

3

The Open-Access Problem

From an artisanal Romanticism, culture has moved into a postmodern aesthetic of publicly-financed collaborative creativity. Against this background, open access raises specific and sometimes technical issues. The question of who pays for scholarship and science looms above it all. Still, differences also distinguish the various tribes of content producers in terms of their workflow, how collectively they labor, and the formats they present their findings in.

Digital technologies and their vast potential are to be thanked for our current dilemma—for both good and bad. In the analog era, we had broadly achieved the access to knowledge then possible. The most important research libraries, vast shrines of learning, had been built up over the eighteenth and nineteenth centuries, often nationalizing what had once been royal collections. Deliberately burning the Library of Congress in Washington during the War of 1812, the British targeted the hoard of knowledge that was recognized as crucial even then.

Libraries eventually dotted the landscape, but not every one was open to everyone. Only faculty and students could use university libraries, with occasional visiting scholars allowed in. A few public libraries could compete in size and holdings: the New York, Boston,

and Chicago Public Libraries which were open to anyone crossing their thresholds. The British Library in London and the Bibliothèque Nationale in Paris were scholarly institutions requiring application and credentials for entry, but the determined and persistent could usually surmount such obstacles. Marx wrote much of his oeuvre at the British Library, although an impecunious, disheveled foreigner. To judge from the rules at the old Bibliothèque Nationale in the Rue de Richelieu, which forbade patrons to rest their heads on the tables, admission was sufficiently accommodating that not all visitors came intent on reading.

For the analog era, these great libraries were the state of the art. Microfilm and microfiche allowed newer institutions to copy otherwise unobtainable works. In the 1930s, American librarians sought to mass-microfilm European holdings, duplicating Old World repositories.¹ Interlibrary loans spared institutions the need to buy the long tail of rarely used materials—at the cost of transport and logistics. Beyond this, it was hard to see how the existing technology could have been much improved.

Digitality upended all this. It has spoiled us, raising the bar and leaving the old system woefully inadequate. That is what progress means—when we come to take a novelty for granted, indeed when we make it a necessity, and then drive down its cost to become a commodity and no longer a luxury. Combined with the internet, digitality expanded access in two directions: to sources beyond those available in the local library. At the same time, these sources were now usable not just in institutions but wherever readers could access the web.

There lay the rub. In the analog universe, a library could grant access to anyone it chose. Public and national libraries served a broad constituency—all readers, or at least all scholars. University libraries were generally restricted to faculty and students. The first-sale doctrine (copyright exhaustion) specifies that having

bought a book, owners (such as a library) are at liberty to do what they please—read, destroy, resell, rent, or lend it. Some things are exempted: in many countries, sound recordings, films, and sometimes music scores may not be lent.² But on the whole, analog content was as available as it could be in libraries.

Publishers have always regarded libraries ambiguously, much as they do secondhand bookstores. Both cut into primary sales by creating secondary markets, either in lends or used copies. In a publisher-dominated world, both would be outlawed. But in the interests of education and the spread of knowledge, publishers were forced to tolerate them. Books borrowed from libraries were potential sales forgone. But, at least there were physical limits to how patrons could use works—one at a time. And after several lends, books in high demand fell apart, needing replacement. Libraries did not just compete with bookstores for readers but were also among publishers' biggest customers. A *modus vivendi* emerged: so long as libraries continued to buy books, publishers would not object to seeing them lent.

Digitality undermined this analog ecosystem of secondary uses—reusing and lending works. Digital works were unlike their analog predecessors. Copies were easy to make and perfect replicas, indistinguishable from the original. A single copy could spawn infinite new ones, none ever wearing out. Lending in the analog sense of passing on something that was then no longer possessed became impossible. Readers enjoyed the promise of digital copying, while publishers feared it. In the worst case, a single digital lend of an indiscriminately copiable work could end a publisher's sales altogether.

Some publishers therefore reneged entirely on the fundamental compact between publishers and libraries. Amazon, the largest issuer of self-published books, refused to allow library loans at all.³ During the Covid epidemic, however, it was criticized for its

apostasy.⁴ In 2021, it struck a licensing agreement with the Digital Public Library of America.⁵ But even less intransigent publishers made things difficult for libraries.

Digital works, either born that way or format-shifted from the analog, were a potential delight for consumers and a nightmare for disseminators. Only by clipping their wings, making them resemble their analog predecessors, could disseminators market works without fear of piracy. The legacy publishers embraced the digital revolution by seeking to undermine most of its advantages. Fearful of piracy, they hobbled digital works to make them more like their analog equivalents—hard or impossible to copy. From this sprang digital rights management (DRM) software, which hampered digital works' free distribution and copying, limiting who could consume the works and how.

Beyond such precautions for the retail market, libraries' rights to lend digital materials were further tailored to analog precedents. After a certain number of lends, paper books wore out, spines broke, and pages dog-eared, so that eventually they had to be replaced. When publishers agreed to permit libraries to lend digital works, they often charged more than for paper copies. They also insisted that they be paid repeatedly, even though digital works did not wear out. Unsurprisingly, publishers and libraries differed on how many lends physical books could withstand before degrading—twenty-six, according to HarperCollins, many thousands in the experience of most libraries.⁶ This analogy was now carried into the digital age. With e-books, libraries have to license rights for a certain number of lends, then pay anew once that is reached.

If lending digital works was complicated, reselling them was worse. Secondary markets for digital content threw up problems. Unlike with analog items, sharing, selling, and copying digital works tend to be the same act. But copying digital works was not allowed, even under the first-sale doctrine. How, then, do you ensure that works are sold in the old-fashioned sense of a singular

copy transferred between users and not in the digital sense of being copied? With resold digital works, how can you be certain that only one copy exists and that each sale does not just multiply copies?

Transferring a digital work from one hard disk to another was not like passing along a book, where the object indisputably changed hands. In digital transfer, a new copy was created while the initial one remained untouched at its source. To sidestep this problem, software was developed to ensure that the original was deleted upon transfer. The company behind a program intended to allow the sale of secondhand digital music recordings argued that as the files were transferred to its servers and then to the buyer's hard disk, they were deleted on the seller's. What appeared to be a series of copies was, in fact, a transfer or migration. That logic was rejected when challenged. The court found that, inasmuch as a digital transfer meant copying, it was not protected by first-sale.⁷ First-sale limited the rights-holders' distribution rights, but not their reproduction rights.

