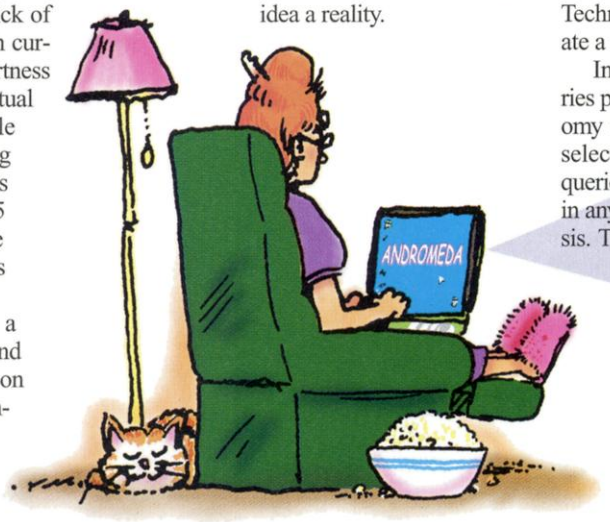
ASTRONOMY

The Virtual Observatory Moves Closer to Reality

Data from decades of observations by dozens of instruments may soon be accessible, changing the way that astronomy is done around the world

If Alex Szalay has his way, the next astronomical observatory won't be built on a remote mountaintop. It won't even have a telescope. Instead, says Szalay of Johns Hopkins University in Baltimore, Maryland, the National Virtual Observatory (NVO) will be an electronic web that gives astronomers access to terabytes of celestial data with the click of a mouse. In addition to eliminating such current occupational hazards as jet lag, shortness of breath, and long, cold nights, the virtual observatory promises to make possible new analyses of the heavens by weaving together information from facilities around the world—and in space. "In 5 years' time, we will have a complete view of the sky in 15 wavelengths," says Szalay. "It will be just so different."

Seven years ago, when Szalay and a few colleagues first started kicking around the idea of a virtual observatory, being on site was still the key to success in astronomy. The Internet was in its infancy, and the concept of an electronic warehouse that would fulfill customer requests hadn't even occurred to Amazon.com founder Jeff Bezos, much less to most of the scientific community. Since then, however, a number of sky surveys have accumulated masses of data, others are well under way, computer and data-handling technology have improved significantly, and pilot projects are laying the foundations for the real thing. As a result, the astronomical version of Szalay's idea has gone big time. This spring the latest report by the National Academy of Sciences (NAS) on priorities for U.S. astronomy over the next decade (*Science*, 26 May, p. 1310) ranked a virtual observatory as the most important small project in astronomy, and last month some 160 astronomers and computer scientists spent 4 days* talking about the nuts and bolts of how to make that idea a reality.



Easy viewing. The virtual observatory will allow astronomers to go online to view the Andromeda galaxy (images, facing page), or any other celestial object, at a variety of wavelengths.

The aim of the NVO will be to pull together data from dozens of telescopes and decades of collecting time. Its use of huge, standardized data sets, collected by survey programs at all wavelengths from radio waves to gamma rays, will make possible large-scale statistical studies. No longer will astronomers have to cope with disconnected hit-and-run observations, all carried out with different instruments and different procedures. "Finally, we will be able to compare apples with apples," says Szalay. Mining large data sets is also expected to turn up the oddballs of the universe—objects so rare that you'd be unlikely to run into them by chance, but all the more interesting for that. In addition, virtual astronomy will allow scientists to study interesting subsets of celestial bodies, such as all objects that are bright at infrared wavelengths, appear fuzzy in visible light, and are not detected as x-ray sources. The combination of these qualities, says Robert Brunner of the California Institute of Technology (Caltech) in Pasadena, "will create a mind shift in astronomy."

Indeed, enthusiasts say virtual observatories promise to change forever the way astronomy is done. They will allow astronomers to select any part of the sky, run complicated queries, and download raw observational data in any particular wavelength for further analysis. This could all be done from an office or home computer, as well as on the road, with a laptop. "It will enable you to do first-rate science, even without access to a large telescope," says Brunner. "This will lead to a true democratization of astronomy."

A driving force behind an NVO is the exponential growth in digital data being collected by numerous sky surveys. Those surveys, in turn, are fueled largely by continuing improvements in charge-coupled device (CCD) detectors, which are getting larger, more efficient, and cheaper. "The number of CCD pixels observing the universe is doubling every 20 months," says Szalay.

Five years ago, a terabyte of data (1 million megabytes, or enough information to fill 1500 CD-ROMs) was huge. Today, it's no big deal. For instance, the Sloan Digital Sky Survey, a U.S.–Japanese effort to map 100 mil-

* Virtual Observatories of the Future, 13 to 16 June, Caltech.

