Space Scientists Split on Proprietary Data Rights

The Hubble Deep Field wowed the world on 16 January 1996. The speedy release to the public of the image, only 17 days after it had been obtained, is still held up as evidence that making data available is good for science.

The question of how long a scientist should have exclusive access to satellite data has been knocking around the astrophysics community for about 20 years. In the early days of space science, researchers got to analyze their data indefinitely before making it public, but now, "one year is about as long as NASA entertains for that," says Alan Bunner, who is in charge of one of the agency's four space science programs. There's an increasing realization, he notes, that the public—including the broader space science community—"deserves to share in discoveries as they occur," and to see where its tax dollars are going.

But conflicting needs make the issue contentious: Individual scientists may lose out by data going public quickly, whereas science as a whole may benefit. Within the space science community, says Goddard Space Flight Center astrophysicist Rob Petre, "it's a sore issue. There are two legitimate sides to the argument, and NASA's problem is finding the middle ground."

The first crack

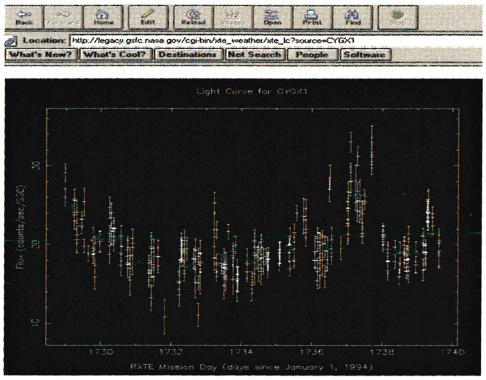
The scientists who think up an experiment and then spend years, even careers, designing and building the instruments should get the first crack at analyzing the data. So goes the argument for granting a proprietary period. Scratching it altogether, says Pennsylvania State University's David Burrows, "would be bad for hardware builders like me." Not much gets published during the building phase, he continues, and with exclusive access, "there was a flood of papers once a mission flew. It's not so clear what will happen in this new regime. And it's a problem for someone trying to get tenure."

Instrument builders are driven by their interest in the science, adds University of Maryland physicist Glenn Mason, "and if you take away the incentive by giving the results to someone else, people won't work as hard." The major danger, agrees Christine Jones, of the Harvard-Smithsonian Center for Astrophysics, "is that people might say, Why should I build this?"

How soon after scientific satellite data are acquired should they be made public?

Other arguments for keeping an ample proprietary period include the fear that accuracy will suffer if researchers race to publish, and the comfort of being able to give a graduate student a project without worrying that he or she will be scooped. And without it, says Fred Walter, an astrophysicist at the State University of New York at Stony Brook, "people at smaller institutions could be hurt. I don't have a big team. The analysis takes awhile."

But that doesn't mean Walter sits on his data: "We scientists thrive on adulation, and that only comes from reporting exciting results. If we find something, and we don't publish, it's awfully difficult to get time for follow-up observations."



A GREAT DEAL OF ASTROPHYSICAL DATA is available on the World Wide Web: The x-ray flux from the black hole binary Cygnus X-1, obtained from the Rossi X-Ray Timing Explorer, for example, is posted several times daily.

Adds Burrows, "I would argue that there is no irredeemable damage done to science by having publication delayed by a year."

Community goods

It's easy to see Burrows's point. But most astrophysicists also appreciate the advantages of quicker data release. They get upset when data is hoarded—as occurred with the Einstein X-Ray Observatory, launched in 1978, and with ROSAT's All-Sky Survey, completed in 1991 (see next story). When France Cordova was NASA's chief scientist, from 1993 to 1996, she spearheaded efforts to write a policy for the agency on proprietary rights. "We said, All things being equal, the data should get out as fast as possible."

The prompt release of data serves science by allowing one set of measurements to guide the next. It can also increase the scientific impact, particularly for surveys and other large data sets, notes Bob Williams, who as director of the Space Telescope Science Institute was the one who released the Hubble Deep Field. Moreover, adds Cordova, who is now at the University of California, Santa Barbara, "when you or your student can't get ahold of data, the advance of scientific knowledge can be stymied."

Cordova also argues that eliminating proprietary time isn't so bad for instrument builders (she is one), who have some other built-in advantages: They know best the subtleties of their instruments, and they typically win funding to analyze the data. "I truly believe that we don't build instruments for ourselves alone, we build them to be part of a larger community," she adds.

To be sure, the degree to which early release of data either helps science or is a threat to the principal investigator depends on the experiment. Sometimes a single image *is* the cutting-edge science, says Maryland's Mason, "such as the first pictures of volcanoes on Jupiter's moon Io. But most of the time, getting publishable scientific results takes a lot of work and a fair amount of money, so it's not really true that putting the data out will lead to getting scooped." Ohio State University's Brad Peterson notes that with sequential measurements, for example, releasing data as they become available, "is likely to frustrate the scientific goal," if people do a rush job on the analysis.

Raw data, notes Ray Arvidson, who runs NASA's planetary and geoscience archiving site at Washington University, are generally "not useful to the scientific community and uninterpretable for the public." And readying data for release, which can include calibrating instrument responses, determining coordinate systems and performing spectroscopic or other analyses, can take three to six months or even longer, he adds. "Take Mars Pathfinder. Images and chemical analyses were posted on the Web immediately, but the data were archived just the other day. It took more than a year, but not because the team was sitting on it. Those are complex data sets." In some other areas of space science, such as high-energy astrophysics, archiving is done by mission teams. Says Stony Brook's Walter, "It is up to the observer to do things like background subtraction. And data are not routinely verified before release because of insufficient personnel."

