

4

Information on Wings: The History of Open Access

“Information wants to be free” is the open-access movement’s mantra. But does it? And can it be? If by free we mean gratis, then the expectation may be exaggerated. The Budapest Open Access Initiative’s declaration from 2002 expected the costs to be “far lower” than for conventional publication.¹ That likely overestimated the possible savings. But the grain of truth here is that, with upfront expenses met, the last marginal copy of a digital work is almost costless. Yet getting to that point takes largely as much work as for an article or book on paper. Only by letting the standards slip or by achieving new efficiencies would digitality help bring down overall publication costs—and not just those of dissemination. Digitality has, in fact, accomplished much of that already. Everyone can produce what would earlier have counted as camera-ready copy on their laptops. The web allows us to post anything for the world to see. If we are willing to call that publishing, then we are all publishers now.

Whether information wants to be free, as in universally knowable, is another matter. As we have seen, knowledge can easily remain a secret so long as it is not imparted. A proper secret—one undivulged—remains with its knower. Even a dangerous, forbidden,

or heretical thought does no harm if thus kept. Nor does it rebound against its knower so long as they refrain from its utterance. Even when the Church vigorously pursued wrong thinking, it excepted heretics' unspoken cogitations. In 1484, the Spanish Inquisitor Tomás de Torquemada decreed that those who had never revealed their apostasy should be allowed to abjure and do penance in secret, thereby never being exposed as heretics.²

Some information wants to remain secret. Not just what we hope to keep from the authorities, but more generally, our private data. As the personal becomes harder to hide, splashed across social media, we self-contradictorily jealously seek to guard it. To this comes the value of turning our daily behavior into data. Acts whose documentation we scarcely realized served any purpose turn out to interest companies eager to parse our profiles as consumers.

The value of some information is enhanced by not being widely known. Whether medieval occultism or modern trade secrets, their worth lies in scarcity. The advantages of hiddenness are exploited, whether for mercenary gain or for the appearance of insight, wisdom, healing powers, sorcery, or technical expertise it gives its bearer. While the public is kept in the dark, the circle of initiates has to be large enough to keep the show running. At some point, the knowledge has to be transmitted or inherited. While the recipe for Coca-Cola is locked in an Atlanta bank vault, the formula for Chartreuse remains in the heads of only three Carthusian monks, who daily drive a jalopy together along a winding mountain road to the distillery.³ How medieval glassmakers produced their stained windows' deep rich blue remains lost. The method the Saxons used to manufacture white porcelain was secreted in the Green Vault of the Dresden palace.⁴

Some information remains secret despite our best attempts to know it. Black box algorithms generate conclusions beyond what humans can yet follow. DeepMind's AlphaFold software predicts the structure of proteins with unexpected accuracy, based on

insights into biological fundamentals that we still do not fathom.⁵ Explainable Artificial Intelligence (XAI) seeks to pry open the black box, allowing humans to grasp the reasoning behind particular outcomes.⁶ A new EU directive requires artificial intelligence's conclusions to be comprehensible to natural persons.⁷ Without human oversight, algorithms may draw accurate results for reasons we do not understand but they may also jump to unjustifiable conclusions.

Deep learning neural networks can now locate abnormalities on medical scans as well as the average radiologist.⁸ Other times, the gains in knowledge are less clear. An algorithm claims to discern people's sexuality from photographs on dating sites better than humans can.⁹ The programmers focused on grooming choices and facial features, which may have reflected hormonal effects in utero. But when other studies achieved comparable predictions even when blurring the faces, it became unsettlingly unclear what the algorithms based their forecasts on—lighting, color, or other factors associated with sexuality in ways we do not yet understand?¹⁰

Beyond speech, writing has vastly expanded our ability to transmit knowledge, but not necessarily broadly. At first, so few people could read that the medium itself was restricted to insiders. As the spread of literacy expanded such limits, secret languages, codes, or encryption again allowed transmission to a select audience. Using Native Americans speaking their tongues for clandestine radio transmissions during the Second World War relied on rarity to keep secrets. Through the Middle Ages, science remained the preserve of the elect. Hippocrates wanted to retain holy things, such as nature's secrets, for the initiated.¹¹ Renaissance scholars reserved knowledge for themselves. Guilds hid craft techniques.¹² Before the law began protecting intellectual property, monetizing knowledge was often accomplished by keeping others in the dark.

Only with the Scientific Revolution in the seventeenth century did the idea emerge that science was something inherently public.

To the discoverer went the intellectual credit; to the inventor, eventually the patent. In both cases, the knowledge was public and could be elaborated and expanded by others. As a scarce commodity, hidden among the initiated, knowledge was a dead end. Barring some inscrutable source of occult insight, how much could any isolated person or small group ever hope to achieve? Breaking open nature's secrets was accomplished better when more participants were engaged. Science is, plainly, a social enterprise.

With the Scientific Revolution, openness and transmissibility were recognized as independently improving information's quality. Widely available knowledge had been tested by other minds. The larger the network it was broadcast across, the more eyes scrutinized it. An avalanche of print inundated the world. Universal compulsory schooling created a huge audience of readers. Newspapers and periodicals, along with cheap new books, sated their curiosity. Libraries supplied readers with costless content. The mail system ensured information's steady flow, with discounted postage subsidizing the spread of printed matter.

To encourage the exchange of news and ideas, newspaper editors could send their latest edition to colleagues across the US for free. By the 1840s, each editor received an average of 4,300 different exchange newspapers annually, a dozen every day.¹³ Scissors ready, they cut and pasted text into their own periodicals—an analog foreshadowing of today's blogs. We now fret as the web effortlessly spreads bad information along with the good. True, but keep in mind that a similar explosion of ephemera rained down on readers in the eighteenth century as pamphlets, broadsheets, and newspapers.¹⁴ The web has amplified but not created data surfeit. The difference is of degree, not kind.

Knowledge that is open and accessible has long been recognized as a virtue. Locked inside our skulls, information is worthless. As perishable as we are, uncommunicated knowledge is also uncorroborated, unchecked, and unverified by interaction with others.

Besides not being put to use, unshared knowledge is of lesser intellectual worth. Individual knowers cannot determine the quality of their insight without correlating it with others or testing it against the outside world. Only knowledge that passes intact through the crucible of verification can claim a truth recognizable to others. But once outside our heads, information can be no more than a social secret, guarded for a while among a few select knowers, yet ever-leakable to others.

The Royal Society of Arts, founded in England in 1754, regarded patents as an illegitimate monopoly of information that ought to be freely available. For a while, it refused its prizes to anyone with a patented invention.¹⁵ That was somewhat unfair to patents. Their aim, after all, was to grant inventors means of profiting from discoveries without having to keep them secret. In exchange for revealing insights, rights-holders were rewarded with temporary monopolies. Everyone gained—inventors earned profits, society reaped new knowledge. However distant this may now seem, the concept of intellectual property was part of broadening access.

Intellectual property was a trade-off between creators and the community. The invention of printing turned ideas and their articulation into potential commodities. That allowed their theft. With writings and books worth something, their content became valuable. Kings and governments granted publishers monopolies on their works. Booksellers had contracts with their authors for exclusive printing rights. But neither grants nor agreements kept competition at bay. Piracy ensued, with works reprinted, repackaged, and repurposed at will by rogue publishers outside the net of enforcement.

That problem was addressed by inventing property rights in works, starting with copyright in the English Statute of Anne in 1710. Law now granted authors a claim to something so intangible that enforceable ownership would otherwise have been impossible. Along with patents and eventually trademarks, this made

intellectual matters property. Creators could briefly exploit their works. Ideas were no longer kept secret or limited to a few initiates. Still, they were restricted to those who paid the price of admission by buying a legitimate copy or joining a lending library.

Independent of intellectual property, the university's prestige economy made secret knowledge a self-defeating ambition. Submitting truth claims to the scrutiny of peers honed their accuracy. Equally important as any profits, publication established priority and credit. To the first discoverer, inventor, or formulator went the spoils. Keeping knowledge under wraps was for alchemists, druids, and other mystifiers.

If secret knowledge was largely worthless, it also became increasingly clear that what an individual could achieve alone was un compelling. It stands to reason that big projects and some forms of intellectual activity must be collaborative: encyclopedias, massive particle accelerators, opera, film, and the like. Turning collaboration from a necessity into a virtue takes the insight a step further. Open software enjoyed the benefits of many participants, all working toward common goals. Such advantages were revealed by open software projects such as the operating system Linux. With that apparent, corporations often unblocked projects that had begun closed. Microsoft opened its once-proprietary software in 2014 to reap the benefits of Linus's law: "Given enough eyeballs, all bugs are shallow."¹⁶

Broadcast television and radio, paperbacks, subscription periodicals, and public libraries: that was state of the art for spreading enlightenment in the analog era. Compared to the year before Gutenberg's invention, an unimaginable cornucopia of content has long awaited any willing reader able to access media. The progress toward more information and better availability was steady, interrupted only rarely. Under totalitarianism, things took a step backward. The East Bloc nations had functioning publishing industries, but their output was official rubbish, read by no one.¹⁷ The

“Classics” sections of East German bookstores were well stocked with Marx and Engels and little else.

