Databases and Inter-Connectivity in Ground-Based Astronomy

T. von Hippel and C. M. Mountain
Gemini Observatory, 670 N. A'ohoku Place, Hilo HI 96720

Gemini Preprint #68

DATABASES AND INTER-CONNECTIVITY IN GROUND-BASED ASTRONOMY

T. von Hippel, C. M. Mountain*

Gemini Observatory, 670 N. A'ohoku Place, Hilo, Hawai'i, USA, 96720. Emails tvonhippel@gemini.edu, mmountain@gemini.edu.

ABSTRACT

Optical and infrared ground-based astronomy is undergoing a renaissance. Advances in material technology, system modeling, and the ability to correct atmospheric distortions in real time have produced a new generation of powerful, large telescopes. An equally profound revolution stems from the availability of large observational databases that span the electromagnetic spectrum. The increased use of such databases as well as the need to operate the new telescopes efficiently requires the development of a National or International Virtual Observatory to set standards for astronomical database formats, data quality assurance, and access protocols, and also to provide all-inclusive centers for data products.

1 VIRTUAL OBSERVATORY GOALS & OPPORTUNITIES

We outline how present developments in observational astronomy lead inevitably to an advanced astronomical data archive, the Virtual Observatory. Although the Virtual Observatory (VO) is logically inevitable, careful technical and scientific planning are required to build a useful and scalable VO. In particular, a VO requires open and easy access to astronomical data by researchers, educators, and students throughout the world. We also highlight the many advantages to research, education, and international cooperation that will follow directly from a strong Virtual Observatory.

2 LOGICALLY INEVITABLE: LARGE, ADVANCED TELESCOPES

As a first step in the logical progression towards the Virtual Observatory, we briefly describe the current tools of observational astronomy. In the past decade advances in materials technology, systems modeling, observatory site modeling, internet connectivity, and active[1] and adaptive[2] optics have motivated the construction of a dozen 8-10 meter diameter telescopes at a total cost of approximately $2 billion (US). These technological advances have also shown us the path towards 30 to 100 meter diameter telescopes, which we expect to cost perhaps $1 billion (US) each. Telescope time on the current 8-10 meter telescopes is worth ~$1/second. Yet the weather at even the best observing sites leaves at most 50-70% of this time usable, and only perhaps 10% of this time is truly excellent. In addition to the cost of telescope time, preparations for the large telescopes are also expensive, often requiring months to years of work on smaller telescopes. As a result astronomers need and funding agencies demand that the new telescopes operate efficiently.

In Figures 1, 2a, and 2b we show the Gemini North 8-m telescope (Gillett et al. 1996) inside its enclosure atop Mauna Kea, Hawai'i. Figure 3 shows the Gemini North primary mirror. The size of the person inspecting the mirror gives a sense of scale. Figures 4a and 4b show the vast improvement in image quality now obtainable with adaptive optics. Figure 4a shows an optical (0.55 micron) image taken by Gemini of the globular star cluster NGC 6934 obtained under good atmospheric conditions, which astronomers characterize by measuring the sharpness of the stars, here 0.6 arc seconds. (Five years ago these would have been considered outstanding

---

* This work was supported by the Gemini Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the Particle Physics and Astronomy Research Council (United Kingdom), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), CNPq (Brazil) and CONICET (Argentina).

[1] Previously telescopes were limited to 4-5 meter diameter mirrors as it was too difficult to make larger mirrors which could support their own weight. Active optics is the name for the computer-controlled system of actuators which maintain the proper shape for modern flexible and light weight primary mirrors.

[2] Adaptive optics performs high-speed optical compensation of the light collected by the telescope in order to remove the distortions due to the earth's atmosphere. The resulting improved images allow astronomers to see both much fainter and much finer details in complicated astronomical objects.

conditions!) Figure 4b shows the central region of the same star cluster obtained with the Gemini North Telescope and the University of Hawai'i Hokupa`a adaptive optics system in the infrared (2.2 microns), now a factor of 7 times sharper (0.09 arc seconds)! The globular cluster, which contains so many stars as to appear cloudy in Figure 4a, appears transparent between the stars in Figure 4b.

Figure 1. The Gemini North Telescope.

