Data-driven education and public outreach with the Sloan Digital Sky Survey

B. Lundgren[1], R. Tojeiro[2], R. L. Beaton[3], M. R. Blanton[4], J. Borissova[5], M. Cano-Díaz[6], K. Grabowski[7], R. Kurtev[5], N. MacDonald[8], S. R. Majewski[9], K. L. Masters[10,11], K. Meredith[12], C. Nitschelm[13], T. O'Reilly[14], J. Raddick[15], D. Skinner[16], A. Thakar[15], A. Weijmans[2], D. G. Whelan[17] & the Sloan Digital Sky Survey IV Collaboration

[1] *Department of Physics, University of North Carolina, USA*
[2] *School of Physics and Astronomy, University of St Andrews, North Haugh, UK*
[3] *Hubble Fellow, Department of Astrophysical Sciences, Princeton University, Princeton, New Jersey, USA.*
[4] *Center for Cosmology and Particle Physics, Department of Physics, New York University, USA*
[5] *Instituto de Física y Astronomía, Universidad de Valparaíso, Chile; Millennium Institute of Astrophysics (MAS), Santiago, Chile*
[6] *Instituto de Astronomía, Universidad Nacional Autónoma de México, Mexico D.F., Mexico*
[7] *Apache Point Observatory, New Mexico, USA*
[8] *University of California Observatories, UC Santa Cruz, USA*
[9] *Department of Astronomy, University of Virginia, Charlottesville, USA*
[10] *Haverford College, Department of Physics and Astronomy, Haverford, USA*
[11] *Institute of Cosmology & Gravitation, Univ. of Portsmouth, UK*
[12] *Geneva Lake Astrophysics and STEAM, Williams Bay, Wisconsin, USA*
[13] *Centro de Astronomía, Universidad de Antofagasta, Chile*
[14] *Department of Astronomy, University of Washington, Seattle, USA*
[15] *Institute for Data Intensive Engineering and Science, Johns Hopkins University, Baltimore, USA*
[16] *Center for Relativistic Astrophysics, School of Physics, Georgia Institute of Technology, Atlanta, USA*
[17] *Department of Physics, Austin College, Sherman, USA*

Contact / blundgre@unca.edu

Resumen / Durante más de quince años, la colaboración internacional de *Sloan Digital Sky Survey* (SDSS) ha puesto a disposición del mundo terabytes de datos astronómicos profesionales. Las vastas imágenes de archivo y los conjuntos de datos espectroscópicos del SDSS, junto con sus múltiples puntos de acceso e interfaces adecuados para los usuarios de todos los niveles de experiencia, proporcionan un novedoso recinto de pruebas para audiencias de una amplia gama de entornos para explorar y relacionarse con los datos mientras cultivan intereses y dominio de la astronomía. Los científicos, educadores, técnicos y observadores de SDSS que participan en el grupo de trabajo Educación y Difusión Pública de la colaboración han colaborado durante la última década para desarrollar un conjunto amplio y diverso de actividades educativas y exposiciones basadas en los datos de SDSS y los conceptos de ingeniería. La reciente expansión de las operaciones SDSS-IV APOGEE-2 al Observatorio Las Campanas en Chile ha impulsado un nuevo énfasis en la ampliación de la accesibilidad de estos recursos educativos para las audiencias de habla hispana.

Abstract / For over fifteen years, the international collaboration of the Sloan Digital Sky Survey (SDSS) has made terabytes of professional astronomical data freely available to the world. The vast archival imaging and spectroscopic datasets of the SDSS, coupled with its multiple access points and interfaces suitable for users of all levels of experience, provide a novel sandbox for audiences from a wide range of backgrounds to explore and engage with the data while cultivating interests and proficiency in astronomy. SDSS scientists, educators, technicians, and observers involved in the collaboration's Education and Public Outreach working group have collaborated over the past decade to develop a large and diverse set of educational activities and exhibits built around SDSS data and engineering concepts. The recent expansion of SDSS-IV APOGEE-2 operations to Las Campanas Observatory in Chile has prompted a new emphasis on broadening the accessibility of these educational resources for Spanish-speaking audiences.