Dissident writers, their voices muffled, had to wield rudimentary tools. Mimeograph duplicators and plain old carbon paper, banged ragged and faint by reuse, allowed dissidents’ manuscripts to circulate clandestinely. Getting to publishers involved perilous smuggling by Western allies—the negatives of Aleksandr Solzhenitsyn’s Nobel speech were carried inside a transistor radio.¹⁸ Even so, the porosity of borders to modern technologies leavened the worst effects. East Germans watched West German TV, except in a few areas too remote for analog signals. Today, the internet firewalls imposed in China and the Middle East have a ferocious bark, but proxy servers and virtual private networks mitigate their bite.

Digitality then expanded the analog universe enormously in two directions. It made producing something worth watching, reading, or hearing less technically demanding. Cinema-ready film could be shot on smartphones, with a bit of postproduction software to sand the rough edges. Laptops became both an orchestra of instruments and their recording studio. Every writer could edit, typeset, and print copy that might not reach the standards of Aldine editions but certainly put to shame 1970s mass-market paperbacks. Dissemination was even more important than this revolution in production. The web allowed any content to be posted on a global bulletin board, readable by anyone, anywhere, anytime. Authors could now be their own agents, publishers, marketers, and distributors. With the work ready, the cost of delivering it to a potentially global audience was negligible, baked into the web’s running costs.

That upped the stakes. In the analog era, cheap media had made works broadly affordable, although not costless. Public libraries had plugged the gap, giving everyone access to most legible content so long as they could make it to a reading room. Libraries did not ignore music and film, but they were not as adequately covered as print for technical and copyright reasons. Museums achieved

something similar for the plastic arts, although dealing with singular objects, their ability to facilitate access was necessarily less ambitious than for reproducible content. Digitality made truly open access technically possible. Content could be available to anyone. The hurdles were no longer how to do this, but at what cost and to whom.

With their material costs, analog media could no more have been given away than food or housing. But with digitality's radical dematerialization, content could now be imagined as something akin to municipal water. Though not free, it could still be provided through taxes, fees, and other means of paying for something accepted as a basic necessity of civilized life, supplied largely at cost to all residents.

The Causes of Open Access

Open access must overcome three hurdles: technical, financial, and legal. Without digitality, global access is unthinkable. Digital editions cost money to set up, but the last marginal copy is practically costless. And finally, copyright gives authors and publishers stakes in content that must be waived or surmounted.

Copyright was invented to give authors who lived from their work a chance to harvest its fruits. That original idea was admirable, and how it was executed, reasonable. After short initial terms, works joined the public domain. Meanwhile, extending copyright's terms has hijacked that intent, leaving the bulk of content in legal limbo. It remains protected, but because most is no longer read, no one benefits—neither authors nor the public. Authors have a stake in protecting their rights, but not in walling off work long after it has lost commercial value.

Intellectual property has also been attacked as socially unjust and contrary to public enlightenment. Some inventors have forgone

patenting their inventions to benefit the public: John Walker, the friction match; Jonas Salk, his polio vaccine; Stanley Dudrick, the technique of direct intestinal feeding of hospital patients; and Tim Berners-Lee, the underpinnings of the web.¹⁹ Authors have released works into the public domain without seeking profit. In contrast to Erasmus, Luther refused all royalties on his massively well-selling works. Tolstoy annoyed his wife by treating his rights carelessly. He hoped to allow anyone to read him after his death.²⁰ Influenced by Tolstoy, Gandhi regarded copyright as hindering the free circulation of ideas. But he also acknowledged its value in attributing works to authors and used it for his own purposes, to curtail exploitation of his writings.²¹ The title page of the 1910 English translation of his *Hind Swaraj (Indian Home Rule)*, issued by Gandhi's International Printing Press, clearly states "No Rights Reserved."²²

Beyond a few high-minded creators, ready to share their work, open access has also become an imperative. Open access was made possible by digitality, but it has become necessary because we now have more scientific research and content than conventional publishers and libraries can handle. Without digitality and its costless marginal copies, growing content could not be made freely available. With paper publication, the costs of each physical copy need to be met. As scientific research expands, it becomes ever less feasible for any one collector to own all content in a system where the reader pays. Not even a large country can shoulder that expanding global load. The dissemination costs must be borne by humanity as a whole—an unlikely prospect barring a global government—or they have to be paid by authors, thus making content free for readers. We return to this basic point in more detail below.

The driving forces behind open access have been several. Without digitality, it would be impossible. In the analog world, the closest approximation would have been a global deposit library, something like the Library of Congress or France's Bibliothèque Nationale, but for the world. Every nation's publishers would have sent their

works, and scholars and readers could then assemble in one place to find everything. Even with costs shared, they would have been insurmountable. And, as with Mecca or other mass pilgrimage destinations, the logistics would have been daunting.

But that moment has passed. Digitality and the potential availability of everything everywhere have changed the game. The global paper deposit library would have brought all the world's scholarly Mohammeds to the mountain, but now the mountain can be everywhere. The initial costs of preparing a work for dissemination remain much the same, although technology has trimmed some charges. Yet, digitality has also added new expenses, such as the bells and whistles of e-editions. Once such up-front costs have been met, however, the coast is clear. In the paper era, there still remained the physical printing, binding, shipping, distributing, displaying, selling, cataloging, and shelving. Much of that has now vanished, replaced in some measure by digital-specific costs, such as tagging with metadata, storing, updating, and maintaining availability.

Digitality was open access's necessary but not sufficient cause. What then prompted the need for it? Foremost, there was copyright's needlessly extended duration. Having been 14 years in the eighteenth century, today it guarantees prolonged rights in works—usually 70 years after the author's death. Even that fails to satisfy the most table-thumping fundamentalists for authors' rights. Why can we not keep our works forever, as we do our houses, they lament? Eternity minus a day was the solution slyly proposed by Jack Valenti, Hollywood's long-time shill in Washington. That would have allowed movie makers to sidestep the Constitution's insistence that copyright was for "limited times."²³

But, in fact, intellectual property was treated more favorably than real estate. Most jurisdictions tax houses, whether they generate income or not. The state thus captures their value every certain number of years. How many depends on the rates, but it is often less

than the life-plus-70 guaranteed for intellectual property. Authors in California, for example, would have to die at age 30 to find their intellectual property treated as badly as their houses.²⁴

Had not nineteenth-century reforms endlessly prolonged copyright terms, this appendix might never have ruptured. Imagine five- or even ten-year copyrights. Few books or articles are still read or bought after a decade, except perhaps by historians. Of the 11,000 US books published in 1950, only 400 were still in print half a century later.²⁵ Shorter copyright terms would have allowed authors and publishers to capture the economic value of academic works. With digitality permitting easy distribution, the swift onset of the public domain would then have released them quickly. Some arrangements would still have been necessary for archiving and updating digital records. A role perhaps for libraries? But that would have been it. Had copyright law been reformed to return it to something like its eighteenth-century origins, we would probably not be having this discussion in the first place. Open access is in large measure a justified reaction to rights-holders' overweening claims.

To copyright elephantiasis then came the challenge of scientific demography, which threw up financing issues. The amount of research now produced globally is simply too much for the inherited system where readers pay. Digitality's efficiencies can alleviate the problem but not solve it. Ultimately, content can no longer be financed by readers and will have to be underwritten either by governments or authors and their funders.

The research establishment has grown continuously, but rarely as fast as in the postwar decades. In the US, the 1944 GI Bill diverted millions of returning soldiers to higher education. At its peak in 1947, recipients made up about half of all college students. Almost eight million new state-funded students sluiced through the massively expanding system.²⁶ As universities multiplied, so did the faculties teaching students and the researchers supplying knowledge to fill their lectures and textbooks. This held equally in the

developing world. No nation has stamped universities out of the ground more rapidly than China. During the first decade of the twenty-first century, its institutions of higher learning doubled.²⁷ Eager to participate both as consumers and contributors, scholars in the Global South have published ever more actively. China issued 21 scholarly journals in 1970, but over 11,000 by 2019.²⁸

Mushrooming in size and number, universities improved in quality, too. Undoubtedly, mindless credentialism, overly specialized research, and a myopic fixation on publication have occasionally been overvalued at the expense of teaching. Much more striking is the overall improvement in the quality of academic institutions, their faculty, and their output. There are more good universities than ever, producing ever-better research. The Ivy League has become the Ivy Plus, and institutions outside its hallowed halls best it at its own game—Berkeley, Chicago, and Stanford, with Duke, UCLA, and Hopkins in the wings. Oxbridge has become the 24 members of the Russell Group. At the top, the best jostle for position in the myriad rankings, whose metrics and their gaming have become university administrators' fervid preoccupation.

Fighting for a handhold at the pinnacle of the greasy pole brings out the worst, and it is easy to mock deans and chancellors endlessly spouting the rhetoric of excellence.²⁹ Higher education inhabits its own Lake Wobegon, where every institution is both above average and in someone's Top Ten. The metrics higher education evaluates itself by often have house-of-mirrors qualities. As always, the measure by which the outcome is judged quickly becomes the goal. If selectivity is prized, the solution is to solicit more applicants to reject.