In Chapter 2, we distinguished between different kinds of works and authors and how each kind raises particular problems. Works in the public domain present the fewest issues. Anyone who wants to pay can digitize them and open them to all. The Google Books project has already accomplished much, and its files are now stored with the HathiTrust.⁸ It has digitized swaths of the world's largest libraries, some 25 million books, at a cost of \$400 million.⁹ Of works not yet in the public domain, only snippets can legally be viewed. But the others can be opened up. Ever more books and periodicals published through the mid-1920s are thus freely available on the web.

Yet the serpent's tail was visible even here. Some libraries have licensed companies to digitize public-domain collections, and then sell subscriptions to other research institutions for use by their patrons.¹⁰ That creates two types of public domain. One is theoretically open to all readers, if they can find a copy they may look at.

Another is available only to those with access to paywalled databases. Of course, digitizing public-domain content is not costless and must be paid for. Privatizing the digitized copies for subscribers only is one less-than-ideal way of recouping costs.

Digitality has also allowed some publishers to discover gold where earlier they saw only dross. Even the public domain—once digitized—revealed an unanticipated value. JSTOR is a database of most Anglophone social science and humanities periodicals from their beginning, often in the nineteenth century.¹¹ Funded at first by the Mellon Foundation, it has been spun off into a separate organization. It is available by subscription to libraries and their patrons. For scholars, it is a lifesaver. Authors once spent countless hours tracking down old journal articles in the compact-shelving basements of the libraries they were lucky enough to access. Today, PDFs rain from the digital heavens like manna.

The journal backlists used to have nothing but academic value. Publishers themselves often did not own a complete run of their journals. Even the most devoted editor might not have realized that there was money to be made from yellowing copies of the *Journal of the County Louth Archaeological and Historical Society*, the *Fairy Tale Review*, *Norwegian American Studies*, the *Journal of Intersectionality*, *Jazz Education in Research and Practice*, *Pacific Coast Philology*, and so forth, in all their stultifying scholarly specificity.

JSTOR awoke publishers to the realization that some customers—research libraries—would pay for these scholarly Cinderellas in sackcloth, rounding out their collections and sparing them the bother and cost of maintaining paper copies. By digitizing the backlist, JSTOR, in effect, created its value. Once this had been done, publishers moved to capture the value for themselves. To suggestions that subscriptions be opened to libraries unable to afford the going rate, or that discounts be offered poorer institutions, or even that JSTOR be thrown open to the public *tout court*, publishers objected, fearful of losing their new income stream. Only slowly and partially

have institutions in the Global South been offered price reductions and sometimes free access.

Except for public-domain material, the digital paradise remains closed to readers without a library card for a major research library.¹² The JSTOR story is repeated for other databases that remain shut to nonsubscribers: HeinOnline, Proquest, EBSCOhost, and the like. Only occasionally have similar collections been opened up to everyone, as with Harvard Law School's Caselaw Access Project. That makes case law up to 2018 freely available, with some limitations that end in 2024 and not counting headnotes for cases after 1922.¹³

With JSTOR and its peers, no reader is worse off than in the analog era. The paper copies remain readable in the few university libraries they have always called home. But that is cold comfort to those without access to digital copies. They must still trek to the host institution, if they can gain admittance, while others read from anywhere. Digitality has thus opened up a divide. Nothing has been taken from those without digital access. But in an age when some gain wondrous new benefits, the excluded can hardly avoid feeling disadvantaged. The haves are now separated from the have-mores.¹⁴ Sometimes, digitized materials are made freely available after a period of exclusive access for research library patrons. If that becomes the norm, the divide may prove to be less of an issue than feared.¹⁵

For in-copyright works, in turn, digital editions have been a boon for publishers and readers. The public can choose between e-books and conventional editions; periodical and newspaper readers can choose between online or app versions and the usual print ones. Whether and how open access can be offered for works in copyright, we take up again later. For works that are still in copyright but out of print, the situation has been ambiguous. Digitizing them remains illegal. But since the publishers and rights-holders for out-of-print works have ceased exploiting them, no one loses by their dissemination in new formats, while readers gain. This situation is explored further in Chapter 7.

One Size Does Not Fit All

For employed authors, especially university academics, variants in approach, medium, and format among fields have created distinct cultures with different means of disseminating work. These have consequences for open access.

First, scientists are not authors in the Romantic sense of people whose works express their personalities and are thus connected to them by ineffable ties of creativity. Not that scientists are not creative, too, but they stake professional recognition on channeling truth in statements that reflect an external reality. If one scientist does not deliver a particular insight, another will. A personal element is unavoidable in expressing their ideas, but it is not the point of the work. Insofar as the personal intrudes, it renders the work less worthwhile as a truth statement.¹⁶

More mundanely, works in the humanities and many social sciences are often published as books, while the hard sciences' primary medium is the article. Yet, the social sciences, too, are shifting toward articles.¹⁷ The humanities, in turn, account for only 9% of scholarly articles.¹⁸ Simplifying somewhat, articles in the humanities are often first or secondary versions of content that later appears in books. Equally often, they are review articles, surveying a field and its literature. Unlike in the sciences, humanities periodicals also publish book reviews, conference notes, letters to the editor, and other incidental work. For most humanities scholars, journals are a sideshow. Careers are not made via articles. A researcher will publish one perhaps every several years, if that. For scientists, in contrast, periodical publication is their lifeblood. A steady stream of output in the most prestigious journals, with its associated cascade of citations, grounds a successful career.

This distinction may seem peripheral to outsiders, but much hangs on it. A book is the consummation of long labor on a subject of considerable import, usually written by one author. Books

tend to be magisterial pronouncements, a high-stakes throw of the dice by scholars who have spent years researching, developing, and refining an argument. They pursue a sustained narrative, arriving at what they intend to be definitive conclusions. At the margins of some humanities fields, mainly in literature, scholars have begun assembling their essays on similar topics, as though they collectively constitute a book.¹⁹ But on the whole, humanities researchers seem content with the significant effort and in-depth scope that anchors the monograph as their medium of choice. Academia's orientation toward the sciences threatens books. In the UK, evaluations of university departments' productivity are used to allocate research monies. These massively discount books, considering them the equivalent of a mere two articles. The evaluating bodies apparently compensate for that unfairness in other ways.²⁰

Once issued, a book may generate reviews in its first year or two. After that, if it resonates in its field, other scholars may respond to it as they, in turn, write their books, taking on board its arguments or disagreeing with its conclusions. The exchange among books takes place over decades. Even humanities articles differ from their scientific peers. Lengthier, more substantive, and written less frequently, they also have a longer shelf life. In contrast, citations of scientific articles peak three years after publication, trailing off after that. The faster a body of work grows, the quicker it ages.²¹

Scientific articles differ from books in several ways. First, they may be ongoing accounts of work as it unfolds. A lab reports on its findings frequently, sometimes weekly. It seeks to establish priority, ensuring that competing teams do not scoop it, and to keep peers up to speed. In many fields, algorithms and specialized software are involved in the writing. Sections are preformatted and literature reviews and citation compilations are partly automated. AI stands poised for authorship.²²

Second, besides serving an informational role, articles may be part of an ongoing discussion among colleagues, rather than the

definitive statement of years of reading, research, and reflection by one author found in a book. An article aimed at the interchange of ideas invites a response. It is a thrust or parry in a longer intellectual duel, one voice in a dialogue. Since the article is just one contribution in a rapidly moving exchange, its value is quickly exhausted, except as future historians of science later take an interest. In some fields, articles are much like blogs, a dialogue in text, where no one link is particularly interesting, well crafted, or well researched, and where the value lies in the act of participating.