Watch This Space!

Right now, if you type www.skyserver.org into your Web browser, you get an "Under Construction" message. But later this year, the site hopes to put the sky at your fingertips. Initially aimed at a more general audience, SkyServer might become part of a National Virtual Observatory.

The idea is to create a seamless mosaic of the sky at many wavelengths and to provide people with a variety of views of every part of the heavens. SkyServer is being developed by the same Microsoft team that built TerraServer (www.terraserver.org), a Web site giving access to satellite images of Earth. In fact, SkyServer has been described as the TerraServer looking up rather than down. "Astronomy has a special public appeal," says Alex Szalay of Johns Hopkins University in Baltimore, Maryland. "It is an incredibly interesting challenge to develop tools that might excite high school kids."

While TerraServer contains fewer than 2 terabytes of data, SkyServer would be much larger. "It would be easy to fill it up to tens of terabytes," says Szalay. He expects a small prototype of SkyServer to be up and running by the end of the year.

The European ASTROVIRTEL project (www.stecf.org/astrovirtel) will also have educational value, says Piero Benvenuti of the European Southern Observatory. "We could offer the digital sky to any school class," he says. "Students could use it as a telescope and carry out their own projects." Benvenuti plans to pitch the idea to several European funding agencies. **–G.S.**

ILLUSTRATION BY JOE SUTLIFF

lion stars, galaxies, and quasars in five different wavelengths, will eventually contain some 40 terabytes of data. The future Large-Aperture Synoptic Survey Telescope, another priority project in the NAS decadal report, will produce 10 terabytes of data per day. (By comparison, the amount of information in the human genome is a mere 10 gigabytes, or 0.01 terabytes.)

Right now, however, no infrastructure exists to access and combine all those disconnected archives. That's the challenge facing a virtual observatory. It would consist mainly of dedicated software and analysis tools, designed to search through vast amounts of astronomical data, as well as a certain amount of computer hardware, such as high-speed networks. The various archives will be accessed remotely, rather than brought together on a central computer. "The data has to be processed where it is sitting," says Szalay, "since moving 40 terabytes of data over a typical Internet connection takes a couple of years." To devise more efficient schemes to navigate and access the enormous amount of data, Szalay has teamed up with computer scientist Jim Gray of Microsoft in Redmond, Washington, one of the designers of TerraServer, a huge database of satellite images of Earth (see sidebar).

Astronomers say it could cost as much as $100 million over 10 years to create this virtual astronomical warehouse. NASA and the National Science Foundation (NSF), which have historically supported ground- and space-based telescopes, are the most likely federal sources of funding. Both have already sponsored pilot projects. And although the NVO isn't expected to be fully operational until 2005, preliminary versions may be imminent, says Brunner. "We won't close the doors, come out 5 years later, and say: 'Here's the Virtual Observatory, use it,' " he says. Adds Szalay: "I would be surprised if there weren't something up and running by the end of this year."

Although the academy report backs a national project, the initiative lends itself to a worldwide partnership. Already, astronomers at the headquarters of the European Southern Observatory (ESO) in Garching, Germany, are working on a similar project. "This has to be global," says Piero Benvenuti, head of the Space Telescope European Coordinating Facility (ST-ECF) at ESO. "We have already started discussions with our American colleagues." Many European scientists attended the Caltech conference, and next month in Munich ESO will host another international conference, entitled "Mining the Sky."

Benvenuti coordinates the joint ESO/ST-ECF ASTROVIRTEL project, a pilot project funded for 3 years by the European Commission. Through ASTROVIRTEL, European astronomers can access a huge data archive that contains observations from the Hubble Space Telescope, ESO's Very Large Telescope at Paranal, Chile, and the Wide Field Imager at ESO's La Silla Observatory, also in Chile. The first round of "virtual observing" proposals will be selected within the next couple of months.