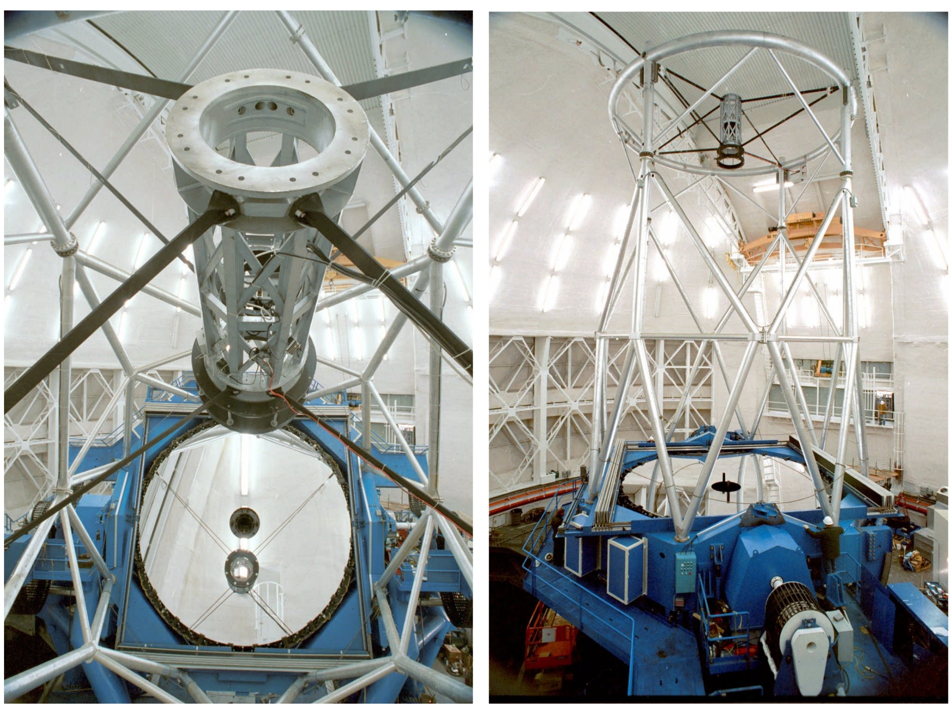

Figure 2. The Gemini North Telescope.

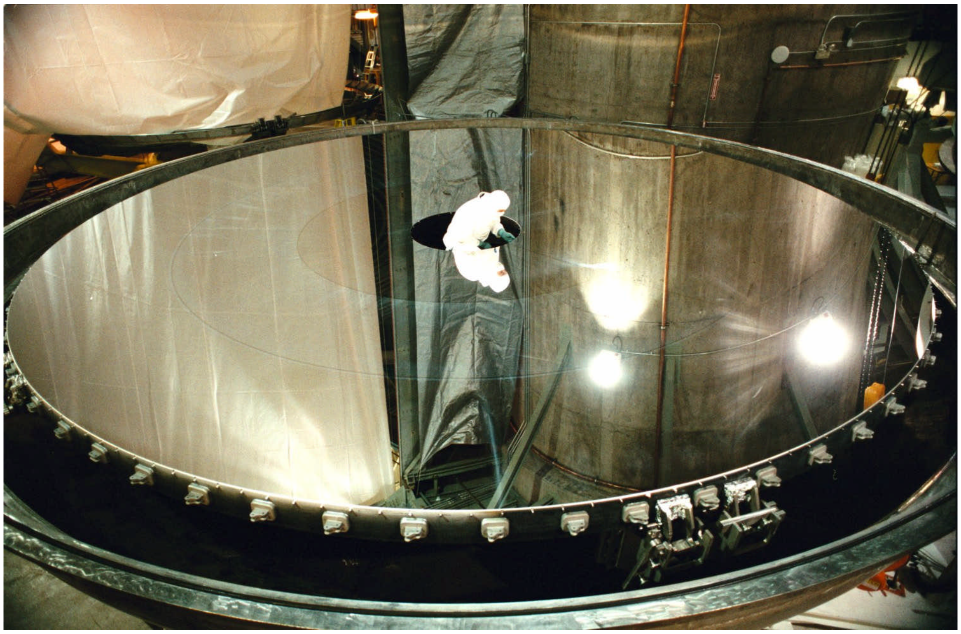

Figure 3. The Gemini primary mirror.