Scientific articles have also become collective efforts, the report of many coauthors, not the fruit of solitary creators. At the start of the twentieth century, most articles had solo authors, but at its end, an average of two to seven collaborators.²³ That affects the concept of authorship. Not only do science articles often have multiple coauthors, but many scientists, with their collaborators' aid, are also wildly prolific by humanities standards. Some 8,000 physicists publish at least 72 papers annually—on average, one every five days. Life scientists are the next most fertile.²⁴ The Soviet chemist Yuri Struchkov churned out one article every three days in the 1980s.²⁵ Tasawar Hayat, professor of mathematics at Pakistan's Quaid-i-Azam University, was once the world's most fecund scholar, issuing almost 1,000 articles between 2016 and 2018. By contrast, Gregory Lip, professor of cardiovascular medicine at the University of Liverpool, coauthored a mere 548 articles during the same time.²⁶

In these fields, the concepts *article* and *author* clearly differ from their definitions in the humanities. Of course, it may be that geniuses in other fields manage to achieve weekly what takes humanities scholars years. More likely, the standards of researching, conceiving, writing, and publishing vary widely. Scientific articles are more a constant trickle of reports from the research front than considered articulations of a summated conclusion. An experiment is performed, written up, and published. The equivalent for humanities scholars would be to issue field notes as they dig in the archives

or read sources—something unlikely to happen, even if they kept such annotations.

What it takes to be an author in the sciences also differs from the humanities. In 1988, the International Committee of Medical Journal Editors set criteria for who counts as an author.²⁷ These expressly excluded merely supervising or mentoring other authors or ensuring funding for a study. In other words, scientists were known to have claimed authorship on such spindly bases. Now, according to the new criteria, to count as an author, a scientist should have done all of four things: participated in designing or conducting experiments or processing results, helped write or revise the manuscript, approved the published version, and taken responsibility for the article's contents. In no humanities field would such criteria need to be spelled out. Clinical cardiologists have miraculously become up to 80 times as prolific when promoted to full professor, department chair, or institute director—just when additional administrative burdens might have been expected to curtail their research. In medicine, one-quarter of purported authors admit to making no substantial contribution.²⁸ Clearly, the criteria of authorship in such fields are not those of others.²⁹

Such differences between book- and article-driven fields are not water-tight. The distinction between articles and books is not always clear, least of all if length is the sole criterion. Articles in law journals routinely rival books in size. Many are hundreds of pages.³⁰ So acute is the problem that in 2005, major law reviews announced they would discourage submissions longer than 70 pages and then proceeded blithely to ignore their own strictures.³¹ Book-length works in other fields are also sometimes published as articles.³² Conversely, short books are becoming a genre in themselves.³³ In online publication, size does not matter, either way. Nonetheless, such distinctions influence how various scholarly communities approach open access.

If nothing else, books and articles differ as aesthetic outputs. A single-author book is a more personal product than a group-written

article reporting the results of last week's experiment in the technical prose affected by the sciences. Humanities authors will have conceived and executed the project largely on their own. They will have revised each sentence multiple times, seeking to instill a distinct and personal voice in their prose. They will have strong opinions about seemingly trivial aesthetic choices, such as typography, page layout, margin size, and placement of the notes (end of the book or bottom of the page).

Let us linger briefly over footnotes. Social science citations are often minimal—in the text, a parenthetical author-date reference to entries in a bibliography at the end of the work, often without page numbers. The results can be cryptic and weirdly anachronistic: “Reality does not necessarily correspond to our image of it (Plato, 1978).” Conversely, law review articles fetishize the footnote. Law professors often take their footnoting obligations to extremes. Respected law school journals (*Harvard Law Review* and the like) are not peer-reviewed and are edited by second-year students, who evaluate and publish their teachers' manuscripts. There are few barriers to entry for a select group of authors, and contributors may feel pressured to demonstrate their scholarly credibility.

At most, half the average law journal article is text, while the rest consists of notes that reference each assertion, however trivial or obvious. The standard guide to legal footnoting insists that every statement outside the author's “reasoning process” be attributed. The author may not assert so much as “The sun rises in the east” without citing Copernicus, as one legal scholar puts it.³⁴ The implication is that, since legal scholars have no ideas themselves, everything they write must be attributed to its source.³⁵ In defense of legal footnoting, the convention of indicating each cited work's argument provides a useful running historiographical commentary.

In contrast to the humanities' preoccupation with presentation, aesthetic concerns are little evident in science articles. They are, in effect, bulletins, written in (at best) serviceable prose, with their

technical vocabulary incomprehensible to outsiders, and printed in dense multicolumned pages to shoehorn in maximum content. In some disciplines, periodicals have taken a back seat to websites that post preprints, the manuscripts submitted by their authors. Mathematicians and physicists already work mainly through prepublication sites, such as arXiv.

The sciences and humanities also differ in citation intensities. This, too, may seem arcane, but prestige and visibility, therefore promotion and salary, hinge on how often work is referred to by scholarly peers. The life sciences are the most assiduous citers. Since each researcher publishes hundreds of articles, eagerly flagging each other, the outcome is a tangled morass of mutual reference. Books cite other tomes much as elephants lumber through the landscape—slowly, stately, and at a remove. The reciprocal citation of innumerable articles, in contrast, resembles Brownian motion. By the law of networks, the bigger the field, the more intense the overall citation rate.

