

WHY THERE'S NO "FREE LUNCH" ON THE INTERNET:

Two stories from academe: Take Two

Stanley N. Katz

Speech to the annual meeting of the Association of American University Presses

Le Centre Sheraton, Montreal

28 June 2008

On February 18th of this year *The Daily Princetonian*, our student newspaper, ran a lead editorial entitled "Make 'public goods' public," which argued that Princeton and other universities are "the producers of 'public goods' in the form of papers, books, reports, classes and lectures. While some of these are highly technical and of benefit mostly to experts . . . , other goods have direct value to the general public." The editorial went on to complain that "few of our eminent faculty make their papers available except in costly scholarly journals," and suggested that we participate in *iTunesU* as "a good way to organize and disseminate Princeton content. This, our students thought, would help our University fulfill Woodrow Wilson's notion of "Princeton in the nation's service."

Well, I would like to complain that our student journalists were being daringly original, but I fear that they were in fact Harvard wannabes. The previous week that school in Cambridge (my *alma mater*) had announced that its faculty had voted on February 12 (Lincoln's birthday, no less) to "give the University a worldwide license to make each faculty member's scholarly articles available and to exercise the copyright in the articles. Provided that the articles are not sold for a profit." I swear that this is the text post on the *Harvard Gazette Online*, although it is clear that no IP lawyer had ever vetted this incomprehensible text. Still, when I read it I thought I could guess what Harvard had in mind.

It turned out that the genius behind this effort was a Professor of Computer Science named Stuart M. Shieber. Are we surprised that in most universities it is the computer scientists who are prepared to give up print and copyright? Not in my experience, at Princeton and elsewhere, since computer scientists consider acceptance of papers for important conferences, not publication in journals, to be their form of peer review. Shieber was quoted as saying the “This is a large and very important step for scholars throughout the country. It should be a very powerful message to the academic community that we want and should have more control over how our work is used and disseminated.” Harvard’s Provost was similarly exuberant: “Today’s action . . . will promote free and open access to significant, ongoing research. It is a first step in the creation of an open-access environment for current research that may one day provide the widest possible dissemination of Harvard’s distinguished Faculties’ work.” And Harvard’s new Librarian, my former Princeton History Department colleague Bob Darnton, an expert on the early modern history of the book, exulted that “By working, as individual faculties and together as a single University, we can all promote the free communication of knowledge.” Darnton was quoted in the *New York Times* as saying that “In place of a closed, privileged and costly system, it will help open up the world of learning to everyone who wants to learn,” “It will be a first step toward freeing scholarship from the stranglehold of commercial publishers by making it freely available on our own university repository.” Probably the most worrisome Harvard development was its recent announcement of the appointment of Prof. Shieber to run its Office of Scholarly Communication: “HUL created the OSC to enable individual faculty members to distribute their scholarly writings in keeping with the University’s long-standing policy that “when entering into agreements for the publication and distribution of copyrighted materials individuals will make arrangements that best serve the public interest.”

“Harvard in the public’s interest”? Bravo, fellow Harvards? Perhaps. The proposal, even five months later, has still not been clearly articulated, and the kindest thing I can say about is that I do not

understand how it will work. Or to put it differently, I do not think it can or should work. The gimmick that Prof. Shieber and his friends have resorted to in order to gain acceptance for the proposal is that it has an “opt out” feature – any Harvard faculty member can refuse to have his article posted on the public access website. This has a sort of Google Book 2.0 feel to it, but it is not clear (unlike Google Book 1.0) how the material gets posted in the first place. This requires a judgment as to when an “article” has been completed and is eligible for posting – presumably only the author will be able to make that determination (since I am assuming that some Hal-like central Harvard computer is not going to capture and post all writings more than, say, 5000 words long). But such “articles” are not even what we normally call preprints, which require the peer review that scholarly journals provide. It sounds as if Harvard will post anything a professor calls an article without subjecting it to prepublication critique of any kind. Garbage in, garbage out, even at Harvard. Many of us post unpublished papers on our web pages – I certainly do, and this talk will be posted on my website, clearly labeled as “unpublished”). It will be freely available through Google and other search engines. So at one level all Harvard is doing is requiring every faculty member to be like me, and have an up to date website.

But the stated rationale for the Harvard open access website is to do an end-around what Shieber, Darnton and their colleagues consider the stranglehold that high-priced commercial journal publishers have over the dissemination of scholarly information – no names are used, but we all know they are talking about Elsevier, John Wiley and other primarily STM publishers. The unspoken premise of the Harvard plan is that such journals (and publishers) add no value to the articles created by Harvard faculty members. If that is correct, Harvard has a point. But if the journals do add value of various kinds, then one has to ask whether scholarship is being promoted or hindered by the Harvard website. I assume that everyone in this audience will think that publishers do add value, and ought to be compensated for it. I think so, too. But of course that does not settle the issue, for if for various reasons

the publishers are charging unfairly high, oligopolistic, prices for their journals, then university libraries have a valid complaint.