Keywords / editorials, notices — miscellaneous — astronomical databases: miscellaneous

1. Introduction

For nearly two decades the Sloan Digital Sky Survey (SDSS; York & SDSS Collaboration, 2000) has collected the images and spectra of millions of stars, galaxies, and quasars. Through regular data releases, the fully reduced astronomical data and value-added catalogs have been published and made freely available to the world via `http://www.sdss.org`. The sheer scale of the data has enabled new statistical understandings of the physical processes shaping the evolution of galaxies and the

large-scale structure of the universe.

The SDSS Collaboration, now in its fourth generation (SDSS-IV; Blanton et al., 2017), is composed of hundreds of astronomers, engineers, technicians, and observers in dozens of institutions around the world. As an indication of the survey's global impact, the SDSS was cited over 700,000 times by authors in 52 countries in its first 10 years of operations (Chen et al., 2009). Over the past decade, the SDSS data and related products have become even more prolific, with research applications extending well beyond the international membership of the SDSS collaboration.

The SDSS Education and Public Outreach (E/PO) working group is charged with advancing education and public outreach in astronomy through SDSS-related activities and, increasingly, the direct engagement of the public with SDSS data. The SDSS E/PO group has leveraged the SDSS data products and research tools to produce inquiry-based activities for thousands of students learning science, technology, engineering and mathematics (STEM) subjects, from elementary through graduate school, and through after-school and informal education programs. SDSS data have been used to enhance the understanding of science among the general public as well, by enabling citizen science projects like Galaxy Zoo (Lintott et al., 2011) and appearing in museum exhibits and planetaria around the world.

2. SDSS Data as a Tool for Astronomy Education

While astronomy has always been at the forefront of data-intensive research, instrumental and computational advances over the past two decades have made terabyte-scale datasets the norm. With even more expansive surveys on the horizon (from e.g., the Large Synoptic Survey Telescope and the Wide Field Infrared Survey Telescope), astronomy − along with other data-rich areas of science and industry − will increasingly rely on individuals who are equipped with the skills to visualize and analyze data of ever-increasing scale. As such, a modern astronomy education should include computational methods as an integral part of the experimental and theoretical core curriculum. Teaching these methods with the use of authentic scientific data can further enhance the learning process and positively affect student attitudes and persistence in STEM (Nagda et al., 1998; Russell et al., 2007)

2.1. SDSS *SkyServer* and *Voyages*

The SDSS *SkyServer*[1] provides the astronomy community and the public with access to the latest SDSS data releases, which currently include:

- Full color images and photometric measurements for >930 million unique objects
- Optical spectra of >4 million stars, galaxies, and quasars
- Infrared spectra of ∼263 000 stars
- Spatially resolved optical spectra of nearly 5000 galaxies in the nearby universe

Since the earliest SDSS data releases, educational materials have been developed and made freely available as part of the *SkyServer*, with the aim of helping students, teachers, and the public explore the growing SDSS dataset and develop a better understanding and appreciation for astronomy. The intentionally broad scope of the SDSS survey has accommodated the development of a diverse set of educational activities, with topics ranging from the structure of our solar system and the Milky Way to the properties of distant quasars and the large-scale structure of the universe.

SkyServer activities have been developed for a variety of educational levels, spanning elementary (i.e., primary) school through introductory college courses. The activities are designed to give educators flexibility to implement them in ways that will fit the time constraints of their classes and the abilities of their students. To assist with lesson planning, teaching guides have also been made available to accompany the *SkyServer* activities.

Since 2012, significant efforts have been committed to developing a new set of online activities with an updated pedagogy and aesthetic. The new flagship SDSS education site, *Voyages*[2], houses a collection of new educational activities and explorations based on the latest SDSS survey data, as well as an updated set of the original *SkyServer* educational activities. The aim of *Voyages* is to provide the pathways and supporting resources to enable student-led discovery of a variety of astronomical phenomena using the same data utilized by professional astronomers. Use of the site has steadily grown since its launch in mid-2017. Currently, ∼900 users interact with the site each month, 25 % of whom reside outside of the U.S.

2.1.1. Translation Efforts

Given the international membership of the SDSS Collaboration and its intent to reach a global user base, the first generation of the SDSS undertook a substantial effort to translate the *SkyServer* website into multiple languages. Versions on the web published prior to the Data Release 8 can be accessed in English, German, Hungarian, Japanese, Spanish, and Portuguese. Limited translations of a some original *SkyServer* educational activities were also published as part of the Data Release 7.

