

helped to organize a meeting at Argonne National Laboratory which triggered subsequent meetings at NIH and elsewhere that resulted in the PSI.

All three agencies support the Protein Data Bank, which is managed by the Research Consortium for Structural Bioinformatics to provide access to the more than 10,000 known protein structures. The addition of 10,000 new structures through the PSI will provide the information needed to define the several thousand key protein folds that, in turn, should enable classification of proteins into functional categories. The contributions of each of these agencies (and those of several non-federal organizations) will be needed to enable the NIGMS initiative to succeed.

Marvin Cassman*, **Ari Patrinos†**

*Director, National Institute of General Medical Sciences, National Institutes of Health, Bethesda, Maryland 20892-6200, USA

†Associate Director for Biological and Environmental Research, Department of Energy, 19901 Germantown Road, Germantown, Maryland 20874-1290, USA

Search is on for better search engines

Sir—Steve Lawrence and C. Lee Giles stated in their Commentary that most of the popular search engines index only about 7–16 per cent of the World-Wide Web¹. This is alarming, as many scientific web pages containing important data may never be discovered. As the web grows it is going to become increasingly difficult for general search engines to give comprehensive coverage. The answer to the problem could be the development of subject-specific search engines able to cover most of the contents within that subject.

Most currently available subject-specific lists and indexes are maintained by humans. Many of them are merely collections of web addresses and lack context-based relevance ranking and retrieval of results in multivariate combinations. What is needed are search engines that could traverse through pages at the last level in a subject-specific website. They would be able to do this as the numbers of such sites would be within manageable limits. Crawlers or robots traversing through such a subject-specific web subset could build up a comprehensive and complete bank of keywords. In turn, such keyword-mounted crawlers would efficiently and more frequently screen the last-level page of the site.

Subject-specific search engines would be able to maintain the freshness of the hits, as the crawlers would check a manageable number of specific web pages more

frequently than they would by moving through the entire web.

Mike Gardner² has rightly suggested that we need science-oriented search engines with sets of scientific metadata, as metadata are the key to better searching. Together with that proposal, an approach similar to peer-reviewing of scientific publications could be applied for categorizing and evaluating web pages based on their content, quality and subject-specificity. It would be feasible to use algorithms and rule-based expert systems to check content-richness, subject-specificity and freshness-based context-relevance ranking for retrieved results.

Millions of dollars need to be invested in developing search engines. This investment could be cost-effective if it resulted in an almost zero noise-to-signal ratio and precise but comprehensive subject-relevant hits. The development work should be done in the academic sector, but the completed search engines would have commercial potential, and would generate more revenue than general search engines because of their subject-specificity.

Development of subject-specific search engines would satisfy the growing demand for the latest, precise, value-added, noise-free hits with a high level of subject relevance. Many such search engines together would be able to index most of the World-Wide Web.

Vishwas Chavan

Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India

1. Lawrence, S. & Giles, C. L. *Nature* **400**, 107–109 (1999).
2. Gardner, M. *Nature* **401**, 111 (1999).

Japan builds bridges to rest of the world

Sir—Your Opinion and News article about research in Japan raise an important and timely issue (*Nature* **401**, 309 & 314; 1999). A focus on higher quality rather than quantity of scientific output, and more emphasis on cooperation with researchers abroad, are indeed important goals for Japanese science. But unfortunately the articles did not compare the Japanese situation with that elsewhere.

You report, for example, that 40 per cent of researchers submit papers only to Japanese journals, which is said to reflect insularity. But how can the reader draw such a conclusion without knowing comparable numbers for other countries? I would not be surprised if 40 per cent of researchers in the United States, the United Kingdom or Germany published only in national journals. And you did not mention that a considerable number of Japanese journals publish articles only in

English, which is surely an attempt to reach readers abroad.

It is interesting to learn that 40 per cent of researchers submit papers only to Japanese publications, while 34.1 per cent also submit to international journals. Does that mean that 25.9 per cent only publish in overseas journals? If true, this would show that there are many internationally minded Japanese researchers.

You do a good job of pointing out weaknesses and important goals for Japanese research. But your omission of comparative numbers for elsewhere in the world does not allow readers to see the problem in its true proportions. It could even be the case that your articles inadvertently promote prejudices about Japanese science, which would be unfortunate at a time when many Japanese research teams are striving to attract foreign scientists.

Philippe Buhlmann

Department of Chemistry, The University of Tokyo, Hongo 7-3-1 Bunkyo-ku, Tokyo 113-0033, Japan

Galileo had accurate vision of the Moon

Sir—Martin Kemp writes that “the details of Galileo’s cratered Moons are difficult to align precisely with actual features”¹. I would like to draw readers’ attention to the work of Ewen Whitaker², a selenographer at the University of Arizona. Whitaker reviewed earlier efforts at identifying the features drawn by Galileo, and took special note of the important contributions of Guglielmo Righini, Owen Gingerich and Stillman Drake. He provided side-by-side comparisons of Galileo’s drawings with modern photographs taken at the same lunar phases, and these provide striking support for his contention that Galileo did indeed observe, and record, very accurately.

Whitaker had examined the copperplate engravings of the first edition of Galileo’s *Siderius Nuncius* and seven manuscript images, and he commented on “previously unnoticed differences between the manuscript and printed versions of *Siderius Nuncius*... In the original edition these engravings present a reasonably well-executed appearance, but subsequent editions utilize woodcuts, and the quality deteriorates very rapidly to the point where they are virtually unrecognizable as Moon images. Some of the disparaging remarks made about the drawings undoubtedly stem from examinations of these cruder images”.

M. W. Friedlander

Department of Physics, Washington University, St Louis, Missouri 63130, USA

1. Kemp, M. *Nature* **401**, 116 (1999).
2. Whitaker, E. J. *Hist. Astron.* **9**, 155–169 (1978).