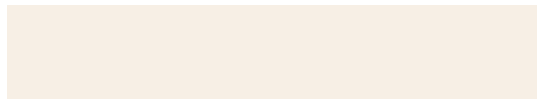# The Occasional Pamphlet

on scholarly communication

An efficient journal

March 6th, 2012

Aficionados of open access should know about the *Journal of Machine Learning Research* (*JMLR*), an open-access journal in my own research field of artificial intelligence, a subfield of computer science concerned with the computational implementation and understanding of behaviors that in humans are considered intelligent. The journal became the topic of some dispute in a conversation that took place a few months ago in the comment stream of the Scholarly Kitchen blog between computer science professor Yann LeCun and scholarly journal publisher Kent Anderson, with LeCun stating that "The best publications in my field are not only open access, but completely free to the readers *and* to the authors." He used *JMLR* as the exemplar. Anderson expressed incredulity:

I'm not entirely clear how *JMLR* is supported, but there is financial and infrastructure support going on, most likely from MIT. The servers are not "marginal cost = 0″ — as a computer scientist, you surely understand the 20-25% annual maintenance costs for computer systems (upgrades, repairs, expansion, updates). MIT is probably footing the bill for this. The journal has a 27% acceptance rate, so there is definitely a selection process going on. There is an EIC, a managing editor, and a production editor, all likely paid positions. There is a Webmaster. I think your understanding of *JMLR*'s financing is only slightly worse than mine — I don't understand how it's financed, but I know it's financed somehow. You seem to believe in fairies.

Since I have some pretty substantial knowledge of *JMLR* and how it works, I thought I'd comment on the facts of the matter.

First, some history. *JMLR* was founded when most of the editorial board of the Kluwer journal *Machine Learning* (now a Springer journal) resigned to establish JMLR, Inc., a nonprofit to develop and publish the new journal on an open access model. The first editor-in-chief was Leslie Kaelbling, a computer science professor at MIT. The journal's first papers appeared in October 2000. Its twelfth annual volume just completed this past December.

One of the main things that journal publishers do is manage the logistics of the peer review and filtering of submitted articles. Starting with the former *Machine Learning* team, the journal put together an editorial board and a cadre of action editors to handle the reviewing process. At the time the journal was launched, there weren't the abundance of open-source journal management

[platforms](#) that are now available. Being computer scientists, the editorial board took the expedient of implementing their own, a custom system that they still use. Much of the clerical effort of tracking the peer review process — assigning papers to action editors, engaging reviewers, tracking reviews, acceptances and rejections, and the like — is automated by the platform. Of course, these days, the platform situation has [eased considerably](#).

Almost immediately, the journal was appreciated as being of top quality. The number of articles it published increased quickly over the first few years, its [illustrious editorial team](#) serving to convince prospective authors of its seriousness. Its first year in ISI's rankings, it had the highest [Impact Factor](#) of any journal in its Web of Science subject category ("computer science, artificial intelligence"). It is currently ranked eighth (of 108 journals) by Impact Factor and fourth by [Eigenfactor](#) and Article Influence. *Machine Learning* is down to 33rd. If you're [into that kind of thing](#).

The journal does not charge any submission or publication fees and has never done so. It has never taken any advertising. Indeed, it has never had any direct revenue at all. In fact, JMLR, Inc. didn't even have a bank account until recently; there was no need.

Of course, there are costs, but they are all provided through in-kind support. By far the largest costs are the labor required for peer reviewing and its management by the editorial board, but this is all volunteer effort as in most all scholarly journals. The primary people involved, the editor-in-chief, managing editor, and production editor, are all unpaid, contra Anderson's conjecture. They volunteer for *JMLR* in their spare time away from their day jobs as computer science professors. MIT implicitly underwrites some clerical help, since Kaelbling's administrative assistant at MIT does a small amount of work for the journal, amounting to a few hours per year.

The webmaster is a student volunteer. Anderson is right that MIT provides the web server, saving *JMLR* the [tens of dollars per month](#) they would otherwise have to pay for commercial hosting. Kaelbling has paid for the domain name [jmlr.org](#) out of her own pocket. The [going rate for .org domains](#) is about $15 per year.

In addition to management of the peer review process, publishers provide production services as well, such as copy-editing and typesetting. One of the main motivations for *JMLR* leaving Kluwer was the sense that the help they were supposed to be providing was sparse and better avoided. Kluwer did no copy-editing of articles. *JMLR* relies on reviewers for the kind of light copy-editing they always have done in the normal course of reviewing. For accepted articles that require large amounts of language help, the authors are requested to find copy-editing help at their expense; such cases are extremely rare. Other than that, no copy-editing is done. It doesn't seem to have harmed the journal's perceived quality.

