The Speed of Write

by Gary Stix, *staff writer*

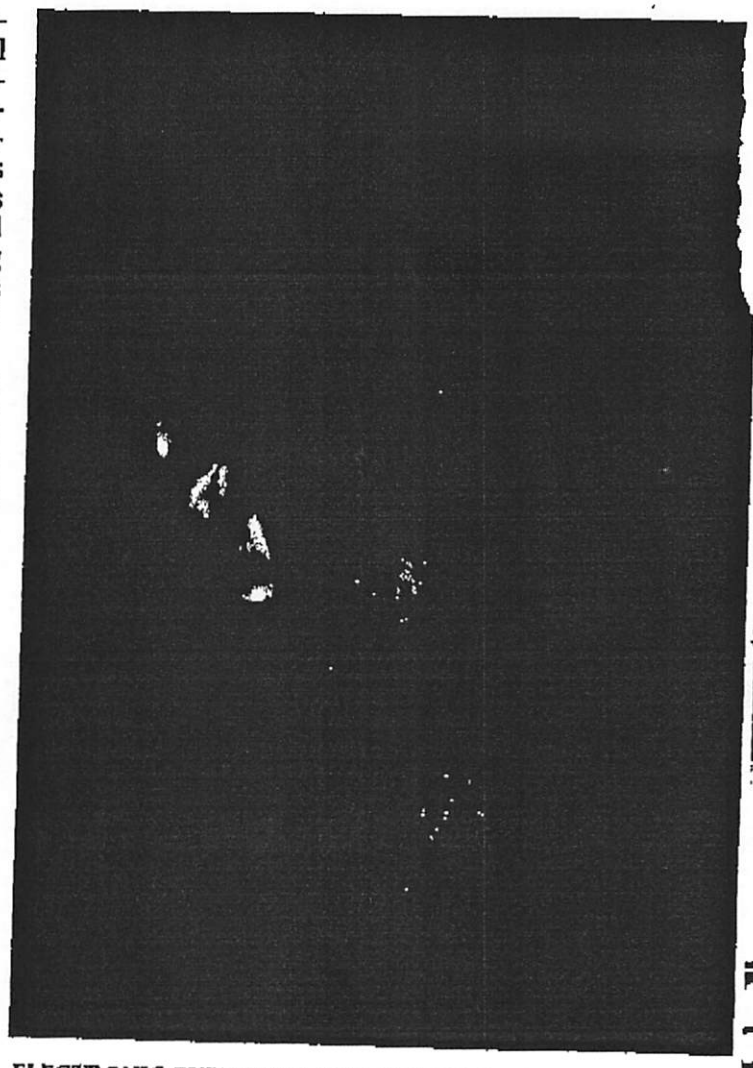
Under a table in a corner of Paul H. Ginsparg's office at Los Alamos National Laboratory, surrounded by piles of scientific papers, a squash racket, dated magazines and the occasional soda can, sits a machine that is quietly changing how science is being done. To the casual observer, it looks like a spare workstation, a Hewlett Packard 9000 not unlike those found in engineering laboratories throughout the country. Yet tens of thousands of scientists around the world pay attention to it every day.

Running on this machine is the product of a project that the bearded 38-year-old physicist has been working on in his spare time for the past three years. Ginsparg, whose loves in life extend to string theory and two-dimensional gravity (reduce space to one dimension, add gravity and stir), has written software that lets other researchers transfer to his computer by way of the Internet the unrefereed copies of research papers, or preprints.

That computer, known by its cryptic Internet address, "xxx.lanl.gov," has become, in effect, a daily wire service for high-energy physics theorists as well as researchers from more than 10 other disciplines, primarily in the physical sciences and mathematics. Every day 20,000 or so electronic-mail messages carry the abstracts of new papers stored in the computer's databases to more than 60 countries. Readers of the summaries then download thousands of copies of the full papers. "It has completely changed how people in the field exchange information," says Steven B. Giddings, an associate professor of physics at the University of California at Santa Barbara. "The only time I look at the published journals is for articles that predate the Los Alamos physics databases."

