

Scientific Publication

[Eric Hehner](#)

Introduction

I have recently retired from a long and interesting career as a professor of computer science at the University of Toronto. An important part of my job was the publication of research papers; both my career advancement and the success of my research grant applications depended upon my publication record. My publishing experience comes from three different vantage points. The first is as an author: two books, and many journal and conference papers. The second is as a referee: countless times for several journals, and program committee member for many conferences. The last is as an editor of three journals: *Information Processing Letters*, *Acta Informatica*, and *Formal Aspects of Computing*. From my experience, I want to tell you what I think is wrong with the research publishing industry, and how it should change. I will also say what implications this has for the way universities and granting agencies measure research success.

Background

The publishing industry developed in a very different technological age. Each publishing company served all the following functions:

- soliciting submissions
- deciding which submissions are worth publishing
- sometimes helping the author to write (sometimes to the extent of doing all the writing)
- editing what the author wrote for content (checking facts, checking consistency)
- editing what the author wrote for style (grammar, spelling, appropriate choice of words)
- choosing or creating artwork and diagrams
- choosing the physical size and paper and placement of the artwork
- typesetting (usually hiring a typesetting company)
- proofreading
- printing (usually hiring a printing company)
- binding (usually hiring a binding company)
- warehousing
- distribution to bookstores, libraries, and subscribers
- advertising
- accounting, receiving and paying money
- protecting the copyright

Publishing was complicated. Scientists need to publish their research, and they needed all the expertise and resources of a publishing company to do so.

Times have changed. Publishing is easy. Almost all authors do their own typesetting, find or create their own artwork and diagrams, do their own spell-checking, and put their papers on their own website where everyone in the world can find them and read them. There is no need for printing, binding, warehousing, and physical distribution. But there are still services that publishers perform: deciding which submissions are worth publishing, making editorial suggestions, and protecting copyright.

Problems

In days gone by, a published paper was a physical, stable object. But now, a paper published by its author on the author's website can be changed by the author at any time. This is both good and bad. On one hand, if author B finds an error in a paper by author A, and publishes a paper citing and quoting author A's paper, author B needs to be assured that the error and the quotations will remain unchanged, for anyone else to see. On the other hand, author A would like to correct the error, which may be done by an easy edit, so that all future readers of the paper will get the corrected version. Author A is ethically obliged to thank author B in the corrected version, but that still leaves author B's paper stranded without a referent.

One of my books, *a Practical Theory of Programming*, was published by Springer in 1993, a couple of years before the internet became widely available. By 2002 I had accumulated a list of improvements and updates and some new material, and it was time for a second edition. By then, the internet was well established, and I wanted to make my book freely available on the internet. Springer, who owned the copyright, refused. How can they and I make money if our product is freely available? To an author of an advanced-level book, the money is not significant; it can never repay the work of writing the book. I didn't give up the fight, and I had some powerful allies, so in the end Springer allowed me to put the book on the web. To everyone's surprise, Springer's sales went up. The numbers are not statistically significant, so my only conclusion is that you can't predict what will happen. Apparently, people browsed the free online version to see if they wanted the book, and if they did, in 2002–2005 they considered that a book was still made of paper, so they bought it. From my point of view, the main benefit was the ability to make changes. Anytime I saw a way to improve an explanation, or to shorten a proof, I made the change that same day. I always want my book to be the best I can make it today, not just the best I could make it ten years ago. I stopped calling editions “first”, “second”, and so on, and started calling them by year-month-day. I also keep a change log, available for all to see. Fortunately, no-one has written any papers about errors in my book, but if they do so, they will just have to refer to the edition date.

Perhaps the most important remaining function of scientific publishers is to decide which books and papers are sound and well-written enough to be worth publishing. From the publishers' point of view, this function is essential because they cannot afford to waste their efforts and pay the costs of printing and distributing books and journals that will not be bought. University administrations judge their scientists based on how many books they have published and how many papers they have in high quality journals; therefore the publishers determine the scientific merit and career advancement of academic scientists. Scientists gain some assurance that a published paper is worth reading; if the publishers were to publish all submitted papers, the good ones would be lost among the junk. No-one has the time to read through all the junk to find the good papers.