Citation intensities vary accordingly. Unsurprisingly, they also fluctuate depending on who is counting. Web of Science, Thomson Reuters's citation index, tallies only articles cited in journals from its database, mostly Anglophone. Google Scholar, in contrast, also indexes books, chapters, dissertations, theses, working papers, reports, conference papers, and articles from non-Thomson journals. The effects can be dramatic, with Web of Science ignoring entire fields. One prominent computer scientist, whose work appears mainly in conference proceedings, had over 20,000 citations in Google Scholar, a mere 240 in Web of Science.³⁶

The most influential writers have been cited over a million times. In Google Scholar, Michel Foucault ranks highest in this pantheon. Freud comes in at about half that. More unexpectedly, Marx clocks in well below, at about 300,000 citations. Sex and power now trump the means of production. Of the top ten cited scholars across all fields, ranked in this case by their h-index (an author-level metric

of productivity discussed below) as well as total citations, seven are currently active medical researchers, while only three are long-departed luminaries from other fields (Foucault, Freud, and Pierre Bourdieu, the sociologist). After that, health and engineering professors dominate until we get to Heidegger at number 24 and then again to number 47, the economist Joseph Stiglitz, the next entry from outside the hard sciences.³⁷

Citation intensities allegedly measure a work's influence. Those most referred to have had the greatest impact. A cynic might argue that scholars write less to be read than cited.³⁸ And doubtless, many logs roll as colleagues in allied fields cite each other. Why is a work cited? All publicity may be good publicity, but perhaps not all citation. The most direct route to heavy citation is to assert something plausible but wrong. Fellow scholars will be encouraged to note in order to refute, burnishing their reputation for eagle-eyed perspicacity while inadvertently boosting the vanquished offender's score.

Conversely, those with the most influential ideas may become victims of their own success. Their thoughts are channeled by others, who are then cited instead. Or they are incorporated into the common stock of knowledge without specific attribution.³⁹ What citation intensity measures is thus less the quality of the work than the attention it has garnered—whatever the motive—among other authors.⁴⁰ But in the academic world, as elsewhere, publicity is generally welcomed. As the Irish writer Brendan Behan once said, the only bad press is an obituary.

The h-index measures how many papers a scholar has published that have been cited that many times. Ten articles cited ten times each give an h-value of 10. But one article mentioned a thousand times is but a tenth of that. Authors who write one highly influential work fare worse than those with many of middling impact. The aim is to identify durable and compelling careers, not one-off flukes. Scholars in collaborative fields with multiple coauthors inherently outperform solitary worker bees.⁴¹ And naturally, the

results for book-driven disciplines, whose scholars enter the citation market only once every decade or so, are so modest that the h-index plays no role for them.⁴² In 2019, the health sciences' total citations in a sample of journals were twelve million, for history, one-hundredth of that.⁴³

Yet, such methodological and presentational differences separating the humanities from the sciences are pipsqueaks compared to the elephant in the room—funding. The sciences are underwritten by direct government or other agency and university monies. Of humanities papers in the US, under 4% receive government funding, in the UK, about 5%. For the social sciences, it is about 20%, for the life sciences, in the high 70s, and in the natural sciences, the high 60s.⁴⁴ Given ample financing, dissemination costs are but an afterthought for the sciences. Dividing total research funding by the number of articles published annually suggests that, on average, each is supported by \$290,000.⁴⁵ Compared to that, an article processing charge (APC) of \$3,000 or \$4,000 is but a rounding error.

Some humanities and social science research is also government funded, but much less so. On the whole, humanities scholarship requires less infrastructure. As the story has it, a dean imposing budget cuts demands savings from the departmental chairs. First, the sciences explain how they would trim lab costs. Then the chair of mathematics points out that they are already frugal, needing only paper, pencils, and a wastebasket. Finally, comes the philosopher, who wonders, why the wastebasket?

Yet, even though the humanities seem like a bargain compared to high-energy particle physics, they still require costly investment. The Large Hadron Collider cost \$4.75 billion. The Library of Congress's Jefferson building cost \$6.5 million in 1897. Inflation-adjusted, that is a bit over \$200 million. At \$100 per volume to buy, catalog, and store its 32 million books, add another \$3.2 billion. Globally, another half-dozen institutions have cost similar amounts. But on the whole, and in any given funding year,

biologists, chemists, and physicists require more equipment than philosophers and philologists.

Humanities scholars enjoy sabbatical funding and, if lucky, research grants to pay for travel to archives and the like. Their salaries derive from some combination of student tuition (or its equivalent in state subsidy), university endowments, and the overhead charged on science funding that deans redistribute to other fields. Humanities research is largely self-funded, insofar as scholars spend the time on it that they are not preparing or teaching classes. Modest amounts of research monies sometimes can be included here. For those few able and lucky enough to tap into an audience, there may be additional income from royalties on books or from journalism.

Scientists, in contrast, live on the public's largesse. Their labs and staff require huge budgets, derived from government and other funders, and only secondarily from their universities. Scientists often teach less than their humanities colleagues, buying out their pedagogical duties through outside funding. In the US, where a Humboldtian belief in the unity of teaching and research remains strong, that is less true than elsewhere. Distinguished scientists can still be found teaching intro courses to undergraduates. In Europe, that is rarer, especially where, as in France and Germany, research has been broadly shifted from the universities to specialized institutions, such as the Max Planck Institute and the CNRS. Here, the pedagogical function is hived off to the universities while the institute researchers take on graduate training, integrating advanced students into their labs.

Such differences affect how the fruits of research are made public. Humanities scholars prove their academic credentials over decades by publishing monographs on specialized topics. Their audience is mainly colleagues, interested amateurs, and students. As they advance up the ranks, they are often tempted to write works of synthesis aimed at a broader public. If successful, they may end up with a bestseller, possibly even narrating a TV series. Popularizers can be found in the humanities, authors without academic positions or

even credentials who write books of general interest. Many presidential biographies, indeed biographies in general, are written by such authors. That is not surprising since biographies are one of the few nonfiction genres that sell, where a diligent author can hope to make a living.⁴⁶

But overall, the scholarly and popularizing functions tend to be filled by the same researchers, though at different points in their careers. Popularization is often a phase of the humanities scholar's professional life cycle. Besides monographs, professors sometimes write books aimed at a broad audience, hoping for popularity and royalties. A few scholarly authors, such as Andrew Roberts, Frances Fitzgerald, and Daniel Goldhagen, have even managed to live off their royalties or other income, including journalism or whatever their families have squirreled away. They follow the gentlemen scholars familiar from the nineteenth century and earlier.

Scientists resist such temptations. For them to publish a book, or more generally to write for a nonscholarly audience, usually marks a career's conclusion, not—as in the humanities—its pinnacle. Popularizing science typically falls to specialized journalists trained to make otherwise impenetrable research accessible to the public. *Haute vulgarisation* (“lofty popularization”) the French call it. Rarely do top scientists write popular books. Stephen Hawking was an exception. Popularization has been baked into the university curriculum in the UK through professorial chairs intended for such writers. More a polemicist than a scholar, Richard Dawkins held one of the first such positions in the Public Understanding of Science at Oxford.