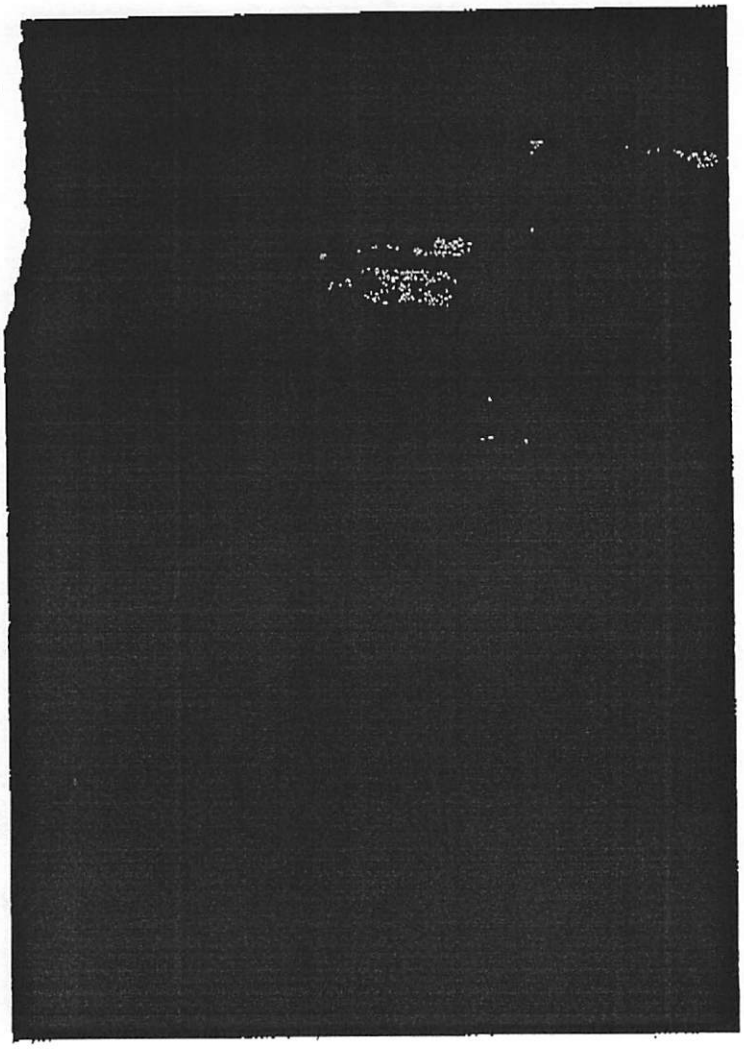
The under-the-desktop machine at Los Alamos represents just one manifestation of what may be a major shift in the way scientists and engineers communicate with one another. The Internet—comprising about 40,000 connected computer networks—has become the world's biggest blackboard.

On subjects from cold fusion to a purported proof for Fermat's Last Theorem, the Internet serves as the medium of choice for conveying news flashes, gossip and quibbles about scientific findings both serious and whimsical. More than just scuttlebutt and self-advertisements travel the network. The Los Alamos machine may represent the future of scientific and technical publishing. If Ginsparg can channel the flood of preprints through some form of peer-review process—one of his next extracurricular undertakings—the traditional paperbound physics journal could become an anachronism.



ELECTRONIC PUBLISHER AND EDITOR Stevan Harnad has become a proponent of low-cost dissemina-

Scientists now transmit reports of their research --from first inspiration to final result-- over electronic networks. Even live experiments can be witnessed on-line. Publishers and libraries may never be the same



***tion of scientific and other scholarly work over the Internet. His electronic journal is* Psycoloquy.**

Established publishers could find their livelihoods threatened, if they fail to react quickly.

These developments suggest that scientific communication is becoming less a historical account and more a live record of how thinking on a research problem evolves. It can transcend the bureaucratic publication process of the paper journal in which an article appears months or even years after researchers have moved on to their next project. "Paper is necessary and important but depressingly and maladaptively slow," says cognitive psychologist and electronic journal publisher and editor Stevan Harnad. "A scholar's work could be inspired to much greater heights if it had a faster response, closer to the speed of thought."

In this view, software running on geographically dispersed computers will be able to link together an entire research corpus, from electronic-mail notification of first results to the actual peer-reviewed article to any commentary that follows.

As this phenomenon emerges, the definition of scientific collaboration may change. Commentators on articles virtually become members of research teams. The network even opens the possibility of broader participation in experimental activity itself. Investigators from throughout a discipline can witness an experiment as it takes place—and register their comments for future perusal by other workers.