In my opinion, publishers do a miserably poor job of selecting the best books and papers from among those submitted. One problem is the system of refereeing. An editor must find at least three referees for each paper submitted. A referee must check that the paper is original, check the facts, arguments, and proofs in the paper, and make a judgement about the importance of the results and the quality of presentation. A referee writes a report for the author and editor to see, and optionally some comments for only the editor to see. A referee is unknown to the author, so that negative comments and judgements will not incur retaliation if ever the author and referee switch roles, as is likely in a specialized area.

The first problem with refereeing is that it is voluntary, with no compensation: not even academic credit. Some scientists don't do their share, leaving the more conscientious scientists to do more than their share. The referees available may not include the best choices for the

subject area. Since refereeing is both anonymous and uncredited, a referee may not give the paper as much time as needed to do a good job of refereeing. A referee who does a good job suffers by taking time away from their paid, credited activities; a referee who does a bad job suffers no ill. Furthermore, referees often have their own research territory and results to protect, and may give a negative report on competing or conflicting research in their own self-interest. And when the author sends their rejected paper to another journal, the probability is quite high, in a narrow specialty, that the same referee will be used. The scarcity of willing, qualified referees too often means that a paper's fate is determined by fewer than three referees who do an incomplete and less than careful reading.

Rejected authors often feel that the publishers are preventing them from being heard, and if they could just get published, the larger community would see the value of their work. Rejected work ranges from terrible to great. One of my favorite books, *the Experts Speak* by Cerf and Navasky, has a long list of major scientific achievements that were rejected for publication on the first, second, and subsequent attempts. The comments of the referees (these are the “experts”) reveal how wrong referees can be. Here's an example I witnessed. In 2009, a new prize, the John Stewart Bell Prize, was established for the best work in quantum physics, and its first award was made at the University of Toronto in the building next to mine, so I decided to attend. The first winner was Professor Nicolas Gisin of the Université de Genève. Like any winner of a top award, he graciously thanked the awarders, and his co-workers, and he talked about the significance of his work. What was surprising about his acceptance speech was its bitter tone. He had done the work for which he was given the award forty years earlier, and for many years he could not get it published. The entrenched physics establishment was unwilling to hear new results and theories that would make their own theories and textbooks obsolete. If Professor Gisin's work had been accepted early in his career, he would have had a very different career. Publishers should not have the power to censor ideas.

Rejection is not always so nefarious as Gisin described. If a paper is too radical, too great an improvement on past work, referees may not recognize that it is good work in their cursory reading, because it does not agree with what they think is true. The safest kind of paper to write, most likely to be accepted, is a paper that agrees with the prevailing theories and opinions, and offers a small improvement or addition. But there is an opposite danger, as pointed out recently (2012) by Nobel prize-winner Daniel Kahneman and others. To be published, results must be at least a little surprising. If results are just what everyone expected, they aren't published. In psychology (Kahneman's field), experimental results can vary a lot from experiment to experiment. If only the results in the surprising direction are published, and the more numerous results in the opposite unsurprising direction are not, the surprising results will be accorded much more significance than they deserve. That is why Kahneman has called parts of psychology a “train wreck”, and called for a restoration of “credibility” by publishing all results.

As an editor of three journals, I handled the papers of an author named Subramani. In his area, the best places to publish for the quality of publication (low acceptance rate), and for visibility of the work, and for speed of publication, were certain conferences, and so he published there. But his university administration counted only journal publications, not conferences. Journals refuse to publish work that has already appeared in a conference, or even a close variant. He found himself in a bind created by the inflexible policies of his administration and the restrictive policies of journals.

The importance of publication count to scientists causes large groups of them to add each other as authors to all their papers. There are authors who haven't even read the papers they are supposed to have written. In chemistry, the practice is so extreme that the author list may be longer than the paper. Some university administrations count a full credit, rather than a fractional credit, for a fractional contribution. But any publication count, even with fractions, is

a very poor measure of a scientist's contribution because it does not consider the quality of the papers.

