SPORE* SERIES WINNER

The Universe Online

Data from the Sloan Digital Sky Survey can be used by students, teachers, and the public to contribute to scientific research.

M. Jordan Raddick† and Alexander S. Szalay

Modern science is advancing at an unprecedented rate, and the amount of scientific data is doubling every year (*1*). These data have sparked a revolution in the way astronomy is practiced. No longer are scientists forced to wait months for access to a telescope to learn about the night sky; instead, entire research projects can be accomplished with online data sources. Representing modern astronomy, the Sloan Digital Sky Survey (SDSS) has made its entire data set available through an online portal for public use as an educational resource and to invite volunteer contributions to scientific research.

The SDSS (*2*, *3*) has worked since 1991 to create a map of the universe and is the astronomy equivalent of the Human Genome Project. The dedicated 2.5-m-diameter telescope in New Mexico used a 120-megapixel camera to image more than one-quarter of the entire night sky, 1.5 square degrees of sky at a time, about eight times the area of the full Moon, both inside and outside of the Milky Way, and has created a three-dimensional (3D) map of the brightest one million galaxies and quasars. A pair of spectrographs fed by optical fibers measured spectra of, and hence distances to, more than 600 galaxies and quasars in a single observation. A custom-designed set of software pipelines kept pace with the enormous data flow from the telescope (see the first figure).

The SDSS's Web site (www.sdss.org) provides the main access portal to all the information about the project. It contains an introduction to the science enabled by the survey, such as the discovery of the most distant quasars and the creation of the most detailed map of the galaxy distribution, as well as a series of press releases describing key new results.

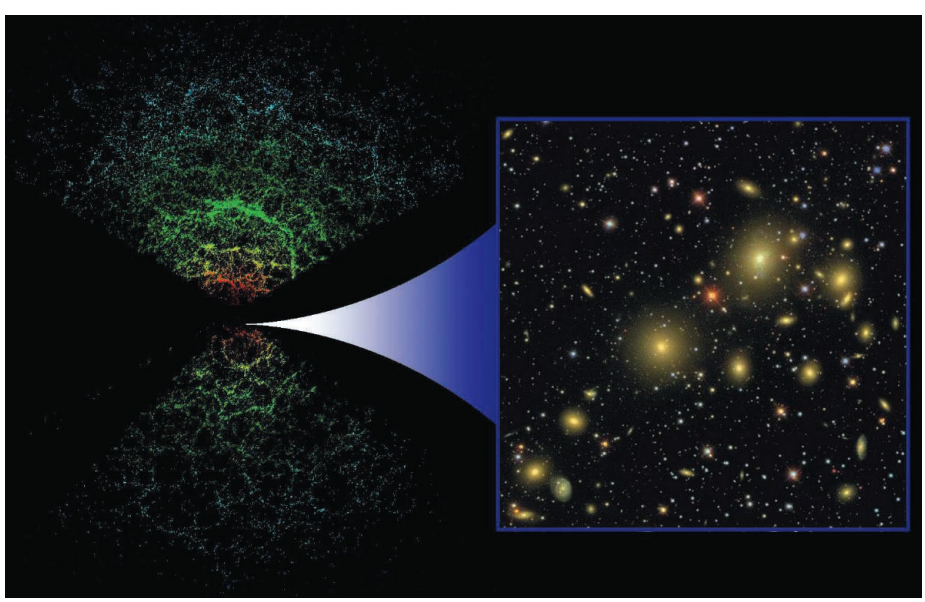
The SDSS has been successful in generating new scientific discoveries, including measurements of thousands of asteroids (*4*), maps of the complicated merger history of the outer Milky Way (*5*), and the first detection of the baryon acoustic peak; a measurement of how structure formed in the early universe (*6*). A search of the Astrophysics Data System (http://adsabs.harvard.edu) shows over 3000 papers with "Sloan Digital Sky Survey" in the title or abstract, resulting in more than 100,000 citations. The series of discoveries resulting from the SDSS will continue for many years to come as an extension of the survey has been implemented.

All data used to make these discoveries, describing more than 350 million stars and galaxies, are available to students, teachers, and the public through the SDSS's SkyServer database and its Web-based interface (http://skyserver.sdss.org). The site includes many different tools for visually exploring and searching the data. The Navigate tool allows users to navigate through the sky in a Mapquest-like interface. Users can pan through the sky, zoom in and out, and click on stars and galaxies for more information, such as that star's or galaxy's position in the sky and its magnitude (brightness). Links to the "Quick Look" and "Explore" tools from the Navigate tool allow users to examine complete SDSS data on any object. The images used by these tools were built from the 2.5 trillion pixels of original raw data.

