Pixel Dust: Illusions of Innovation in Scholarly Publishing

By Johanna Drucker



JANUARY 16, 2014

PRONOUNCEMENTS ABOUT THE CRISES in academic publishing have been with us for decades, and the underlying culprits have remained basically unchanged: shrinking library budgets, more supply (books) than demand



(buyers), increased production costs, continued pressure on university presses to play an unpaid but crucial role in the tenure process, diminished subventions from sponsoring institutions — and more recently, the exorbitant and skyrocketing subscription costs for scientific journals.[1] The proposed cure for this crisis is innovation in the form of digital platforms and novelties.

Why should this crisis matter to a nonacademic? Because academe provides a gold standard of scholarship. Because its ideas filter down, sometimes through unsuspecting channels, and stimulate thought in virtually every field of human endeavor. Because the fate of the humanities is being influenced by a campaign of misinformation (numbers of students enrolled in humanities courses are actually up, for instance, not down). Because the hyped myths about digital publishing are far from reality. Because much of the knowledge currently being produced is at risk of being locked into silos and kept behind firewalls. Because large publicly supported projects (e.g., the Digital Public Library of America) need to succeed to keep intellectual life and public discourse vital.

To make sure humanities scholarship thrives, it is crucial that we cut through the fog of pixel dust–induced illusion to the practical realities of what digital technology offers to scholarship. Among the prevailing misconceptions about digital production of any kind is that it is cheap, permanent yet somehow immaterial, and that it is done by “machines” — that is, with little human labor. We could add to this

another pervasive two-part misperception, that "everything" is digitized and that everything digital is available on Google. Each of these views is profoundly inaccurate. Costs of production and maintenance (or, to use the current grant-required buzzword, *sustainability*) are much greater with digital objects than print. Every aspect of the old-school publishing work cycle — acquisition (highly skilled and highly valued/paid labor), editing (ditto), reviewing, fact-checking, design, production, promotion, and distribution (all ditto) — remains in place in the digital environment. The only change is in the form of the final production, which becomes a matter of servers, licenses, files, delivery, and platform-specific or platform-agnostic design instead of presses, paper, binding, and so on. The fact that the print object is removed means the single obvious revenue-producing part of the work cycle is eliminated, replaced with a dubious business model of digital sales that as yet doesn't seem to work well for most authors, and even less well for scholarly monographs.

Not cheap, either, are the costs of ongoing maintenance. A book once finished sits on the shelf, opens without electricity or upgrades to its operating system or to the environment in which it is stored. Five hundred years from now? The complexly layered and interdependent material conditions that support digital storage, access, and use have an unprecedented rate of obsolescence, and half a millennium is unlikely to extend the shelf life of files that aren't backward compatible across two or three versions of software upgrades. As for permanent — the fact is that every use of a file degrades and changes it, that "bit rot" sets in as soon as a file is made, and that no two copies of any file are ever precisely the same as another. In the archival

community, the common wisdom is that "Lots of Copies Keep Stuff Safe," (LOCKSS) — a formulation premised on the recognition that permanence is as elusive in the digital world as in the mortal coil. As for human labor, the serious business of digital work (the task of analysis and interpretation that add value in a digital environment) is extremely intensive, demanding, and skilled. It can't be outsourced, and it can't be automated — that is, if by "it" we mean tasks such as adding metadata, "mark-up" or other tagging that cross-references the contents of texts, images, sound, and video files, and which add intellectual value and functionality. As for Google, the firewalls, licensing agreements, hidden collections, and un-indexed resources all constitute subjects for treatises about what does not get accessed through that or any other search engine.

These basic myths are merely the fundamentals, the first principles on which a host of other delusions are being built. In the mad rush to throw the humanities into the digital breach, the current trend is to invoke the general rubric of "design innovation" — a confusion of novelties. The much-touted "nonlinear" approach to composition is a choose-your-own-adventure model for grown-ups, and the desktop embedding of multimedia encourages all manner of fantasies about crowdsourced, participatory knowledge generation that would essentially de-professionalize knowledge production: "getting students to do it," and other naiveties, such as marketing data sets or primary materials, or staging debates among prominent academics using a combination of antique road show and reality show techniques, or fantasies of free online textbooks served up to hordes of formerly disenfranchised youth — every bit of the imaginary bouillabaisse except for the crowning glory, a

magnificent engagement with what can only be called *magic-onomics*: a business model in which publishing thrives without a revenue stream.

