A reprint from

American Scientist
the magazine of Sigma Xi, The Scientific Research Society

Open Access and the Progress of Science

Alma Swan

THERE'S AN OLD JOKE about asking the way to somewhere and being told it would be best not to start from where you are. It's a good way to frame some thoughts about whether our present system of scholarly communication aids the progress of science or gets in the way.

If we could start now, equipped with the World Wide Web, computers in every laboratory or institution and a global view of the scientific research effort, would we come up with the system for communicating knowledge that we have today? The system we have, which originated as an exchange of letters and lectures among scattered peers, does some things well. But in its current form—a leviathan feeding on an interaction of market forces within and outside science—one can hardly argue that the system satisfies the needs of a modern scientific community. And new developments in the way science is done will make it even less fit for its original purpose in the years ahead.

No, we would think of a new way, one that would provide for rapid dissemination of results that any scientist could access, easily and without barriers of cost. We might debate how to implement quality control, how to ensure that originators of ideas or findings are given their proper due, how our new and better system should be paid for and how to deal with bandwidth constraints in some parts of the world. But no one would say, "Hey, why don't we only let some researchers see this stuff and see how science gets on?" Yet that is precisely where we are today, in a system where gateways limit access to research results, and as a consequence only a small fraction of the world's research libraries subscribe to some journals. The gentleman's club survives, if only as metaphor.

The power to transform research communication may be at each scientist's fingertips

For the past decade or so, a number of scientists have argued that the World Wide Web offers a way to unlock the gates that was not possible when scientific results were conveyed solely by print-on-paper. Advocates of "open access" argue that research results must be made available such that all scientists can see them and use them, for free, via the Web.

Other arguments in favor of open access come from different perspectives. Early calls for publishing reform cited rapid rises in the cost of journals and the ensuing "serials crisis," wherein libraries have been forced into repeated rounds of subscription cancellations. Others focused on the plight of developing-world scientists and their difficulty in accessing journals (at all, in some cases). Commercial and scholarly-society publishers responded with initiatives that addressed these issues in specific ways, while sticking largely to the subscription-based "toll gate" models of literature access that have been dominant during the growth of international science publishing.

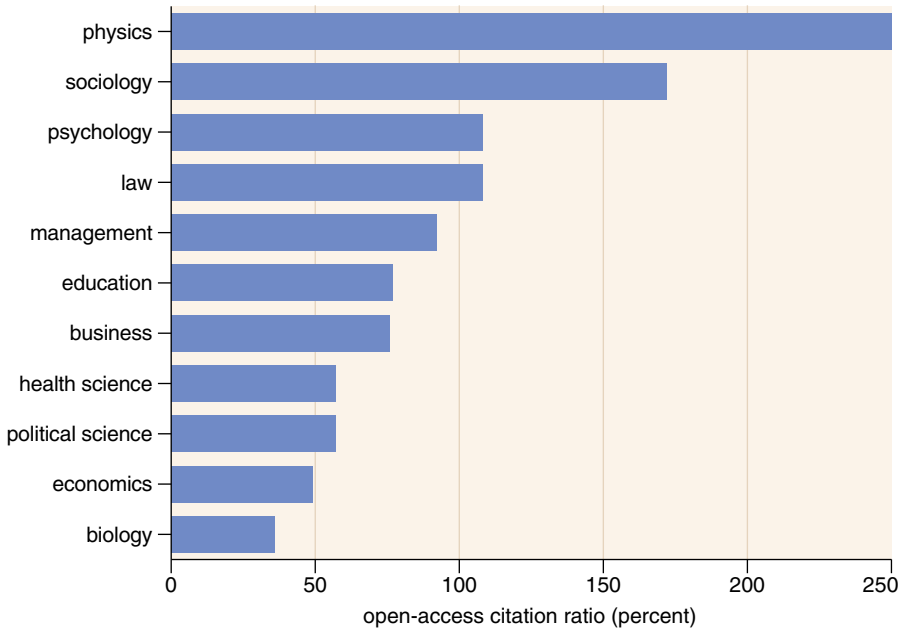
Today an entire "who will pay, and how much?" debate swirls around the question of access to literature. The bickering over varied business models, and the side arguments over public access to publicly funded results, obscure a larger, more important question: Can open access—the fundamental change to a system where scientists no longer face barriers to accessing others' work (or their own)—advance science? My work involves measuring, analyzing and assessing developments in scholarly communication. From that perspective I argue that the answer is yes, and that the advance of science is the prime reason that access is an imperative.

