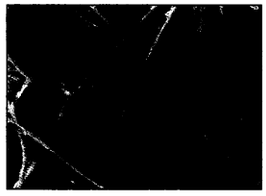
various overlapping degrees, on basic research advances in materials science, immunology, and developmental biology, including the biology of stem cells, cell adhesion, and molecular signals regulating cell-cycle control, differentiation, and pattern formation.



D. MOONEY/PHOTO RESEARCHERS

Bioengineered collagen

Despite the tremendous interdisciplinary nature of bioengineering, most bioengineering programs are built almost exclusively around an engineering–medical school relationship. While this combination has been a natural one, it tends to ignore important research and educational contributions made by investigators in other basic science departments. One example of this has been the absence from many bioengineering programs of research on models of animals that have unique powers of regeneration, such as flatworms and amphibians. Basic research on these animal models needs to be brought into the mix, because it will contribute to our ability to achieve a wider range and greater degree of regeneration of mammalian tissues in vivo, as well as help provide answers to fundamental biological questions. And we need a wider vision of bioengineering collaborations that includes biology, chemistry, and physics departments as full partners if we are to maximize our success. We will be more successful in this endeavor if we do not create a separate bioengineering institute within the National Institute of Engineering.

David L. Slocum
School of Science, Indiana University–Purdue University, Indianapolis, IN 46202–3272, USA

References

1. D. L. Slocum, *Science* **276**, 15 (1997); in *Cellular and Molecular Basis for Regeneration: From Invertebrates to Huyiztitis*, P. Ferretti and J. Geraudic, Eds. (Wiley, Chichester, UK, 1998), pp. 411–450.

Searching the Web, Continued

With regard to the report "Searching the World Wide Web" (3 Apr., p. 98) by Steve Lawrence and C. Lee Giles, I think it's important to quantify—with the use of well-designed studies—the growth of the Web and the effectiveness (or not) of the popular search engines.

Lawrence and Giles do not ask what might be a key question: Exactly what makes up the "indexable Web"? When they say that "Scientific information retrieval and literature search, previously dominated by librarians, is now directly available to a widespread group of scientists," they do not acknowledge that most of the scientific literature is copyrighted by the authors or by the publishers of the major journals and that copyrighted material is not generally available in the "publicly indexable Web." The full text of their report, for example, is not available at *Science* Online, unless the user (or an institution) pays for an electronic subscription to *Science*. Most of the full-text science literature now on the Web is accessible through fee-based systems only, such as PROQUEST DIRECT (UMI Corp.), Web of Science (Institute for Scientific Information), Academic Press Electronic Journals, and so forth. The contents of these database products (thousands of citations, abstracts, and full-text journal articles) are not accessible to the commercial Web search engines (such as Altavista or Hotbot).

Many researchers may self-publish preliminary or "unofficial" versions of their articles on their personal Web pages, but I wonder if these versions are citable by those who may come across such a paper while searching the "indexable Web." Do authors sign agreements with journal publishers that restrict them from also making

their research available elsewhere on the Web? Would a professor accept a student paper with only Web citations, or must the student seek out the published, peer-reviewed literature?

The Web definitely has the potential to make scientific literature much more accessible and affordable to researchers and institutions, but in order for that to happen, the criteria used to evaluate faculty will have to expand beyond the requirements to publish in specific journals to include electronic publications. Even then, one should not expect that copyrighted materials will be freely available to the public (should they be?) through the Web and that they will be retrievable by means of the major commercial search engines.

The report quantifies what the Web actually "is" at a specific point in time. It is important, however, to ask questions about Web content and why something is or is not there at the time of a search. To do this, we must consider political, economic, social, and cultural factors—things that are hard to quantify. The Web has done a bit to raise consciousness of the socio-economic aspects of information, but much Internet hype also creates false expectations about the availability and accessibilty of information. Scientific and technical literature has always been the most expensive (and thus the hardest to get) information sector. The Web alone cannot change this. But maybe conscientous researchers, authors, and librarians can.

Molly Molloy
New Mexico State University Library, Las Cruces, NM 88003, USA, E-mail: mmolloy@lib.nmsu.edu

Response:

Molloy raises a number of issues about the availability of scientific information on the "publicly indexable Web" (pages that the major Web search engines index). We note that scientific information encompasses more than journal publications. A diverse range of scientific material is available online, often freely, including scientist, university, and project homepages, preprints, technical reports, databases (for example, gene sequences, molecular structures, and image libraries), and teaching resources.

All will agree that the amount of scientific material and the number of scientific and technical journals available on the Web is growing, not diminishing. The number of (citable) scientific articles available on the publicly indexable Web is also growing. The Los Alamos e-Print archive (http://xxx.lanl.gov/) is an active site for physicists to distribute information. Some journals owned by societies such as the Institute of Electrical and Electronics Engineers (the largest technical-scientific society—more than 300,000 members), the Association for Computing Machinery (ACM), and the Society for Industrial and Applied Mathematics are permitting papers to be placed on authors' Web sites with the proper copyright notices. Other publishers—MIT Press, for example—are doing the same. Some commercial publishers permit prepublication Web posting of manuscripts with appropriate copyright information, but do not permit posting of the final published version.

Although our report is not directly available on the publicly indexable Web because of *Science* policy, it is possible to obtain a copy by searching the publicly indexable Web. We searched for it on 10 May 1998 (+"searching the world wide web"+ lawrence+giles). Two of the six engines we studied returned links to one of the two pages that provide an e-mail address for requesting a reprint.

Many articles are not available on the publicly indexable Web, which raises several questions. How does one find out about scientific information that is not part of the publicly indexable Web? Will scientists start to publish in mostly those journals that give authors Web permissions, thus allowing wider availability of their work? Will more journals give authors limited self-use? What is interesting is that many scholars are ignoring restrictions and making their papers freely available on the publicly indexable Web anyway. In the long run, this may contribute to the demise or loss of impact of certain journals.

Access to scientific articles on the publicly indexable Web is improving. However, comprehensive search of online articles is limited by the fact that many articles are in formats other than HTML, and many are only available through specialized search engines. A valuable tool would be a scientific search engine that indexes non-HTML documents on the Web, such as Postscript or PDF articles [such a tool, which further incorporates autonomous citation indexing, is being developed at the NEC Research Institute (*1*)]. Another would be a "scientific metasearch engine," similar to tools like MetaCrawler, which searches scientific databases on the Web. Such tools could work in agreement with publishers, allowing the indexing of material that is not freely available.

Steve Lawrence
C. Lee Giles
NEC Research Institute, Princeton, NJ 08540, USA, http://www.neci.nj.nec.com/

References

1. C. L. Giles, K. Bollacker, S. Lawrence, "CiteSeer: An automatic citation indexing system" (Third ACM Conference on Digital Libraries, ACM, New York, 1998), pp. 89–98.