All of the terms on which books —with their melded content and object-ness — worked as commodities (subject to terms of first sale, resale, circulation, and so on) have had to be rethought in a new and even more complicated system of licenses. Is a file stored on a publisher's server and merely accessed by a link? Accessed by one reader at a time from any institution with a subscription? By multiple readers? Is it available for download but programmed to self-destruct after a reading or lock after a certain date? Is it guaranteed to be available in five, 10, 15 years as was the case with bound journals and physical books? These are all issues librarians are trying to work out with publishers. But the agreements being pounded out are all hazard calculations, all built around assumptions of piratical abuse.

¤

Be that as it may, digital platforms will be part of distribution from now on. Hard, serious, life-long dedication to scholarship, the actual professional work of experts in a field, will remain at the center of knowledge production. The hype around digital alternatives to print sometimes seems like a distraction to blind us from the need to address that reality, as if it will take care of itself if we just find enough apps and devices and platforms to play with. The technical issues involved are resolvable, the business questions harder, and this is why not just the publication outcomes of its practitioners, but also the system that supports humanistic knowledge itself is at risk.

The stumpers for innovation like to describe books as "linear" and argue that analog formats constrain the reader. Since the days of the CD-ROM and hypertext, we've been sold, as a great advantage of the digital, an explosion of the design constraints of the conventional book. The problem with this story is that it is historically wrong and theoretically misguided. Texts were never linear; their structures of repetition and refrain create meaning that builds across the temporal and spatial unfolding in their composition. What seems like the strict linearity of the physical book — or of the unrolling scroll of antiquity — should not obscure the multivalent properties of a text. The *Song of Songs*, the *Bhagavad Gita*, *Gilgamesh*, and the *Iliad* are hardly linear. Their poetical and philosophical complexity is structured into their language, not their letter-by-letter sequence. The invention of the codex in the third and fourth centuries of the Common Era brought a random-access device into active use, and the spatial as well as temporal dimensions of the form make for multiple paths and points of entry. The medieval European codex was a complex apparatus specifically engineered to support discontinuous reading, not linear in the least.[2] Those who advocate electronic formats as the answer to the linearity of print need to realize that a screen is a flattened space, with limited real estate, whose spatial dimensions are diminished compared to the book, not expanded, in spite of the illusion of endless connectivity to all and every reference, association, or source. Texts based on the modular structure of an underlying database have required elaborate acrobatics to produce readable texts. The really innovative aspects of digital work are under the hood and in the nitty-gritty of what we call "back-end" work as much as in the front-end

user experience, the multiple pathways or complicated routes through a site or collection of materials. In fact the most exciting and innovative aspects of digital presentation are the ways in which structured data — texts with humanly embedded organization — can be searched and analyzed, not the way they are presented.

But the search for innovation is everywhere, and the wow factor of newly hatched visualizations or diagrams often grabs headlines, making the connection between familiar texts and digital displays seem like the next best thing. The inventory of current solutions includes all kinds of work that really have little to do with scholarship, but instead crowdsource, crowd-please, entertain, blur the lines, and prep material for consumption. What matters most in scholarly publishing is scholarship. With this comes peer review, selection and elimination, editing, branding, vetting, curating, and pulling new content into being in publishable form. Only then do these ancillary solutions come into play.

The most awkward aspect of the hunt for novel solutions to economic and cultural problems is that the cures are more expensive and difficult than the traditional methods. At a recent conference, for instance, I listened to peers suggest that multimedia productions of digital video to produce long-form arguments would be a way to replace and/or supplement traditional texts. Interviews and sound files were also promoted as novel inserts into the conventional page, now migrated onto the screen. But what could be more linear than these time-based media? Writing and thinking in database formats poses other challenges, intellectual as well as technical. Trying to design arguments in modular chunks, with many and varied links between the

elements, requires a whole new rhetoric. For the viewer, in turn, these multiply branching artifacts pose other difficulties. We have yet to begin to organize ways of reading these works, chart a path, know how much of a project or argument one has encountered or missed.