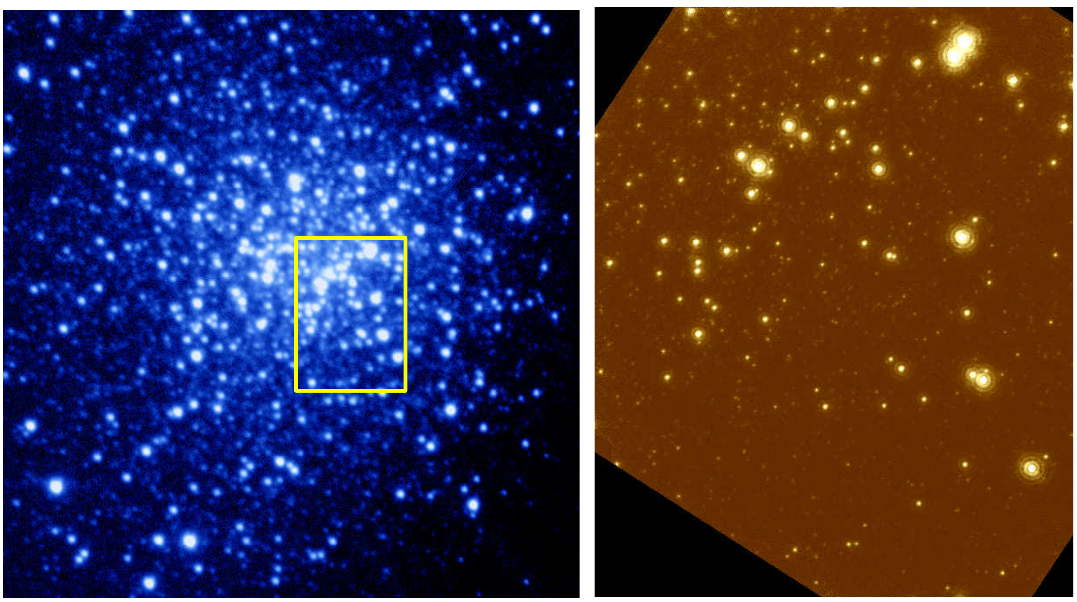

Figure 4. The globular cluster NGC 6934. 4a (left) shows an optical (0.55 micron) image taken by Gemini; 4b (right) shows the central region of the same star cluster obtained with the Gemini North Telescope and the University of Hawai'i Hokupa`a adaptive optics system in the infrared (2.2 microns).

3 LOGICALLY INEVITABLE: QUEUE OBSERVING

Large telescopes like Gemini are supported by a considerable observing infrastructure. Software tools and existing data archives are used for scheduling and preparing observations (e.g., see Figures 5 & 6), and for near real-time quality assessment. Site monitoring tools (Simons 1995) are used to measure atmospheric quality for real-time queue scheduling decisions. Weather prediction (Figures 7 & 8, see also www.soest.hawaii.edu) also greatly aids scheduling for queue observing. All of these tools, along with the drive for efficiency, the desire to match observations to the most appropriate observing conditions, and the need to schedule observations at the right time of night and season, have led most large telescopes to operate queue observing programs. Queue

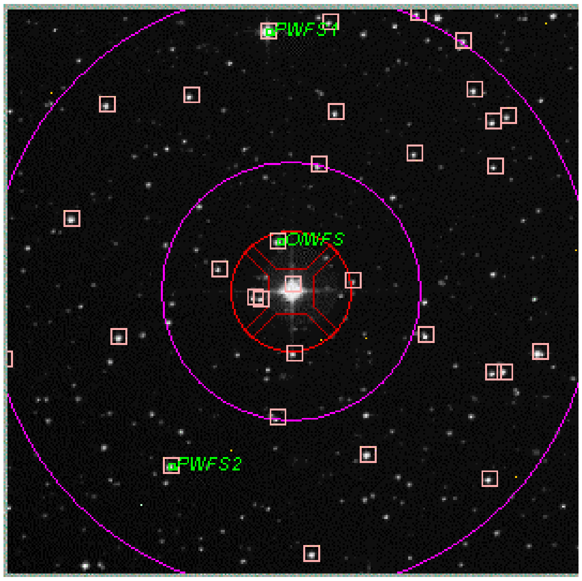

Figure 5. A sky visualization tool used to prepare observations (Wampler et al. 1997).

Figure 6. A real-time atmospheric monitoring instrument.

observing (Boroson, Davies & Robson 1996, Puxley 1997) is a technique whereby observatory staff astronomers, aided by software, place observations in a queue based on the above criteria and priorities. Experience to date suggests that queue observing is ~2 times as efficient as classical observing for the large telescopes. This is the first major logical development from modern telescopes to the VO: the expense of large telescopes and the capabilities of modern observing infrastructure have led to the development of queue observing.