The growth of the Internet as a means for researchers to distribute their findings is particularly fortunate in light of what some observers think may be a crisis in the traditional system for scientific publication. As scientific disciplines grow and splinter into subdisciplines, each new area of study requires its own journal. Libraries are having trouble keeping up with this information glut.

The wealth of scientific information doubles about every 12 years; stacks of journals can dwarf even a national monument [*see illustration on next page*]. In a widely distributed electronic paper on the demise of the traditional journal, Bell Laboratories mathematician Andrew M. Odlyzko makes the extraordinary estimate that nearly half of the one million papers ever published in mathematics have found their way into print only during the past 10 years.

The costs of producing the journals to hold all these papers are recovered from a tiny subscriber base that consists mostly of university libraries. There are sometimes only a few hundred subscribers, so prices remain high. Journal publication, moreover, can be painfully slow: it can take a year or more to get into print.

The cost and time lag have become a bottleneck for scholarly communication. Subscription

prices for journals more than doubled from the 1985-86 to the 1992-93 academic year [*see right illustration below*]. Percentage increases well outpaced both inflation and library budgets. A year of some journals can now cost about as much as a Ford Escort. Librarians have cut back on subscriptions over the objections of faculty and graduate students. "There is a process of deacquisition by dozens of libraries," says Nobelist Joshua Lederberg of the Rockefeller University, a longtime proponent of electronic distribution of scientific literature. "Soon, one institution will get a publication, and everyone else will borrow it by interlibrary loan. It's an absurdity. There'll be just one copy, and you'll have to pay $1 million for a subscription."

An alternative has emerged. Professors and researchers can compose and typeset their own papers and then distribute a periodical directly through their institution's connections to the Internet, thereby circumventing traditional publishers. The 1994 edition of the *Directory of Electronic Journals, Newsletters and Academic Discussion Lists*, published by the Association of Research Libraries, listed 440 electronic journals and newsletters, up from 110 in 1991, for the first edition of the directory. During the past two years, the number of peer-reviewed titles has quadrupled to about 100. "The figures tell us that in the real world where 'hard choices' must be made, scholars and scientists are in fact mastering the challenge of providing high-quality scholarly information to a very wide audience at no or virtually no charge to the end user," writes Ann Okerson of the Association of Research Libraries.

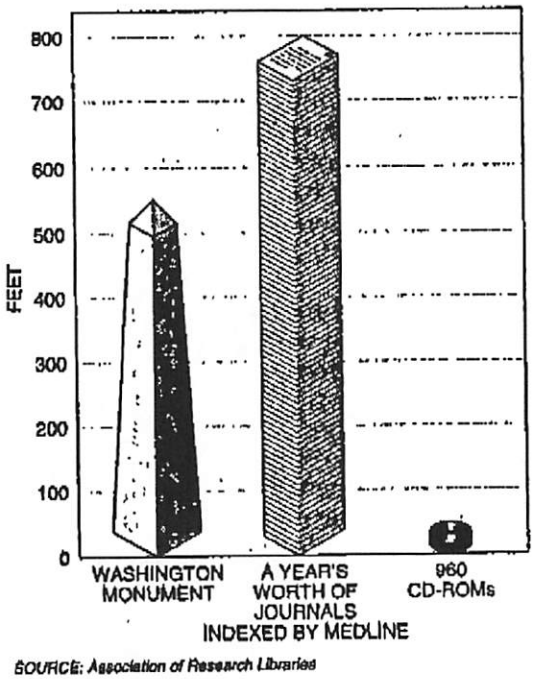
With names like *Postmodern Culture*, most of these journals address themes in the humanities, not the sciences. But scientists and engineers are starting to take their turn. Neil J. Calkin, a professor of mathematics at the Georgia Institute of Technology, and Herbert S. Wilf, a professor of mathematics at the University of Pennsylvania, decided to become members of this new breed of publisher after Wilf read Odlyzko's writing on the demise of traditional journals. The peer-reviewed *Electronic Journal of Combinatorics* has published since April, gaining enthusiastic backing from a number of prominent mathematicians. Calkin estimates that besides his time, which might be donated anyway if he were running a print journal, the project consumes only 25 megabytes of computer hard-disk storage space (worth less than $50).