Publishing as a Profit Center

Digitality has accentuated the distinctions between science and humanities publishing. Research libraries are the biggest customers for scholarly work. In the analog era, their budgets paid for

both humanities and science output. Until recently, the distribution of resources between fields was broadly equitable. The sciences broadcast mainly in periodicals, the humanities in both books and journals. Libraries bought them all. But during the 1980s, scientific publishers began putting the system under pressure. They jacked up the cost of existing journals and issued many new ones. Academic journals' prices increased more than eightfold in the quarter-century after 1984, while inflation merely doubled.⁴⁷ Some scholarly publishers expanded enormously.

In 2009, the world had 53,000 commercially licensed academic journals, some 24,000 from large scientific publishers and 28,000 from small ones. The large houses published hundreds of distinct titles, the biggest, thousands. By offering libraries expansive portfolios of standardized-format content in package deals, these publishers gained the upper hand.⁴⁸ The commercial publishers also took over journals from scholarly societies. Earlier, the societies had often published the main journals in their fields, sold inexpensively with subscriptions usually baked into membership fees. Now, as scientific publication became big business, the for-profit houses bought these journals for their own stables.⁴⁹ Today, commercial houses issue 64% of journals, societies 30%.⁵⁰

In 1960, there were some 30 English-language economics journals, half published by nonprofit organizations. By 2000, of the 300 economics journals now in print, two-thirds were issued by commercial publishers.⁵¹ The commercial publishers both multiplied and consolidated. Fewer and bigger houses dominated the field. The five largest academic publishing houses now bring out over half of all natural science and medical research: Reed-Elsevier, Wiley-Blackwell, Springer, Taylor & Francis, and, depending on the metric, either the American Chemical Society or SAGE Publishing. The social sciences are even worse off. In 1973, the big five published one in ten articles; now, it is more than half. They bring out 71% of all psychology papers.⁵²

More and more specialized journals emerged, each becoming a must-have for libraries. Peer-reviewed periodicals increased by a third over the twentieth century's last decade.⁵³ Thanks to the academic market's peculiarities, having more journals did not lead to competition and lower prices. Quite the contrary. Rather than selecting among journals, libraries' clients insist they buy them all. Scientific publishers' goods are not substitutable. If readers cannot afford a specific article in one physics journal, they cannot just read another competing, cheaper one instead—or do without.

Yet, as an aside, in theory, academic publications are not wholly unsubstitutable. Where the subject is the invariable physical world, information not available from one source could potentially be found in reasonable approximation elsewhere. Studying *Drosophila* for a particular purpose in one lab will lead to conclusions—if the science is solid—similar to those drawn in another. But suppose the study involves others' interpretations of events, or specific content, as is often true in the humanities. Then those precise materials must be available to scholars interested in the subject. Studying Shakespeare's sonnets requires reading them and the associated secondary literature.

In any case, to avoid substitutability, scholars usually conduct exhaustive literature searches to ensure that no one else has already invented their particular wheel. That is becoming easier with digital search tools and as the scientific world adopts one lingua franca. Even so, Bradford's law of information scattering holds that searching for relevant literature outside a core of pertinent sources means trawling through exponentially growing numbers of unfamiliar journals.⁵⁴ A literature search that entirely eliminates the chance of replication may be more time-consuming than simply performing the work in the first place. Otherwise, scholars might end like Poincaré's would-be traveler, packing and repacking a suitcase for a journey never undertaken.

Nor is every kind of content unsubstitutable. No doubt, some romance novels are better than others. But readers looking for a

good cry will find that many do the trick. A similar effect probably holds for porn. Even something as seemingly academic as textbooks is substitutable. Textbooks for schools and universities are by-products of the scholarly enterprise, thus tangential to our concerns, but they soak up professorial time and energy. With prices rocketing, they, too, have become unaffordable, like scientific journals. Yet, because they are substitutable, they are subject to competition, and the market has begun providing alternatives.

When a scholar writes an article, anyone seriously studying the subject must read it, however expensive it is. No other one on the same or similar topics will do. Not so with textbooks. One decent one will serve as nicely as another. Teachers will have preferences, but their students will object when prices hit the pain threshold. The rat race of producing ever-updated editions of existing textbooks, compelling successive cohorts of students to buy new ones, can be derailed by competition from open-access or other cheaper content. Rice University's OpenStax program has developed a suite of textbooks that are free or cheap.⁵⁵ MIT has something similar, and other programs exist elsewhere.⁵⁶ From a low base, open textbooks are becoming more widely adopted.⁵⁷ Two statistics are worth noting. Overall expenditure on educational books rose from \$8.5 billion in 1999 to \$11.6 billion in 2019.⁵⁸ Meanwhile, students have responded by reducing their spending on textbooks per capita from \$700 in 2007 to \$415 in 2018.⁵⁹

For most serious academic work, however, substitutability is not an option. Journals multiply and cannot be interchanged. Nor is there downward pressure on prices. Scholars are indifferent to costs. They do not pay for subscriptions and care little how expensive they become. Only exceptionally do authors bear costs. The *Proceedings of the National Academy of Sciences* levies both article charges on writers and subscriptions on readers. As a quirk of the discipline, economics journals often have submission fees, not refundable even in case of rejection. But on the whole, the costs are shouldered

by the consumer, which is to say, by libraries. They have to buy ever more expensive journals.

Among the peculiarities of scholarly publishing is that presses receive their content largely free. Conventional periodicals are staffed by salaried writers and editors with their associated costs. Freelancers are paid. Scholarly journals' articles, in contrast, are researched, written, vetted, and largely edited by their academic authors, paid for by science funders.⁶⁰ The product is then delivered for free to the journals, ready for their finishing touches: organizing—but not paying for—peer review, some editorial work, and typesetting. In the pre-digital days, they had to print and distribute, but now they maintain sites and storage. Thanks to this lopsided division of labor, the scientific publishers' profit margins are substantial.

