"Accepted Community Standards"

Lately we have been hearing quite a lot about the understandings scientists have regarding rules for the publication of research results. Some of those who have communicated with us refer to their particular principles as "accepted community standards" (hereafter ACSs). That interests us, because some have charged us with departing from theirs; and it interests scientists, who must try to hold to a consensus of some sort in a world that is changing rapidly and becoming increasingly proprietary.

In the course of negotiating the submission of papers reporting the human genome sequence, we were told, with a conviction sometimes bordering on passion, about an extension of a principle we thought had general acceptance—an extension for which ACS status was claimed. The principle, of course, is that persons reading a scientific paper should have access to the materials and data needed to verify the conclusions. No problem there. The extension, communicated to us by a number of scientific leaders, is that these must be made available free, with no charge to potential commercial accessors.

Is this standard alive and well? To find out, we need to look at the world as it is, without hopeful presumptions. In today's world, an increasing proportion of basic research is being done in for-profit organizations, and an increasing number of academic institutions now profit from faculty innovations. In response, publication policies have already been transformed. As a result, in the case of this particular extension of the open access principle, the horse has already left the barn. Many recently published papers (in journals such as the *Proceedings of the National Academy of Sciences*, *Nature*, and *Nature Genetics*) have guided the reader to Web sites offering critical data or methods without charge to academic scientists, but available to for-profit users only for a subscription fee.

In the real world, there are even exceptions to the disclosure of primary data or methods, based on proprietary restrictions. We even made such an exception ourselves. In 1987, *Science* published a paper in which proprietary seismographic and other data compiled by Exxon were used to construct the long-term history of Earth's sea level. Because the subject was relevant to the emerging climate change issue and the analysis helped spark new research to improve our understanding of sedimentary basins, we thought it made sense to grant a singular exception. Did that establish a precedent? We didn't think so, but one can imagine conditions under which it might happen again. Furthermore, many papers have been published reporting results of numerical modeling or analysis—in economics, physics, engineering, climate analysis, and even biology—by both public and private groups, in which the general approach and parameters are published, but the raw code is not provided or is not available at all.

We need to generate consensus around a set of standards.

It is clear, moreover, that ACSs are not the same across disciplines. Research in economics, an increasingly mathematical science, depends on large sets of quantitative data. The majority of economics journals do not require that authors make the data on which their analysis is based available to others; even, in many cases, to reviewers or commentators. Some government data sources restrict distribution to certain favored institutions or authors. The *American Economic Review* requires authors to share data with requesters, but its instructions to authors note that when such data are proprietary, the editor is prepared to negotiate that provision.

A rather different extension of the principle has been offered by members of the bioinformatics community, who equate free acccess with unrestricted use. Their proposition—that all sequence data must be deposited in GenBank—ignores a problem well illustrated by the letter published on p. 827 of this issue of *Science*. There is a growing resentment between sequencers and bioinformatic sequence users, fueled by the ambiguous status of stored data: Is it accessible and usable, or merely accessible? In the public domain, or merely public? One community's ACS may thus mean trouble for another.

All this confusion and inconsistency has led the U.S. National Academy of Sciences to propose a study of the entire issue. We think that is a good idea. But the analysis needs to begin with the world as it is, not with a world imagined from wishful ACSs. The exercise will be well worth the effort if it generates consensus around a set of standards. But if it does, we will all then have to do what we're not doing now: live by them.

Donald Kennedy
Editor-in-Chief