Solution

The solution to all these problems is easy. First, let everyone publish their books and papers on their own website; they do anyway. This allows everyone everywhere to access all papers and books freely. No other form of publication is necessary. Second, they can send their (electronic) books and papers to an archive service; its only purpose is to date and store books and papers in case anyone needs to prove that a book or paper said some particular thing at some particular time. Or perhaps an archival web-crawler does that automatically. And finally, there must be independent reviewers. They are just like the book reviewers in the current newspaper book review column, or the movie reviewers in newspapers and on television. A reviewer is not anonymous; they sign their name to their reviews, and they make their reputation on the honesty and thoroughness of their reviews. They are paid by selling their reviews to subscribers. Authors and anyone else may send them papers and books in the hope of having them reviewed, but with no hope of influencing the review. A scientist gets guidance as to which papers and books of other authors are worth reading. A scientist gets academic credit by getting good reviews on their own papers and books from reputable reviewers.

Other Publishing

I have been discussing scientific research publication, an activity I have participated in for forty years in various roles. But it seems to me that much of what I have said applies also to the wider world of non-scientific book publishing, and even to music publishing and movie publishing. These activities used to require expensive equipment and specialized expertise; now anyone can record, edit, and publish books, music, and movies, almost for free.

The music-makers and movie-makers have a problem not faced by scientists: getting paid. Scientists are paid by their companies or universities or government departments, and publishing is part of the job they are paid for doing, whereas other book authors and musicians and movie-makers are paid by the people who read/hear/watch their products. Music and movie publishers sell copies of music and movies, and try to protect their livelihood by copyright law. It is a losing battle. More music and movies are copied and sold illegally by "pirates" than legally by the copyright owners. The copyright notice at the beginning of each movie, which threatens that the FBI will throw us in prison, is just laughable, especially outside the USA. The record labels and movie distributors are still trying to enforce their past business practices, as if technology had not changed. Meanwhile iTunes has a business practice that works: it sells songs and movies so conveniently and at such a small price, with zero overhead, that it's not worth a buyer's trouble to find a free version to copy. If the labels and distributors would just get out of the way, iTunes can sell on behalf of musicians and film-makers directly.

Scientists want their papers to be read, even for free. It is not in the scientists' interest, nor in the interest of science generally, to reduce the potential readership to just those who pay. The publishers' interest and the scientists' interest conflict. But publishers, who were essential to scientific publication in the past, are now superfluous, clinging to territory they once owned. Copyright protects publishers; it too should disappear. There is no reason to restrict copying of published scientific works, nor is it possible to do so. We just need to insist (by law) that an author is always given credit for their work.

The publishing of written material differs from the publishing of music and movies in another way: the importance and treatment of authors. Movies prominently advertise their

director and actors, but the writer who created the content is buried in the credits at the end of the movie, along with the gaffer and caterers. Popular songs are found all over the internet, listed as being “by” a person or group who perform the song, and its author is not even mentioned. I have even seen song listings on the internet showing the song title, and below it, the word “by” followed by the name of the person who posted the song on the internet. Radio stations broadcasting on cable and satellite list the song title, the artist performing the song, the album it comes from, and the publisher, but not the song's author. iTunes lists the the name of the piece, its duration, the artist performing it, the album it comes from, and the genre of music; but no author. Can you imagine seeing an advertisement for *Harry Potter* by Bloomsbury Publishing, with no mention of J.K.Rowling? (By the way, many publishers rejected *Harry Potter* before Bloomsbury accepted it.) How about the *Ninth Symphony* by the Toronto Symphony Orchestra, with no mention of Beethoven? McLean's Magazine made a list of great Canadian songs, and included “Four Strong Winds” by Neil Young. A reader wrote a letter to say that “Four Strong Winds” was by Ian Tyson. Below the letter, the editor replied that the song they wanted was the one by Neil Young. (It was written by Ian Tyson.) Popular music and movie publishers and distributors do not consider their authors important enough to mention. It's time to turn that around: the authors needn't consider the publishers and distributors important enough to use them anymore.

Conclusion

Publishing companies are a relic of the past. At present, they are an irritant. In the future, I wouldn't mourn their disappearance.

[other essays](#)