Still, it seems churlish to lament a fundamentally desirable process. Universities compete to be good, better, or at least improving at their mission: delivering well-educated and well-socialized young adults, shedding light on social problems, and probing what science documentaries portentously call the mysteries of the universe. The

patronizing idea that newcomer institutions will never join today's elite universities ignores strivers from the lower ranks and especially the developing world's aspirants.³⁰ There is plenty of room at the top.

The number of students mushroomed as more 18-year-olds entered tertiary education. As a matter of simple arithmetic, a burgeoning denominator has dropped the overall averages of many outcome metrics, compared to the era when only a few of an already select elite's offspring attended university. It has led to more stratified systems, with specialized institutions aiming at various clienteles—polytechnics, ag schools, liberal arts colleges, research universities, junior and community colleges, *Gesamthochschulen*, and the like. But in fact, more well-prepared and thriving students attend higher education than ever.

Grade inflation may be a problem. But at least it testifies to the urgent desire of Stakhanovite undergraduates for recognition of their hard work and the world's acknowledgment that university study is desirable. Compare that to yesterday's gentlemen's C tradition, which exemplified the indifference to education among well-heeled students. We have gone from *Brideshead Revisited* (Oxford undergraduates do everything but study) to *The Paper Chase* (Harvard law students work too hard).

As well-trained PhDs flooded the market, institutions that earlier had not presumed to demand advanced training or independent research from their faculty now had their pick of young scholars with reputations, publications, and ambitions. Franklin Ford, Harvard history professor and dean, once received an honorary degree at a small southern college. During the ceremony, almost everyone was addressed as Professor; he, however, as Dr. At some point, one of the locals on the stage leaned over and explained: We have lots of professors here, but not many PhDs.

This has changed in the meantime. A massive leveling-up has raised many ships in academe's flotilla. Universities have bootstrapped themselves up the league tables. Barriers that had once kept

outsiders from the gentlemanly club of self-financing amateur scholars fell away as the universities expanded. The obstacles were gradually cleared that had excluded women, Jews, ethnic minorities, and working-class students. They now made careers of what had earlier been reserved for those who could take it up as a hobby or a calling.

Added to this expansion in the West was the growth of science in the developing world over the past half-century. America once sent its scholars to Europe for training. In the twentieth century, American institutions overtook the motherland's, reversing the brain drain. That is happening again. Chinese and Indian students who once flocked to the US and Europe are now returning as mature scholars to institutions at home with first-rate facilities and infrastructure.³¹ Their students will soon no longer need to make an intellectual pilgrimage abroad. These recent arrivals to scholarship now find their place in the pantheon of publication, reception, and recognition. The scientific world has doubled if not tripled over the last half-century.

In step with the massing of faculty and researcher ranks has come an explosion in productivity and publishing. Professionalizing criteria for advancement and promotion, universities have emphasized scholarly output, painstakingly measuring quantity and quality.

Readers of David Lodge's *Changing Places*, set in the late 1960s, remember his description of the literature don, Philip Swallow. Mildly interested in many things, Swallow had no academic specialty, had published no more than a few book reviews and an occasional article. His main function seems to have been emotionally holding the hands of troubled undergraduates in his seminars. Publish or perish was the new hurdle that caught the Philip Swallows of his generation's university dons by surprise. Although an American invention, it spread widely. Swallow was Lodge's representative of British universities, anno 1969—the foil to hard-charging Morris Zapp from Euphoric State University on America's West Coast.

Zapp's ambition was a series of exhaustive commentaries on classic authors, starting with Jane Austen. Having covered every conceivable critical angle, he intended to shut down the need for future works on these subjects, putting his fellow literary scholars out of business. Meanwhile, having once dallied like amateur scholars, British academics are now held to more stringent publishing requirements than in the US. Britain's repeatedly renamed research assessment system has unleashed publish or perish to new ferocity. University departments' financing now depends on faculty productivity—submitted to labor-intensive assessments by government funders—in a way the decentralized US system could never enforce.

Intellectual productivity has accelerated. Far more books have been issued in the past century than ever before. Three-quarters of the titles in Google Books were published after the Second World War, half after 1974.³² Total global book production as of 1911 was slightly over 10 million, as of 1940, about 15 million.³³ That leaves the remaining 165 million books in existence today, some 92% of all, having been issued in the subsequent 75 years.

Eviscerating the Libraries

So much for the quantum of intellectual output, increasing as more researchers notch up their productivity. To that comes the cost of acquiring it. In the paper era, university libraries were largely the only customer for academic periodicals and the main one for monographs. Occasionally, a wealthy collector—Aby Warburg—could be a player, and a few private libraries remain: the London Library and the Mechanics' Institute in San Francisco. Some scholars and professionals may take periodicals in their fields, and corporations with research departments subscribe, but for scholarly journals and monographs, the locomotive is the university library.

For academic monographs, libraries account for about half of sales. They buy one-quarter of university press books more generally, which today include many titles marketed to a broader audience.³⁴ Late in the twentieth century, before the worst of the library budget cuts, a university press could count on selling out an edition of 1,500 to research libraries alone, recouping its expenses. Today, average sales of a humanities monograph (half perhaps to libraries) are 600 or 700 five years after release.³⁵

Libraries once footed the cost of academic publishing. Scholars may have been the main readers but not the primary customers. Little did they care how expensive a journal or book was, so long as they could expect to find it on the library shelf. Nor would they accept substitutes. Each article or monograph is unique and equally necessary for scholars hoping to write authoritatively and exhaustively.

Few criticisms are more damning than suggesting that a scholar has missed or neglected a crucial source. The defense that the local library could not afford it would be laughed out of court. One way or another, the presumption is that serious scholars have access to everything in the field. Only laziness or ignorance could explain not consulting a source. Librarians are thus mandated by their primary clientele to buy whatever the publishers issue. No matter that the vast majority of works in any given library sit unread for decades, usually forever. In a medium-size US university library, only 20% of books are checked out even once.³⁶ On the off chance that someday a scholar may amble the stacks looking for this obscure article or that dusty tome, it must be ready for its Cinderella moment. For research libraries, tails are very, very long.

The need to be comprehensive left academic libraries vulnerable. The scientific publishers identified a captive clientele required to buy their offerings at whatever price. Committed to exhaustive coverage, librarians could not be discerning consumers. Media entrepreneurs found themselves in a business fantasy. There were no

market failures, no misjudging customer wishes, no New-Coke-style fumbles: their products were guaranteed a market at the price they chose. Not since Soviet tractor monopolies did producers have it so cushy, and even the machinery manufacturers had to scramble for raw materials and energy. With academic publishing, the materials delivered themselves for free in ever-growing quantities to the receiving bays. All that was required was sorting it a bit, slapping on a fresh coat of paint, and deciding at what price to push it out the door.

The potential for exploiting this situation was not new. Academic libraries and publishers had long danced a pas de deux. But their interactions had worked because the scholarly offerings had not yet exploded and the suppliers were self-restrained by a courteous gentlemanly understanding that publishers and librarians were in the same business of providing for scholars. Publishers neither oversupplied nor overcharged, and librarians bought most of their goods. The new media entrepreneurs, in contrast, were prepared first to fatten up, then eviscerate the system.

Starting in the seventeenth century with the first scientific academies and their meetings, papers, proceedings, and reports, scientists worked for free, motivated by truth and recognition. Their output was not regarded as their intellectual property, except to vouchsafe their moral claims to attribution and recognition. If they discovered something that could be patented, that was one thing, but for basic research and scholarship, there was little market. Scholarship rested on an economy of symbolic exchange of prestige and recognition.

That did not change until science assumed geopolitical implications and governments began underwriting research. States had long supported technologies of warfare, mining, exploration, and other endeavors, allowing them a leg up. Early academies and universities had received government funding for such pursuits.

The increasing cost of labor and the difficulty of sacrificing citizen-soldiers, whose families voted, made warfare even more

expensive as democracy spread. Medieval armor had been the preserve of the aristocracy. Military technologies that both attacked and protected even common soldiers, such as tanks, did not come cheap, nor did those allowing combatants to fight at arm's length, sparing personnel—bombers, rockets, missiles, drones. Most expensive of all were those—atomic—armaments whose logic was that they never be used at all. Warfare without killing costs more than the deadly kind. Nor did state-funded explorations come cheap. Columbus received the equivalent of one million dollars in royal support for his voyages.³⁷ By modern standards, that was a pittance. At \$160 billion and counting, the International Space Station is the most expensive object ever built.³⁸ The pyramids were a bargain by comparison—perhaps a billion dollars in construction costs each.³⁹

The Second World War accelerated government funding of useful R&D, which continued into the Cold War, with the arms race and competition in space. With big science came big universities and big collaborative projects: the Manhattan Project, Sputnik, Collider Detector experiments at Fermilab, the Hubble space telescope, CERN, the Human Genome Project, the National Nanotechnology Initiative, and the like. Government funding went to universities and other research institutions, such as Bell Labs (13 Nobel laureates). The universities' growth undermined their own monopoly, producing numerous graduates who went to work for new competing institutions.⁴⁰ As the supercharged engines of research began spitting out results, the inherited system of scholarly dissemination faltered.