Other goods also enjoy a similar quasi-monopoly position, such as first-class postage or electricity. Hence they are carefully regulated.⁶¹ Not the publishers. They have exploited their hammerlock—monopoly without regulation—by increasing the subscription costs of scientific periodicals at triple the consumer price index over recent decades.⁶² As the commercial publishers tilled their new-found turf, prices skyrocketed. In the US, research library spending on periodicals during the 1980s and 1990s increased by over 200%, while it *decreased* by 21% for monographs.⁶³ In Australia, the cost of journals quintupled between 1986 and 1998, while monograph prices increased merely 50%.⁶⁴ As libraries shifted resources in their direction, buying fewer books, journals devoured the lion's share of budgets. In the mid-1980s, research collections spent half their money on books, but by 2011 three-quarters went to subscriptions.⁶⁵ In the UK, books cost 9% of research library budgets in 2008, but serials almost a quarter.⁶⁶

Everyone has their favorite example of outrageous scientific journal prices. In Elsevier's stable, *Tetrahedron Letters* costs \$20,960 per annum and the *Journal of Chromatography A*, \$22,025.⁶⁷ Before we

let indignation get the better of us, recall that these journals' prices reflect the patentable market value of their content—the chemical structure of new molecules of potential industrial use.⁶⁸ If *Harvard Business Review* published investment advice and stock tips, it, too, would charge more than its current subscription (\$120), much as no one expects a Bloomberg terminal to be cheap (\$24,000). But even apart from such outliers, average library subscription rates for chemistry journals were almost \$5,000 in 2019. By comparison, the cost of humanities journals is modest, even touchingly cheap. History journals cost \$472, and music periodicals are practically giveaways at \$332.⁶⁹ For decades, libraries have patiently paid the scientific journals' charges, skimping on acquisitions in other fields and formats as their budgets drained into the pockets of the major academic publishers.

Some humanities journals are subsidized by the professional societies that publish them; others calibrate prices at the modest level their markets will bear. Institutional subscriptions for journals such as the *American Historical Review* and the *Proceedings of the Modern Language Association* almost never exceed three figures. Despite their lower prices, humanities journals are arguably more expensive to produce. They get more editorial attention and include material not found in their scientific counterparts, such as book reviews and conference proceedings.⁷⁰ They are often less specialized and receive more submissions per published article than scientific journals, thus requiring more sorting and filtering. Their average article is longer.⁷¹

Also worth noting is that until recently, few libraries paid retail. Such eye-wateringly expensive journals have been rolled into big-deal bundles bought by libraries for lump sums. The official sticker price of each subscription has been rising, but overall, the cost per journal and per article in these bundles has dropped.⁷² Whether the price per read, download, or citation has also fallen is another question. More information may have been available, but was it used less? While the journals held by one group of US university libraries

in the 1980s and 1990s quadrupled, the number cited only doubled.⁷³ Large consumers seem to have done best from bundles. Their cost per article download is below \$1.⁷⁴ But the average was over \$4, and publishers formulating their business plans counted on libraries being willing to pay up to \$25 per download.⁷⁵ Whether that holds across the board is unclear.

Much-used journals are bundled together with obscure ones that libraries might not subscribe to if left to their own devices. Like music albums long ago, you must buy it all to get the hits you want. The publishers load the smorgasbord ever higher with intellectual herring and lutefisk. Even though libraries no longer pay à la carte, the all-you-can-eat price has risen as the assortment has broadened. Thanks to burgeoning content combined with ever-higher overall costs, acquisitions budgets continue to be drained. Hence libraries have recently sought to unbundle their big deals, a process to which we return later.

This serials crisis of ever-increasing subscription prices predated digitality and open access. Both could have helped alleviate the problem, but the scientific publishers have not allowed that. True, the costs of publishing are often underestimated and the promise of digitality in delivering content cheaply has likely been exaggerated. Digitality does vaporize many of conventional publication's expenses: printing, warehousing, distributing. It also lowers libraries' costs in receiving, cataloging, binding, storing, shelving, reshelving, preserving, and replacing paper copies. And open access eliminates expenses from the old system, such as managing copy-right and subscriptions.

But even digital publishing is not gratis. The output can be treated more or less thoroughly. But some degree of vetting, evaluating, copywriting, and editing makes it more presentable. Content must be tagged and metadata entered to make it discoverable. PDFs are serviceable, and not every text needs an e-book's bells and whistles. But optical character recognition to allow searchability is crucial and can be done at varying levels of accuracy and cost. Servers

need maintenance and software updating, and backward compatibility must be ensured after major upgrades.

We take the coddling required by paper media for granted—binding and rebinding, humidity and temperature control, not to mention intensive measures for ancient and priceless texts. And we often ignore how old texts survived only as they were upgraded to new media, copied into later works that in turn came into our hands. Papyrus lasts 200 years, parchment, many centuries.⁷⁶ Papyrus writings had to make it to parchment to survive for us. The equivalent problems for digital content are not unusually demanding. Nor are they peculiar to open access. They are an issue for any form of digital record-keeping, whether government or corporate, and must be solved by society as a whole. Digitality is marvelously efficient and scalable, but its content does not fall like manna from heaven.

Open access has brought some relief to library budgets, which no longer have to pay for certain publications. That has been counterbalanced by university administrations cutting acquisitions appropriations in tune with their new obligation to pay publication fees for their faculty's writings instead. Sometimes the trade-offs have been explicit. The Sponsoring Consortium for Open Access Publishing in Particle Physics (SCOAP³) channels funds released from library budgets by open access to pay instead for article charges.⁷⁷ In the same spirit, the Open Library of the Humanities uses contributions from libraries to sponsor open journals, thus lessening the burden of subscriptions.⁷⁸

Varieties of Open Access

Open access comes in various colors. Strictly speaking, gold open access means journals (or books) that offer immediate access upon publication to the final, typeset iteration of the author's manuscript, usually known as the version of record. How this is paid for

varies. Some journals rely on article processing charges or publishing fees paid by authors or funders. But other sources of finance are also possible. Presses that can waive publishing fees thanks to independent sources supply diamond open access, and this is technically speaking a subset of gold. Hybrid journals offer some content to subscribers while other articles have been paid for and are available to all. Read-and-publish agreements between publishers and research libraries combine payments for reading access and the cost of disseminating in a lump sum, allowing staff to publish without further charge and everyone to read the resulting articles. This, too, is a gold variant.

Green open access releases a version that is degraded and thus not in direct competition with conventional subscription publications or gold editions. This is a parallel form of green access that runs in tandem with an otherwise unreformed publication system. The degradation can be either in the quality of the version or in the delay before the version of record is available. In the latter case, access to the version of record is granted, but only after an embargo. In the first, inferior prepublication versions are offered freely when a subscription journal is issued. Immediately available variants are stunted to hobble competition with paid versions. They may convey the gist but are less polished and—lacking the typeset publications' pagination—can be cited only imprecisely. Access can be permitted on a journal's website or via aggregators like JSTOR or by authors' self-archiving.