But we (and I count myself among the guilty) have all jumped onto the innovation bandwagon. The crowdsourcing frenzy alone is enough to cause uneasiness — the costs of editing, fact-checking, keeping spam bots and hackers at bay is the intellectual equivalent of being a traffic cop in Midtown Manhattan on a day when a major intersection signal is out of order from a water main break. The overhead that would be required to maintain the flow of information in a massive crowdsourced project is mind-boggling, a kind of 24-7 attention to a gazillion details. A handful of projects, like the Jeremy Bentham transcription, or the New York Public Library's menu decipherment, were expertly designed, highly constrained, and made effective use of contributions by the public. The redesign of scholarship to allow for participation is an enormous undertaking, not yet much beyond prototypes, none of which have yet proved fully viable except the wiki. And the difference between a book chapter that lays out a well informed and studied discussion of new research and a set of guided activities for the acquisition of that knowledge is the difference between research and pedagogy. They perform different roles.

A few facts about academic publishing might be refreshing as an antidote here. Between the 1960s and the 1990s, the number of university presses increased dramatically (from 60 in 1960 to 96 in 2003), and so did the output (on average, from 41 to 88 titles per press per year), until things began to

flatten (steady at 8,500-9,000 titles a year) and droop (sales went from 1,500 a title to 200 on average) as collateral damage in the boom and bust cycles of the education industry.[3] In the larger economic ecology, devaluation of higher education has become a rote justification for eroding support for perceived frivolities — like the study of the classics, critical race theory, literary studies, poetics, or historiography. The larger cause of this is the retreat (in the United States) from the leveling of playing fields that was the hallmark of post-WWII education initiatives, not least of which was the G.I. Bill.

But the history of the humanities and the modern university is not my focus here, even if recognizing the larger frameworks within which the recent "crisis" of scholarly publishing is being staged demonstrates that the implications go beyond mere tempest-in-a-teapot concerns for the professoriat (who look more and more like an on-the-verge-of-extinction species stranded on ice floes of an unsustainable environment) or their ilk. What risks being lost in the current climate of change is the legacy of humanistic thought and the methodological richness it brings to public discourse. If the humanities lose their cultural authority in the process of becoming digital, becoming managed quanta, or superficial entertainment, then what is the point? Citizen science, like public humanities, is a fine idea, but no substitute for work whose validity is guaranteed by professionals. The humanities have a role to play in demonstrating that knowledge is historically and culturally situated, which though it might be part of the *method* of science, is the *substance* of humanistic work, which is always consciously embedded in the social, psychological, and cultural conditions that

produce communities of discourse through exchange. Novelty and insight are effects of reading, not byproducts or packaging.

In 2003, the eminent American Council of Learned Societies published a still-definitive paper co-authored (in four discrete sections) by individuals who were at the forefront of digital initiatives. They had been asked, explicitly, to address "the problem" of academic publishing. Their analysis of the crisis matched the outline above, and their envisioned solutions sketched an array of reasoned propositions.[4] These were very sophisticated folks, and their observations far from foolish. John Unsworth suggested alternatives to the monograph in the form of themed research collections (to solve the tenure problem). Lynne Withey sought a hybrid between journal and book publishing, electronic and print (to solve the work-flow problem). Cathy Davidson described the complexity of issues and ways costs of publishing might be distributed (to solve part of the cost problem). Carlos Alonso discussed subvention possibilities and professional repositories (to solve the distribution problem).

The immediate precedent for the ACLS piece was an article by Stephen Greenblatt, whose 2002 "Call for Action," written on behalf of the Modern Language Association's Executive Council, generated considerable response because of the status of its author and the proposal to rethink the grounds of tenure on the basis of articles, not books.[5] But making the deck chairs smaller would not have saved the Titanic. What separated the ACLS authors from Greenblatt was that they were part of an emerging "digital humanities" cohort, and all saw various forms of electronic publication as at least a partial solution to the pressures faced by

university presses. The intervening decade has intensified the problems, exacerbated the apocalyptic rhetoric, and, above all, raised the level of hype touting digital technologies as a panacea. University presses are facing real problems. They were supported by a combination of subventions from their institutions and revenue from book sales, with a balanced calibration of best-selling reference or crossover titles and works of scholarship with limited but defined audiences. This business model worked pretty well until recently.

