metrics," *Nature*, vol. 520, no. 7548, pp. 429–431, Apr. 23, 2015, doi: 10.1038/520429a.
 22. "San Francisco declaration on research assessment," DORA, 2012. <https://sfdora.org/read/>
 23. A. Hatch and R. Schmidt, "5 common Myths about evaluation, rethinking research assessment: Ideas for action," DORA, 2020. https://sfdora.org/wp-content/uploads/2020/05/DORA_IdeasForAction.pdf
 24. J. Wilsdon *et al.*, "The metric tide: Report of the independent review of the role of metrics in research assessment and management," HEFCE, 2015. [Online]. Available: https://kar.kent.ac.uk/81123/1/Metric_Tide_main_report.pdf, doi: 10.13140/RG.2.1.4929.1363.
 25. "The metric tide: Correlation analysis of REF2014 scores and metrics (Supplementary Report II to the Independent Review of the Role of Metrics in Research Assessment and Management)," HEFCE, 2015. [Online]. Available: http://www.dcsience.net/2015_metrictideS2.pdf, doi: 10.13140/RG.2.1.3362.4162.
 26. P. Wouters *et al.*, "The metric tide: Literature review (Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management)," HEFCE, 2015. [Online]. Available: http://www.dcsience.net/2015_metrictideS1.pdf, doi: 0.13140/RG.2.1.5066.3520.

HAL BERGHEL is a Fellow of IEEE and ACM and a professor of computer science at the University of Nevada, Las Vegas, Nevada, 89154, USA. Contact him at hbl@computer.org.

SUBMIT TODAY

IEEE TRANSACTIONS ON SUSTAINABLE COMPUTING

► **SUBSCRIBE AND SUBMIT**

For more information on paper submission, featured articles, calls for papers, and subscription links visit: www.computer.org/tsusc