The Search Form tool has a more traditional approach; users select from a series of menus to search the database. For example, users can search for stars at a given position in the sky, or they can search for galaxies brighter than a certain limit. Users can also enter queries to the database in Structured Query Language directly, which allows more flexible and sophisticated searches.

SkyServer also has a multiuser collaborative environment, called CasJobs (http://casjobs.sdss.org/casjobs), which enables users to launch extensive analyses. Users can store and share their (sometimes very large) intermediate results on the SDSS server and also form efficient collaborations. CasJobs also makes automatic notes on the ad hoc analysis steps the users performed, which can later be recalled. Much of the technical descriptions and facts about the SDSS and the data presented are closely linked to the data itself by an extensive automated documentation framework that describes the details of the data, the units, and the links between different data items. This documentation can be found on both SkyServer and on the SDSS's main Web site.

The data and access that SkyServer offers present an opportunity for students, teachers, and the public to learn about astronomy and other sciences. This has implications for the way students learn about science, as education research studies show that engagement with the authentic process of science is a key component of science learning (*7*, *8*). Because of the expense and paucity of access to the large



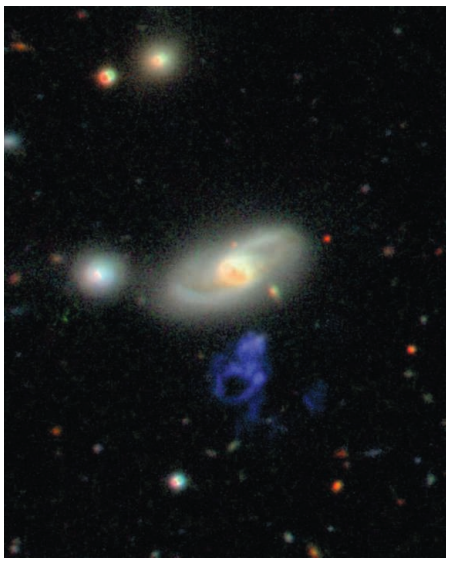
The SDSS is two separate surveys in one. Galaxies are identified in 2D images (right), then have their distances determined from their spectra to create a 2 billion light-years-deep 3D map (left) where each galaxy is shown as a single point, its color representing its luminosity.

CREDIT: BILL KEEL

Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA.

*SPORE, *Science* Prize for Online Resources in Education; www.sciencemag.org/special/spore/. †Author for correspondence. E-mail: raddick@jhu.edu

Published by AAAS



Hanny's Voorwerp. The mass (shown here in purple) is a new class of cosmic object discovered by a Dutch schoolteacher, an astronomy novice, while using Galaxy Zoo.

telescopes, telescope time has usually been reserved for leading scientists, and learners have largely been excluded from direct and active engagement with astronomy. But now, with access to the visual data that SkyServer offers, anyone with Internet access can interact with the data just as the scientists do.

To help enable the use of scientific data in the classroom, SkyServer also features a series of science projects for middle school, high school, and college students (*9*). With these projects, students can re-create major discoveries from the history of astronomy: They can understand why stars have different colors, can classify stellar spectra, or can create a Hubble diagram to show that the universe is expanding. Most projects include Excel spreadsheet templates so that students can analyze the data they find with tools they already know how to use.

The projects were designed by the science writer who developed the rest of the Web site, in collaboration with middle and high school science teachers, in order for the projects to have the same look and feel as the rest of the site. Each project took 2 to 4 months to create. Projects were developed iteratively; the design cycle began with discussions between the teachers and the science writer, and then they would trade drafts of the lesson plans. Web developers also worked to improve online tools using feedback from teachers. The projects require extensive use of the SDSS and various tools for analysis. All projects are linked to Project 2061 *Benchmarks for Science Literacy* (*10*).

SDSS science projects have been used by high schools, colleges, and self-paced learners throughout the world. As of 2006, the projects had registered 297,000 Web sessions, defined as a continuous stream of mouse clicks for more than 30 s, equivalent to providing 47,000 hours of instruction (*11*). As of June 2010, the projects have more than 25 million Web hits and more than a million different individual devices (Internet Protocol addresses) were accessing the data.

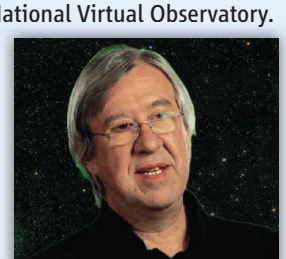
Access to SDSS data is not limited to the SDSS and SkyServer Web sites, however. The Web service that the Navigate tool is based on, ImgCutout (*12*), delivers SDSS data to a variety of other tools, including Microsoft's World Wide Telescope (www.worldwidetelescope.org) and Google Sky (www.google.com/sky). Both tools allow users to start with a view of the entire night sky and zoom in to a specified object, such as the Eagle Nebula, and include many other data sources besides the SDSS.