We Cite What We See

How does science measure the worth of a published piece of work? The standard metric today is the citation: Highly cited articles (and journals) have measurable impact. As open-access publishing experiments are moving forward, they are beginning to rack up numbers. By definition an open-access article has greater visibility, and it's becoming evident that scientists do take the opportunity to read and use what they would otherwise not have seen. The bar chart on the next page shows that across a range of scholarly disciplines, opening access to articles increases their citation rate. Behind the numbers are the new collaborations that result when scientists who don't know of one another's work discover synergies that can be exploited. Science needs open access to facilitate that process.

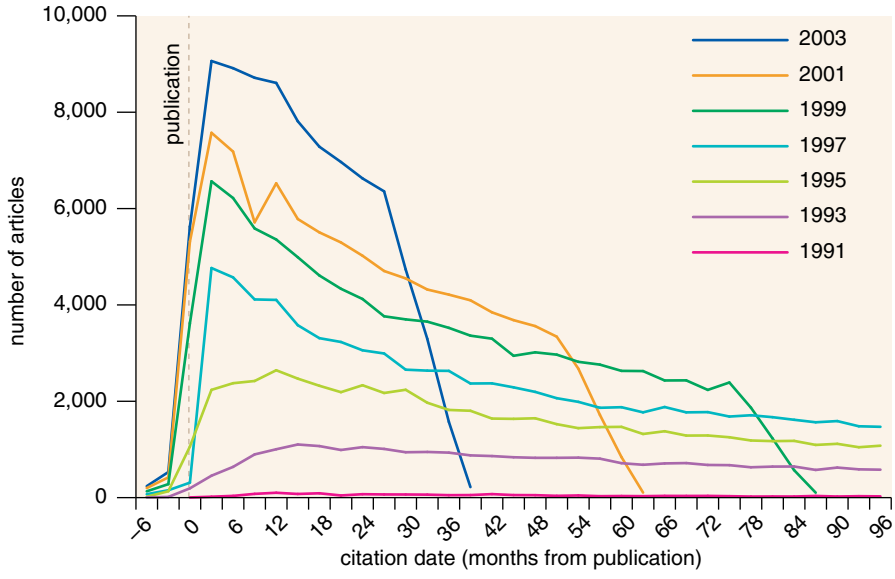
Open access can advance science in another way, by accelerating the speed at which science moves. In most fields,

In 1996, after more than two decades in medical cell biology research and scholarly publishing, Alma Swan cofounded Key Perspectives Ltd., a consultancy in the area of scholarly communication. She holds graduate degrees in cell biology (Ph.D., Southampton University) and business administration (MBA, Warwick University). She has served on the faculty of Leicester University and on the staff of Pergamon Press/Elsevier Science, where she was senior managing editor. Address: Key Perspectives Ltd., 48 Old Coach Road, Playing Place, Truro, TR3 6ET, UK. Internet: www.keyperspectives.co.uk



Across many fields, journal articles made openly available on the Internet are more heavily cited than those that remain behind subscription barriers, evidence that open-access articles have a greater impact on research. This chart shows results from a 10-year tracking of citations. Shown is the ratio of citations of open-access articles to citations of closed-access articles published in the same issue of a given journal, averaged by discipline. (Data from Hajjem, Harnad and Gingras 2005.)

open access is still a rarity rather than the norm, but in some fields of physics (high-energy, condensed matter and astrophysics) it has been commonplace for more than a decade. The arXiv, an open-access archive now maintained at Cornell University, contains copies of almost every article published in these disciplines, deposited by the authors for all to use. Tim Brody of Southampton University has measured the time between when articles are deposited in arXiv and when citations to those articles begin to appear. Over the years, this interval has been shrinking as the arXiv has come into near-universal use as a repository and as physicists have taken advantage of the fact that early posting of preprints allows them immediate access to others' results. In other words, a system built on open access is shortening the research cycle in these disciplines, accelerating progress and increasing efficiency in physics.