As the survey diversified and broadened its scope in its third and fourth generations, the effort of translating the frequently revised and updated *SkyServer* website was not actively maintained. The current version of the *SkyServer* website (serving Data Release 15) is available only in English, although thanks to the the improved automatic translations of modern web browsers (e.g., Google Chrome), the site can still be broadly accessed by an international audience.

The expansion of SDSS-IV APOGEE-2 operations to Las Campanas Observatory in Chile provided new motivation for making the SDSS educational resources

[1] `http://skyserver.sdss.org`

[2] `http://voyages.sdss.org/`

```
slice1_red = np.where( (all_gals.z_spec > 0.02) & (all_gals.z_spec < 0.03) & (all_gals['u'] - all_gals['g'] > red))[0]
slice1_blue = np.where( (all_gals.z_spec > 0.02) & (all_gals.z_spec < 0.03) & (all_gals['u'] - all_gals['g'] < blue))[0]

plt.figure(figsize=(20,8))
plt.subplot(121)
plt.scatter(all_gals.loc[slice1_blue]['ra'], all_gals.loc[slice1_blue]['dec'],  marker='.', s=10, color='blue')
plt.xlabel('RA')
plt.ylabel('Dec')
plt.ylim(-5,70)

plt.subplot(122)
plt.scatter(all_gals.loc[slice1_red]['ra'], all_gals.loc[slice1_red]['dec'],  marker='.', s=10, color='red')
plt.xlabel('RA')
plt.ylabel('Dec')
plt.ylim(-5,70)
```

```
(-5, 70)
```

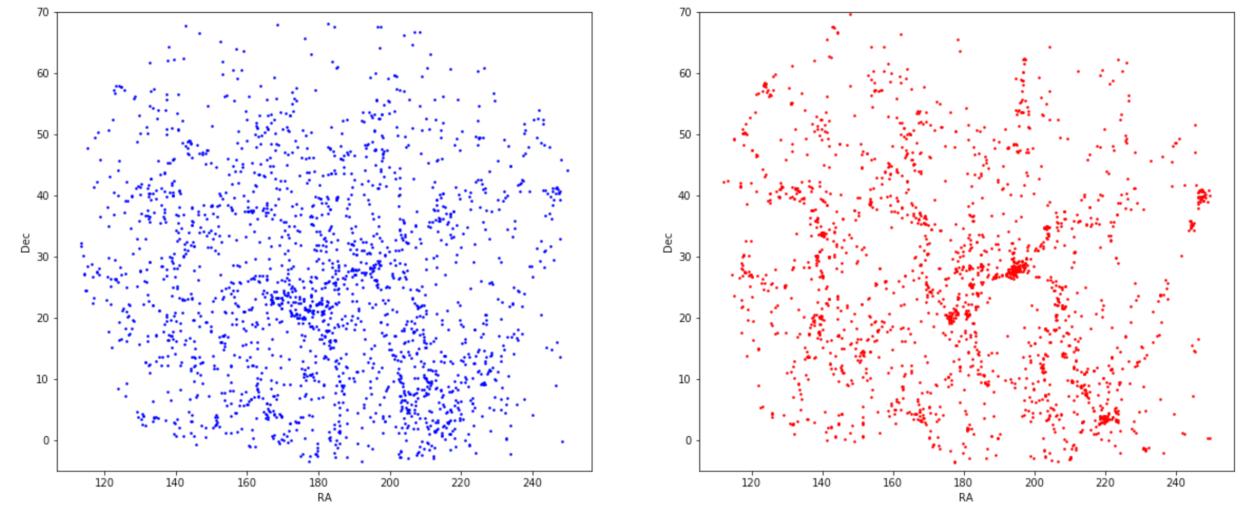

Figure 1: An example cell from a SciServer notebook activity exploring the color-density relation of SDSS galaxies.

accessible to Spanish-speaking audiences with more carefully curated translations. Through the volunteer efforts of Spanish-speaking SDSS astronomers, the SDSS *Voyages* website was translated and launched in a Spanish version[3] in 2017. Further development of the Spanish site is still underway, in parallel with continued updates to the English site.

2.2. SciServer

Basic competency in computer programming has become an essential skill required of any aspiring astronomer. Still, there remains a dearth of accessible university-level activities that teach coding within a context of real astronomical investigations.

