CENTRE FOR GLOBAL HIGHER EDUCATION

Centre for Global Higher Education Working Paper series

An Index, A Publisher and An Unequal Global Research Economy

David Mills

Working paper no. 91
July 2023

UKRI Economic and Social Research Council

UKRI Research England

Published by the Centre for Global Higher Education,
Department of Education, University of Oxford
15 Norham Gardens, Oxford, OX2 6PY
www.researchcghe.org

ISSN 2398-564X

The Centre for Global Higher Education (CGHE) is an international research centre focused on higher education and its future development. Our research aims to inform and improve higher education policy and practice.

CGHE is a research partnership of 10 UK and international universities, funded by the Economic and Social Research Council, with support from Research England.

An Index, A Publisher and An Unequal Global Research Economy

David Mills

Contents

An Index, A Publisher and An Unequal Global Research Economy

David Mills

David Mills is Associate Professor (Pedagogy and the Social Sciences) at the Department of Education at Oxford. He is Deputy Director of the Oxford Centre for Global Higher Education (CGHE) and a Co-Investigator on CGHE's Research Programme on Supranational Higher Education. He is also Vice President of Kellogg College, Treasurer of the European Association of Social Anthropologists, and a director of the African Books Collective.

David uses ethnographic methods to study higher education and academic publishing. Trained in anthropology, his research interests include the colonial history of disciplines and universities, research methods, and the political economy of the global science communication system. He is currently writing about the impact of the global research and publishing economy on institutional research and publishing cultures in African universities.

His books include Ethnography and Education (Sage, 2013) and *Difficult Folk: A Political History of Social Anthropology* (Berghahn, 2008). His 2023 book, 'Who Counts: Ghanaian Academic Publishing and Global Science' is co-written with colleagues in Ghana and Oxford, has been published Open Access with African Minds.

Abstract

This is the story of how a publisher and a citation index turned the science communication system into a highly profitable global industry. Over the course of seventy years, academic journal articles have become commodities, and their meta-data a further source of revenue. It begins in Washington at the end of a second World War, when the US Government agrees a massive increase in funding for research, after Vannevar Bush champions basic research as the ‘pacemaker of technological progress’. The resulting post-war growth in scientific publishing creates opportunities for information scientists and publishers alike. During the 1950s, two men – Robert Maxwell and Eugene Garfield – begin to experiment with their blueprint for the research economy. Maxwell created an ‘international’ publisher – Pergamon Press – charming the editors of elite, not-for-profit society journals into signing commercial contracts. Garfield invented the science citation index to help librarians manage this growing flow of knowledge. Over time, the index gradually became commercially viable as universities and publishers used it to measure the ‘impact’ of their researchers and journals.

Sixty years later, the global science system has become a citation economy, with academic credibility mediated by the currency produced by the two dominant commercial citation indexes: Elsevier’s Scopus and Clarivate’s Web of Science. The reach of these citation indexes and their data analytics is amplified by digitisation, computing power and financial investment. Scholarly reputation is now increasingly measured by journal rankings, ‘impact factors’ and ‘h-indexes’. Non-Anglophone journals are disproportionately excluded from these indexes, reinforcing the stratification of academic credibility geographies and endangering long established knowledge ecosystems. Researchers in the majority world are left marginalised and have no choice but to go ever faster, resorting to research productivism to keep up. The result is an integrity-technology ‘arms race’. Responding to media stories about a crisis of scientific fraud, publishers and indexes turn to AI tools to deal with what is seen as an epidemic of academic ‘gaming’ and manipulation.

Does the unfettered growth in publishing ‘outputs’, moral panics over research integrity and widening global divides signal a science system in crisis? And is the ‘Open Science’ vision under threat, as the ‘author-pays’ publishing business model becomes dominant? With the scientific commons now largely reliant on citations as its currency, the future of science communication is far from certain.

Keywords: Academic publishing, Science communication, Citations, Citation index, Open Access, Bibliodiversity, Inequality, Research economy

Acknowledgment: This Working Paper is being published by the ESRC/RE Centre for Global Higher Education, funded by the U.K. Economic and Social Research Council (award numbers ES/M010082/1, ES/M010082/2 and ES/T014768/1).

The rise and fall of Robert Maxwell

Robert Maxwell's body was found floating off Gran Canaria on 6th November 1991. He had disappeared overnight from his luxury motor yacht, the Lady Ghislaine. Amidst fevered speculation about the cause of death, attention focused on the huge debts facing his media and business empire. Two weeks later, the Mirror newspaper, having initially run with the headline 'The man who saved the Mirror', revealed that he had stolen £526 million from his Mirror Group of companies: most of this was from the pension fund. Maxwell is now remembered for his ambition, his ego and his fraud. Less recognised is that his sprawling business empire was built on the profits and success of Pergamon Press, the academic publishing venture he began in 1951.