Among the first publishers to scent blood was Robert Maxwell and his Pergamon Press. In the immediate postwar years, Maxwell worked for the press control division of the British occupying army in Berlin. In command of paper supplies, he was lobbied by publishers and struck up a relationship with Ferdinand Springer, owner of a once-powerful scientific house. As with so many other German intellectual and cultural achievements, the Nazis also vaporized

academic publishing. Learning the ropes from Springer, Maxwell was astounded to discover a unique business model. Unlike other content industries, scientific publishers did not need to create the product or pay others for it. Academics delivered it for free, libraries paid to take it off publishers' hands. Only a bit of sprucing up was needed before the finished output could be sold.

Maxwell returned to Britain, having acquired worldwide distribution rights to Springer's huge backlog of journals and other scientific output. In addition to trainloads of periodicals and books came seven rail carriages of manuscripts awaiting the light.⁴¹ Maxwell had worked as a commodities broker, selling a Noah's ark of products. If commodities are goods distinguished primarily by price (pork bellies, iron ore, wheat), scientific articles were the opposite—something whose provenance was determinative, that one could charge almost any price for without killing the market. Maxwell had spotted the opportunity nestled where increasing scientific output outstripped the conventional channels' ability to keep up.⁴²

In the early 1950s, Maxwell entered publishing directly, founding Pergamon to repeat Springer's prewar success.⁴³ With a keen eye for developing new academic fields, he positioned Pergamon as the premier English-language supplier of scientific content in an era when the research world was becoming Anglophone. He recognized the allure of slapping "International Journal of" before almost any field or discipline, from "Solids and Structures" to "Parasitology," from "Educational Research" to "Applied Radiation and Isotopes."

Maxwell and his competitors identified a value in academic publishing distinct from any marketable discoveries it might contain. As state financing for research increased, disseminating it became a juicy prize. Even if it cost only a few percent of total research funding to issue the results, the sums were impressive. As citations began to be tracked and journal impact factors measured, the market became less elastic. Certain highly cited journals became the core of collections, must-have acquisitions for librarians. Maxwell and

other scientific publishers raised the price of journals libraries could not forgo.⁴⁴ The postwar university boom produced thousands of new researchers itching to publish and hundreds of new libraries to fill. By itself, Pergamon issued 700 new journals.⁴⁵

The commercial scientific publishers did not cause the demographic revolution in science, the thousands of new scholars and their output, but they did exploit its opportunities. Expanding the number of journals was necessary, jacking up their prices less so. The explosion of scientific research would eventually have posed an unsustainable expense for libraries even had it not been exacerbated by rent-seeking publishers driving up costs as well. Their greed merely brought forward the moment of reckoning. Independent of price, the research explosion tolled an end to the old system of subscription journals (and retail books, to a lesser extent).

The Serials Crisis

Something as seemingly mundane as the increasing price of scientific periodicals precipitated the open-access movement. This crisis of serials subscriptions sprang from the postwar divergence between supply and outlet for scientific research. Researchers writing articles multiplied, yet publication venues failed to expand in tandem. Into the breach stepped the commercial houses. Before the war, professional societies and their journals had issued most scientific research. Their prices were reasonable, the pace sedate, the volume adequate. But when government research funding continued into peacetime, the old system failed to keep up. The new commercial journals took up the slack. At first, library budgets increased along with general research financing, and subscriptions to new journals were part of a rising tide.

All was well until late in the century.⁴⁶ University library budgets could accommodate the increased volume, and prices remained

reasonable even as quantity increased. But as of the 1980s, the shoe began pinching. Cutbacks in university financing left libraries unable to keep pace. Worse, the publishers got greedy. Vast new numbers of specialty journals proliferated, their subscription prices arching ever skyward, far beyond inflation or costs. Libraries' expenditure on serials now escalated at triple the consumer price index.⁴⁷

At the same time, the big commercial publishers consolidated into a handful of giant firms running huge stables of titles. In 1991, for almost half a billion pounds, Maxwell sold Pergamon and its four hundred titles to Elsevier, then a small Dutch publisher of technical journals.⁴⁸ From the 1980s onward, three publishers (Reed Elsevier, Springer, and Wiley) enjoyed the lion's share (42%) of the 25,000 leading English-language scientific periodicals. No other house controlled more than 3% of the market.⁴⁹ When the libraries' shrieks of despair could no longer be ignored in the late 1990s, the publishers switched tactics.

Eyeing the burgeoning internet and its threat of easy and low-cost dissemination, they realized that the libraries' golden geese had perhaps been pushed to the limit. Rather than having to order à la carte, libraries would now be offered all-you-can-eat buffets. Big-deal packages gave them access to a publisher's entire list for flat-rate prices. Smaller presses lacking sufficient content to stock bundles on their own joined aggregators, such as Aggregagent, BioOne, or Project MUSE.⁵⁰ Eventually, book publishers extended similar deals for their monograph lists.

The big-deal packages dropped the per-article or per-read costs but imposed new expenses on libraries. It locked them into multi-million-dollar annual contracts for content they neither chose nor controlled. If they balked, large fractions of their collections would go dark. Prices were calculated not on publishers' costs but on the libraries' purchasing history. Rates were set slightly above what libraries had earlier paid for all their individual subscriptions. That increased the publishers' intake and cemented in the libraries

at already prohibitive expenditure levels. The now locked-in prices ratcheted up at rates set by the publishers.

Big deals did nothing to alleviate costs and froze the status quo of funneling library budgets to the publishers' bottom line. They did supply the full range of publishers' output for those institutions able to keep pace on the big-deal treadmill. Above all, they made the publishers indispensable. By 2015, Elsevier, now one of the largest, owned a quarter of scientific journals. Anyone who dared monkey with its big deals—as a few libraries valiantly tried—had to be prepared to ask faculty to do without (or find other access to) a quarter of all content.⁵¹

Despite the Covid pandemic and numerous boycotts of its journals, Elsevier posted robust results for 2020—down from 2019, but still with profits of over £2 billion on revenues of £7 billion.⁵² Big deals also undermined small publishers and scholarly societies, with their few or sometimes single journals. With their budgets locked up by big deals, libraries had little money left for anyone but the major players. To survive, small journals often sold themselves to larger publishers and were folded into their bundles.

Scientific publishers have also played hardball with libraries in at least two other ways. TV viewers will recognize the first strategy from how cable companies package channels. Since no one wants more than a few in each bundle, customers are encouraged to pay for several. Publishers also sell collections of journals, insisting that each be bought in toto and making it hard for libraries to pick and choose. In desperation, libraries have begun unbundling their subscriptions. They pay *à la carte* for the journals that faculty demand most, relying on interlibrary loans or individual purchases to plug gaps.

Second, confidentiality agreements forbid disclosure of publishers' arrangements with each library. Pricing is thus what the market can bear, with richer institutions and nations generally paying more than others. Grotesquely, sometimes developing countries,

such as South Africa, have been billed more than rich ones, such as Germany.⁵³ Ignorant of what others are charged, libraries cannot strike cooperative purchasing agreements and the publishers are spared competition.⁵⁴

From all this flowed enviable profit margins, an astounding 35% to 40%. Few other businesses outside software and pharmaceuticals could promise anything close. Besides banknotes, the *Frankfurter Allgemeine Zeitung* said in 2012, scientific papers were the most lucrative item you could print.⁵⁵

Open Access Takes Off

Against this background, open access seemed like a solution not only to the serials crisis and squeezed library budgets but also to broader problems of scholarly publishing. Digitality promised to trim dissemination costs in general. If control of content could be wrested from commercial publishers, exorbitant profit margins could perhaps be pared back. Yet, while digitality broached the opportunity of making everything available, it did not specify how to meet costs.

There were two aspects to open access, embodied respectively in each of its terms. *Open* came from the open-source software and free culture movements of the 1980s. Emphasizing a do-it-yourself and anticorporate ethos, this strain sought not just access, but for productive and creative processes to be reclaimed from the institutions thought to have monopolized them—libraries, think tanks, journals. Those attracted by the openness of the new technologies expected them to unleash new possibilities. Anyone could create near-perfect output on ubiquitous devices. Cutting out the middle people, creators could speak directly to their audiences. The web linked to everyone and all but eliminated connection costs, allowing new collaborations to emerge spontaneously. Unprecedented

information resources could be assembled—vast collections of photographs or the miracle that is Wikipedia.⁵⁶

For the *access* enthusiasts, in contrast, the issue was less whence and how content emerged but that it was available to consumers without unreasonable impediment.⁵⁷ The inherited institutions did not necessarily have to change so long as end consumers could freely use their output. Over what that meant, opinions could differ. Two decades ago, during the web's cowboy days, Pirate Bay, Napster, MegaUpload, and similar sites posted bootlegged copyrighted content that could be downloaded or viewed without payment. Differing in the technicalities, they all claimed to allow private file-sharing among users without centralized storage or reproduction of content, therefore not in breach of copyright law.