Over the past two years the SDSS E/PO working group has been partnering with the developers of SciServer[4], a science platform of the SDSS, to produce innovative and accessible Python-based data explorations for astronomy classrooms and labs, with a focus on undergraduate-level students. Operated by the Institute for Data Intensive Engineering and Science at Johns Hopkins University, SciServer is a fully integrated cyberinfrastructure system encompassing related tools and services to enable researchers to cope with the challenges of big data. SciServer enables a new approach that allows researchers to work with terabytes or petabytes of scientific data via web-based notebooks, without downloading any large datasets or installing any software. This flexibility is particularly useful in educational settings that have traditionally lacked the computational resources to take advantage of large-scale datasets like the SDSS.

The university-level SciServer curriculum developed by the SDSS E/PO team consists of free, highly accessible, and modernized activities using SDSS data for a broad community of astronomy educators. A set of introductory-level activities has been piloted with 80 students at the University of St Andrews, in the UK, and a complementary laboratory curriculum for upper-level astronomy undergraduates has been piloted at the University of North Carolina Asheville in the U.S.

These pilot programs have demonstrated the power of SciServer in undergraduate teaching at all levels. Using web-based notebooks, the first-year class at the University of St Andrews was able to discover the large-scale structure of the Universe and color-density relation of galaxies using only very simple Python statements (e.g., Fig. 1). The new 5-hour lab was a significant departure from teaching methods traditionally used in that course, but it was a success. A large majority of the class completed bonus exercises in their own time, driven by interest and enabled by the free and easy access to SciServer. Following the pilot, the SciServer lab was adopted as an integral part of the course, and has run for a second year with equal success. SciServer-based teaching has

[3] `http://voyages.sdss.org/es/`
[4] `http://www.sciserver.org/`

also since been adopted in later years of the undergraduate curriculum at the same institution.

At more advanced levels, SciServer equips students to engage in meaningful astronomy research projects with ease. In upper-level observational astronomy courses at the University of North Carolina Asheville and the University of St Andrews, SDSS SciServer activities are providing students with a strong foundation of Python-based data analysis, within the context of SDSS science. Some students who completed these pilot courses have already gone on to produce original independent research, which has been presented at national meetings of the American Astronomical Society.

The first release of the SDSS SciServer activities was made available to participants at a half-day workshop at the 2018 Winter Meeting of the American Astronomical Society in National Harbor, Maryland. These activities, along with new curricula currently in development, are available on *github*[5][6]. SciServer's latest release (v2.0.0; Betelgeuse) in July 2018 solidifies the team's commitment to classroom use, via the introduction on classroom-friendly features. Educators can now easily set up a classroom in SciServer Compute[7] for each of their classes using the new group sharing capabilities. A customizable Jupyter notebook is also available to automate setting up a user space, group configuration, and sharing privileges for a class. These features allow for the easy set-up and sharing of a library of notebooks with a course or workshop, along with the ability to add or modify notebooks as needed during a course.

2.3. Plates for Education

Millions of SDSS spectroscopic observations have been acquired to date using tens of thousands of aluminum plates. Each plate has as many as 1000 fiber-optic cables plugged to transmit light from astronomical sources captured by the telescope to the spectrographs. Each plate is custom-designed to correspond to objects that are visible in a specific area of the sky at a certain time of night for optimal observation. Once the affiliated observations are complete, the plates are retired. In 2015, the SDSS began distributing the growing number of retired plates to educators through *Plates for Education* workshops. Attending educators each receive a plate and a poster of matching scale, which shows either a single filter image from the SDSS or a 2MASS image of the unique region of the affiliated with the plate (an example is shown in Fig 2). Supporting exercises hosted on *Voyages* offer ways for students to explore the images and spectra corresponding to objects on each plate in greater depth.