Observers of the contemporary global higher education landscape tend to focus on the latest fast-moving developments. But the commercial landscape of today's science communication system can be traced back to the policy foresight of Vannevar Bush, the deal-making of Robert Maxwell and the data skills of Eugene Garfield. Bush made the case for sustained government funding for basic research. Maxwell was the first to realise just how profitable scientific publishing could be, and seemingly no limits to the potential for scaling up journal outputs. Garfield's initial attempts to measure and quantify research, and then to make money out of these measurements, similarly moulded the unequal and stratified research terrain we now inhabit.

The first chapter of this story starts at the end of the second world war, and the influence of Vannevar Bush on US Government policy. Bush was an American engineer, inventor and science administrator, who during World War II had helped set up the U.S. Office of Scientific Research and Development (OSRD). Bush oversaw US wartime military R&D, including research on radar and the Manhattan Project. He pioneered digital circuit theory and ideas about hypertext through his concept of the 'memex', expanded memory, in his famous piece 'how humans think' (Bush 1945a).

Asked by Franklin D Roosevelt to develop a vision for the future of science, Bush wrote 'Science, the endless frontier' (1945b). Declaring that 'the pioneer spirit is still vigorous within this nation…science offers a largely unexplored hinterland for the pioneer', he played on the American settler-colonial vision of science as a constant war against disease and aggression. He saw research as the 'pacemaker' that underpins scientific progress and needed dedicated funding and support. His tenure and influence was marked by a massive expansion in US science funding, and a few years after the war, the National Science Foundation was launched.

Vannevar was a skilled administrator, but not an entrepeneur. To understand the commercial opportunities built into this emerging science system we need to turn to first Robert Maxwell and then Eugene Garfield. Maxwell was born to a poor family in Eastern Czechoslavia, and after escaping the Nazi occupation, joined the Czechoslovak Army in exile during World War II. He later won a military cross for active service in the British Army, and subsequently styled himself 'Captain' Robert Maxwell. He was based in Berlin after the war as a British military attache and later was revealed to have been a Russian double agent. As the war came to an end, the allied powers were keen to profit from German scientific knowledge. Maxwell used his Soviet army contacts to obtain copies of secret Soviet documents about every important German industrial plant along with scientific material. The plan was to strip much of this material and remove it to the Soviet Union. During this time, he made the most of his business and government contacts to help the German publisher Springer get their journals out of Berlin, providing them with paper and fuel to restart their business.

Until the second world war, academic publishing was primarily viewed as a service provided by university presses and scholarly societies to their members. Whilst a few Victorian popular science serials did develop large readerships (Brock 1980), *Nature* was unique in sustaining its scholarly credibility and commercial success (Baldwin 2015). Then geopolitics reshaped the European scientific landscape. Long established German publishing houses struggled to survive the convulsions created by the rise of Nazism. During the 1930s the struggling Dutch publisher Elsevier benefited from the emigration of experienced Germany editors and publishing staff to the Netherlands. Elsevier made the most of this technical and science publishing expertise, publishing more English and German scholarly texts. After the war, commercial academic publishers in the US, Holland and Britain all benefited both from the post-war revival of international scientific collaboration, and the tough restrictions placed on both German publishers (Brown 1947).

Robert Maxwell competed with Elsevier to dominate this emerging market (Cox 2002). Making the most of his Berlin contacts, in 1951 Maxwell paid £13,000 to buy UK distribution rights for Springer Verlag publications: six science journals and two textbook series. By 1960 his new company, Pergamon Press, was distributing 59 'international' scientific journals, and circulation grew at 5-10 percent each year. Working closely with ambitious academic editors, he rapidly expanded Pergamon: its profits underpinned the broader Maxwell publishing empire. Pergamon aggressively launched new journals from the profits of existing serials, sold textbooks throughout the world, and developed a highly profitable series of encyclopaedias. Maxwell was proud of his relationships with senior journal editors, offering them favourable contracts to secure their business. He wooed scholarly associations and journal editors with extravagant holidays and lavish parties at Headington Hall, where Maxwell based his companies. According to one colleague, Maxwell was smart because 'he knew just what to offer to buy a person – fame or money' (Preston, 2021). In the early years, he also benefited from Cold war paranoia, landing a lucrative US state department contract to translate huge numbers of Russian scientific papers.

Maxwell's recipe involved a combination of journal expansions and acquisitions, skilful marketing and creative new business models. Journals that previously had a national remit became 'internationalised' with new editorial boards and titles. Emerging sub-disciplines were also targeted, with Maxwell claiming that there were endless opportunities for journals to support ever smaller specialist fields. Both Pergamon and Elsevier focused on growing institutional rather than individual subscriptions. Many American and European scientific societies were persuaded to outsource their journal publishing to commercial 'partners', attracted by the income it would provide for conferences and membership benefits. In the 1950s and 1960s few academics could have foreseen the consequences of this new publishing economy. English replaced German as the international language of science (Gordin 2015), and ever more English-language science journals were launched fostering new international research communities (Meadows 1980). By the time Pergamon was sold to Elsevier in 1991 for £440 million, it had published 7000 monographs and launched 700 journals, of which more than 400 were still active. Cox argues that Maxwell had a ‘profound effect’ on scientific publishing, which the debacle of his death, his debts and his misuse of the Mirror’s pension funds ‘eclipsed from history’ (Cox 2002:274).