However, such claims were ruled disingenuous, and the sites were pursued and curbed. In their place, a regularized ecosystem of legal streaming services—Spotify, Apple Music, Hulu, Netflix—now offers much the same at reasonable prices. An Amazon Prime subscription at \$13 per month is not open access in the strict sense of content available costlessly. But the opportunity to watch some 13,000 films for the price of one trip to the cinema monthly is a good deal as lending libraries go. Does that count?

Today, open access is dominated by scientists. Initially, however, the humanities and social sciences sparked the initiative. As with Tolstoy and Gandhi, copyright has long been feared as stifling the free use of ideas. Nor is the celebration of reusing content for new creations an exclusive hallmark of contemporary culture. Modernity threw off the Romantics' exalted view of individual creativity. Mutual indebtedness and influence, use of others' works in unabashed bricolage and pastiche, have become commonplace. As early as 1819, Washington Irving had argued that literary creation was parasitical. Like birds, who served nature's intent by excreting fruit seed, authors were but a means of conveyance, passing along ideas from old works into the present.⁵⁸

Creativity seen as a collective effort was the soil that nourished more mundane ambitions for bypassing the publishing establishment to open up content. The old regime of paper and ink had limited information's spread, but once digitality sprang the technical barriers, new vistas opened up. In the university world, an online journal appeared for the first time in 1987, with the perhaps less-than-pulse-quickening first issue of *New Horizons in Adult Education*, hosted by Syracuse University.⁵⁹ The 1980s and 1990s brought other journals dedicated to accessible content: *Surfaces*, *CTheory*, *Postmodern Culture*, *Music Theory Online*, *EJournal*, *Journal of Political Ecology*, *Electronic Journal of Differential Equations*, and the *Bryn Mawr Classical Review*.⁶⁰

Meanwhile, a series of foundational meetings elaborated the emerging view of open access. As activists formulated their ideas during the first years of the new millennium, ambitions expanded. In 2002, George Soros's Open Society Foundation hosted a meeting in Budapest. Of its 16 participants, most had been trained in the humanities or social sciences, only four in the hard sciences. Its declaration is commonly taken to be the starting gun of the movement, but as we will see, it was pipped to the gate by developments in Latin America.

The Budapest declaration took the conventional format of peer-reviewed scientific articles as its primary concern.⁶¹ It defined *open* as the right to read, download, copy, distribute, print, search, or link the articles, to crawl them for indexing, and to pass them as data to software, or any other lawful use. The only restrictive author's rights it recognized were to be acknowledged and cited and to control the work's integrity. It accepted digital self-archiving of (refereed journal) articles as a way of disseminating them. With a shopping list of "many alternative sources of funds," it foresaw the founding of new journals that charge neither subscriptions nor publishing fees. In its bootstrapping vision, these were solutions that scholars themselves could effect, without relying on changes in legislation or the existing publishing industry.

The humanities' influence quickly evaporated, however. The following year, in April 2003, a meeting was held in Chevy Chase, Maryland, but its declaration was named after neighboring Bethesda. Attendance was heavily from medicine and focused on the "biomedical research community" and its "primary scientific literature."⁶² The Bethesda Statement expanded the Budapest principles. It required that for content to be considered open, its reuse for derivative works must also be allowed. Curiously, despite its expansive definition of *open*, it required only digital uses; therefore, the statement had to tack on a separate right to make limited numbers of printed copies for personal use. The statement also demanded immediate deposit of works upon publication in repositories run by academic institutions, scholarly societies, government agencies, or other well-established organizations engaged in open access. If the work was not published openly, it had to be simultaneously made freely available, with no embargo. Nor would works be considered open unless they were put in a suitable noncommercial repository. Publication in a commercial open-access journal was insufficient.

In the autumn of 2003, a further meeting at the Max Planck Society in Berlin—Germany's largest and most prestigious network of scientific research institutes—brought together over a hundred organizations from far and wide. It defined open access as "a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community."⁶³ Oddly, that shifted the focus from the work's content to its origins. Stranger still, it restricted open access to material evaluated by the scientific community, even though the concept of cultural heritage surely went beyond that. It also expanded the scope of the openly available from scientific articles to include raw data and metadata, source materials, digital representations of pictorial and graphic works, and scholarly multimedia material. It loosened requirements for the attribution right by leaving it to "community standards" for enforcement.

The Publishers Capture Open Access

Early plans for open access seemed to threaten the scientific publishers. Demands that content be available for free did not sound encouraging. Nor the expectation that digitality would trim costs. For scholars to begin disseminating their work threatened publishers' business. Advocates for open access did not just want the commercial houses to take on the enterprise. They hoped that hitching digitality to open access would allow a refounding of the business altogether. "It is time to return control of scholarly publishing to the scholars," the Max Planck Society announced 10 years after the Berlin Declaration.⁶⁴ New noncommercial journals were launched in this spirit. *BioMed Central*, begun in 2000, was a for-profit venture, while the *Public Library of Science* (PLOS) from 2003 was not. To cover costs, *BioMed Central* pioneered article processing charges, paid by authors or their funders. In 2002, these were a modest \$500 per accepted article, with waivers for hardship and for authors from developing countries.

But expecting the commercial houses not to crowd in was unrealistic. Libraries had financed subscription journals. The publishers had learned how to turn acquisitions budgets to their ends. Was that now to evaporate? Open accessors eyed library budgets as a potential funding source for new scholar-led journals. If the commercial publishers could also tap such sources for open access, would their exorbitant profit margins from the subscription days go unchallenged?

The shift from reader- to author-side financing is a story in progress, and its ultimate outcome is still unknown. So far, the commercial publishers appear to have turned it to their advantage. Despite a decade of protest, boycotts, funder grumbling, threats of legislation, investor pull-back, and other obstacles, their profit margins remain substantial. In-copyright digitized content can either be opened up or closed down. Digitality allows both liberation and

control.⁶⁵ Just being online does not mean it is freely available. That is among the considerations that have helped commercial publishers make their peace with open access.

In 1991, Elsevier launched The University Licensing Program (TULIP). After digitizing back issues of some 40 science journals, Elsevier made available to university libraries online versions of content most of them already owned on paper. Elsevier did not charge fees, but participating libraries had to invest in hard and software and log in via various then-novel security mechanisms to ensure a walled-garden approach to access.⁶⁶ A similar project launched the following year by Springer, the Red Sage Digital Journal Library, reinforced the publishers' conclusion that digitized content was useful and had a market, and that academic libraries were willing to cooperate by keeping it squirreled behind authentication walls.⁶⁷

As born-open journals issued forth, how did existing subscription periodicals manage the transition? The scholarly societies, publishers of many traditional journals, were in a difficult position. Unlike the commercial rivals, their prices were not extravagant. Many, especially in the humanities, would have found it hard to flip to open since their members could not pay publishing charges.⁶⁸ Their earnings (over £100 million in the UK, much of it from foreign subscribers) were plowed back into their scholarly mission.⁶⁹ They had not been part of the serials crisis aggravated by their commercial rivals. Many felt unfairly burdened by the new demands for access to which they had never presented more than modest barriers. Yet, to be fair, nor had they done much to help disseminate the increasing content from the postwar research explosion. They had allowed the vacuum that commercial publishers rushed to fill.

In turn, commercial publishers quickly learned to live with open access. With the green version, they made their peace by insisting on long embargo periods and that repositories not post the version of record. Gold open access, in turn, opened new vistas.

Cleverly, some publishers turned their journals hybrid. They included some paid-for articles, freely available, while others were reserved for subscribers only. The most prestigious outlets were still the venerable subscription journals. Scientists with sufficient funding who wanted the best of both worlds sought broad access to their writings together with the prestige of the respected venues. The hybrid journals accommodated both. Articles for which publishing fees were paid appeared for the world to read; others remained behind paywalls or were restricted to subscribers. To supply readers with everything, libraries had no choice but to continue paying subscriptions, even as some content was open. The publishers double-dipped, the libraries were double-charged. Hybrid journals should have offered discounts in proportion to the content already paid for via fees, but the overall cost of subscriptions showed no signs of declining.

Hybrid publishing had begun in 1998 at the *Florida Journal of Entomology*. In 2004, Springer began experimenting with the idea, collecting fees in addition to subscriptions.⁷⁰ What was there not to like? To be fair, hybrid had been accepted from the start in the open access declarations and was not just a cynical publisher's ploy to have it both ways. The Bethesda Statement specified that open access was a property of individual works, not necessarily journals or publishers. And it allowed for an "open access option" for articles published within conventional journals.⁷¹ The Open Knowledge Foundation's definition forbade restrictions (such as requiring accessibility) on other works published together with open ones.⁷²

From hybrid, it was only a small step for publishers to plunge fully into the open-access stream. Springer bought *BioMed Central* in 2008. Other commercial publishers followed.⁷³ Springer, owner of the venerable subscription journal *Nature*, now publishes *Nature Communications* and *Scientific Reports*, charging several thousand dollars for each article, somewhat less for reports. One calculation estimated the total income for both journals in 2016 at \$50

million.⁷⁴ Gold access spares publishers some expenses (sales, licensing, marketing, copyright, and subscription management), which once accounted for about a third of total costs. Investment bankers have concluded that, with article charges of \$3,000 per article, publishers are unlikely to lose revenue.⁷⁵

Fully open-access publishers like Biomed Central and Hindawi enjoy juicy profit margins (27% and 43% in 2008 and 2011, respectively), comparable to their conventional competitors.⁷⁶ Hindawi cashed in when Wiley bought it for almost \$300 million in January 2021. Elsevier has bought Mendeley, an academic social network, and the Social Science Research Network (SSRN), a prepublication repository. The latter allows it to guide readers to its open-access versions of papers that may later appear paywalled with competing houses.⁷⁷ Gold access has increasingly become the preserve of the dominant conventional publishers, able to afford such acquisitions: Springer, Wiley, and Elsevier. In November 2020, Springer announced that *Nature* would flip to open. Tooting its own horn as a “progressive publisher and innovator in open access,” it also announced that the fees for each article would be over \$11,000 (€9,500).⁷⁸ The gift of open access kept on giving—to the publishers.