The distribution of retired plates is typically conducted by astronomers at SDSS partner institutions, who host plate workshops for educators in their local communities. Workshops have also been held in affiliation with SDSS collaboration meetings and at meetings of professional societies. Through these workshops

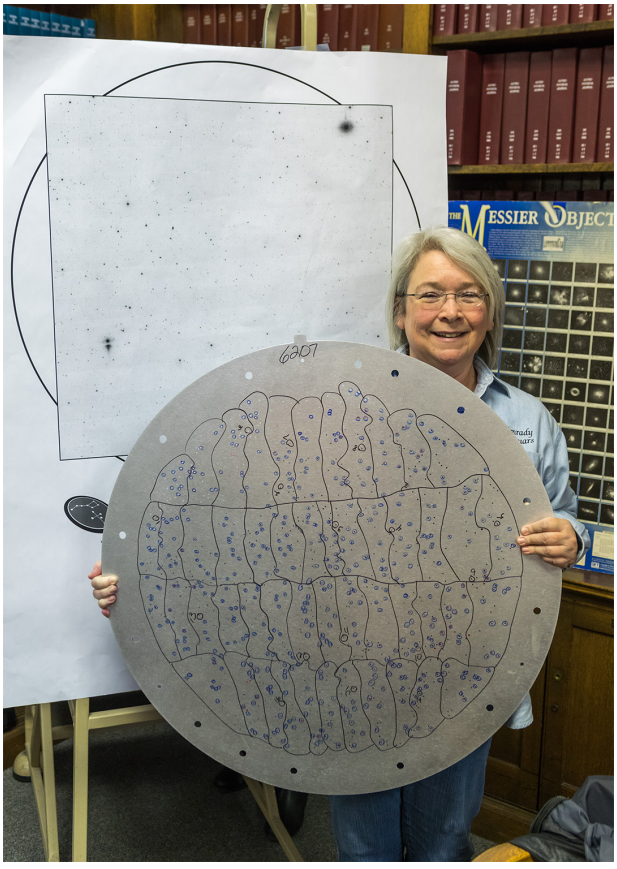

Figure 2: Chicago-area teacher and SDSS-IV pre-college educational consultant Peggy Piper stands with a retired SDSS plate and its accompanying poster, two of the primary resources distributed to teachers as part of the SDSS-IV Plates for Education Program.

teachers and informal educators are connected with mentors in the SDSS, who can provide further assistance with using the plate and associated SkyServer tools.

A new Scottish-based plate distribution program led by the University of St Andrews is also being undertaken to train secondary science teachers to use SDSS educational activities, data, and science in their classrooms. Retired plates and associated materials are being made available for lending in libraries across Scotland. Special training sessions are also being organized to help increase teacher confidence in specialist subjects and concepts. The program aims to promote engagement in science, and astronomy in particular, by giving students a backdrop of exciting activities, built around the engineering concept of the SDSS plates.

To date, plate distribution has occurred in the United States and Europe, and efforts are underway to begin distributing APOGEE-2S plates in South America after they are retired from Las Campanas Observatory. A survey of thirty-seven U.S.-based plate recipients in 2017 indicated that over ninety-five percent of respondents had used the plate in an education lesson using SDSS data and that the lessons were effective in connecting the SkyServer resources with the survey's engineering concepts and data collection methods. An equal

[5] `https://github.com/brittlundgren/SDSS-EPO`
[6] `https://github.com/ritatojeiro/SDSSEPO`
[7] `https://apps.sciserver.org/compute/`

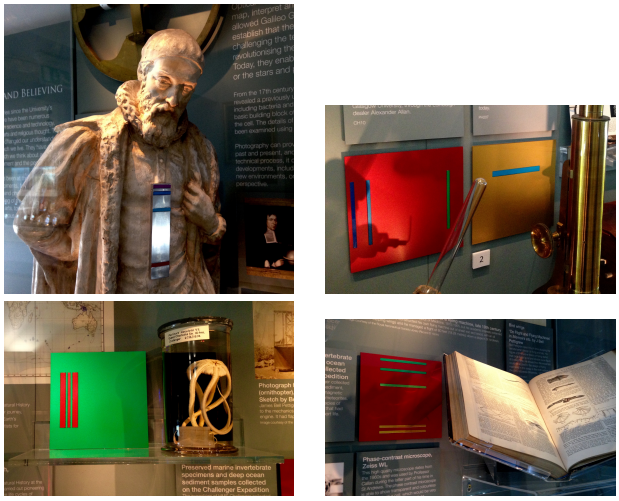

Figure 3: Examples of artwork inspired by SDSS spectra, from the exhibit 'Code for Everything' (2018) in the Museum of the University of St Andrews, by SDSS artist in residence Tim Fitzpatrick. For this exhibit, the artwork was embedded within the permanent collection of the museum. Image credit: T. Fitzpatrick.

number of respondents said the program was a useful engagement tool and effective in generating student questions. Most respondents also planned to continue to use the plate in the future.