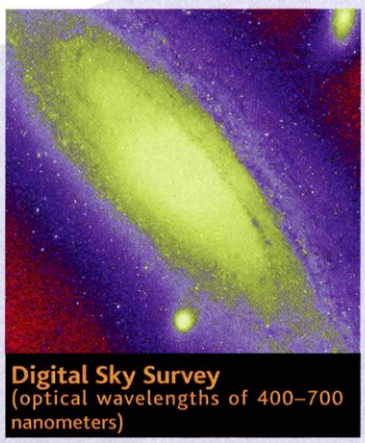
"Our approach is to start with the proposals," says Benvenuti, "and build the software tools from that." A review committee will choose on the basis not just of scientific merit but also on whether the question is general enough to warrant the development of a specific query tool. "In the end," says Benvenuti, "we want our users to think of using the database as if it were an astronomical facility. You can't operate it all by yourself, just as you can't run the Hubble Space Telescope or the Very Large Telescope without expert help." At Caltech, Brunner and Tom Prince have developed their own program called the Digital Sky Project. Funded by NSF and Sun Microsystems, it acts as a small-scale technology demonstrator for the NVO. "We've learned a lot about connecting and searching large databases," says Brunner. "The National Virtual Observatory will certainly be founded on this experience." Another prototype of a virtual telescope is SkyView (skyview.gsfc.nasa.gov), which lets the user look at a particular part of the sky in different wavelengths.

The next step in creating a virtual observatory is working out the structural details. Caltech's George Djorgovski, who together with Szalay, Prince, Brunner, and others wrote a white paper on NVO for the NAS panel, foresees a series of small, focused workshops over the next year or so to address topics such as establishing new data archives, writing standard protocols for storage and query of data and images, and developing a uniform Web-based interface. "We also need to build real partnerships with computer scientists," he says. "In the future, we might end up using virtual reality software techniques that are being developed right now for the next generation of Sony's PlayStation."

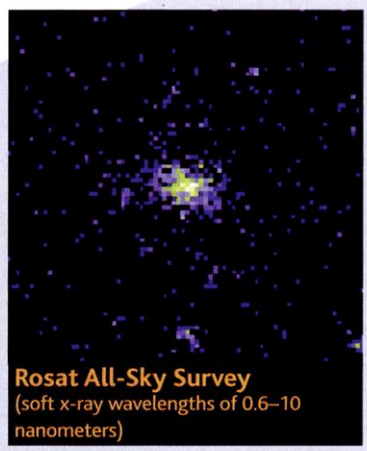
Nobody is predicting that virtual reality will supplant real telescopes, however. After all, an archive has limited value in the study of transient phenomena such as supernovas, variable stars, and asteroids. "The big scopes will not be replaced," he says. "But I think a lot of routine footwork will be shifted from relatively small telescopes to virtual observatories."

–GOVERT SCHILLING

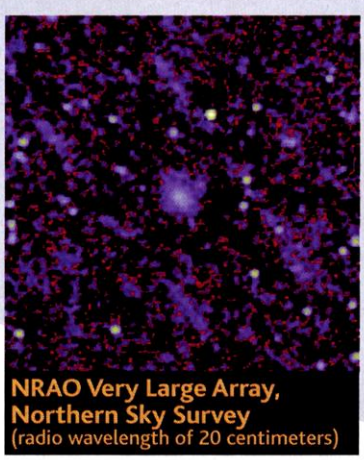
Govert Schilling is an astronomy writer in Utrecht, the Netherlands.



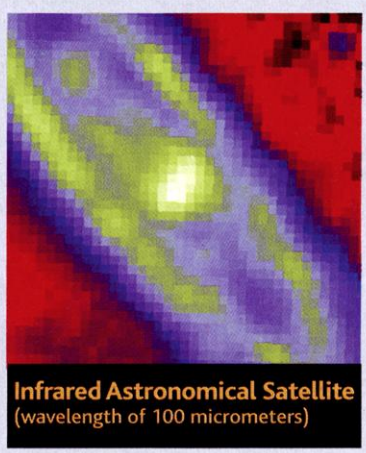
Digital Sky Survey
(optical wavelengths of 400–700 nanometers)



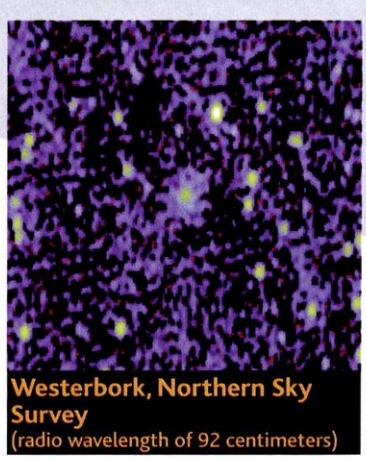
Rosat All-Sky Survey
(soft x-ray wavelengths of 0.6–10 nanometers)



NRAO Very Large Array, Northern Sky Survey
(radio wavelength of 20 centimeters)



Infrared Astronomical Satellite
(wavelength of 100 micrometers)



Westerbork, Northern Sky Survey
(radio wavelength of 92 centimeters)

CREDITS (CLOCKWISE FROM TOP RIGHT): ROSAT TEAM; IPAC/JPL; WENSS TEAM; NRAO; STSCI