Session 7: Big Data in Astronomy

https://doi.org/10.1017/S1743921321000594 Published online by Cambridge University Press

Education and Heritage in the era of Big Data in Astronomy
Proceedings IAU Symposium No. 367, 2020
R. M. Ros, B. Garcia, S. R. Gullberg, J. Moldon & P. Rojo, eds.
doi:

Introduction

Introduction

This section for Session 7 on the fourth day of the symposium begins with the invited talk "The vigorous development of data driven astronomy education and public outreach (DAEPO)" by Shanshan Li, Chenzhou Cui, Cuilan Qiao, Dongwei Fan, Changhua Li, Yunfei Xu, Linying Mi, Hanxi Yang, Jun Han, Yihan Tao, Boliang He, and Sisi Yang. The authors stress the importance of education and public outreach (EPO) for astronomy and highlight data driven astronomy education and public outreach (DAEPO) as an important offshoot from the advances in Big Data and Internet technology. They describe the development of DAEPO, outline its many utilities, and stress its potential for the future.

After the presentation Anahí Caldú asked:

How do we address the large differences to technology access? I am thinking about Mexico, in which many regions have really limited access to internet or technology. Can we adapt big data projects to "simpler" versions?

Shanshan replied:

When we have this situation, we will bring data with us. Pre-download some data, install WWT on the laptop. But if the internet really limited, just download some picture and video will be great.

Juan Calos Terrazas asked:

For many students it is important to know the numerical-statistical methods with which the data can be processed. Are there seminars for data analysis?, for example with Python, R, C, etc. Thanks

Shanshan answered:

We just finish a Python seminars in China, It can be seen online but it's in Chinese. There are training last year also, you can check https://asaip.psu.edu/ this website.

"The role of Big Data in Astronomy Education" is a discussion of a primary theme of the symposium that is written by Areg Mickaelian and Gor Mikayelyan. They mention the range of astronomical observation over the entire electro-magnetic spectrum and that databases make such information available. They describe astrophysical Virtual Observatories (VOs) and an International Virtual Observatory Alliance (IVOA) to coordinate development of such technology. They point out that many astronomical education tools now use Big Data and that this greatly benefits the field.

"GALAXY CRUISE: Accessible Big Data of the Subaru Telescope for Citizen Astronomers" was written by Kumiko Usuda-Sato, Masayuki Tanaka, Michitaro Koike, Junko Shibata, Seiichiro Naito, and Hitoshi Yamaoka. They describe the GALAXY CRUISE project conducted by the National Astronomical Observatory of Japan (NAOJ) and that with it citizen astronomers classify and identify interacting galaxies on a computer screen. They point out some of the details and outline the project's value.

Following the talk Kentaro Yaji asked:

Why so many users of Galaxy Cruise in Russia and Ukraine? Are they the public people or young students?

Kumiko answered:

Articles were published on some online news sites. In addition, an influencial Russian YouTuber introduced GALAXY CRUISE.

Yusuke Tampo asked:

My question seems to related to every speaker, How the data calibration should have done for education and citizen science? Same as the data release for astronomers, or need more specified calibrations?

Baerbel Koribalski responded:

We use the same data, same calibration. What differs is the presentation of the data

and Kumiko answered: *We use the public released data, already calibrated by researchers who are in charge of the survey.*

The final paper in this section is "Creation of a MOOC and an Augmented Reality application on exoplanets for Exoplanets-A project" by Raphaël de Assis Peralta, Vincent Minier, and Pierre-Olivier Lagage. The authors describe the Exoplanets-A project and how results have been provided through the creation of a MOOC and an augmented reality application. They describe both and discuss the value for education.

https://doi.org/10.1017/S1743921321000594 Published online by Cambridge University Press

Education and Heritage in the era of Big Data in Astronomy
Proceedings IAU Symposium No. 367, 2020
R. M. Ros, B. Garcia, S. R. Gullberg, J. Moldon & P. Rojo, eds.
doi:10.1017/S1743921321000594

The vigorous development of data driven astronomy education and public outreach (DAEPO)