Gold access aims to prevent readers being excluded by subscriptions or paywalls. Yet, in doing so, it solves one accessibility problem by creating a new one. Both gold and green permit anyone to read content. But the author-side fees required by gold keep out would-be contributors who cannot afford them. Paywalls become playwalls.⁷⁹ Only if other resources to allow diamond publishing are available or if largely costless green approaches via self-archiving are accepted can such authors participate. Of the two accessibility

problems—for readers and authors—only one has been resolved. Not everyone can afford article charges, much less their heftier, although more intermittent, book equivalents.

The idea that digitality would make dissemination cheaper worked only if publishers charged for actual costs. That is what funders sought, but not what they got.⁸⁰ They miscalculated on two counts. First, publishers billed according to the price structure inherited from conventional subscriptions, thus high. Second, they charged what the market would bear.⁸¹ The more prestigious the venue, the costlier. Publishing charges therefore vary widely. The Mellon Foundation calculated the sums required to sustain open journals, arriving at between \$960 and \$1,622 per article. Building in a reasonable profit, they concluded that charges for First World periodicals should lie between \$1,103 and \$2,566.⁸²

In Latin America's well-developed open-access landscape, publishing charges are reasonable, between \$200 and \$600.⁸³ In the industrialized world, the average is \$900, but top-ranked periodicals demand much more, from \$2,500 to \$5,000.⁸⁴ Squeezing its prestige and position until the pips squeak, *Nature* now charges over \$11,000.⁸⁵ The price for issuing open books (which we return to) ranges from \$5,000 to \$15,000. Even though we are early in the transition to open publishing, the sums reaped are impressive. Publishing charges for open periodicals generated almost half a billion dollars in 2017, expected to grow steeply. The leading conventional journal publishers have established their command also of open access.⁸⁶

From these publishers' vantage, whether expenses are recouped beforehand as publishing charges or afterward as subscriptions or books' wholesale prices is irrelevant so long as the overall sums remain comparable. If anything, book publishers benefit from upfront money, much as subscriptions pay journals before they deliver the product. True, when all the publisher gets for a book is the publishing charge, they miss the windfalls from unexpected bestsellers—unless many buyers splash out for physical editions

rather than just downloading free e-versions. Indeed, the term *best-seller* becomes something of a misnomer in the open-access world. We need a new word—*best-read* or something to that effect. (Royalties for Amazon's e-books already depend on the number of Kindle pages read, rather than just downloads.⁸⁷)

But publishers reap the charges also for books that turn out to be turkeys and, in the conventional model, would lose money. Authors now bear the risks, while readers enjoy the rewards. There are few economic windfalls in open publishing, only fame and reputation for authors whose books go viral. Journals, in contrast, do not suffer the same volatility as books. Subscriptions gradually rise or fall, but runaway successes or abject failures are less common.

Gold open access is like reversing a river, as in 1900 when the Chicago River was re-engineered to flow into the Mississippi rather than Lake Michigan. The direction changes, but the river remains. Green open access, in contrast, poses a bigger threat to entrenched publishing interests. But that depends on what kind of green accessibility. In the usual formulation, it changes little. It patches up the existing subscription model, supplying an overlay of good-enough, freely available versions. That provides a work-around within a largely unchanged conventional system. Indeed, this variety of green arguably allows the old system to continue, preventing the emergence of fully open journals. Subscription periodicals remain, but various forms of access are permitted. The final version is opened up after an embargo period, allowing publishers to skim off most of its commercial value. In effect, this merely accelerates what copyright was originally intended to be—monopolies of short duration followed by the public domain's timeless embrace. It returns us to the original laws' abbreviated terms—but even briefer.

How long an embargo between publication and the freely available version of record should be is hotly debated. Subscription publishers favor lengthy embargos, extending their monopoly; readers and funders prefer shorter ones. Embargos affect the sciences most,

where progress moves rapidly and oven-warm articles command a premium. The humanities and social sciences require less speed. Embargos hurt more when the topic is a promising new treatment for childhood leukemia than if a history of the dietary habits of French peasants in sixteenth-century Dordogne. The half-life of humanities and social science articles (the time to reach half an article's total downloads or citations) is much longer than in the hard sciences, especially medicine.⁸⁸ Longer embargos may therefore be justified for humanities subjects than some sciences.

Yet, green open access that relies on access to a hobbled version and/or embargos is not the only way. Rather than keeping subscription publishers in their accustomed position, depository green open access potentially undermines inherited procedures. If depositing content in the cloud is considered the equivalent of publication, new vistas open up. Digitality will then finally have contributed to making publishing affordable and widely available.

Green access usually means conventional subscription publication leavened by preprint dissemination. But if, instead, it disseminates content via online repositories, new possibilities arise. Imagine that the version posted is not just a typescript but approximates what once would have been published—copyedited, typeset, proofread, and otherwise like an e-pub. (We leave aside the selection, filtering, and attention-drawing functions of publication for later discussion.) That would increase costs compared to merely posting raw preprints. Yet, judging from the experience of preprint repositories, it would be much cheaper than open-access journals' fees. And for books, as we will see in discussing Amazon self-publication, the expenses could be a small fraction of what open-access houses demand.

Repository-based green access would not require the parallel realm of subscription journals or the existing scholarly book publishers. Indeed, it might well undermine both, though one could imagine a continued role for publishers-on-paper for trade books

and works with sufficient appeal to percolate beyond the repository. A continued demand for books-as-our-ancestors-knew-them needs fulfillment. But for academic research, conventional publishers may be superfluous. We return to repository open access in Chapter 8.

Green access as currently practiced threatens little. Parking preprints or embargoed versions of record in repositories leaves conventional subscription and gold journals free to roam. This version of green continues the inherited system while giving those readers hindered by subscriptions or paywalls a slightly inferior but fundamentally equivalent substitute. That largely solves the access issue. But it does nothing for affordability. Indeed, it worsens that problem by adding the expense of repositories to exploding subscription costs. Gold access suits scientists well. Academic publishers also prefer gold, having learned to profit from the system. But the coexistence of gold and green generates friction.

Suppose some nations follow a gold route and others green. As readers, the latter free ride off the publishing charges paid by the first, while the gold players get no more than the preprints or embargoed versions allowed in green. The UK, a research-intensive nation, has worried that striking out along a gold route would leave it shouldering disproportionate costs as it paid for both its authors' publishing fees and subscriptions to foreign journals.⁸⁹ For research-intensive countries, the old subscription model (combined with green open access) is preferable because it spreads dissemination costs to those who do more reading than writing. The same holds for fields where corporations often take subscriptions, distributing publishing costs beyond the scholarly world. And for those, such as psychology or medicine, where practitioners subscribe, shifting burdens beyond academia.