But, as some Harvard faculty and other commentators have pointed out, many publishers are far smaller and poorer than Elsevier and Wiley – some are non-profits (especially academic presses and learned societies) and others are much smaller-scale commercial publishers. Open access websites like Harvard's, if they come to be commonplace, may well put smaller publishers out of business – and a special concern for academics like me is the threat to university presses and the publications of scholarly societies. They provide services not readily available at individual universities, and essential to the creation and dissemination of scholarship. There is a big difference between the *Journal of the History of Ideas* and *Nature* or *Cell*. One size does not fit all in academic publishing. One way to characterize this difference is to say that the open access movement was and should be a response to the STM publishing problem for academic libraries. Another is to say that so far as I can tell, scholarship in the humanities and social sciences is now (and is likely to continue to be) dependent upon the capacity of nonprofit publishers to recover cost.

The larger threat is of course that open access websites like the one proposed by Harvard will eliminate the publisher's function altogether – the whole bundle of skills and services provided to authors by both commercial and nonprofit publishers. I think it may be the case (as I gather it is in the field of computer science) that formal peer review of the traditional journal-based type is not necessary for the evaluation of rapidly created scientific knowledge (I am not expert enough to say). But I can tell you that the publisher's function is still central to the creation of knowledge in the humanities and social sciences. There is a good reason why university departments, in evaluating scholarly work, ask which examples of a candidate's published work are peer reviewed and which are not. That distinction

disappears in an unmediated open access website world, and it is one that scholars should fight to preserve.

And I do not need to tell this audience that the legal mechanism that makes the traditional publishing modality work is the law of intellectual property. I do not feel the need to defend everything in the Digital Millennium Copyright Act of 1998 (anymore than I feel the need to defend everything that publishers do), but copyright is in many ways essential to the best aspects of our system of intellectual communication. As I have said elsewhere, I think there are some serious problems with the extension of 1976 copyright notions in 1998 in a “medium neutral” manner, for I think that the digital world is in important ways orthogonal to the analog copyright world. It is a system that can and must be improved, though I do not kid myself about how hard that will be. I should also say here that I think there are attractive aspects of the emerging system for open access ideas and new forms of licensing. I was, by the way, very favorably impressed by your February, 2007 *Statement on Open Access* (about which more later). One of the most important problems with the rhetoric of copyright these days are its tendency to Manicheism – either one defends traditional copyright or one proclaims that “information wants to be free.” We all want to be free, but the underlying intuition of the modern democratic state is that social function involves trade-offs between liberty and order. And intellectual property is no different.

I think I can make this point best by turning to a small but interesting controversy in which I am currently engaged, relating to a publishing project that some consider arcane, the Papers of the Founding Fathers. I am, for my sins, the Chairman of the Board of the Papers of the Founding Fathers, Inc.! Bear with me for a few minutes while I explain.

Modern documentary editing of historical sources in the United States began during World War II, when the Princeton Librarian, Julian Boyd, undertook a new, annotated edition of the papers of

Thomas Jefferson. Boyd's methodology, which has come to be the model for all subsequent projects of this type, involved the letterpress publication of a careful selection of both incoming and outgoing correspondence, enclosures and related documents created by or received by Jefferson. Boyd's edition, of which the first volume was published in 1950, has now reached volume 34, which deals with four months of the first year of Jefferson's presidency, 1801. The Princeton University Press is the publisher of the series, which is compromised of fat (767 pages for #34), heavily annotated volumes that appear these days at the rate of about one a year.

In the years following World War II a great many comparable editing projects were begun, usually based at universities and always published by academic presses. Among the most notable of the editions are those of Benjamin Franklin, George Washington, John Adams and James Madison. These, along with Jefferson, constitute the FFP, Inc. group, formed in 1981 to raise funds to relieve the editors of the time-consuming process of applying for grants by raising funds centrally to support all five projects. The Franklin Papers are based at Yale and published by the Yale University Press, the Adams Papers are at the Massachusetts Historical Society and published by the Harvard University Press, while both the Washington and Madison Papers are at the University of Virginia, and are published by that University's press. I should mention here that, the rightsholders for these works are not always the projects themselves. MHS owns the copyright to the Adams Papers, Yale University and the American Philosophical Society hold copyright for the Franklin Papers, while the Princeton University Press and the University of Virginia Press hold the rights to the Jefferson, Washington and Madison Papers. A total of 217 volumes have thus far been published by the FFP projects, but it will take an estimated seventeen years more to complete the letterpress editions.