How a photocopied pamphlet changed the world of science

Commercial publishing is only one part of the story of contemporary science. The other is the challenge of managing and measuring this huge new flow of information. One could equally argue that the contemporary global research infrastructure has its roots in a photocopied pamphlet, entitled *Current Contents*, initially printed in a hen coop by an entrepreneurial young US information scientist – Eugene Garfield.

Born in the Bronx to second-generation Lithuanian immigrants, the young Garfield was inspired by the science fiction of HG Wells and his vision of a ‘World Brain’. For his doctoral degree in Chemistry and Library science, he developed an algorithm for converting chemical nomenclature into formulas. Garfield’s vision was of a new field of ‘information science’. Frustrated at the conservatism of traditional abstracting services, he wanted to make research knowledge accessible. Garfield felt that research funding was not being matched by financing for research communication, and that new technologies of data management could help create ‘efficient’ information systems.

Garfield’s first big innovation was refreshingly low-tech. Realising how hard it was for librarians to keep abreast of new research, he started sending out a weekly photocopied and stapled pamphlet of the contents pages of 150 life-science journals. Printed on cheap airmail paper in a converted chicken co-op, it became essential reading for librarians, sparing them having to browse through individual journals. *Current Contents*, as it was known, started in the life sciences in 1958 with 150 journals and demand rapidly grew. By 1967 *Current Contents* covered 1,500 journals in physics, chemistry and the life sciences.

Setting up the Institute for Scientific Information (ISI) as a company in 1955, naming it after a Moscow research institute, Garfield also provided reprinting and alerting services. His first customers included major pharmaceutical companies. Corporate subscriptions enabled ISI to expand, but also ensured that ISI's offerings were responsive to commercial needs. By the late 1970s, *Current Contents* was indexing more than 4500 journals. With new journals and fields lobbying to be included, Garfield was ambitious to make it as comprehensive as possible.

Garfield's most influential idea was equally straightforward: the concept of a citation index. He was fascinated by finding ways to assess the utility of research, and concerned about the citation of 'fraudulent, incomplete or obsolete' data. The idea emerged from a US legal paper-based research tool called Shepherd's citation that allowed lawyers to research case law and track precedent. Garfield felt that, in the same vein, scholars should also know about the existing citations of an article they were also citing, and that links to earlier work to help them to understand the 'transmission of ideas' and the intellectual structure of thought. The total number of citations could be counted, so that scientists could thus measure the 'impact', and hence importance of published work. So was launched scientometrics, the science of measuring and tracking the circulation and citation of scholarly knowledge (Garfield 1955). Garfield struggled to get research funding to develop his ideas, but in 1959 the US Air Force gave him a five-year contract to trial a prototype (Aronova 2021).

Garfield recognised that a comprehensive citation index ideally needed to cover all published scientific journals, but he recognised the economic and logistical impracticality of this. He turned to Bradford's law of scattering, named after a British mathematician, that held that that the most important literature in any scientific field is published only in a narrow group of journals. Pareto's law, or what is called the 80/20 distribution, allowed Garfield to make the case for a very focused selection of what he called the most 'significant' journals. Garfield cited one study that showed that 75% of references in the life sciences were to fewer than 1000 'core' journals, and 84% were to just 2000 journals. This justified an index based on the most influential and important journals in each field (Garfield 1955). It was also an astute commercial decision, given the huge logistical challenges and costs of indexing a potentially endless number of citations, with only the most basic of computing facilities.

After several trials, the first with just three genetics journals, Garfield's prototype index, published in 1963, assembled citation data from 560 scientific journals, with 70% published from the US or UK, and nearly all the rest from Europe. Garfield's selection of 'key' journals drew on his US-centred knowledge of the journal landscape, and primarily on the contents of Current Contents, which in turn had evolved to meet the needs of commercial subscribers. The first Science Citation Index (SCI), published in 1966, was similarly heavily reliant on the US-based research ecosystem and the offerings of commercial publishers. The academic geography of a Euro-American publishing economy was hard-wired into the index from the very start. Two Chinese journals were included, but none from Africa.

The index's rapid growth paralleled that of *Current Contents*. In 1966 SCI included more than 1150 journals, and by 1968 covered 2000 journals. Gradually more non-European journals were indexed, but their overall proportion remained very small, given the parallel growth in US and European serials, a topic I return to below. Garfield made the most of emergent computer technology to reduce the costs of indexing, and ISI employed a huge team of 100 data operators adding data to a central mainframe via desk tapes. Working two shifts five days a week, they were able to process 25,000 references a day (Garfield 1979). Like Maxwell, Garfield also benefited from Cold War tensions. He had been inspired by the possibilities for data management enabled by a centralised state, and later developed close links to Russian science administrators and the scientometrician Vasilii Nalimov. He helped broker a major contract selling IBM computers him to Russian ministries, profitably wrapping in a 10-year subscription to SCI services (Aronova 2021).