The journals realized that it made little difference whether you called their monies subscriptions or article charges. It was merely a question of flipping the funding stream from consumer to producer, from reader to author. The source was ultimately the same, the government or private science funders. And for them, the costs were so small that reversing the current meant little.

Nor did publishers meet much resistance from scientists. As publishers and funders learned not to be fussed by gold access, the scientists climbed aboard. Keen to preserve the inherited ecosystem of prestige, few minded continuing the subscription model or the emergence of hybrid publications. The Bethesda Statement had been issued by a group of researchers mainly from biomedicine. Their meeting was held at the Howard Hughes Medical Institute,

the second biggest US philanthropy and the second richest medical research institution globally. It had blithely stated that since publication was an essential part of science, such costs were part of research expenses, and it promised to help meet them.⁷⁹

The monies were already in the system. It mattered little precisely how they were deployed. For scientists, gold or diamond access left inherited methods of communication and credentialing largely untouched. That was the attitude given voice by the Max Planck Society in 2015. The monies were already present, they argued, locked into library subscriptions. Only repurposing them was needed.

In their doubtless well-intentioned approach, the scientists were oblivious to how the serials crisis had already allowed their journals to devour the bulk of library budgets, undercutting humanities publications. Maintaining “the established service levels” scientists had become accustomed to was the goal, plundering library budgets as “the ultimate reservoir for enabling the transformation without financial or other risks,” the means.⁸⁰ Had such discussions been held before the serials crisis, not in its bleak aftermath, it may have been better received among other scholars. As things stood, to flip the funding stream from subscriptions to publishing charges while leaving library budgets to bear the brunt meant locking in a status quo that was unacceptable to all but the hard sciences.

Many governments aligned themselves with gold access. In the UK, the Finch report in 2012 embraced this route forward.⁸¹ Ignoring Britain’s well-developed green repositories, it plumped for gold.⁸² Whether fully open access or hybrid did not concern it, but publication charges should fund the transition. Embargo periods should not be overly shortened since that endangered traditional subscription journals. The report recognized that during the transition period, funding would be duplicated, with both subscriptions and publishing charges, not to mention the cost of repositories. It therefore called for more government money and savings from

“other features of the research process” as well as efficiencies in publishing. Boats were not to be rocked. The Finch committee had been instructed not to endanger the British publishing industry.⁸³ Its report followed through, insisting that “the underpinning of high-quality publishing channels” should not be put at risk.⁸⁴ In effect, the Finch report sought to preserve the commercial publishing model, not change it.⁸⁵

On this march toward gold access in the sciences, Plan S has been the latest installment. Using the tortured acronyming popular in open-access circles, a group of European private and governmental funding agencies named “cOAlition S” launched Plan S in 2018. As of 2021, it required all scholarly work they underwrote to be published in open journals or to have the accepted version of the manuscript made available in repositories without embargo.⁸⁶ This undercut green access, which relied on the accessible version being handicapped compared to the published edition. Plan S thus suffered the faults of gold access without solving any of them except by demanding compliance. The affordability problem for underfunded researchers hoping to publish remained untouched.⁸⁷ Nor did the concerns of humanities and social science scholars without funds receive attention.

By pushing a narrow view of open access, Plan S also alienated conventional scientists, who were unhappy that anyone accepting such funding was barred from subscription journals. Eighteen hundred chemists harrumphed that this was an attack on their academic freedom.⁸⁸ Nor has the Global South been impressed. Plan S struck them as merely restating gold access. It allowed readers in, while doing little to help scholars in developing nations to publish.⁸⁹

In response to objections, Plan S adjusted course. It introduced a Rights Retention Strategy that allowed authors to be compliant by self-archiving the final version of their manuscript.⁹⁰ Those who published in subscription periodicals could thus remain within the pale so long as journals accepted that authors posted competing

versions of papers. Only the journals that flatly refused to accept work that had also been self-archived would not be compliant.

The success of this attempt to swing the guiding assumptions of science toward gold open access remains unclear. As of 2021, Plan S mandates extended only to those who accept research monies from a limited set of European funders. Outside of Europe, only three funders had signed on: Howard Hughes, the Templeton World Charity Foundation, and the Gates Foundation.⁹¹ Some European funders had exited: Sweden's National Bank Jubilee Fund and the European Research Council. German and Swiss funders had also jumped ship; the Spanish and Belgians never joined. So far, the outcome has been a two-tier system, obliging some European researchers to publish in compliant journals while their colleagues elsewhere remain free to go where they can. By dropping its initial price cap on article publishing charges, Plan S did nothing to restrain costs.⁹²

Financing beyond Gold

The analog world's most far-reaching ambition could not be more than national deposit libraries. Once digitality broached the opportunity for all content to be available everywhere, a new financing model was required. Making readers pay, whether directly or through libraries, no longer sufficed. No consumer could afford the world's output. But if content's dissemination costs were prepaid, then everyone could have access.

So much for the theory. In practice, the developed world's scientists have hijacked the existing system via gold access. For this to change requires at least one of two scenarios: either the science funders revolt or copyright has its terms shortened by legislation to make the prize less worth fighting over.

Many funders have long insisted that results of their support appear openly. Governments were first off the mark. As of 2005,

the US National Institutes of Health demanded that its sponsored research be made publicly available within a year after publication.⁹³ The Research Councils UK and the European Research Council followed suit.⁹⁴ In the US, private funders, such as the Gates Foundation and Howard Hughes, insist on it.⁹⁵ In the UK, the Wellcome Trust requires articles to be open access (tolerating no more than a six-month embargo) and offers to pay fees.⁹⁶ Other European funders have similar requirements.⁹⁷

State authorities have been especially likely to insist on open access. As stewards of taxpayer monies, their conclusion that public funding should entail public availability is direct and obvious. Where research funding is centralized, compliance can be more easily extracted. The National Institutes of Health requires open access. But its remit is more limited in federalized America than for its equivalents in Europe or China's more centralized system, where the state funds most research. Discussions have therefore gone further in the UK than in the US. Even Britain's humanities and social sciences are now being squeezed into the open-access mold. Universities' research funding in the UK hinges on their faculties' openly available productivity. Monographs, not just articles, are being brought under the umbrella. Austrian, Dutch, and Swiss science agencies have also recently begun requiring accessible monographs.⁹⁸

Universities, too, enforce open access, although without the same leverage. In 2008, Harvard's faculty agreed to require posting articles in a university repository, while granting a nonexclusive copyright license to archive and distribute them.⁹⁹ Even that, however, allowed faculty to request exceptions. Enforcement is reputed to be indifferent at best. Other American universities' policies appear to be similar.¹⁰⁰ Comparable conditions also hold in Europe.¹⁰¹ British universities, in contrast, have put more muscle into deposit requirements. Thanks to centralized research funding, the government exerts leverage over faculty.¹⁰² Because textbooks do not count as

research for the assessments, scholars are now discouraged from writing them.¹⁰³ The threat to withhold monies has concentrated the professoriate's mind. Cambridge's Apollo website, for example, is a model of organization and clarity.¹⁰⁴

Yet, enforcement is not easy. In the humanities, most research is self-financed by scholars, and outside funding rarely covers more than some costs. Should a grant for a summer archive visit entail that the ensuing book appear with an open publisher a decade later? Could an author repay the grant with royalties from a trade edition? What if, later on, the scholar writes a popular trade book using some of the research funded with an open-access requirement? Other researchers can refer to and use their colleagues' open-access-mandated results without themselves being similarly obligated. Why not the original author? Does a statute of limitations eventually release authors from funders' strictures?