3. Collaborations with Artists

Collaborations between scientists and artists can be a fruitful means for reaching new audiences that traditionally have not been engaged with science and technology. Art and music offer new ways to explore complex scientific data and concepts, and provide a different context for audiences to learn more about the underlying ideas.

The SDSS has inspired a number of artists to design artwork around its scientific ideas and data collection; examples include *Transient Objects* by British artist, Xavier Poultney (exhibited in Prague in 2014), *The Universe* by Chinese artist Jian Yan (exhibited in Beijing, 2016) and SDSS plates incorporated into artwork by US artists Adrienne Outlaw (exhibited in Nashville, 2015) and Sarah Ruether (exhibited in Seattle, 2014).[8]

As of 2017, the SDSS has an official artist in residence, Tim Fitzpatrick. Tim Fitzpatrick is based in Scotland and works with astronomers at the University of St Andrews on the *Shine* project:[9] exploring light through science, music and art. His work based on SDSS spectra has been on display in the UK in St Andrews, Dundee, Hull and Oxford, and was presented at the Sharing One Sky conference in Santiago, Chile, in 2017. Currently Fitzpatrick is preparing SDSS art for the Las Cruces Space Festival in New Mexico (which will feature a visit to the SDSS Telescope at Apache Point Observatory) and is developing an SDSS collaboration-

[8]For more examples see: `https://blog.sdss.org/2014/09/19/sdss-plates/`

[9]`http://shine.wp.st-andrews.ac.uk/`

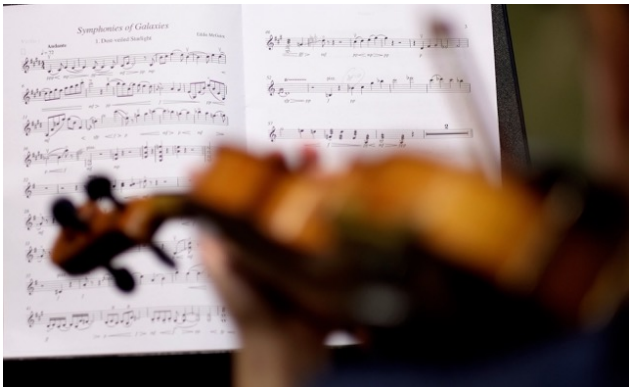

Figure 4: First rehearsal of the MaNGA Galaxy Symphonies by Scottish composer Eddie McGuire. The piece was performed in St Andrews by the New Music Ensemble of the Music Centre at the University of St Andrews, under direction of Bede Williams. Image credit: T. Fitzpatrick.

wide art project around the periodic table. Examples of his spectra inspired work are shown in Fig. 3.

As part of *Shine*, and to celebrate the International Year of Light 2015, the School of Physics and Astronomy at the University of St Andrews commissioned Glasgow-based composer Eddie McGuire to write a short musical piece on galaxies. McGuire composed the 15 minute suite 'MaNGA Symphonies of Galaxies' (see Fig. 4), which was premiered in November 2015 in St Andrews. Since then, the music has also been performed in Antofagasta as part of the Chilean National Astronomy Meeting 2016. A description of this work and its first performance is given by (Williams et al., 2016), and the work is available for hire in the Scottish Music Library.

4. Sharing One Sky

The conference series *Sharing One Sky: SDSS, APOGEE, and Astronomy Outreach*, now planning for its third gathering, was initiated to promote public dissemination of SDSS science (APOGEE in particular) and foster the sharing of methods for education and public outreach with special attention on cross-cultural exchange, particularly within and across Latin America. Given APOGEE's strong presence in Chile, the *Sharing One Sky* meetings have all been held there, and have included a special, though not exclusive, emphasis on astronomy outreach within the Chilean community, with participants from all levels of Chilean STEM education. Begun with contributing funds from the University of Virginia's Center for Global Inquiry & Innovation and the Universidad de Valparaíso, *Sharing One Sky* seeks to bridge the language gap between North and South America and provide contributing partners with new ideas for teaching astronomy to school-age children as well as the general public.