Whilst Garfield's original aim may have been to facilitate information searching, the index quickly defined 'reputable' academic knowledge. Inclusion mattered for journals, and publishers were prepared to pay the hefty subscription fees. With ever more 'international' journals being launched by two rapidly expanding commercial publishers - Pergamon Press and Elsevier - the index began to hold a powerful gatekeeping role. In the subsequent two decades it doubled in size, and by 1990 was indexing around 4000 journals. Garfield was skilful at promoting sales of SCI across the world, requiring indexed journals to subscribe (Garfield 1972).

Many were critical. Some mocked the idea that objectivity could be achieved by 'not reading the literature' (Oliver 1970). Sociologists and science scholars questioned the global coverage of the index (Narin,1976, Frame et. al. 1977, Rabkin and Inhaber 1979, but see Garfield 1983), and the meaningfulness of the data for different disciplines and regions (Cole and Cole 1971). A statistical critique of SCI's systematic discrimination against third world journals was published in Scientific American (Gibb 1995), leading to a strong riposte by Garfield (1997).

Garfield may not have anticipated that universities, academics and publishers would use the index to compete. Yet citation data allowed users to score and rank journals based on their citation 'impact factor'. Unwittingly or not, Garfield had created the tools for academic game-playing and institutional performativity. The shift was from maps to counts, from 'descriptive to evaluative' (Biagioli 2018:250). Csiszar (2020:51) describes an encounter where this risk is spelt out. In a packed Palo Alto seminar room at the first-ever conference on science indicators in 1974, Merton warned Garfield that 'whenever an indicator comes to be used in the reward system of an organization or institutional domain, there develop tendencies to manipulate the indicator so that it no longer indicates what it once did.' Merton had coined what became known as Goodhart's law.

Initially the index had little commercial value, despite Garfield's sales pitches, and income from Current Contents subsidised the index until the late 1970s. One obituarist described Garfield as 'visionary' rather than 'book-keeper' (Wouters 2017). The index had become a major drain on ISI resources, and in 1988 his business was bought out by another publisher for $24 million. The indexes were fully digitised and sold on to Thompson in 1991 for $210 million. Their commercial potential only became clear when the first global rankings of universities were launched in the early 2000s, many of which used citation data to assess academics. In 2011 Thompson sold the business for $3.5 billion to Clarivate.

Clarivate's Web of Science 'core collection' now covers more than 21,000 journals within four different indexes: Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI), Arts & Humanities Citation Index

(A&HCI) and the Emerging Sources Citation Index (ESCI). Elsevier, the largest of the academic publishing houses, based in Amsterdam and London, launched a rival index, Scopus, in 2004. The latter indexes around 20% more journals and has a more international profile, as its subscription services are actively marketed to universities globally.

Both indexes have exacting metrics-based selection thresholds and evaluation policies, indexing at the most 5-6% of all active academic journals. They publish broad guidelines on their evaluation and selection procedures for new journals. Web of Science[1], the more selective of the two databases, employs an 'in-house' editorial board of seven to adjudicate on journal inclusion decisions. Candidate journals for the Web of Science core collections first must meet a minimum set of quality and compliance standards. The 24 quality criteria include adhering to community standards, a distributed set of authors, the composition of editorial boards, and 'appropriate citations to the literature'. New journals are evaluated on these quality criteria before being accepted into the ESCI. The impact criteria include assessment of content significance and three citation-based metrics: including analysis of author citations, editorial board citations, and comparative citation data. This citation data is used to select and promote the most influential journals in their fields from ESCI into the SCIE, SSCI or AHCI. Journals can also be demoted from the three databases to ESCI if they lose impact.

Scopus journal selection is overseen by a group of Elsevier-appointed external experts called the Content Selection and Advisory Board (CSAB)[2]. CSAB consists of 17 Subject Chairs, representing different scientific fields. These researchers, scientists and librarians, all with university affiliations, are responsible for reviewing all titles proposed for inclusion into the Scopus database. Furthermore, the CSAB also provides recommendations to Scopus about new priorities based on Scopus user data. To expand its collections in

---

[1] https://clarivate.com/products/scientific-and-academic-research/research-discovery-and-workflow-solutions/web-of-science/core-collection/editorial-selection-process/editorial-selection-process/

[2] https://www.elsevier.com/solutions/scopus/how-scopus-works/content/scopus-content-selection-and-advisory-board

non-English speaking regions, Scopus has in recent years created four local Expert Content Selection and Advisory Committees (ECSAC) in Russia, Thailand, South Korea and China. Each of these is tasked with advancing 'the overall standards and quality of journals published in non-English speaking countries.' The rubric for these boards includes the aim 'that titles published primarily for a local audience but deserving of international attention' are included in Scopus. Elsevier's assumption seems to be that journals not published in English are by definition 'local' and aimed at national audiences.