The push from NASA headquarters for faster release of data, adds Arvidson, "is becoming louder and stronger." That's clear from NASA's recent calls for proposals. For example, bidders for time on the Hubble Space Telescope can now choose to make their data public in less than the usual year, which, Williams says, "might enhance the scientific value of a particular proposal and give it an edge" in the tight competition. For experiments on some of NASA's other spacecraft, the agency's policy now states that: "data are to be released as soon as possible after a brief validation period appropriate for the mission."

And a squeeze on guaranteed experiment time for the scientists who design and build a satellite is also part of NASA's trend toward openness and fairness. Since the early 1980s, beginning with the Einstein X-Ray Observatory, most NASA missions have been open to guest observers, but now there is incentive to waive guaranteed time completely. That's what was done by the planners of the Rossi X-Ray Timing Explorer, a satellite used to study plasmas, neutron stars and other high-energy x-ray sources. Explains MIT's Hale Bradt, one of RXTE's leaders, "Part of the motivation was political, Did the mission really look like it was serving the community? Every drop in the bucket counts when you are trying to get a mission approved." Our policy, Bradt adds, "led to a psychology of openness and sharing, with little distinction between guest and instrument-team scientists. We did this with significant misgivings, but it has proven to be enormously successful scientifically."

Users of RXTE get a year of exclusive access to their data. That, says Bradt, "is as short as one can argue for, I would think. Only for certain special types of data like surveys does immediate release seem important."

Getting a grip

Williams, who favors making data public, still "wouldn't advocate the proprietary period disappearing altogether. We live in a world where rewards are a motivating factor, and people would like to think, if they are going to have creative ideas, they should have a chance to follow them through." What he does advocate is that astrophysicists "make a concerted effort to address what is best for science and scientists in terms of proprietary rights. The community needs to discuss the issue, and get a grip on it."

NASA, adds Ohio State's Peterson, who chairs the agency's Astrophysics Working Group, an informal advisory body, "is very reluctant to make a blanket statement [on these issues]. That's good. Each mission has to be dealt with differently. If you can avoid making sweeping rules, you are more likely to be able to do something sensible."

Toni Feder

Sloan Survey Spurs ROSAT to Release All-Sky Survey Data

No contract has been breached, no agreement has been broken and no one disputes the scientific value of the reams of x-ray data collected by ROSAT (short for roentgen, or x-ray, satellite). But the German team that designed and built the satellite has frustrated the international astrophysics community by not releasing its All-Sky Survey data, which were collected more than seven years ago during ROSAT's first six months in orbit.

The inaccessibility of the data, says Columbia University's David Helfand, "continues to have serious consequences, especially for designing follow-up studies and x-ray satellites." And the University of Alabama's Ray White, who studies galaxy clusters, adds that "without the survey data, it's tough to estimate the x-ray background accurately for objects that are big compared to the field of view."

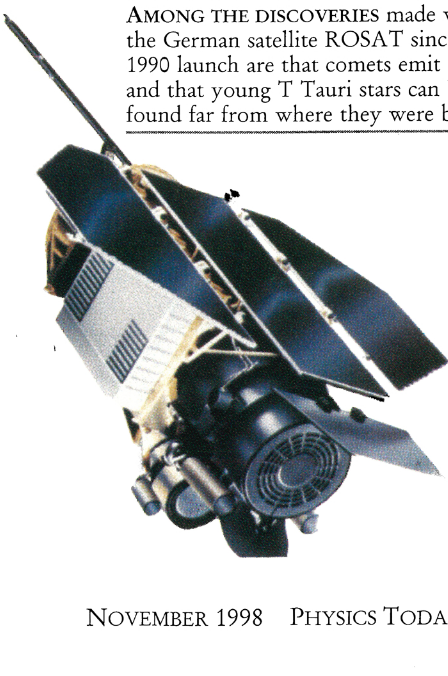
The leader of the All-Sky Survey team, Joachim Trümper, of the Max Planck Institute for Extraterrestrial Physics (MPE) in Garching, Germany, downplays the delays in making the survey data public. A lot has already been released, including a catalog of bright sources, he notes, adding that a faint-source catalog will probably come out next year. The MPE has limited manpower, Trümper emphasizes, and priority goes to new ROSAT data—taken by scientists applying through the American (50%), German (38%) and British (12%) space agencies—which must become public after a year.

In effect, the MPE prefers to take heat for delaying release of its data than for errors that might result from doing so prematurely. Wolfgang Voges, who is in charge of analyzing the survey's results, admits, "People want to see the photon-by-photon data, so they can use their own criteria for selecting sources to study. But we discovered that certain fields had lower exposures than expected. We want to make the data more homogeneous before we release it, so that people won't come to the wrong conclusions."

The data are close to ready, Voges adds, "and Sloan is the one that could force us" to finish up faster. In a proposed collaboration, the Sloan Digital Sky Survey (SDSS)—a vast optical and infrared photometric and spectroscopic survey (see PHYSICS TODAY, February, page 61)—would take spectra to identify x-ray sources discovered by ROSAT's All-Sky Survey. "It is tremendous what we will gain. We will learn so much about large-scale structure," says Voges. Coupling the data, he adds, means that "so many questions can be attacked much better." Per field of view, 20–30 of SDSS's 640 optical fibers would be trained on ROSAT sources.

Bruce Margon, the University of Washington astrophysicist negotiating for SDSS, refuses to comment, but the

AMONG THE DISCOVERIES made with the German satellite ROSAT since its 1990 launch are that comets emit x rays and that young T Tauri stars can be found far from where they were born.



DAIMLER-BENZ AEROSPACE