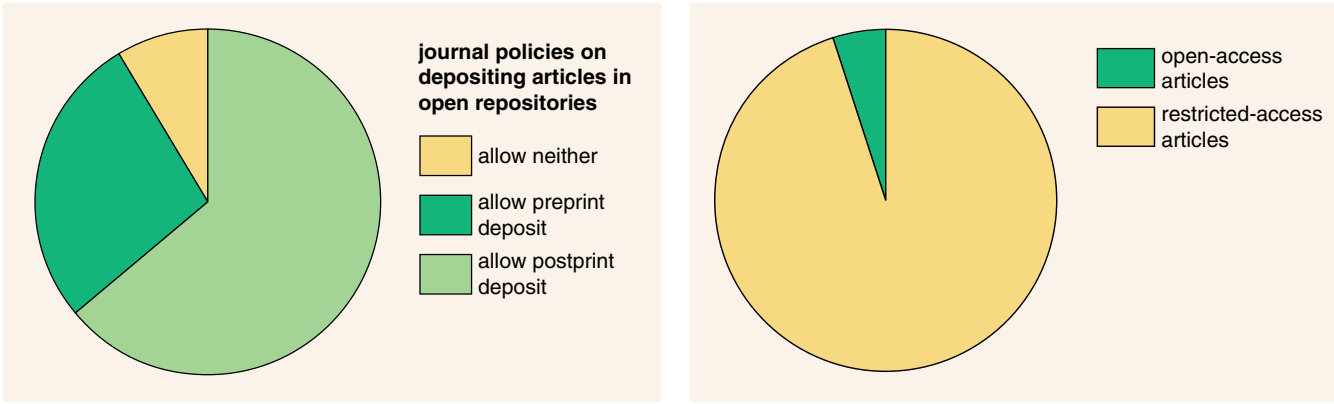


Articles in the arXiv repository are being cited more and more rapidly, suggesting that this open-access database may be playing a role in accelerating the pace of research in physics and the other fields it serves. In some cases, preprints placed in the arXiv are cited even before they are published.

Open access can also advance science by enabling semantic computer technologies to work more effectively on the research record. Such advanced software technologies already exist, awaiting a larger corpus because they need the full text of scientific articles to work on, not just the abstract. Semantic technologies can do two things. First, they hold out the promise of being able to integrate different types of research output—articles, databases and other digital material—to form a single, integrated information resource and to create new, meaningful and useful information from it. An early example of this sort of knowledge creation is the Neurocommons, a project of the ScienceCommons organization. Second, Web 2.0 technologies, the set of tools that aid collaborative effort (including social tagging and filtering and weblogs), can help scientists in their work by offering personalization mechanisms that enable them to tailor and enhance what information they access and share, saving time and effort.

Open access also enables a different kind of software tool to aid the management of science. Such tools search full-text articles and index the references they contain—the citations to other articles. They can thus calculate the impact of an individual article (the number of times it is cited) and do the same for its author, and for her research group, department or institution if required. They can track the evolution of ideas, topics and fields and facilitate trends analysis, enabling better prediction of which research areas are waxing and waning. The value of such tools to research managers, policymakers and funders will be enormous, enabling better funding and planning decisions to be made in the interest of scientific progress. To work, though, they need access to the full-text of research articles—an open literature.