The minimum criteria for inclusion in Scopus include peer-reviewing, journal registration, statements on publication ethics, and the requirement to 'have content that is relevant for and readable by an international audience' including English language abstracts and titles. There are five further categories under which journals are assessed: journal policy, journal content, journal standing, publishing regularity and online availability. Each of these is assessed numerically. For example, journal standing is assessed by the 'citedness of journal articles in Scopus', whilst journal policy includes measuring the 'diversity in geographical distribution' of editors and authors. Scopus also uses citation-based peer benchmarks to adjudicate inclusion decisions, including evidence of self-citation (greater than 200% higher than the average) and where citation rates, numbers of articles, and number of clicks on Scopus are all less than 50% of the average amongst peer journals. All these metrics discriminate against small journals, those published in languages other than English, in the global South or catering primarily for national and regional scholarly ecosystems beyond Europe.

Growing pains

According to UlrichsWeb periodical database, there are now more than 100,000 academic journals published worldwide. This is likely to be an underestimate, as globally, higher education research continues to expand. Back in 1961 the mathematician Derek de Solla Price predicted that science would continue to grow exponentially, and that by 2000 there would be 1 million journals (Price 1961). He was broadly right about consistent growth: this is currently around 5.4% each year, according to the latest estimates (Bornmann et. al. 2021). Yet he could not have foreseen the rise of mega-journals, pre-prints, and a multiplicity of other ways of sharing research knowledge.

Four multinational companies dominate, each publishing more than 2000 journals each - Springer Nature, Elsevier, Wiley-Blackwell, and Taylor and Francis. They are based in London, Amsterdam, Hoboken (New Jersey) and Oxford, from where they seek to manage their global profiles. Together, they publish more than 70 percent of all social science journals, and 50 percent of journals in the natural sciences. Sage is in fifth place with more than 900 journals.

In a growing global tertiary education sector, new market opportunities constantly emerge. For example, Hindawi, was founded in Egypt in 1997 and became an innovative publisher of 230 Open Access journals. It later moved to London and was bought by Wiley in 2021 for $300 million. MDPI was launched in 1996 and Frontiers in 1997. Working from bases in Switzerland, both offer similar gold Open Access publishing opportunities. They champion ‘customer’ service and rapid editorial decision making. Some journals review and publish accepted submissions within a few weeks. They all require accepted authors to pay article publishing charges (APCs), unless they qualify for, or are granted, waivers on grounds of geography, career stage or institutional affiliation. MDPI charges an average APC of £1,900, but, for now, most of its journals (especially in the social sciences) waive between 70-100% of these fees. Frontiers – partly owned by the major shareholder in Springer Nature - charges APCs between £1,000 and £2,500, depending on the funding available in the field. In 2021, Frontiers published 85,000 articles in its 140 journals, and was ranked the third most cited publisher, whilst MDPI doubled its output to 235,000 articles. Both make extensive use of special issues, with MDPI publishing more than 6,700 in 2020 (Crossetto 2021). At this rate, MDPI’s article output will soon rival the 430,000 articles published annually by Elsevier, which remains the largest of the established commercial publishers.

Meanwhile, elite journal ‘brands’ have become tradeable marketing tools for their commercial owners. Where there was once one *Lancet*, there are now 22 Lancet-branded journals. Springer-Nature’s ‘brand expansion’ strategy has meant there are now more than 30 journals within its portfolio, all with *Nature* in their title. *Nature* publishes the very strongest submissions it receives, but the publisher ‘cascades’ rejected articles to other *Nature*-branded journals, including to an Open Access journals with high publication fees. *Nature* has an 8% acceptance rate, *Nature research* journals have a 10% acceptance rate, *Nature Communications* has a 20% acceptance rate (and a $5400 APC), and *Scientific Reports*, Nature’s OA mega-journal, has a 60% acceptance rate.

Springer-Nature have also created their own journal ranking index, publishing an increasing number of branded supplements, special sections and ‘advertorials’. It increasingly seems that scientific publishers are ‘responding more to the logic of a market than to that of a community’ (Khelfaoui and Gingras 2022, 196). The skills is to achieve commercial success and sustain reputational credibility, with the latter measured largely by citations. The journal impact factors of several *Nature* journals have increased by almost 50 percent over two decades. At the same time, several long-established – and formerly prestigious – scientific journals owned by professional societies have seen their status, submissions and income decline.

In an unequal global research system, acceleration and productivity become survival strategies. Universities are ever more focused on their national and international rankings as they compete for students, funding patronage and reputation. Many incentivize their staff to publish through financial incentives and promotion pressures, changing academic practice. The commercial academic publishing model requires growth to sustain profits whether from publishing more in each issue, soliciting more special issues, or launching more journals. Subsidies and institutional expectations foster an acceleration of the research publication cycle. The logics of reputational stratification across a hierarchical global science system require those at the peripheries (especially precarious junior and adjunct staff) to publish more and faster to stay visible, putting yet more pressure on the system.