More than a decade ago we realized that we needed a way to publish these materials electronically, but this posed a very large problem for us, since the great majority of the volumes were

printed before the editors created digital printing files. The problem has been solved by the University of Virginia Press, which has created an electronic imprint (Rotunda) which among other things specializes in documentary sources for the Founding Era of American history. With a handsome subvention from the Andrew W. Mellon Foundation, Rotunda is engaged in a retrospective conversion of the hundreds of analog volumes, and it will mount the new letterpress volumes two years after they appear (to give the publishers an opportunity to market the print version). The digital volumes will be part of a single database, so that users can search across all the Rotunda editions, thus presenting an unparalleled new method of scholarly research on the Founding Fathers.

We have also been concerned to make the heretofore unpublished papers available to scholars (and more readily accessible to the editors). Each of the project offices is full of photocopies of the writings of its Founding Father, painstakingly collected over the years – only the Adams Papers, with its headquarters in the MHS, towns a great many of the original documents. We decided some years ago to keyboard and digitize these unedited materials, and the Packard Humanities Institute agreed to partner with us to fund (and process) the material. PHI has now completed the Franklin Papers, which are now available on the Web www.franklinpapers.org, and we are working on the other projects.

I should mention here that FFP, Inc. has raised funds entirely from the private sector, and especially from the Pew Charitable Trusts in Philadelphia and the Andrew W. Mellon foundation in New York. Each of the projects has raised money from other, mostly smaller, foundations and from private individuals. But over the years nearly half of their funding has come from the public sector, from the National Endowment for the Humanities and the National Historic Publications and Records Commission, each of which has programs to support scholarly editions – and the Commonwealth of Virginia.

This is the background for the story I am about to tell. For several years there has been private sector pressure, principally from the Pew Charitable Trusts, to complete the letterpress editions more quickly. The President of the Trusts, Rebecca Rimel, has repeatedly urged that the editions be completed within her lifetime, though why Ms. Rimel's longevity should create the standard is not clear to me. There have also been complaints to NEH and NHPRC that the FFP volumes were appearing too slowly. More recently there has been a growing demand for the rapid electronic publication of the volumes – and of the heretofore unpublished documents. At least to some extent this pressure can be attributed to Ms. Rimel's Pew, which has spent \$190,000 over the past year or so to hire a lobbyist to place pressure on NHPRC to demand that the FFP volumes appear more quickly and online.

The lobbyist succeeded in getting report language added to the NHPRC appropriations bill this year that disparages “the lengthy amount of time currently required to complete the publication of the Founding Fathers historical papers projects,” and directs the Archivist of the United States (who chairs NHPRC) to develop a plan for “online electronic publication, within a reasonable timeframe, of the papers of the Founding Fathers.” The staff of the Commission then proceeded to develop such a plan, which is currently before the Appropriations Committee. And quite apart from this development, Senator Leahy of the Senate Judiciary Committee (which has no jurisdiction over the matter) scheduled a public hearing on the publication of the papers of the Founding Fathers early last February. President Rimel was among those testifying (as was I), and she once again made her “in my lifetime” speech.

Why should you be interested in this boring little academic story? Well, of course you are the publishers of the FFP volumes, and in some cases you are also the rightsholders. And the larger problem is that the pressure on the Hill is not simply for online publication, but for *free* online publication. This is of course a very small instance of what was proposed in the Federal Research Public Access Act of 2006 (FRPAA), but it represents a Congressional mindset that is not likely to disappear anytime soon.

Although the Senators can see that we are currently publishing “the books” online, and we plan to mount the unannotated papers online, they do not see why both of these projects should not be mounted on the “free” Library of Congress public access website.

Why shouldn't they be, if “information wants to be free”? Well, I think it is fair to say that both the editors and publishers of the papers of the Founding Fathers would like their data to be widely accessible. You are all nonprofit organizations, but we all know that even nonprofits have a bottom line, for at the very least they have to recover their own costs. This is why Rotunda negotiated a two year “moving wall” for newly published volumes – the bulk of print sales for reference works occur within that timeframe. But this is also why Rotunda licenses its material to libraries and other distributors of information – it needs to recover cost for the production of the digital volumes, for marketing, for the maintenance of its website, and for many other activities that you as publishers will understand. Here I want to remind you of Katz's Law, which is critical in the electronic environment – it costs more to maintain a website than to create it. In other words, Rotunda has had to create a very careful business plan to in order to publish the FFP database, despite the fact that they had substantial philanthropic support in order to undertake retrospective conversion of the early printed volumes.