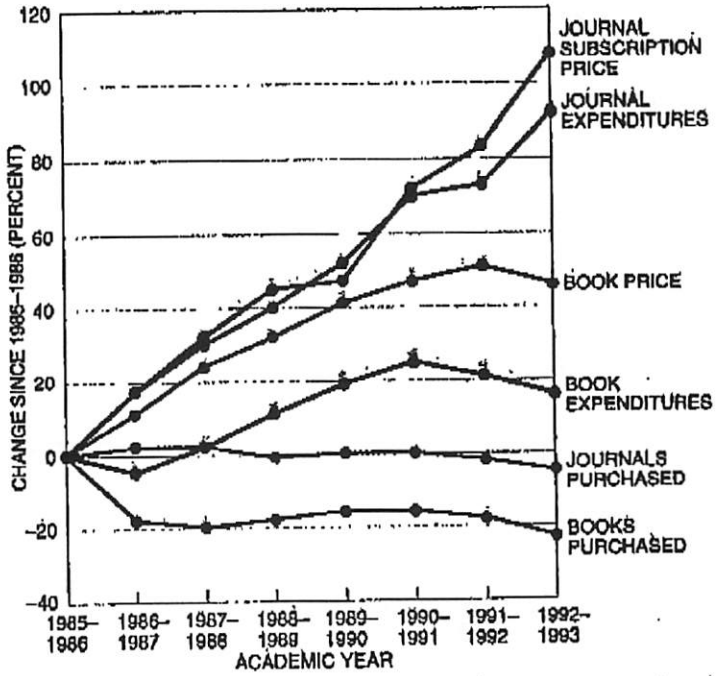
Electronic publications benefit as well from a growing number of public-domain software tools that make it simpler to display complex equations, chemical models and other graphics. Three English researchers, for example, have proposed a set of technical standards that would let pictures of molecules be transmitted over the Internet as readily as an E-mail text message.

Henry Rzepa, a computational chemist at Imperial College in London, Benjamin J. Whitaker of the University of Leeds and Peter Murray-Rust of the pharmaceutical maker Glaxo have worked on this dial-a-molecule project. Rzepa made his contribution because of the difficulty he had in trying to act as a referee of scientific papers without being able to visualize the results of the work presented.

The proposed standards, called Multipurpose Internet Mail Extensions, or more simply MIMEs, specify a uniform format for sending chemical data over a network. Examples of these transmittals might be the spatial coordinates for display of a molecule or the mathematical variables for its spectral analysis. A chemical-modeling software package can then use a MIME to display, manipulate and even annotate a molecule. "Chemists and biologists can now exchange molecules in a virtual sense, rather than just words," Rzepa says.

The so-called E-journal movement



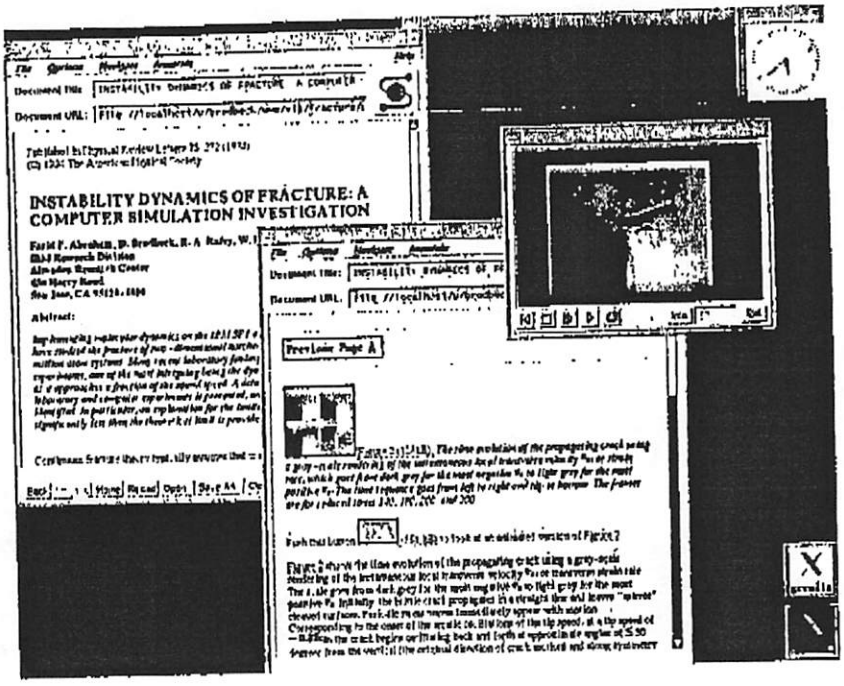


SOURCE: Association of Research Libraries

TOWER OF JOURNALS beside the Washington Monument shows a pile of these publications referenced in the Medline bibliographic database of biomedical literature during just one year. Capturing this material on CD-ROMs or in an on-line database could help libraries cope with the deluge of information.