In this context, it is not surprising that some academics take short-cuts to survive. In China, as in other emerging economies, doctors and other professionals need to have academic publications in ‘top’ journals to get promoted. If they have no research or writing experience, the chances of getting their work into SCI journals is slim. The only option may be to purchase authorship, and there is a burgeoning demand for brokers and agents who can help with this process. This has led to a series of high-profile mass retractions. Science sleuths, aided by investigative media watchdogs such as Retraction Watch, uncover problematic cases of indexed journals publishing huge special issues, but whose contents are out of scope, plagiarized or just plain nonsense.

When academic publishers talk about the importance of integrity and trust, this is because they are acutely aware that academic credibility and reputation is a precious asset. Major scandals - such as that which hit Wiley's Hindawi journals in 2022 - directly impact share prices. It is little wonder that publishers promote the burgeoning scholarly literature on so-called 'predatory publishing' (Inouye and Mills 2021) or amplify media caricatures of Chinese 'paper mills'. Fraud gets portrayed as an existential threat to the integrity and future of science (Mills et al 2021).

In response, many publishers are introducing elaborate AI-driven detection tools as well as relying on citation benchmarks and metrics indicators to detect journal 'outliers'. This reliance on technology and aggregated publication metrics means that journals with more distinctive profiles risk being seen as potentially fraudulent or fake. There is less thought and reflection from the publishing community on whether the tactics used by those on the margins are just more extreme forms of the 'gaming' that is required by a metrics-based system. Few ask if the accusations of fraud and 'gaming' are appropriate, given that all the actors within this communication system are caught within an integrity-technology 'arms race'. Focused on tracking down misconduct, publishers and integrity watchdogs dwell less on the systemic features of mainstream science that are generating these mimetic practices and copies (Jacob 2020, 256). For Griesemer (2020), to even frame the 'gaming' of metrics as academic misconduct is to accept these metrics as appropriate normative standards. He describes the problem as being 'a prime mode of escalation in a metrics arms race between standards imposers and gamers' (ibid, 79). More broadly, the discourse of 'fraudulent' science also serves to reassert the boundaries of genuine science, and to shore up the exchange value of its key currency: citation data.

What gets left out?

The unequal geographical representation of scholarly journals by the Science Citation Index was first pointed out more than 50 years ago. Today, thanks to the business models developed by Maxwell and his rivals, the hold of commercial publishers is stronger still. Despite calls to decolonise Open Access and promote bibliodiversity, the two commercial citation indexes cast a long shadow across academic publishing in the global South.

Toluwase Asubiaro is an activist Nigerian information scientist who works to document the impact on the visibility and status of African academic publishing. Inspired by his doctoral supervisor's long campaign to create an African citation index (Nwagwu and Ahmed 2009), he has documented the impact of international collaborations on the visibility of African research (Asubiaro 2018, Asubiaro and Badmus 2020). Based in Canada, he has set up the African Research Visibility Initiative, and uses his bibliometric skills to evaluate these indices, and to find alternative ways to measure and assess African research, such as the use of altmetrics, google scholar or Crossref.

Asubiaro and Onalaopo (2023) use Ulrichsweb and AJOL data to estimate there are currently around 2200 active journals published in Sub-Saharan Africa. Of these, only 166 were indexed in Web of Science (and 174 in Scopus), around 7.5% of such journals. Of the 166 in Web of Science, around 75% were published from South Africa. This means that only around 50 journals are indexed in Web of Science from across the rest of Sub-Saharan Africa. Many African countries have no journals in the index. Only 21 Nigerian published journals are indexed, four from Ghana, and five each from Ethiopia and Kenya. Very few journals from Francophone Africa are indexed. This shows just how much knowledge is ignored and effaced by these indices.

Is Open Science the answer or the question?

The modern Open Science movement begins in the early years of the internet. Initiatives such as Project Gutenberg sought to make research publications more widely available, whilst a number of publishers launched free to read digital journals. A landmark 2001 conference in Budapest organised by the Open Society Foundation set out a vision of Open journals that would make no charge for the reader to access. The Budapest Open Access Initiative, as it was called, led to further policy initiatives, such as the European Plan S in 2018.

This obliged scientists and researchers to publish their work under an open Creative Commons license and in Open Access repositories and journals.

Responding to Plan S mandates, the majority of commercial publishers have developed 'transformative agreements', transitioning to Open Access funded by Article Processing charges. Most have seen their profits grow through this model, even if many are failing to meet Plan S timelines. Far from helping to decolonise the publishing ecosystem, this commoditised model of Open Science seems to be strengthening the position of commercial publishers, raising fears of academic 'platform capitalism' (Meagher 2021, Knochelmann 2021, Mirowski 2018). It also sustains the marginalisation of researchers in the majority world. When citations become the dominant currency of acemic credibility and reputation, those at the peripheries are often forced to resort to acceleration and productivism as survival strategies.

Today, Open Science is an increasingly contested concept. The UNESCO 2021 Open Science recommendation envisions research infrastructures that are 'organized and financed upon an essentially not-for-profit and long-term vision, which enhance open science practices and guarantee permanent and unrestricted access to all, to the largest extent possible' (UNESCO 2021). In May 2023, the European Council recommended that European member states 'step up support' for the development of a not-for-profit publishing platform free to both authors and readers (so-called Diamond OA). A series of Horizon Europe funding projects, including DIAMAS and OPERAS have been tasked with building a high quality, sustainable and community-owned scholarly communication system, including a set of institution-funded technological infrastructures and common standards for Open Access scholarly journals.