As for the typesetting of articles, computer science authors typically use the open-source [LaTeX](#) typesetting system for writing their articles, a system designed for beautiful typesetting of mathematical material and far better for mathematical typesetting than the typical systems publishers are accustomed to. The process of retypesetting that many journals have historically performed inevitably introduces errors, leading to a product inferior to that computer science authors typically provide. *JMLR* used an approach where authors submit camera-ready copy based on a publisher-supplied LaTeX style file. By dropping the retypesetting with an inferior system, errors in the process are eliminated and the quality of typesetting improved. Increasingly, journals in computer science and related fields (mathematics, physics) are moving to this system. In fact, *Machine Learning* itself accepts LaTeX submissions and provides an appropriate LaTeX style file for authors to use. Thus, the

total cost to *JMLR* for copy-editing and typesetting is zero.

The biggest expense, it turns out paradoxically, is paying a tax accountant. Kaelbling explained the problem to me:

We have to file a bunch of annoying forms to maintain tax exempt status, etc. I have paid for the original incorporation and some amount of the accountant out of my pocket. But I have gotten a couple of donations (totaling $7K) which I have also used for that stuff. It wouldn't need to be so expensive, except I'm too disorganized and late to keep on top of it myself.

JMLR has always appeared both free online and by subscription in print. The print edition was originally intended to satisfy the desires of authors who hung onto a view that online-only journals may not be viewed as "serious", but also has the advantage of substantially solving the digital preservation problem for the journal. The print edition of the first four volumes was published by MIT Press, at first quarterly, then semi-quarterly as submissions grew and more articles were accepted. *JMLR* received no revenue from the print edition and paid no subvention to MIT Press. MIT Press handled all aspects of fulfilling the print subscriptions and kept all the revenues from a quite reasonable subscription fee of just under 30 cents per page. From the fifth volume on, the print edition was taken over by [Microtome Publishing](#) under the same zero-zero arrangement. Under Microtome Publishing's approach, which [leverages important aspects of the print editions specific to open-access journals](#), the subscription cost decreased dramatically over the next few volumes, settling at a steady state of 8 cents per page for the last several volumes.

Adding it all up, a reasonable imputed estimate for *JMLR*'s total direct costs other than the volunteered labor (that is, tax accountant, web hosting, domain names, clerical work, etc.) is less than $10,000, covering the almost 1,000 articles the journal has published since its founding — about $10 per article. With regard to whose understanding of *JMLR*'s financing is better than whose, Yann LeCun I think comes out on top.

[**Update 3/18/12:** In the comments section, Leslie Kaelbling corrects her estimate of outside donations to $3,500, so I should revise my estimate of *JMLR*'s cost per article to be about $6.50 per article.]

How do I know all this about *JMLR*? Because (full disclosure alert) *[I am Microtome Publishing](#)*. Microtome is a sole proprietorship providing "publishing services in support of open access to the scholarly literature." I've worked with *JMLR* for many years now, and consequently have gained a good understanding of all aspects of its operations and of the operations of a subscription-based print journal as well. I don't pretend to have all of the knowledge of a professional publisher by any means. On the other hand, I don't believe in fairies.

Does *JMLR*'s success and efficiency mean that all journals could run this way? Of course not. First, computer science journals are in a particularly good situation for being operated at low cost. Computer scientists possess all of the technological expertise required to efficiently manage and operate an online journal. Journal publishing is an information industry and computer scientists are specialists in information processing. Second, the level of volunteerism that *JMLR* relies on is atypical for the entire spectrum of journals. Paid editorial positions for computer science journals are exceptionally rare; we're used to the volunteerism of running a journal. As authors, computer scientists are accustomed to performing their own typesetting and we prefer to do it ourselves. *JMLR* reviewers are relied on for whatever copy-editing is

done. Paying professional copy-editors if that was desired would add more to the cost per page (though apparently not even *Machine Learning*'s commercial publisher was doing so when the board left). Third, some of the costs of operating a journal are the overhead costs that are being absorbed by various institutions. An independent publisher would have to pay for office space for staff, for instance, whereas the primary editors use their homes or offices, hiding that cost.

Nonetheless, the success of *JMLR* does provide a clue that the cost of running a premier journal might be far less than publishers imply, if they were to rethink the process substantially — maybe not $10 per article, but surely far less than the $5,000 average revenue per article that scholarly publishers currently receive. This expectation is borne out by the several non-profit and commercial open-access journal publishers that are able to operate in the black with publication fees a fraction of that average.

Anderson closes his comments on *JMLR* with these recommendations for LeCun:

You should look at yourself in the mirror, and ask why you don't understand even the most basic financial realities (computers cost money to run, editors get paid, and webmasters get paid), why you don't understand how *JMLR* is funded, how much you've benefited from tuition/fee increases foisted on students at +395% over the past decade,[1] and why you feel compelled to argue points you haven't adequately examined (you tell me how *JMLR* is funded, and you'll have much better face validity).

The call not to argue points one hasn't adequately examined is surely apt.

---

[1]With regard to tuition hikes foisted on students see my earlier post.