UK Research and Innovation, the body responsible for government research strategy and funding in Britain, tied itself into knots in 2021 when trying to formulate a policy for books. Trade books it defined as ones that appeal to a broader audience. Had the research been funded with its monies, they fell under its requirement to be open. But then it turned on a dime and declared that trade books need not be compliant, leaving that decision to author and publisher discretion. Also excepted were textbooks, fiction, and books resulting from dissertations.¹⁰⁵

As things stand, requiring open access in the humanities accomplishes little other than putting scholars in a bind if their work does not, or cannot, appear freely available. Even in the sciences, which are much more heavily dependent on outside financing, enforcement is lax. Short of threatening never again to underwrite those who ignore instructions, funders have little leverage. The Wellcome's generous policy is complied with only a bit more than half the time. The National Institutes of Health has achieved only 60% compliance with its deposit requirement.¹⁰⁶

All this leaves unaddressed the elephant in the room, the cost of dissemination. UNESCO estimated global gross expenditure on R&D at almost \$1.5 trillion in 2013. Most of that was corporate spending, but about a third came from government and a fifth from universities, \$800 billion in sum. The expense of disseminating that research is difficult to pin down. Annual revenues from English-language science journal publishing are thought to have been \$10 billion in 2017, with a global market estimated at \$25.7 billion for all forms of scientific, technical, and medical information.¹⁰⁷ If so, then dissemination costs some 3% of government and university research spending, which tallies with other estimates of between 1% and 2%.¹⁰⁸ Disseminating thus swallows only a small fraction of research outlays. With little skin in this game, funders are unlikely to press publishers to trim their profit margins or to cut costs. Scientists themselves resist change to their inherited ecosystem of prestige and reward. The existing commercial system works well for them.

Reforming copyright legislation, in turn, means limiting not just the rights to academic content but to all works. Either that, or it requires making distinctions between scientific and other content that will be difficult to define and police. Only a fraction of all content is tax-funded academic work, for which open access is justifiable. How to distinguish between it and other nonfiction that may be equally serious and valuable? Reforms that shorten copyright terms for all content would face fierce opposition from rights-holders. If, instead, reforms aimed only at academic content, endless disputes over which side of the line particular works fell on would result. Thoroughgoing copyright reform therefore seems a distant possibility.

The most likely outcome is that, having won the battle, the science publishers will not be dislodged from their position astraddle the main funding channels. Other disciplines and the developing world will have to exploit digitality's efficiencies to establish new

forms of dissemination and seek other financing. How the landscape of alternative models may evolve is only dimly perceptible now. But we can point to current experiments that may turn out to have shown the way.

Someone other than consumers must pay if content is to be freely available. Gold open access can be only a partial solution, as shown by its success in the sciences. Having authors shoulder the burden excludes creators who cannot afford fees. To rectify that, governments could finance dissemination costs, as they do for much research. As we see below, that approach has been taken in Latin America. Even without state funds, other groups can achieve similar results.

Readers can band together to underwrite the price of opening work, bootstrapping it for the world. The brilliance of digitality is that bulk cooperative purchasing not only gives immediate buyers a discount but also provides the goods gratis for all consumers. The analog era knew such paying in advance, too. Books were subscribed to and published only once the requisite sums were collected. *Festschriften* were volumes of essays issued to celebrate a distinguished colleague on retirement or a round-number birthday. They were often financed by collecting monies from the contributors, who would receive a copy in return. Weighted down by paper and binding, such volumes naturally had to be conventionally bought by others who wanted one.

Digitality has simplified such techniques. Knowledge Unlatched, for example, organizes university libraries to subscribe to books.¹⁰⁹ Once sufficient funds have been collected, works are released to the participating institutions and thus to the world. In 2014, 28 new books were unlatched, at an average cost to pledging libraries of \$43.¹¹⁰ The Knowledge Unlatched website claims a total of over 2,700 such works. In the meantime, however, Knowledge Unlatched became a commercial enterprise incorporated in Germany. It charged libraries for its amalgamation of open books,

which largely duplicated what is available for free at the Directory of Open Access Books.¹¹¹ Then, in 2021, Knowledge Unlatched was bought by Wiley.¹¹² *Toward an Open Monograph Ecosystem* is similar.¹¹³ Other publishers, too, have joined in such subscribe-to-open ventures.¹¹⁴ Having assembled 160 participating libraries, MIT Press published much of its Spring 2022 list as open access.¹¹⁵ Such funding experiments collaborate with existing open publishers, providing the fees normally collected from authors.

Unglue.it crowdfunds books.¹¹⁶ With the necessary sums in hand, the volumes are released. This works retrospectively, too, allowing publishers another bite of the apple. In 2013, De Gruyter applied crowdfunding to previously published monographs. Once \$2,100 had been collected, the title was released under a Creative Commons noncommercial, no-derivatives license. Only if the campaign was successful were pledgers charged and the book opened.¹¹⁷

The tactic works with journals, too.¹¹⁸ PeerJ offers plans allowing authors to publish one article a year for \$99, or as many as wanted for \$299.¹¹⁹ In 2020, the Public Library of Science (PLOS), a prominent series of open journals, scaled up this logic. It issues some of its titles by collecting annual flat fees from institutions.¹²⁰ Affiliated researchers can publish without separate charges in three journals. Those without an institutional association pay fees per article, almost double the current ones. The goal is to lower expenses, to collect no more than necessary (fees beyond what is needed are returned), and to distribute costs equitably among institutions by staggering payments according to their research intensiveness. Third World institutions would automatically be members without fees.

Open journals do not need payments from authors if they have other financing. Many scholarly societies and other organizations sponsor diamond journals that impose no author payments. Most open periodicals charge no fees—some 12,000 out of 17,000 in the 2021 Directory of Open Access Journals.¹²¹ But counting journals

paints a rosier picture than articles. Of scholarly articles published annually, some 8% to 9% are diamond and 10% to 11% are gold.¹²² And, as seen, prestigious journals charge the most eye-watering fees.

Resistance

The march of open access has not gone unchallenged. Sadly, the movement has gained a martyr, Aaron Swartz. He saw the fight in moral terms, taking from greedy corporations and giving to the dataless.¹²³ As a 24-year-old, he was arrested in 2011, charged with electronic and computer fraud for having downloaded almost five million documents from JSTOR via the MIT library.¹²⁴ However well-intentioned, his downloading violated MIT's license terms, prompting the supplier to shut down campus access for several days.¹²⁵ Had he released the 80% of its archive he had downloaded, none of which was JSTOR's property, the damage would have been significant. Nor was JSTOR, as a nonprofit digital distributor of social science periodicals to libraries, an obvious villain. To set an example, the authorities sought a long sentence but also offered Swartz a zero-to-six-month plea bargain, substantially discounting the seven years he could expect if convicted. Tragically, he found his predicament so intolerable that he committed suicide.¹²⁶

Swartz's solution was much like Sci-Hub, Z-Library, and other pirate sites—a calculated violation of copyright law. In his case, within the law's reach, he would have suffered the consequences. The extraterritorial pirate sites, in contrast, continue in rude health, useful for the cause in pressuring publishers to temper their demands. Yet, such approaches are no long-term solution. Offshore tax havens are being pressed into the corners by international fiscal regulation. Eventually, a similar fate will overtake pirate sites, as has already happened for those that dared challenge the music and film industries.

The 1960s and 1970s were the golden era of European national radio systems. A handful of government stations in each nation, sometimes only one or two, offered a restricted menu of listening options to captive audiences of tax- and fee-paying customers. Pirate radio stations flourished in this stifling atmosphere, operating from ships in international waters or renegade jurisdictions, such as Luxembourg. Dangerous content—more than an hour of jazz or even rock music—thus leached into the European soundscape. Today, such breaches of official channels seem merely quaint. Very little content cannot be found somewhere on the web. In an age of surfeit, pirate stations have little use. They are likely to offer the most telling historical analogy to the pirate sites' eventual fate.

Beyond the pirates and Swartz's sadly quixotic attempt to right the wrongs, others have resisted the commercial publishers. A "Cost of Knowledge" boycott of Elsevier was launched in February 2012.¹²⁷ To date, it has been joined by 19,000 researchers from many disciplines and institutions around the world. Its leader, the Cambridge mathematician Timothy Gowers, cited the exorbitant subscriptions and big-deal bundles that libraries have to swallow whole.

Large consumers of journals have also boycotted publishers. Among the notable is the University of California (UC), one of the largest research institutions, with a huge collective acquisitions budget. Starting in 2019, UC boycotted Elsevier for two years.¹²⁸ In the interim, faculty and students made do with alternative sources of articles and interlibrary loans. UC announced its aim to make its research immediately available at no cost to readers. Elsevier was holding out for a double-dip, both publishing charges and subscriptions. UC wanted the sums paid at either end in effect to count as publishing charges, eliminating subscriptions.¹²⁹

The outcome of the UC boycott testified again to the publishers' clout. In March 2021, the parties signed a four-year agreement. UC researchers' articles in Elsevier journals would be gold access. The UC libraries paid the first \$1,000 of the fee. Authors

would be responsible for the rest if they had funds, otherwise the libraries would pay. Elsevier extended a 10% or 15% discount on the usual charges.¹³⁰ For journals whose fees often amounted to \$8,000 or \$9,000, the main novelty was a modest discount, with publishers still charging as usual, and those costs now shared—for UC members—between libraries and authors. In any case, authors could opt out if they wanted to or could not afford the fee despite the \$1,000 library contribution.

UC-authored articles now appeared freely readable by anyone. That was a clear advantage. But the journals issuing them remained as before, often hybrid, and therefore still substantially locked behind paywalls. Nor did it lessen the subscription burden of hybrid journals for the UC libraries. While this outcome may have worked for a reasonably flush and research-intensive institution like UC, the signal sent to those unable to negotiate with publishers was discouraging.