These policy visions are ambitious and idealistic, given that they pose a direct challenge to commercial interests by redirecting resource flows to community-owned infrastrucutures. The ideas are still embryonic, and scaling up a funder-owned Open Source publishing infrastructure would need huge political will and deep pockets, given that Elsevier alone spends billions each year on technological development (Esposito and Clarke 2023).

The European Council's vision of equity and sustainability is informed by influential debates around 'bibliodiversity'. The underlying rationale is that equitable and sustainable community-ownership is the best way to promote a diverse range of regional initiatives, publishing infrastructures and knowledge ecosystems (Berger 2021). As Shearer et al (2020, 1) note in their call for action, 'diversity in services and platforms, funding mechanisms, and evaluation measures will allow the scholarly communication system to accommodate the different workflows, languages, publication outputs, and research topics that support the needs and epistemic pluralism of different research communities'. Latin America offers an exemplar: a strong regional Portuguese and Spanish-publishing research ecosystem supported by the community-owned SciELO database and associated publishing infrastructures. There are a growing number of 'diamond' Open Access publishing platforms and experiments, and much policy interest in sustainable Open Science. It is an idealistic vision of a more equitable research world in which, as Arturo Escobar puts it, 'many worlds might fit' (Escobar 2020).

Conclusion: what comes after the citation economy?

Sixty years after Garfield launched his first citation index, and more than seventy since Maxwell founded Pergamon, academic journal publishing has been transformed into a profitable global industry. Commercially intertwined, the indexes and the publishers have together built a citation economy. Today, scholarly reputation and status is measured by journal rankings, 'impact factors' and 'h-indexes'. The reach of these citation indexes and their data has been amplified by digitisation, computing power and financial investment. Citation metrics reinforce existing academic 'credibility economies', built around Euro-American publishing networks and commercial interests. Maxwell and Garfield have a lot to answer for.

This working paper has explored how non-Anglophone and regional journals are rendered invisible by exclusion from these indexes, reinforcing the stratification of academic geographies and undermining long established regional knowledge ecosystems. Across the majority non-Anglophone world, journals excluded from Scopus and WoS face constant challenges to their legitimacy and reputation. The pressures on scholars at the margins have led to academic acceleration and research productivism. This in turn has provoked media concern about scientific fraud and an integrity-technology 'arms race', rather than questions about the sustainability of a commercially-mediated research economy.

Is there a way out of this recursive growth loop? Diamond Open Access advocates, funders and researchers in Europe are beginning to envision a more equitable research system built around community-owned publishing infrastructures and standards. If this is to reach beyond well-resourced European universities, governments across the world will need to adequately fund national and regional research ecosystems. The first step on this journey is helping scholars and universities to divest from the citation economy.

References

Aronova, Elena. 2021. 'Scientometrics with and without Computers: The Cold War Transnational Journeys of the Science Citation Index.' in Mark Solovey and Christian Dayé (eds.), *Cold War Social Science: Transnational Entanglements* (Springer International Publishing: Cham).

Asubiaro, Toluwase. 2018. 'Research collaboration landscape of the university of Ibadan biomedical authors between 2006 and 2015', *African Journal of Library, Archives and Information Science*, 28: 17-31.

Asubiaro, Toluwase, and Oluwole Badmus. 2020. 'Collaboration clusters, interdisciplinarity, scope and subject classification of library and information science research from Africa: An analysis of Web of Science publications from 1996 to 2015', *Journal of Librarianship and Information Science*, 52: 1169-85.

Asubiaro, Toluwase, and Sodiq Onaolapo. 2023. 'A comparative study of the coverage of African journals in Web of Science, Scopus, and CrossRef', *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.24758

Baldwin, Melinda Clare. 2015. *Making nature: the history of a scientific journal* (The University of Chicago Press: Chicago).

Berger, Monica. 2021. 'Bibliodiversity at the Centre: Decolonizing Open Access', *Development and Change*, 52: 383-404.

Biagioli, Mario. 2018. 'Quality to Impact, Text to Metadata: Publication and Evaluation in the Age of Metrics', *KNOW A Journal on the Formation of Knowledge*, 2: 249 - 74.

Bornmann, Lutz, Robin Haunschild, and Rüdiger Mutz. 2021. 'Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases', *Humanities and Social Sciences Communications*, 8: 224.

Brock, W. H. 1980. 'The Development of Commercial Science Publishing in Victorian Britain.' in A. J. Meadows (ed.), *The Development of Science Publishing in Europe* (Elsevier: New York, NY).

Brown, Charles Harvey. 1947. 'Scientific Publishing in Continental Europe: Notes on Its War and Postwar Status', *Science*, 106: 54-58.

Bush, Vannevar. 1945a. 'As we may think', *The Atlantic monthly*, 176: 101-08.