LESS FOR MORE is a persistent trend in subscription and book purchasing at major university libraries. Costs for journals more than doubled during the past eight academic years, whereas the number of subscriptions actually dropped. Yet publishers continue to bring out more journals.

MULTIMEDIA JOURNALS are presaged by projects of individual scientists. Researchers at IBM placed on the Internet an article they had published in *Physical Review Letters*; they added two computer simulations of a crack getting larger.



could readily be dismissed as the biggest vanity press effort of all time. Few on-line journals have yet achieved the prestige and standing that translates into a line on a résumé for securing tenure and promotion. Calkin, in fact, acknowledges that he will also have to publish in well-known print journals to further his career.

The egalitarianism and "anything goes" sensibility of the Internet could work against attempts to raise the status of E-journals. The Usenet, a vast miscellany of bulletin boards accessible to the network's millions of users, represents the antithesis of the selectivity exercised by the editors and referees of a journal. The unfettered discussion there ranges from genetic experiments carried out by UFO aliens to legitimate debate about an advance in robotics or gene therapy. Some academics fear that the sheer volume of literature and a growing inability to distinguish the good from the bad in what gets published may lead to an overall decline in standards. "There will be far more trash than there is today," says Frank S. Quinn, a professor of mathematics at the Virginia Polytechnic Institute and State University.

Imposing a rigorous system of quality control could change these perceptions. Perhaps the most vocal proselytizer of the network as a scholarly vehicle is Stevan Harnad, a Hungarian-born cognitive psychologist who is a profes sor at the University of Southampton in England. He is best known as the founder and editor for more than 15 years of *Behavioral & Brain Sciences* (a journal whose acronym, *BBS*, is appropriately enough the same one used to designate computer bulletin-board systems).

At *BBS*, an article first goes through review by five or so referees drawn from at least three different disciplines, which may include behavioral biology, computer science, linguistics, philosophy, robotics or machine vision. If an article passes muster, it then goes to up to 100 people in those fields for a critique. As many as 30 commentaries appear with the original submission.

BBS bears remarkable similarities to the facile exchanges that take place over computer bulletin boards--with one big difference. Harnad is no populist. Unlike Internet evangelists who view the network as the ultimate equalizer for dismantling hierarchy, Harnad is an unabashed academic snob. The best thinkers in a field, he believes, should have access to one another, undisturbed by the noise of crowds milling outside the ivory tower.

Global Graffiti Board

Despite this elitist stance, Harnad discovered the potential for electronic scholarly discourse during the mid-1980s in, of all places, an electronic bulletin board on artificial intelligence on the Usenet. He found that the frenzy of activity on this "global graffiti board for trivial pursuit" spurred his own thoughts on why the mind does not behave like a computer. This experience resulted in a paper he later published in the journal *Physica D* as "The Symbol Grounding Problem." "I realized that I came on this as a result of jousting with pygmies," Harnad says. "I asked myself what would happen if the people I was talking to had been real scholars."

Harnad coined the phrase "scholarly skywriting" to describe the speed and breadth of dissemination that would result if academics eschewed the paper medium to publish over the network.

He practices what he preaches. Five years ago Harnad started an electronic journal, *Psycoloquy*, that is identical to *BBS* in the topics it covers and in its editorial practices, except that the papers are somewhat shorter. The journal is free; Harnad supports it on a $15,000-a-year stipend from the American Psychological Association that covers the expense of a part-time editorial assistant and a clerk to maintain a mailing list. The journal is sent to computers all over the world and is even available on the Usenet. (Amateurs, Harnad says, should be allowed to observe and emulate scholars' work from the bleachers.)

This past summer Harnad sent to an electronic mailing list a "subversive proposal" that suggested that authors and institutions should try to undermine the present publishing system by posting preprints to the network. This action would, he hopes, provoke the type of mass migration to the electronic medium from throughout academia that has already been experienced in the high-energy physics community.

