

Brooks Hanson is Deputy Editor for physical sciences at *Science*.

Andrew Sugden is Deputy Editor for biological sciences and International Managing Editor at *Science*.

Bruce Alberts is Editor-in-Chief of *Science*.

## Making Data Maximally Available

SCIENCE IS DRIVEN BY DATA. NEW TECHNOLOGIES HAVE VASTLY INCREASED THE EASE OF DATA collection and consequently the amount of data collected, while also enabling data to be independently mined and reanalyzed by others. And society now relies on scientific data of diverse kinds; for example, in responding to disease outbreaks, managing resources, responding to climate change, and improving transportation. It is obvious that making data widely available is an essential element of scientific research. The scientific community strives to meet its basic responsibilities toward transparency, standardization, and data archiving. Yet, as pointed out in a special section of this issue (pp. 692–729), scientists are struggling with the huge amount, complexity, and variety of the data that are now being produced.

Recognizing the long shelf-life of data and their varied applications, and the close relation of data to the integrity of reported results, publishers, including *Science*, have increasingly assumed more responsibility for ensuring that data are archived and available after publication. Thus, *Science* and other journals have strengthened their policies regarding data, and as publishing moved online, added supporting online material (SOM) to expand data presentation and availability. But it is a growing challenge to ensure that data produced during the course of reported research are appropriately described, standardized, archived, and available to all.

*Science's* policy for some time has been that “all data necessary to understand, assess, and extend the conclusions of the manuscript must be available to any reader of *Science*” (see [www.sciencemag.org/site/feature/contribinfo/](http://www.sciencemag.org/site/feature/contribinfo/)). Besides prohibiting references to data in unpublished papers (including those described as “in press”), we have encouraged authors to comply in one of two ways: either by depositing data in public databases that are reliably supported and likely to be maintained or, when such a database is not available, by including their data in the SOM. However, online supplements have too often become unwieldy, and journals are not equipped to curate huge data sets. For very large databases without a plausible home, we have therefore required authors to enter into an archiving agreement, in which the author commits to archive the data on an institutional Web site, with a copy of the data held at *Science*. But such agreements are only a stopgap solution; more support for permanent, community-maintained archives is badly needed.

To address the growing complexity of data and analyses, *Science* is extending our data access requirement listed above to include computer codes involved in the creation or analysis of data. To provide credit and reveal data sources more clearly, we will ask authors to produce a single list that combines references from the main paper and the SOM (this complete list will be available in the online version of the paper). And to improve the SOM, we will provide a template to constrain its content to methods and data descriptions, as an aid to reviewers and readers. We will also ask authors to provide a specific statement regarding the availability and curation of data as part of their acknowledgements, requesting that reviewers consider this a responsibility of the authors. We recognize that exceptions may be needed to these general requirements; for example, to preserve the privacy of individuals, or in some cases when data or materials are obtained from third parties, and/or for security reasons. But we expect these exceptions to be rare.

As gatekeepers to publication, journals clearly have an important part to play in making data publicly and permanently available. But the most important steps for improving the way that science is practiced and conveyed must come from the wider scientific community. Scientists play critical roles in the leadership of journals and societies, as reviewers for papers and grants, and as authors themselves. We must all accept that science is data and that data are science, and thus provide for, and justify the need for the support of, much-improved data curation.

— Brooks Hanson, Andrew Sugden, Bruce Alberts

10.1126/science.1203354