The subscription model also redistributes costs between read-intensive and research-intensive academic institutions. Libraries in liberal arts colleges and teaching-intensive institutions help

pay the costs generated by research universities. In contrast, with gold access, high-research nations and institutions bear the brunt of expenses alone.⁹⁰ With subscriptions, research-intensive libraries shoulder only an average burden, proportional to their purchasing but not to the publishing of their faculty. As these institutions now face publishing charges for their prolific scholars, they seek supplemental monies from elsewhere. Grant funding is sometimes used to pay publication costs.⁹¹ Authors might also have to participate directly, giving them skin in the game and helping contain costs.

Such tensions threaten to split the world between green and gold approaches, with gold regions curbing access to the information they have paid for.⁹² Jean-Claude Burgelman, the European Commission's open-access envoy, argued that a "geospecific access model" would twist arms in those areas not signed up for gold by locking them out.⁹³ A partially gold system threatens to treat research outputs not as public goods but as club goods—available only to the in-group that has underwritten them.⁹⁴ Such balkanized geo-walling of open access undermines hopes of solving the accessibility problem. Those outside the golden bubble would be no better off than in the days of subscriptions.

The Version of Record

Among the disputed issues between the humanities and sciences and gold and green approaches comes also the version of record. The version of record is the publishing equivalent to standard units of length and weight—meters, pounds, and so forth. Once, these were metal facsimiles housed in controlled circumstances, but they have since become linked to physical constants of the universe.⁹⁵ The version of record is the canonical variant of the work, inscribed on metaphorical stone, against which others are judged. Scientific publishers like to emphasize the importance of a version of record

since that makes readers dependent on their edition. Elsevier, for example, uses Crossmark, a system to track variants and alert readers to changes made to published content.⁹⁶ Publishers appreciate how it uses their version as the benchmark to which readers are pointed.⁹⁷

The version of record is a consequence of paper media, the idea of a singular, static, official rendition. Has that become outmoded in the digital age? Digitality introduces a certain fluidity to texts. They now permit variation, emendation, and revision. They can be linked to other web content that is not part of the text itself but forms a larger ecosystem of meaning and reference. Yet, digitality also tracks each change and variation and thus accounts for alterations. Wikipedia articles, for example, are often edited, sometimes with the changes hotly debated. Each modification can be identified, attributed, tracked, revised, and reversed if necessary. The text is both more malleable and more stable than the illusory singularity of paper print. Faculty of 1000 Research, for example, is a platform that encourages multiple, revised incarnations of articles, all linked and independently citable, thus blurring the line between versions of record and others.⁹⁸

A version of record could instead be replaced, it has been suggested, by a record of versions—a trackable catalogue raisonné of the text as it is refined, edited, and revised.⁹⁹ There is something to this. Fetishizing a static version of record ignores possibilities raised by digitization. For one thing, thanks to repositories being constantly updated, the prepublication manuscript and the version of record are increasingly indistinguishable.¹⁰⁰ With an article revised, accepted, and published by a journal, the author re-deposits the updated manuscript in the depository so that it now contains an edition identical to the version of record. Some publishers specifically make authors promise not to update their preprints, however little cricket that may be.¹⁰¹ But that is not all. As authors continue to revise and update works, depositing ever-newer editions in the repository, the version of record, supposedly etched in stone, begins

to look more like yesterday's newspaper.¹⁰² Given a choice between the canonical version of yore or the latest, which would you prefer?

Suggesting a move from a version of record to a record of versions is perhaps more than just clever wordplay. Nonetheless, it also ignores the role a stable edition still plays in some fields. It is sometimes claimed that the prepublication version, usually the author's accepted manuscript, is not inferior to the version of record.¹⁰³ That is whistling in the dark. In many ways, it obviously is. For one thing, it is not properly printed. For scientific journals, with their triple columns and minuscule typeface, that may be an advantage. Scientific articles are one of the few instances where publication diminishes readability compared to typescript.

But with their adjustable font sizes and correspondingly fluid—sometimes nonexistent—page numbers, e-publications have raised a relatively trivial problem. How does one quote a specific passage without a stable page number? The hard sciences, publishing brief articles, are not reliant on page numbers to indicate sources of information. The social sciences use in-line references to back-of-the-work bibliographies. They often do not give precise locations, even for quotations and even from lengthy books. The results, as noted, are often imprecise and weirdly anachronistic. But most humanities subjects, usually heavily text-based, must refer to a stable version of the source, accurately indicating the origin of a quotation, reference, or idea. For them, a version of record is indispensable. Digitality's ability to search text provides a solution only for precise quotations, not for paraphrases or references to ideas.

Both e-publications and preprints undercut scholars' reference to stable versions of record. If the freely available version is not identical to the published one, readers can still plumb the gist of the argument, even if quoting chapter and verse is impeded. We could also sidestep this problem by adopting a Wittgensteinian numbering of paragraphs rather than pages—stable across different editions. New editions of classic philosophy texts—like Kant's—usually provide

the pagination of the original or canonical edition in addition to their own, facilitating citation across editions and translations. And of course, the Bible showed the way, organizing references layout-independent of its countless editions. Early church councils organized its division into books, but the numbering of chapters was standardized only by Stephen Langton in 1203 in the Dominican concordance. Verse numbering followed in the sixteenth century.¹⁰⁴

More broadly, being able to cite specific pages (or paragraphs) raises the broader problem of needing a stable version to refer to. In fields that exhaustively parse subtleties of meaning and rely on precise quotation of locatable passages, versions of record remain necessary. The transcendental deduction is one paragraph in Kant's *Critique of Pure Reason* that answers Hume's problem of moving from correlation to causation. Without an agreed-upon text to parse, the scores of books written on the topic would have faced even knottier problems. Scholars of literature confront similar issues. However quaint, a stable referent is not entirely outmoded in the digital age.

Equally crucial: the digital record's fluidity invites retrospective fiddling, undermining faith in any stable referent. Digitality need not be indeterminate. Blockchain techniques or other less complicated means of establishing invariant records allow data to be chiseled in electronic stone.¹⁰⁵ But without some such assurance of invariance, digital texts may be suspect. Dominic Cummings, Boris Johnson's Rasputin-lite special advisor, sought to buff his foresightedness by altering past texts from 2019, making them appear to warn of a Covid-like pandemic soon to befall Britain.¹⁰⁶ He was outed when the original versions were excavated from the Internet Archive's Wayback Machine. Similarly, two historians responded to criticism of their claims that the discipline was losing its influence in the public arena by covertly doctoring their text in its open-access version to present less of a target. Only after sharp-eyed observers brought the alterations to the publisher's attention were they explicitly indicated and acknowledged.¹⁰⁷

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