At about the same time, Harnad, Lederberg and others met with a high-ranking UNESCO official for a broad-ranging discussion of electronic distribution of biomedical information to developing nations. They suggested, among other recommendations, that the agency endorse low-cost electronic publishing of scientific articles. "I think this will get the ball rolling," Lederberg says. "If you get a few institutions doing this, then others will feel that they have to be on the bandwagon."

Commercial scientific publishers are feeling the pressure to enter the electronic arena. Many, however, have yet to determine ways to smooth the transition. Simultaneous print and electronic distribution could, if anything, cause subscription rates to rise. Illustrations, graphs and equations do make scien-

tific publishing more difficult than providing an all-text database. More than a decade after the first market tests of electronic newspapers, many scientific editorial houses are drafting their plans for moving into the new medium. "Publishers are still trying to figure out how they can recover their costs," says David Rodgers, a research scientist at the University of Michigan and the former director of electronic products and services for the American Mathematical Society.

Both professional societies and for-profit science and technical publishers have started to piece out initial strategies. The American Physical Society, for one, has plans to circulate the electronic *Physical Review Letters* along with the printed version.

Reed Elsevier, the $4.3-billion British-Dutch company, recently made a major commitment to electronic publishing by buying the Lexis legal database and the Nexis news service. Yet the company is only beginning to make its 1,100 science and technical journals available on-line. It has an experimental program to provide image files of its 43 materials science journals to nine universities. Next year it will test-market all its journals in an electronic format but as yet has no firm plans as to when its entire suite of publications will become available as fully searchable databases.

In January, Elsevier Science, the company's science journal publishing division, did unveil its first electronic journal. The event was an indirect tribute to the work of Paul Ginsparg. *Nuclear Physics Electronic* (NPE) is an on line compilation of already reviewed papers that are being readied for print in *Nuclear Physics*, a $12,000-a-year set of paper journals.

Elsevier had considered creating this type of on-line system for a decade, given that high-energy physicists were distributing paper preprints of their work through the mail. But physicists told Elsevier they were not interested—until Ginsparg's database began to gain a following. "Clearly, the market was ready; Ginsparg more than proved that," says Karen Hunter, an Elsevier vice president.

Many of the non-peer-reviewed versions of the papers that appeared in *NPE* are available in virtually identical form on the Los Alamos computer long before they appear on the Elsevier system. The range of topics on the Los Alamos databases is much broader.

Elsevier has not yet requested that papers be withdrawn from the Los Alamos databases once they appear in their copyrighted form in *NPE*, an action that would surely alienate physicists. For his part, Ginsparg may eventually need the help of mainstream publishers—he will have difficulty instituting a peer-review system for the papers in his databases while remaining a practicing physicist. He did, in fact, meet with officials from the American Physical Society in mid-October to discuss how a review process for the databases might work. If Ginsparg is brought into the fold, though, journal publishing could still change: costs could remain dramatically lower than for paper journals; reviews could proceed more quickly and with a broader base of contributors to evaluate manuscripts.

The library, the usual repository for journals, would also be affected by these shifts. The computers that store electronic publications may reside in a laboratory, not a library building. The librarian will remain, but the job will not be the same. Librarians, along with publishers, have traditionally been entrusted with the responsibility to make order out of scholarly chaos. Yet so far computer scientists, not librarians, have usually been the ones to create the tools for navigating the Internet.