Journals and monographs are distinct streams of publication — different in nearly every respect (cost, audience, frequency, editorial responsibility). Scientific journals have caused the greatest uproar and generated pushback. They are considered essential for the timely dissemination of findings, and publish work that is often supported by grants from government agencies, require a subvention from their authors, use the more or less free labor of scholars to vet and peer review materials, and then sell back the published content at an exorbitant rate. Library budgets are devastated by the high costs ($40,000/year for a subscription to a single journal, for instance), and the conspicuous appearance of market-cornering and price-fixing by Elsevier, Wiley, and the other members of the top six publishers in this field have come under serious scrutiny and prompted orchestrated resistance (including a boycott led by MIT mathematicians).[6] The details of licensing agreements, firewalls, and use agreements keep library legal staffs busy, since the terms of traditional copyright and intellectual property don't work well for digital materials. These are familiar stories, and more, much more, can and has been said. The move toward Open Access has prompted

the University of California to institute a policy whereby faculty are encouraged to make their scholarship available through a deposit to the digital library. Still far from fully public, the policy at least cracks the vault open a bit.

Not only are journals and books different, but also the humanities and sciences have distinct publishing cultures. University presses publish between 8,000 and 9,000 book titles a year, and, if in the same period, from approximately 50,000 newly minted PhDs produced in North America, even a third are in the humanities or related fields where a monograph is the requirement for tenure, publishing can't keep up[7] Dissemination of humanities scholarship is not simply about tenure and promotion, or it would be just a guild matter. Humanists have a serious engagement with the cultural record and "all that it means to be human." Academics don't make money from their books — in fact, they subsidize them with permission costs, reproduction rights, and their labor (paid for teaching, promoted for research). While some reports suggest that the number of titles published per year is relatively flat, rather than decreasing, areas in which lists are being eliminated show the vulnerability of the humanities (part of a larger trend to diminish the importance of studies that once formed the central core of undergraduate education).

Lists that focus on literary studies, philosophy, foreign language studies, poetry and poetics have been cut, slaughtered in full public view, sacrificial lambs in the scholarly publishing market, as if they might make clear the dramatic sacrifice required to keep university presses going. Going for what? And how? To what end and for what audiences? As with other discussions of the costs of higher

education in current Euro-American culture, the complexities of publishing need to be seen as an ecological system, not a set of discrete decisions about specific practices divorced from the greater political, ideological, and economically justified (at least on the surface) conditions of which they are a part. Or, to put it very simply, cutting humanities lists is like keeping dessert from a stepchild at the family table — it saves very little money and causes lots of distress. But humanities are not a luxury, and to show that they have a substantive contribution to make to the world we live in, we need to demonstrate their relevance to policy, politics, daily life, and business, not just rehash the same old bromides about critical thinking and imaginative life. The vitality of humanities is the lifeblood of culture, its resounding connection to all that is human makes us who and what we are. The preservation of cultural ecologies is akin to preserving ecologies in the natural world, it is, in fact, the human part of them. The humanities are us. Their survival is our survival.

And we can't rely on a purely technological salvation, building houses on the shifting sands of innovative digital platforms, with all the attendant myths and misconceptions. Which aspects of digital publishing are actually promising, useful, and/or usefully innovative for the near and long term? The costs of innovation are high; the illusions promoted are many. The horizon is filled with multiple mirages, of magically conjured multimodal and variably mediated artifacts, where crowdsourced, cloud-based productions synthesize amateur and professional expertise into a happy harmony of intellectual industry that "democratizes" scholarship and popularizes academic practice in ways that might fulfill the long-dead John

Dewey's most ambitious yearnings for a fully participatory community of learning. Among the host of current projects, some are silicon snake oil, some are the solid foundation of futures ahead, and many define the spectrum between reality and illusion. Learning to distinguish among them is crucial.