Finally, the new ways in which science is being done are themselves requiring the culture and norms of open access. Interdisciplinary science, a rapidly growing phenomenon, needs open access because traditional methods do not provide effective ways by which scientists can reach out to those in un-



Most publishers of scholarly journals now have policies that allow authors to post versions of their research papers in online repositories to enable open access. Participation in such "open-access self-archiving" has not been high. The chart at left categorizes the policies of journals listed by the SHERPA/RoMEO service in the U.K. Estimates of the number of scholarly articles now openly accessible on the Internet range from 100,000 to 250,000—a small fraction of the approximately 2 million published articles available online in one way or another.

connected fields. An open literature facilitates the finding and coming together of disparate scientific efforts that in a closed-access world are circumscribed by conventional definitions of topic, field or discipline and isolated from one another in discrete families of journals. The rise of e-science, where global collaborations generate data in vast quantities, demands the means for open and immediate sharing of information. And informal channels such as wikis and blogs that are used for disseminating scientific information that cannot be communicated by journals—including time-critical information—must be accompanied by access to the peer-reviewed literature if scientific information is to be accurately conveyed and interpreted.

So yes, open access can advance science and will do so more and more effectively as more scientists make their work freely available. Moreover, science will not benefit in a vacuum: New work by economist John Houghton and colleagues at the University of Victoria in Melbourne shows that enhanced access to research findings is likely to result in an enhanced return on investment in research and development, something that can benefit every economy in the world. Research is expensive enough that the world can scarcely afford an antiquated, inefficient and high-cost system of information dissemination.

A Way Round

Which access model offers the most promise for advancing science? Open-access journals—numbering around 2,500 (approximately one-tenth of the world's peer-reviewed journals)—provide one option, but they may not offer every scientist the publishing route of choice. Scientists remain under intense pressure to publish in high-impact journals. Most of these are still subscription-access only and continue to find older business models profitable. Moreover, many open-access journals have replaced subscription fees with up-front payments to cover article-processing costs; these pose difficulties for some scientists.

A mechanism may eventually be found to transfer the money currently spent on journal subscriptions into the hands of authors to pay for publication; this question is at the center of current debates on open-access legislation before the U.S. Congress. But such a mechanism is not yet properly in place, and value has still not been driven into the system. There is a simple alternative that rests in the hands of the scientific community itself. Institutions around the world have been building robust research repositories; many of these institutions and their scientists have taken advantage of publishing agreements that enable the posting of postprints in repositories. To provide open access, all that is needed is for each scientist to place a copy of each article, as soon as it has been peer-reviewed, into an open repository at his institution. Known as self-archiving, this act takes a few minutes and costs a scientist nothing.

A global network of institutional open-access repositories is rapidly becoming established. They all expose their content to Google and other search engines, providing worldwide visibility and the immediate opportunity for an article to be read, used and built upon. No subscription-based journal can boast that it has a potential audience of the whole world's scientific community. Self-archiving is growing rapidly. I survey authors periodically to chart their activity. Between the last two surveys, in 2004 and 2005, the percentage of scholars reporting self-archiving activity in some form rose from 23 percent to 49 percent.

At a stroke, by self-archiving, a scientist can banish the threat of that bane of scientific life—obscurity. A few minutes at the keyboard today makes one's work visible to any scientist who might build on it tomorrow. While commercial publishers, scientific societies and librarians struggle over business models and tough longer-term issues such as who will maintain the record of science in a digital age, it remains the individual investigator who has the tools at hand to speed science along.

Bibliography

Hajjem, C., S. Harnad and Y. Gingras. 2005. Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact. IEEE Data Engineering Bulletin 28(4):39–47. http://eprints.ecs.soton.ac.uk/12906/

Houghton, J., and P. Sheehan. 2006. The economic impact of enhanced access to research findings. CSES Working Paper number 23, University of Victoria, Melbourne.

Jeffery, K. 2006. Open access: An introduction. ERCIM News (January). http://www.ercim.org/publication/Ercim_News/enw64/jeffery.html

Suber, P. 2006. Open access overview. http://www.earlham.edu/~peters/fos/overview.htm

Swan, A. 2005. Open access self-archiving: An introduction. http://eprints.ecs.soton.ac.uk/11006/01/jiscsum.pdf