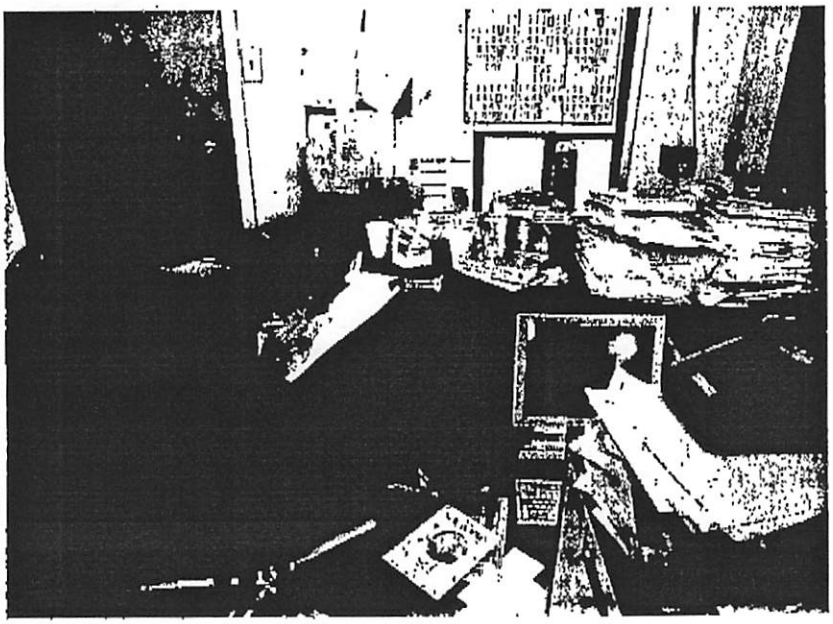
Career Changes

Daniel E. Atkins wants the School of Information and Library Studies at the University of Michigan to turn that trend around. Atkins spent the better part of his career tweaking high powered computers as a professor and dean of engineering before coming to his present position as dean of the library school in 1992. His current mission is to redefine the job of the librarian. Atkins envisages an "information professional" who will combine the skills of the computer scientist, the business graduate and even a little of the old-school librarian.

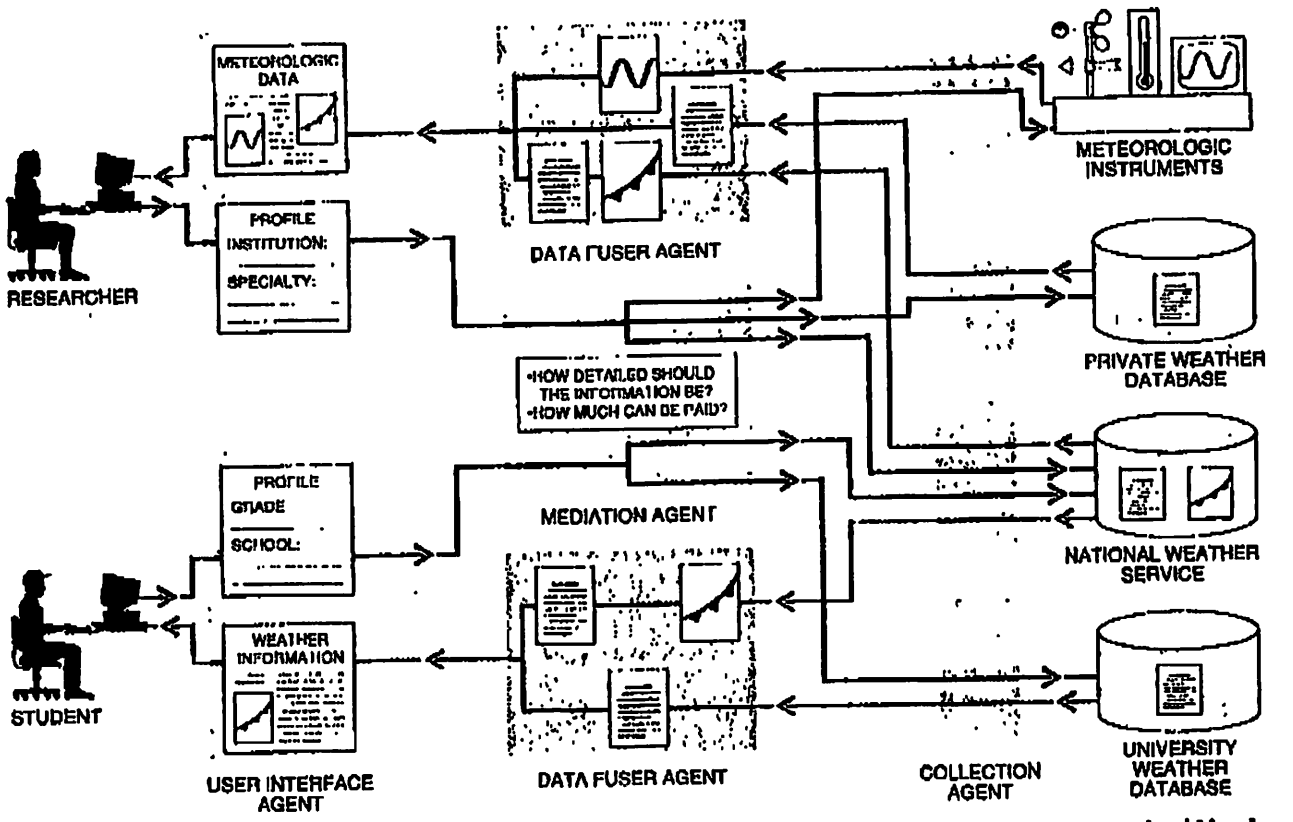
One of the jobs of the new librarian will be to help make sense of the labyrinth of different information sources available on the Internet. "You can waste 24 hours a day browsing," Atkins comments.