The first *Sharing One Sky* meeting took place in Viña del Mar in May 2016, and was attended by an international complement of approximately 40 scientists, undergraduate and graduate students, museum curators, members of the press, and school teachers. The main Chilean astronomical institutions were

Figure 5: SDSS-IV observer Katie Grabowski introduces some newly drilled spectroscopic plug plates for the APOGEE-2 South survey at the 2017 *Sharing One Sky II* conference in Santiago, Chile.

represented: the Millennium Institute of Astrophysics (MAS), CONICYT and the *CONICYT Explora* program[10], the Interactive Mirador Museum (MIM), the Planetarium of Santiago, and all Chilean universities with astronomy programs. Presentations were largely given in English, but translators were on hand to facilitate Spanish-speaking presenters and audience members. The event was reported on by Chilean radio, with various participants contributing interviews.

In addition to scientific talks and demonstrations, the Viña del Mar meeting promoted discussions of a great diversity of outreach methods and techniques, and participants shared their experiences and best practices for communicating science results with secondary-level students and the general public. Members of the SDSS demonstrated various outreach methods and materials that the survey has produced. Astronomy outreach performed by students was a highlight of the meeting. Three examples include Chilean undergraduate students who showcased the primary and secondary school outreach activities of *Bling Bling Universe*[11], which include content provided by the SDSS; the creators of the *Star Tres*[12] program, run by Chilean native astronomy graduate students studying in Europe, who discussed their blog and social media initiatives to engage members of the public, particularly school-aged children; and representatives of the graduate student run *Dark Skies, Bright Kids*[13] program from the University of Virginia, which has been developing bilingual English/Spanish language astronomy materials for use in the United States.

Sharing One Sky II took place on the campus of the Pontifical Catholic University of Chile in July 2017. It was held in conjunction with the SDSS-IV collaboration meeting, and included special talks and a training course aimed at educators interested in including SDSS activities in high school and undergraduate curricula. The SDSS artist-in-residence Tim Fitzpatrick (discussed above in Section 3.) also attended and showcased artwork he created around the idea of atomic spectra.

Sharing One Sky III, which is scheduled for November 2019 in Antofagasta, Chile, will be the third convening of the conference series. It will continue the mission of bringing together SDSS astronomers and educators from English- and Spanish-speaking countries to share their work and promote bilingual collaborations. A natural continuation of the series, the meeting will include a special section for educators connected with astronomical tourism in the region as well as a workshop focused on astronomy outreach for the blind and visually impaired using three-dimensional printed galaxies from *Tactile Universe*[14] (Bonne et al., 2018) as part of the *AstroBVI* initiative[15] in South America.

5. Summary and Future Work

For nearly two decades the SDSS has been promoting public engagement with SDSS data through the development of freely accessible educational resources, workshops, and arts programming. SDSS E/PO resources continue to be developed, modified, and translated, in order to meet the evolving needs of a diverse and international community of astronomy educators, students, and interested members of the public. Future work by the SDSS E/PO group will seek to better evaluate the impact of these activities in all of the diverse settings they aim to reach.

Acknowledgements: Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is `www.sdss.org`.

SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

Sharing One Sky is partially supported by the Ministry for the Economy, Development and Tourism, Programa Iniciativa Científica Milenio grant IC120009, awarded to the Millennium Institute

[10] `https://www.explora.cl`
[11] `https://bbu.cl`
[12] `http://www.startres.net`
[13] `http://http://faculty.virginia.edu/DSBK/`
[14] `https://tactileuniverse.org/`
[15] `https://astrobvi.org/`

of Astrophysics (MAS), the Universidad de Valparaiso, Chile, and a Global Research Program of Distinction grant from the Center for Global Inquiry & Innovation of the University of Virginia to SRM.

References

Blanton M. R., et al., 2017, AJ, 154, 28
Bonne N. J., et al., 2018, Astronomy and Geophysics, 59, 1.30
Chen C., Zhang J., Vogeley M. S., 2009, IEEE Intelligent Systems, 24, 74
Lintott C., et al., 2011, Monthly Notices of the Royal Astronomical Society, 410, 166
Nagda B., et al., 1998, The Review of Higher Education, 22, 55
Russell S. H., Hancock M. P., McCullough J., 2007, Science, 316, 548
Williams B., et al., 2016, Scottish Journal of Performance, 3, 59
York D. G., SDSS Collaboration 2000, AJ, 120, 1579