On a final note and theme, the boundaries between publishers and libraries are being debated in university environments as part of this larger set of conversations. Libraries have a crucial role to play in making data and projects and other scholarly materials available. But they cannot do the expensive and time-consuming work of a publisher. Finding an alternative to Google (or any other private company) as the default support system for access to the broad cultural inventory is essential. The most exciting and potentially transformative initiative in this realm right now might be the DPLA, the Digital Public Library of America.[8] Meant as an integrated portal to online materials stored in academic and public repositories, the project was the brainchild of Robert Darnton and a team of gifted and dedicated scholars, librarians, academics, and other cultural visionaries.[9] Not a publishing venture, it is envisioned as a fully public, completely integrated online library with access to the highest quality of ongoing knowledge production. Currently being developed under the stewardship of Dan Cohen, who during his tenure at George Mason University's Roy Rosenzweig Center for History and New Media was responsible for some of the most reliable noncommercial platforms for online cultural work — Omeka and Zotero — the DPLA is a brilliant and ambitious idea that will only succeed if it captures imaginations in the public, private, and philanthropic sectors. Making scholarship freely and

publicly available for access and use is another matter, one that harks back to the public library movements and the days of Andrew Carnegie and his admirable, successful, and transformative project. Between 1883 and 1929, Carnegie's philanthropic support built 1,689 libraries in the United States (and more than 800 elsewhere) and endowed them with funds for their operation.[10] He used more than $60 million dollars to create a public legacy that was unparalleled. Carnegie was not a saint, far from it, but his model of philanthropic activity deserves emulation among current moguls of the digital world. Humanities scholarship, scientific publishing, the preservation of cultural knowledge and promotion of its activity have to be subsidized. In a world where taxation is demonized and corporate wealth defines the one percent, the responsibility for this work lies squarely with those who have the means individually or collectively. This also suggests advocating tax revenues in support of intellectual infrastructure to serve the public interest.

In the end, no special effects, dazzling displays, augmented realities, or multimodal cross-platform designs substitute for content. Scholarship, good scholarship, the work of a lifetime commitment to working in a field — mapping its references, arguments, scholars, sources, and terrain of discourse — has no substitute. The difference between amateur and professional remains significant. Novelty and special effects are smoke-and-mirror attempts to conceal the fault lines in the cultural infrastructure, a futile attempt at a technical solution to a social problem. Design innovation is necessary, yes, and every new medium develops its aesthetic capacities over time. But we can't design ourselves out of the responsibility for supporting the humanities, or for making

clear the importance of their forms of knowledge to our evolving culture.

---

[1] Carl Schweser, "The Economics of Academic Publishing," *The Journal of Economic Education*, Vol. 14, No. 1 (Winter, 1983), pp. 60-64; Stephen Boyd and Andrew Herkovic, Crisis in Scholarly Publishing, 18 May 1999; Also here.

[2] Peter Stallybrass, "Books and Scrolls: Navigating the Bible," *Books and Readers in Early Modern England*, Jenny Anderson, ed., (Philadelphia: University of Pennsylvania Press, 2002); pp. 42-79.

[3] Lynne Withey, p. 46; figures on titles from Mellon President's report. Carlos J. Alonso, Cathy N. Davidson, John M. Unsworth, and Lynne Withey, "Crises and Opportunities; The Futures of Scholarly Publishing," ACLS Occasional Paper #57, 2003; also New York Times and AAUP.

[4] Carlos J. Alonso, et. al, *ibid*; 2003 outlined the perceived and actual components of the situation in terms that echoed themes from earlier articles by Stephen Greenblatt and others. The trends outlined by Alonso et al. are the same 10 years later — drop in monograph publishing, pressure for publishing for tenure in academic positions, falling library subscriptions and budgets, high costs of publishing.

[5] Stephen Greenblatt, "Call for Action on Problems in

Scholarly Book Publishing: A Special Letter from Stephen Greenblatt." 44 Modern Language Association. May 28, 2002.

[6] MIT Libraries News; also The Guardian.

[7] See the American Association of University Presses report; also Mellon Scholarly Publishing Initiative.