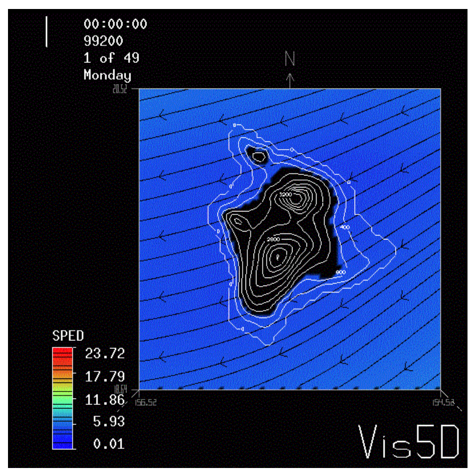

Figure 7. Mauna Kea weather prediction: 48-hour wind forecast model.

4 LOGICALLY INEVITABLE: ARCHIVES

A successful queue observing program requires a clear link among the science, calibration, and site weather data so that astronomers who did not participate in obtaining the data can properly calibrate and analyze their observations. These same calibration requirements form the basis for a good archive, where astronomers years later may retrieve data to answer different questions, apply new techniques, or form supersets of data acquired by others. Archives additionally help observers plan for upcoming observations, for example by making available a wealth of data taken with the same instrument under similar conditions. More generally, the current generation of survey and Hubble Space Telescope archives has already proven invaluable to astronomers preparing observing campaigns for a wide variety of telescopes and instruments. Education and outreach are also greatly facilitated by archives, especially when they are designed with at least some tools for the non-professional user. Experience to date, e.g., with the Hubble Space Telescope, indicates that archives containing high quality data with good search and extraction tools increase the usefulness of the telescope by a factor of at least ~3 (e.g., Albrecht 2000). This increase in telescope usefulness is primarily a quantitative increase in the productivity of the telescope to perform similar types of scientific studies. There is also a qualitative enhancement in the types of science one can derive from archives. Some archives, for example the Sloan Digital Sky Survey (York et al. 2000), can be used to calibrate much of the sky, lessening the need for all future observers who work at similar wavelengths to obtain observations of standard stars. We have just discussed the second major logical development towards the VO: queue observing greatly facilitates good

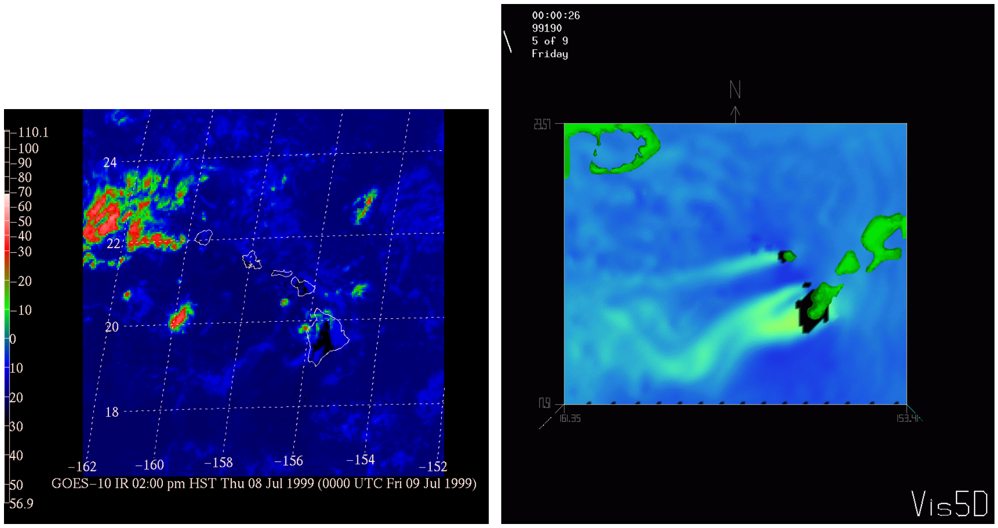

Figure 8. Mauna Kea weather prediction: (l) satellite image from NOAA (www.nws.noaa.gov) and (r) model prediction.

archives. As a closing thought on archives, we note that valuable digital data produced at most telescopes during the 1980s and early 1990s will never be available to future observers since they were never properly archived. There are still some excellent modern telescopes that do not archive their data.