Shanshan Li[1], Chenzhou Cui[1], Cuilan Qiao[2], Dongwei Fan[1], Changhua Li[1], Yunfei Xu[1], Linying Mi[1], Hanxi Yang[1], Jun Han[1], Yihan Tao[1], Boliang He[1] and Sisi Yang[1]

[1]National Astronomical Observatories, CAS,
20A Datun Road, Chaoyang District, Beijing, China, 100101
emails: lishanshan@nao.cas.cn, ccz@nao.cas.cn

[2]Physics Department, Central China Normal University,
NO.152 Luoyu Road, Wuhan, Hubei, China, 430079

Abstract. Astronomy education and public outreach (EPO) is one of the important part of the future development of astronomy. During the past few years, as the rapid evolution of Internet and the continuous change of policy, the breeding environment of science EPO keep improving and the number of related projects show a booming trend. EPO is no longer just a matter of to teachers and science educators but also attracted the attention of professional astronomers. Among all activates of astronomy EPO, the data driven astronomy education and public outreach (abbreviated as DAEPO) is special and important. It benefits from the development of Big Data and Internet technology and is full of flexibility and diversity. We will present the history, definition, best practices and prospective development of DAEPO for better understanding this active field.

Keywords. Astronomical Education, DAEPO, Astronomical data, Education, STEAM, Citizen Science, WorldWide Telescope

1. Introduction

Astronomy is a natural science based on observation. From ancient time, people keep recording the observation results and forming original astronomical data. It seems natural to use these data in astronomical education and public outreach (EPO) activities. Educators have long realized that one picture shows a spectacular galaxy or a short video generated by numerical simulation which describes the evolution of the universe is better than thousands of words for student to understand. In the case of Internet, television, newspaper and other mass media, this effect is particularly obvious. For instance, people may not know what is M16 or the eagle nebula but have seen the picture of "pillars of creation"†, taken by the Hubble Space Telescope (HST). This is an example of astronomical data playing a role in EPO. It makes the public have an intuitive impression of celestial bodies like nebula.

These activities are not the major data driven astronomy education and public outreach (DAEPO) activities we want to discuss. Besides simple display as an image, astronomical data can play a more significant role in EPO activities. However, we still include them

† https://www.nasa.gov/image-feature/the-pillars-of-creation Feb. 23, 2018

in this paper because they impact the most extensive public and are the easiest accessible material for educators. This paper divides other common DAEPO activities into following categories: interactive data platform, scientific projects with amateur participate (mainly citizen science), interdisciplinarity study and activities. In each category, brief description, definition and best cases will be discussed.

2. Development of DAEPO

It is hard to clearly sort out the development of DAEPO, for this kind of astronomy EPO has great flexibility, diversity and contingency. It can be a small project like an interactive course based on astronomy data provided by professional astronomers or educators in local schools and museums. This kind of activities have strong randomness and the number of participants is relatively small. It also can be a multiuser online data analysis platform, with thousands of participants from all over the world.

For example, the national schools' observatory (NSO) from Liverpool John Moores University provides students around the world an opportunity to use the Liverpool Telescope (LT), the world's largest robotic telescope. It has an online interface and users can obtain different access permissions according to their roles. School teachers cooperate with NSO can carry out an interesting lesson and students can participate in the process of data acquisition. Although the threshold is high, NSO has over 16,000 active users.†

Besides the creativity and diversity of DAEPO activities, the rapid development of the communication means and the development of information technology make them even harder to collect and summarize. Even if the collection of these activities is completed right now, the information may be missing, outdated or inaccurate in the next second. The website CitizenScience.gov collectes crowdsourcing and citizen science activates across the USA. Some projects labeled "active" are actually suspended for various reasons. But through some representative events we can understand the general idea of DAEPO.

2.1. *The impact of Big Data on DAEPO*

The development of DAEPO is closely related to the growth of astronomical data volume and data processing requirements. Data records and analysis helps scientists understand many natural laws, like the Kepler's three laws of planetary motion. Compared with ancient time when people observing stars with naked eye, the inventions of astronomical telescope and photographic techniques make the observation more precise. After entering the 21st century, with the development of Big Data and Internet technology, astronomical data can be obtained, stored and transferred much faster than before and the amount of data increases every year.