In September the University of Michigan was awarded one of six grants in a digital library initiative sponsored by the National Science Foundation and two other federal agencies. Michigan's grant is for the development of technology to help students and scholars wade through the limitless store of information available on the earth and space sciences.

The NSF program also includes the New York Public Library, Stuyvesant High School in New York City, the Ann Arbor public school system and a number of corporate participants, among them Elsevier and IBM. Atkins and team will attempt to exploit "agents," the much publicized (though still rarely implemented) software tools from artificial-intelligence research that can retrieve information or perform other automated tasks. These agents would try to replicate in software some of the



DESK-BOTTOM PUBLISHER Paul H. Ginsparg of the Los Alamos National Laboratory sits in his office beside a Hewlett Packard 9000 containing a series of databases that have become, for a growing number of physical scientists, a substitute for the paper journal.



DOING THE BIDDING of research scholar or student, a network of software "agents" under development at the University of Michigan is intended to operate like an experienced reference librarian. After receiving a question—"What is the weather in Travers City?"—the user interface agent extracts a profile of the user from information in its database. The request is channeled to the mediation agent, which evaluates the profile to determine how detailed the information retrieved should be and how much the user can pay. The mediation agent decides what information resources should be allocated to what type of user. The researcher may get access to real-time meteorologic data and both public and private weather databases, whereas the student is relegated to less expensive public information sources. A collection agent then executes an actual search of the relevant databases. A data fuser agent assembles the collected information—text, maps, graphs, raw data—before relaying it to the user agent for display as a document on the student's or the researcher's screen.

functions of a reference librarian [*see box above*].

The digital library project is also exploring ways of combining information retrieval with laboratory tools accessible to researchers over the Internet. Atkins heads another NSF-funded project, the Upper Atmospheric Research Collaboratory (UARC), that has been examining how scientists can make use of research tools located in remote environs. A "collaboratory" uses electronic networks so that investigators around the globe can witness the results of an experiment as it progresses. Besides space physics, collaboratory projects are being pursued in molecular biology and oceanography—the Human Genome Project and a decade-long research effort to characterize the El Niño ocean currents are two examples.

For two years, the UARC has given scientists at various institutions (SRI International, Lockheed, the Danish Meteorological Institute and the universities of Alaska, Maryland and Michigan, among others) around-the-clock access to one another and to readings from instruments on the southwest coast of Greenland for observation of the solar wind. The software lets researchers in these various far-flung locations exchange comments as they observe the trace lines from incoherent scatter radar and other instruments that follow the impact on the atmosphere from the waxing and waning of the solar wind. The NSF digital library project will allow Atkins and his colleagues to expand the UARC. They will evaluate how the massive quantities of instrument readings, bulletin-board conversations and later annotations can be managed and used by the different research teams.

The goal of the librarian in this project will be to organize this information in a form that will let workers replay an experiment six months later. "We're producing a full fidelity, multimedia transcript of the process," Atkins says.

The electronic network blurs boundaries between experimentalist, author, publisher, reviewer and archivist. As the work of Paul Ginsparg and of Daniel Atkins suggests, there is a need for a new institutional framework—both electronic publishers and collaborators in their collaboratories—to cope with the glut of scientific knowledge and the data from which it is derived.

FURTHER READING

DIRECTORY OF ELECTRONIC JOURNALS, NEWSLETTERS AND ACADEMIC DISCUSSION LISTS. Edited by Ann Okerson. Association of Research Libraries, Office of Scientific and Academic Publishing, 1994.

LOS ALAMOS E-PRINT ARCHIVES. Available on the World Wide Web as http://xxx.lanl.gov/

SJEVAN HARNAD'S ELECTRONIC ARCHIVE. Available on the World Wide Web as http://www.princeton.edu/~harnad/