Bush, Vannevar. 1945b. ' Science: The endless frontier', *National Science Foundation–EUA. Washington*.

Esposito, J, and E Clarke. 2023. "Diamond." *The Brief*.

Cole, Jonathan, and Stephen Cole. 1971. 'Measuring the Quality of Sociological Research: Problems in the Use of the "Science Citation Index"', *The American Sociologist*, 6: 23-29.

Cox, Brian. 2002. 'The Pergamon phenomenon 1951–1991: Robert Maxwell and scientific publishing', *Learned Publishing*, 15: 273-78.

Crosetto, Paolo. 2021. 'Is MDPI a predatory publisher?' *Paolo Crosetto*. https://paolocrosetto.wordpress.com/2021/04/12/is-mdpi-a-predatory-publisher/

Csiszar, Alex. 2020. 'Gaming Metrics Before the Game: Citation and the Bureaucratic Virtuoso.' in Mario Biagioli and Alexandra Lippman (eds.), *Gaming the Metrics: misconduct and manipulation in academic research* (MIT Press: Cambridge, MA).

Escobar, Arturo. 2020. *Pluriversal Politics: The Real and the Possible* (Duke University Press: Durham).

Frame, J D, F. Narin, and M. Carpenter. 1977. 'The Distribution of World Science', *Social Studies of Sciences*, 7: 501-16.

Garfield, Eugene. 1955. 'Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas', *Science*, 122: 108.

———. 1972. 'Citation Analysis as a Tool in Journal Evaluation', *Science*, 178: 471-79.

———. 1979. *Citation indexing: Its theory and application in science, technology, and humanities* (Wiley: New York).

———. 1983. 'Mapping science in the Third World', *Science and public policy*, 10: 112-27.

———. 1997. 'A statistically valid definition of bias is needed to determine whether the Science Citation Index discriminates against third world journals', *Current Science*, 73: 639-41.

Gibbs, Wayt W. 1995. 'Lost science in the third world', *Scientific American*, August: 76-83.

Gordin, Michael. 2015. *Scientific Babel: How Science Was Done Before and After Global English* (University of Chicago Press: Chicago).

Griesemer, James. 2020. 'Taking Goodhart's law meta: Gaming, meta-gaming, and hacking academic performance metrics', in Mario Biagioli and Alexandra Lippman (eds.), *Gaming the metrics: misconduct and manipulation in academic research* (MIT Press: MIT).

Inouye, Kelsey, and David Mills. 2021. 'Fear of the academic fake? Journal editorials and the amplification of the "predatory publishing" discourse', *Learned Publishing*, 34: 396-406.

Jacob, Marie-Andrée. 2020. 'Crack open the make-believe: Counterfeit, publication ethics and the global South.' in Mario Biagioli and Alexandra

Lippman (eds.), *Gaming the metrics: misconduct and manipulation in academic research* (MIT Press: MIT).

Khelfaoui, Mahdi, and Yves Gingras. 2022. 'Expanding Nature: Product line and brand extensions of a scientific journal', *Learned Publishing*, 35: 187-97.

Knöchelmann, Marcel. 2021. 'The Democratisation Myth', *Science & Technology Studies*, 34: 65-89.

Meadows, A. J. 1980. *Development of science publishing in Europe* (Elsevier: NY).

Meagher, Kate. 2021. 'Introduction: The Politics of Open Access — Decolonizing Research or Corporate Capture?', *Development and Change*, 52: 340-58.

Mills, David, Abigail Branford, Kelsey Inouye, Natasha Robinson, and Patricia Kingori. 2021. '"Fake" Journals and the Fragility of Authenticity: Citation Indexes, "Predatory" Publishing, and the African Research Ecosystem', *Journal of African Cultural Studies*, 33: 276-96.

Mirowski, Philip. 2018. 'The future(s) of open science', *Social Studies of Science*, 48: 171-203.

Narin, F. 1976. *Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity* (Computer Horizons, Inc: Cherry Hill, NJ).

Nwagwu, Williams E, and Allam Ahmed. 2009. 'Building open access in Africa', *International Journal of Technology Management*, 45: 82-101.

Oliver, P T. 1970. 'Citation Indexing for Studying Science', *Nature*, 227: 870.

Preston, John. 2021. *Fall: The mystery of Robert Maxwell* (Viking: London).

Price, Derek J. de Solla. 1961. *Science since Babylon* (Yale University Press: New Haven).

Rabkin, Y. M., and H. Inhaber. 1979. 'Science on the periphery: a citation study of three less developed countries', *Scientometrics*, 1: 261-74.

Shearer, Kathleen, Leslie Chan, Iryna Kuchma, and Pierre Mounier. 2020. 'Fostering Bibliodiversity in Scholarly Communications: A Call for Action!' https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1153&context=scholcom

UNESCO. 2021. 'UNESCO Recommendation on Open Science.' Paris: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en

Wouters, Paul. 2017. 'Eugene Garfield (1925–2017)', Nature, 543: 492-92.