5 LOGICALLY INEVITABLE: THE VIRTUAL OBSERVATORY

Now that a number of modern, large telescopes are building archives and populating them with excellent data, how do we link these multiple archives containing data obtained under different conditions, with different instruments, and with different techniques into a useful whole? The first criteria, which actually is essential at the archive level, is the free exchange of huge volumes of non-proprietary data. Since every astronomer benefits from the Virtual Observatory, and since every archive becomes substantially more valuable when properly connected to the VO, it is in the interest of the entire astronomical research and education community to make all of the best quality non-proprietary data freely available. Certainly observatories face substantial operational issues in order to ensure that their archived data can be properly linked into the VO. Instrument use, calibration, quality control, and ancillary data must be appropriate to achieve high quality data products in a standard or at least readily understood format. This requires dedicated personnel and equipment to maintain and develop archiving procedures, and to store and provide access to the data. The final link in the logical chain towards the VO, the full use of archives, leads to the need for advanced connectivity tools, the standardization of those tools, and the possible standardization of the calibration techniques. The gain of such a facility is hard to estimate. It is likely to be at least as useful to astronomers as the earlier archive concept, but with the additional expectation of major qualitative changes in the types of studies that can be undertaken with the VO.

6 LOGICALLY INEVITABLE: ANOTHER ROUTE TO THE VIRTUAL OBSERVATORY

A completely different and compelling argument for the Virtual Observatory is the huge data rates that modern instruments are generating. Astronomy is entering the Giga pixel era, when a single detector has approximately one billion pixels and produces ~100 GB per night. The number of pixels in astronomical detectors is following a geometric growth similar to that seen in the increase in computer processing speeds. Fifteen years ago astronomers could examine every pixel of every image they obtained. Such careful scrutiny of every pixel is no longer possible. Current instruments overwhelm human abilities and require standards for data format, quality assurance, and access so that a real-time system can properly process the data. Such huge data rates and real-time systems will allow astronomers to study truly large-scale problems such as the cosmic structures traced by billions of galaxies or the variety and physics behind the expected ~10 billion varying or transient extragalactic sources. Data rights for such large data producing machines destroy the scientific leverage these surveys can achieve. Furthermore, many of the scientific questions addressed by these large surveys are tractable only with the ability to connect current with past observations of the same object or class of objects at all wavelengths and using all observing techniques, i.e., these instruments are only really usable when they feed and are fed by the VO.

7 THE VIRTUAL OBSERVATORY: MORE THAN A PIXEL SERVER

As we have implied throughout this discussion, the Virtual Observatory must do far more than just make data available to users. Certainly sophisticated data products need to be part of the VO. In addition, observations need to be connected to the literature, advanced tools need to be available to provide N-dimensional visualization and analysis, and a mechanism needs to be in place to look for correlations between data properties that are along axes not already built into the data archives.

The Virtual Observatory will change the nature both of what and how astronomers observe as new science becomes possible and old work is more fully leveraged. Telescope assignment committees will expect proposers to be fully aware of past observations by preparing with the VO. In addition, the VO could be both the front and back ends to major projects. Survey work produces data that ends up in the VO, and survey work would start with funding from the VO and telescope time assigned by the VO, thereby avoiding the pitfalls of multiple proposal jeopardy. While it is difficult to estimate the efficiency gain of the VO, we fully expect both quantitative and qualitative scientific gains, and we further expect the VO to be a democratizing force in observational astronomy. While funding may be tied to the country of the user's origin, the data should be made available to any investigator or student, regardless of location. Bandwidth and computer infrastructure issues will remain, but a far greater fraction of the world's scientists and students will be able to participate in and propel the rapid advances in observational astronomy.

8 REFERENCES

Albrecht, R. (2000) Scientific Research using the Electronic Archive of the Hubble Space Telescope, *Astrophysics & Space Sciences*, 273, 127.

Boroson, T., Davies, J., and Robson, I. (1006) New Observing Modes for the Next Century. *Astronomical Society of the Pacific* Conference Series, 87, 13.

Gillett, F. C., Mountain, M., Kurz, R., Simons, D. A., Smith, M. G., and Boroson, T. (1996) The Gemini Telescopes Project, *Revista Mexicana de Astronomía y Astrofísica*, 4, 75.

Puxley, P. (1997) Execution of Queue-scheduled Observations with the Gemini 8m Telescopes, *SPIE Proceedings* 3112, 234.

Simons, D. (1995) Remote Sensing of Atmospheric Emissivity over Mauna Kea Using Satellite Imagery, Gemini Technical Note TN-PS-G0032.

Wampler, S., Gillies, K., Puxley, P., and Walker, S. (1997) Science Planning for the Gemini 8m Telescopes, SPIE Proceedings 3112, 246.

York, D. G., et al. (2000) The Sloan Digital Sky Survey: Technical Summary, *Astronomical Journal*, 120, 1579.