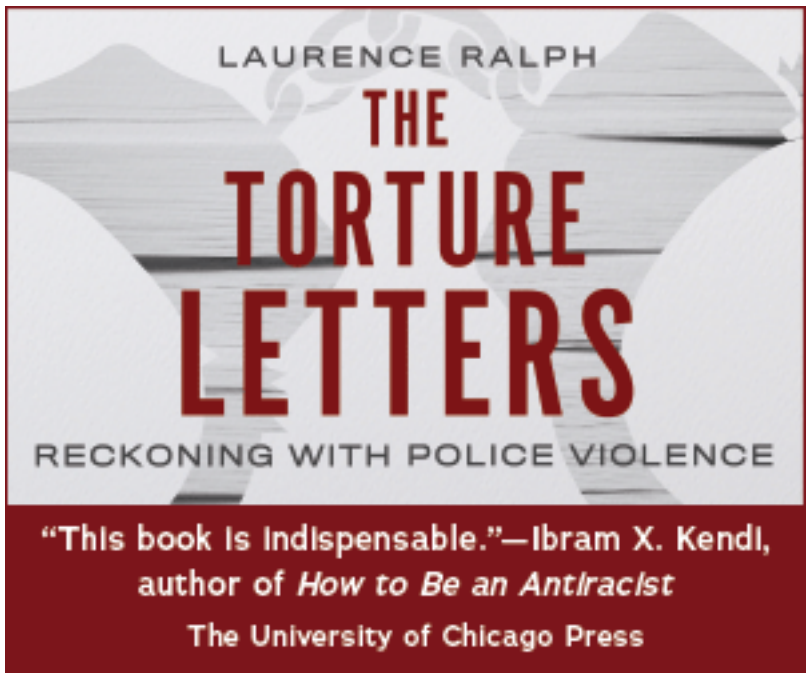
[8] Robert Darnton, "The National Digital Public Library is Launched!" April 25, 2013.

[9] JISC, Inform, "Digital Public Library of America,"

[10] Andrew Carnegie.

¤

Johanna Drucker is the Breslauer Professor of Bibliographical Studies in the Department of Information Studies at UCLA.



7 Comments Los Angeles Review of Books Disqus' Privacy Policy 1 Login

Recommend Tweet Share Sort by Best

Join the discussion…

LOG IN WITH

OR SIGN UP WITH DISQUS ?

Name

Screwtape • 6 years ago • edited

Digital versions don't degrade. That's what makes them digital and not analog. I mean, magnetic tape degrades (but it's not usually used for digital storage anyway) and I guess the plastic in DVDs will age quicker if you don't store them right, but if you back up your data often enough you will experience literally no data degredation, certainly not by making copies. Entropy doesn't apply to math.

"[T]he fact is that every use of a file degrades and changes it, that “bit rot” sets in as soon as a file is made, and that no two copies of any file are ever precisely the same as another." This is utterly wrong.

"Bit rot" refers to the degradation of storage media, not the actual bits—the term is about 40% a joke. No two copies of a vinyl record are exactly the same, but a digital conversion makes the physical grooves in, say, your LP of "Exile on Main Street" into binary. Granted, there's always some degradation the first time something is converted to digital (which is why early CDs sound so terrible), but that just means that what was a sine wave is now a set of staircases with extremely tiny steps. The equation for "Rocks Off" will be the same in 2,000 years as it is today, provided the storage media can be adequately preserved (LOCKSS indeed).

Books LITERALLY rot.

2 | • Reply • Share ›

Jason ➚ Screwtape • 6 years ago

"Digital versions don't degrade." This is actually not true which you sort of agree with. But really, if we were to get philosophical about it, a digital book is not just the code but the device upon which you are reading, the network from which you receive the code and the data center providing the feed. All will fail at some point sooner than later and must be fastidiously mainained. The average useful life of a hard drive in a data center is just 4 years. And every line of code must be managed or else it will become unstructured, risking being essentially lost. Storage and maintenance costs of data are quite high. Storing and babying a terabyte of data from 1988-1998 cost an estimated $500K. And just as the costs of software and hardware have plummeted in the past 5 years, the costs of architecture, security and employee compensation have skyrocketed. If we were to magically eliminate every

computer, take the savings and put them into paper publishing, I imagine we could supply a