The UC deal with Elsevier was an example of read-and-publish or so-called transformative agreements. These were yet another variant on gold. Sometimes subdivided into read-and-publish and publish-and-read variants, their divergent details need not detain us. They sought to grapple with the obvious unfairness imposed on libraries by hybrid journals. Libraries were required to pay twice, both publishing charges for open articles and subscriptions for the rest. In some nations, mainly European, payment streams were split. Libraries continued to pay for read access via subscriptions while funders underwrote write access. Either way, publishers cashed in at both ends, increasing overall costs.

Read-and-publish agreements sought to alleviate this imbalance, consolidating charges. University libraries would pay a lump sum allowing faculty to read a given press's content and publish a certain number of articles in its journals without separate charges.¹³¹ Renewed every so often, read-and-publish agreements were a new form of subscription that libraries were locked into—*déjà vu* all over

again.¹³² Their faculty could read all they wanted, as in the days of subscription journals, and they could publish all they wanted, as in the subscription era. What had changed, other than a smattering of articles now free for outsiders to read? Because only works authored by the pertinent faculty were flipped, not entire journals, the overall effect was modest, even for readers outside the university bubble.

In theory, if every institution did this for every publisher, all content would be both free to read for everyone and every scholar free to publish. Except that researchers without institutions, publishers without agreements, and countries whose authors remained outside would still be excluded. And read-and-publish agreements left publishers where they wanted to be, in much the same position as with subscription journals. Valid for a few years at a time, with built-in price increases, such agreements nailed fast the dominance of the gold approach and incentivized researchers to publish with the journals in question.

Nor did the agreements solve skyrocketing subscription prices. They also loaded research-intensive institutions with most of the burden in a kind of reverse Matthew effect—from those who have, much shall be demanded. British universities signed such an agreement with Wiley in 2020. Its claim that the proportion of open articles by UK researchers would increase from 28% to 85% in year one, potentially going to 100% thereafter, made sense only if all British researchers published in Wiley periodicals.¹³³

Read-and-publish agreements fit the trend of publishers directing the open-access impulse to their ends. The agreements did nothing for overstretched library budgets or the prestige hierarchy that allows sought-after journals to charge above the going rate. They have undercut whatever competition might have been hoped for if authors had taken their publishing funds to where they got most value.¹³⁴ In theory, gold access might have unleashed competition, driving down costs. Subscription models leave authors

price-insensitive, since they do not pay. But with gold access, they are price-insensitive as well, since their funders pay.¹³⁵

One suggested solution has been to make authors pay directly to publish, prompting them to consider costs, select cheaper venues and press prices.¹³⁶ Giving authors skin in the game promised to make them skinflints. The UC multipayer model described above called on grants to help cover charges, not just library budgets—or national research monies, as in Europe.¹³⁷ But insofar as third-party payers still pick up the tab, scholars remain indifferent. Since the libraries signing transformative agreements have sought only to shift funding from readers to authors without using the leverage of their boycotts to drive down overall costs, this failed to address the serials crisis.¹³⁸ But at least such agreements implicitly acknowledged the unreasonableness of publishers' double-dipping. Whatever the final amounts, the monies paid them were now to be conceptually amalgamated as one sum for both reading and publishing.¹³⁹

Latin American Success

The scientific publishers' entrenched position has hindered open access's spread. Creative destruction removes past encrustations, allowing the new to emerge. In Gerschenkron's formulation, the advantages of backwardness meant not needing to destroy before creating headroom for change. This applies to Latin America's publishing industry, which has been less developed than in Europe and the US. Many of the editions sold are imported from its colonial homelands, Spain and Portugal.¹⁴⁰ Even today, authors take publication by a Spanish press as a mark of having arrived.¹⁴¹ And many Latin American markets are dominated by Global North trade houses. Penguin Random House is the largest or second-largest venue in several markets—Chile, Argentina, Colombia, and Mexico.¹⁴² Whatever that implies for Latin cultural development, the

absence of dominant scientific publishers left a blanker slate with room for new experiments.

Although Global North academics are often unaware of it, an entire parallel universe of Hispanophone scholarship has established its own networks and institutions. Journals in the Latin world are usually brought forth by university departments rather than scholarly societies or commercial publishers. The Guadalajara International Book Fair plays a role like Frankfurt's as the largest Spanish-language market. Latindex, created in Mexico in 1997, indexes South American scientific articles.¹⁴³

Diamond open publishing, paid for by governments, has become the customary route for Latin scholarship. In 2011, gold journals made up 74% of all outlets here, compared to 7% in Europe and 5% in North America.¹⁴⁴ Almost a quarter of venues listed in the Directory of Open Access Journals is Latin. Of these, most are diamond, charging no fees.¹⁴⁵ Over 70% of academic output in Latin America is open, while no other region exceeds 20%.¹⁴⁶

Credit for this happy state of affairs goes largely to SciELO, an open publishing and indexing platform started in Brazil in 1998.¹⁴⁷ That put it four years ahead of the Budapest Initiative, which has been conventionally taken as the movement's opening salvo.¹⁴⁸ Latin America's gambit arose not from the Northern movement, but directly from the continent's own hopes of lowering dissemination barriers. With no well-resourced research library system to pay their subscriptions, conventional journals found it hard to take root.¹⁴⁹ Without library budgets pulling the scholarly communication train, commercial publishers had no reason to hijack this mode of locomotion. Like Africa skipping landlines and going straight to mobile phones, Latin America avoided the circuitous detour where libraries financed journals. They cut out the middle person, with governments directly underwriting open periodicals.

SciELO garnered government support in Brazil and expanded first to Chile, then to most other Ibero-American countries as well

as South Africa.¹⁵⁰ Intent on helping compensate for the weak representation of Latin journals in international indexes, it is now the major provider of entries in the Directory of Open Access Journals.¹⁵¹ Similar platforms emulated its success, including RedALyC, CLACSO, AmeliCA, Latindex, and LA Referencia.¹⁵²

However, being theoretically accessible on the web does not also mean being findable. A posted work may hide in plain sight, invisible to all but the most dedicated pursuers. SciELO has tackled such problems, too, seeking to supply the metadata and indexing required to make its content readily apparent to researchers worldwide.¹⁵³ It is now indexed in its own citation index, partnered with Clarivate's Web of Science, and most of its content is in SCOPUS as well.¹⁵⁴ And on Google Scholar, a more ecumenical index, it ranks among the top ten most accessed sites.¹⁵⁵ In terms of indexing, SciELO is no worse off than arXiv, the physics preprint repository. That is also unindexed by the major commercial services, yet is one of the most popular open sites globally.¹⁵⁶ SciELO is not limited to Latin America and includes the Hispanophone world more generally.¹⁵⁷ Brazil provides over five times as many articles as its closest rivals, Colombia, Mexico, and Chile. But, even so, that is only 45% of the total.¹⁵⁸

Despite its flourishing open-access repertoire, Latin America has not escaped arguments like those in the Global North. From the start, SciELO has functioned as an indexer with similar impact metrics as the developed world's.¹⁵⁹ RedALyC, more focused on the social sciences, has, in contrast, signed on to the DORA declaration, discussed in Chapter 5, to downplay the importance of numerical metrics in appointments, promotions, and funding.¹⁶⁰ It sees itself as more aligned with the predicament faced by researchers in the Global South to win recognition for their work.¹⁶¹

This is a section of [doi:10.7551/mitpress/14887.001.0001](https://doi.org/10.7551/mitpress/14887.001.0001)

Athena Unbound

Why and How Scholarly Knowledge Should Be Free for All

By: Peter Baldwin

Citation:

Athena Unbound: Why and How Scholarly Knowledge Should Be Free for All

By: Peter Baldwin

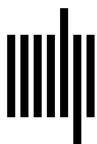
DOI: [10.7551/mitpress/14887.001.0001](https://doi.org/10.7551/mitpress/14887.001.0001)

ISBN (electronic): 9780262373968

Publisher: The MIT Press

Published: 2023

The open access edition of this book was made possible by generous funding and support from the author



The MIT Press

© 2023 Peter Baldwin

This work is subject to a Creative Commons CC-BY-NC-ND license.

Subject to such license, all rights are reserved.



The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in ITC Stone Serif Std and ITC Stone Sans Std by New Best-set Typesetters Ltd.

Library of Congress Cataloging-in-Publication Data

Names: Baldwin, Peter, 1956– author.

Title: Athena unbound : why and how scholarly knowledge should be free for all / Peter Baldwin.

Description: Cambridge, Massachusetts : The MIT Press, [2023] |

Includes bibliographical references and index.

Identifiers: LCCN 2022027103 (print) | LCCN 2022027104 (ebook) |

ISBN 9780262048002 (hardcover) | ISBN 9780262373951 (epub) |

ISBN 9780262373968 (pdf)

Subjects: LCSH: Open access publishing. | Scholarly electronic publishing.

Classification: LCC Z286.O63 B35 2023 (print) | LCC Z286.O63 (ebook) |

DDC 070.5/7973—dc23/eng/20220628

LC record available at <https://lccn.loc.gov/2022027103>

LC ebook record available at <https://lccn.loc.gov/2022027104>

10 9 8 7 6 5 4 3 2 1