What the Senators are now telling me is that the fees charged by Rotunda are too high, and that “the man on the street” will not be able to access the FFP material. Let's set aside the fact that I think few street people are yearning to read James Madison's political writings – and the last time I heard the feds had not imposed price controls on the publishing industry. They are also concerned that the prices for the print volumes are too high, and that too few libraries purchase them. They have a point, of course. I wish that the Jefferson Papers sold for \$10 a volume, and I wish that a Rotunda license was less expensive. But I think I can say with confidence that neither Princeton nor Virginia is enriching itself in the process of making the correspondence of the Founding Fathers available to the public. I think

that NHPRC now understands this, but its “solution” is to call for the private sector (presumably philanthropic foundations) to compensate Rotunda and the other three presses (which derive royalties from the Rotunda license) in order to permit the Library of Congress to mount a mirror site for the Rotunda database. It may be that Ms. Rimel and other philanthropy executives will rally to this flag. But I doubt it – and yet I do not think the demand for “free and open access” for government-funded information will disappear anytime soon.

But, as I have already mentioned, about half of FFP funding has come from public sources. Something like this funding situation probably characterizes much of the publication of the Harvard faculty. But of course that means that half (or more) is funded privately. Does one dollar of federal money entitle the feds to declare that research done on its nickel should be freely accessible online? Ten dollars? One million dollars? It cannot be the quantity of funding that drives this – it has to be the principle. And one of the principles involved is that of intellectual property.

IP comes into the FFP story in another way, as well. Three or four years ago a group calling itself ConSource <consource.org> appeared, having as its aim the construction of a database that would contain all of the writings that contributed to the formation of the United States Constitution in 1787. Their idea, quite attractive to scholars and other potential users, was to mount digitally all of the texts relevant to the writing of the Constitution, and they have been building their database from a variety of sources. These sources include, as it turns out, scanning the texts of Franklin, Washington, Hamilton, and Madison from the letterpress editions of the FFP presses – without permission and without payment, since their contention is that the publishers’ copyright covers only the editors’ annotations, not the actual texts of the Founders. I don’t want to bore you with why I think that establishing the text of a Jefferson letter is different from compiling a telephone directory (*Feist* is the case in point), but let me point out that the editors have to establish which of the various versions of a letter is authoritative,

they have to transcribe the hard to read handwritten original, and they have to convert the text into modern orthography, punctuation and the like. In other words, establishing the text of the letter is a creative scholarly act. But ConSource is actively scanning FFP volumes now, and mounting our information. To be sure, ConSource is a 501(c)(3), so this is not a case of nonprofit vs. commercial usage. But it is a case in which IP is at best being narrowly construed in the name of the public interest. Information wants to be free.

What is at stake in the two stories I have sketched this morning is the question of whether the push for free digital information will further tilt the already shaky balance between creativity and use in the nonprofit publishing sector. In the first story, I fear that one of the consequences of the Harvard open access website (assuming, as I do not, that it will actually work), will be to make it impossible for some humanities and social science scholarly journals to remain in existence. In the second story, the question is whether the university presses who publish the letterpress FFP editions can and will continue to do so, and whether Rotunda can create and sustain a business plan adequate to maintain its website. In both cases, intellectual property requires at least some degree of protection if the public interest in the production and preservation of crucial scholarly information is to be sustained.

I am a supporter of both online scholarship and public access to scholarship, but I am also a supporter of academic presses and nonprofit journals, and I fear that we are seeing the emergence of a movement to privilege access in a manner that will threaten our capacity to freely create and preserve knowledge. What is the true public interest in this situation? I had not read your Open Access statement until Sandy Thatcher provided it to me in preparation for this meeting. It seems to me a very sensible statement, for I agree that the Budapest approach is too linked to the STM problem to be generally applicable, unless, as you suggest, the universities are prepared to pay more to subsidize faculty/student access than they are now paying to support their publishing operations. The open

access model assumes that universities will be willing to transfer the payments they currently make to support their presses (assuming that they are still making such payments) to subvene the cost of open access websites. Perhaps, but I doubt it.

But I fear that, even then, the threat to the sustenance of academic and learned society publishing in the humanities, at least, will deepen. This is not simply a money question, but rather one of reinventing the system of scholarly communications for the Digital Era. That may well be a good idea. But it is one we have been thinking about for two decades without any consensus on what a better system would look like. My personal view is that we should not ask either Harvard or Congress to design the system for us. The urgent need is for scholars, universities, learned societies and publishers to sit down for a systematic discussion of what an entirely different system of scholarly communication would look like. Until then, we need to protect and defend the system currently, but tenuously, in place.