Another project that SDSS data access has enabled is a "citizen science" Web site, Galaxy Zoo (www.galaxyzoo.org), where Internet volunteers have classified galaxies on the basis of their SDSS images (*13*). Galaxy Zoo now uses data from the Hubble Space Telescope. Galaxy Zoo discoveries based on SDSS data have included a determination of the relation between the morphology of galaxies and their environment (*14*) and the observation that galaxies close to one another tend to have aligned spin directions (*15*). None of these discoveries would have been possible without the participation of thousands of Galaxy Zoo volunteers, who have visually classified over 40 million galaxies, a task too large for a small number of individuals. Galaxy Zoo volunteers are also using SDSS data to conduct their own creative scientific research. A comment on the Galaxy Zoo forum by a Dutch schoolteacher led to the identification of a unique object, Hanny's Voorwerp, which proved to be an astronomical object of unknown nature and remains the object of active astrophysical research (*16*) (see the second figure). Further investigation on SkyServer by volunteers led to the discovery of a new class of star-forming galaxies (*17*).

SkyServer tools allow students, teachers, and the public to view astronomy data. SkyServer projects allow learners to use these data to recreate famous discoveries in modern astronomy, such as detecting the expansion of the universe. Other interfaces, such as Google Sky and World Wide Telescope, permit the public to view SDSS data in relation to other astronomical data sets. Galaxy Zoo allows online volunteers to contribute to the advancement of science. Together, these sites are tapping into the hunger and enthusiasm of members of the general public to engage with scientific research, and are helping to make science more democratic.

About the Authors

M. Jordan Raddick is an education and public outreach specialist in the Department of Physics and Astronomy of the Johns Hopkins University. He also maintains the SkyServer Web site of the Sloan Digital Sky Survey and the virtualobservatory.org Web site of the National Virtual Observatory.



Alexander S. Szalay is a professor of Physics and Astronomy of the John Hopkins University with an interest in theoretical astrophysics and galaxy formation.

CREDITS: (TOP) MIKE EVANS (BOTTOM) LEFT, M. JORDAN RADDICK; RIGHT, JOHN PATZER

References and Notes

1. J. Gray *et al.*, *SIGMOD Rec.* **34**, 34 (2005).
2. M. Lemonick, *Discover* (March 2009); http://discovermagazine.com/2009/mar/06-world.s-hardest-working-telescope/.
3. D. York *et al.*, *Astron. J.* **120**, 1579 (2000).
4. Z. Ivezi *et al.*, *Astron. J.* **122**, 2749 (2001).
5. V. Belokurov *et al.*, *Astrophys. J.* **642**, L137 (2006).
6. D. J. Eisenstein *et al.*, *Astrophys. J.* **633**, 560 (2005).
7. J. D. Bransford, A. L. Brown, R. R. Cocking, *How People Learn: Brain, Mind, Experience, and School* (*Expanded Edition*) (National Academies Press, Washington, DC. 2000).
8. S. Michaels, A. W. Shouse, H. A. Schweingruber, *Ready, Set, Science!: Putting Research to Work in K–12 Classrooms* (National Academies Press, Washington, DC, 2008).
9. R. Sparks, C. Stoughton, M. J. Raddick, *Astronom. Soc. Pacific Conf. Ser.* **319**, 394 (2004).
10. American Association for the Advancement of Science, *Benchmarks for Science Literacy: Project 2061* (Oxford Univ. Press, New York, 1993).
11. J. Gray *et al.*, *SIGMOD Rec.* **34**, 34 (2005).
12. M. Nieto-Santisteban, A. Szalay, J. Gray, *ADASS XIII Astronom. Soc. Pacific Conf. Ser.* **314**, 666 (2004).
13. C. J. Lintott *et al.*, *Mon. Not. R. Astron. Soc.* **389**, 1179 (2008).
14. S. P. Bamford *et al.*, *Mon. Not. R. Astron. Soc.* **393**, 1324 (2009).
15. A. Slosar *et al.*, *Mon. Not. R. Astron. Soc.* **392**, 1225 (2009).
16. C. J. Lintott *et al.*, *Mon. Not. R. Astron. Soc.* **399**, 129 (2009).
17. C. Cardamone *et al.*, *Mon. Not. R. Astron. Soc.* **399**, 1191 (2009).

10.1126/science.1186936

Published by AAAS