Currently, astronomer use data from Sloan Digital Sky Survey (SDSS), Gaia mission, Large Sky Area Multi-Object Fiber Spectroscopy Telescope(LAMOST), Very Large Array(VLA) and many other observation facilities. SDSS, for example, Data Release 16, has 273TB in total‡. Gaia mission announced an EDR3 last year. The compressed CSV files is about 1.3TB released to all astronomer around the world (Gaia Collaboration *et al.* 2020). Chinese optical telescope LAMOST, has the highest spectral acquisition rate and the spectrum it has collected so far is more than the sum of other telescopes collected in the world. The Lamost team has already released over 10 million spectra (dr6.lamost.org). Soon, astronomy study will enter a new era of big data, not GB, TB, but PB, or even EB. The data sets acquired by telescopes will be too large to download and analyse using users' own facilities. Like the FAST, 500 meter Aperture Spherical radio Telescope in China began operation in science since 2020, will generate 20PB data

† https://www.schoolsobservatory.org/about, About the NSO project
‡ https://www.sdss.org/dr16/data_access/volume/, SDSS DR16 data volume, Dec. 2019

every year (Qian *et al.* 2019).The Legacy Survey of Space and Time (LSST) of Rubin Observatory plans to acquire 25TB every night (www.lsst.org).

With the data acquisition increase, astronomers do not have enough equipment and manpower to analyse and process all the data. They began to consider distribute some simple work and corresponding data to amateur astronomers, connected them through Internet and try to use their resources and wisdom to help scientific research. Among all these attempts, Internet-base public volunteer computing project SETI@home was one of the most famous. Launched in 1999, created by Berkeley SETI Research Center and hosted by the Space Sciences Laboratory, at the University of California, Berkeley, SETI@home aims to use the computing resources of volunteer to analysis observation data to detect intelligent life outside Earth (setiathome.berkeley.edu). Although until it stop sending out data, no sign of aliens was found, it is a significant attempt for astronomers to cooperate with amateurs. The enthusiasm and the power of amateur astronomers around the world impressed scientists. It also shows the potential of data related EPO activities.

The citizen science project Galaxy Zoo lunched in 2007 is one of the most important milestone of DAEPO. One reason for its design is also the needs for huge astronomical data processing (Masters & Galaxy Zoo Team 2020). Due to the reasonable design, public was enthusiastic with the project and some participants did make real discoveries.

2.2. *Virtual Observatory and DAEPO*

The development of DAEPO projects also benefit from the openness of astronomical data. Different from many other fields, most telescopes and astronomy projects will release their data to all the people around the world. Not only for astronomers, but also for scientists in other fields, engineers, teachers, students, etc. That means these massive data can be perfect material for EPO project.

Virtual Observatory(VO) must be mentioned when discuss the use of astronomical data. VO is a data-intensively online astronomical research and education environment, taking advantages of advanced information technologies to achieve seamless, global access to astronomical information (Cui *et al.* 2017). The resource linked by VO include observational data, computing resources, storage resources, software platforms and even observation equipments. In 2002, the International Virtual Observatory Alliance(IVOA) was formed and started to work on the development of VO standards and protocols. IVOA now has 21 members including VO projects from different countries and data centers, organizations (www.ivoa.net). With the efforts of all members, IVOA has already developed a set of standards and protocols, greatly improved the efficiency of data publishers and users. Major astronomical data centers and most observational projects in the world use IVOA protocol for data sharing and interoperability.

The concepts and protocols of IVOA not only help astronomers to better use data for scientific research, but also push the openness and sharing of astronomical data in more areas and scenarios. Data-intensively online astronomy EPO projects began to appear. Among them, the WorldWide Telescope(WWT) developed by Microsoft Research (Rosenfield *et al.* 2018), ESASky provided by European Space Agency(ESA) (Giordano *et al.* 2018) and the concept of data to dome proposed by International Planetarium Society (IPS: ips-planetarium.org) are relatively influential. Besides, with the continuous development of Big Data technology, such projects show good scalability and portability. For example, the interactive data visualization platform WWT can support normal PC, planetarium and virtual reality equipment. With the help of different media platforms and scenarios, astronomy knowledge will be transmitted more efficiently and accurately. The distance between the public and astronomical research are shortened.

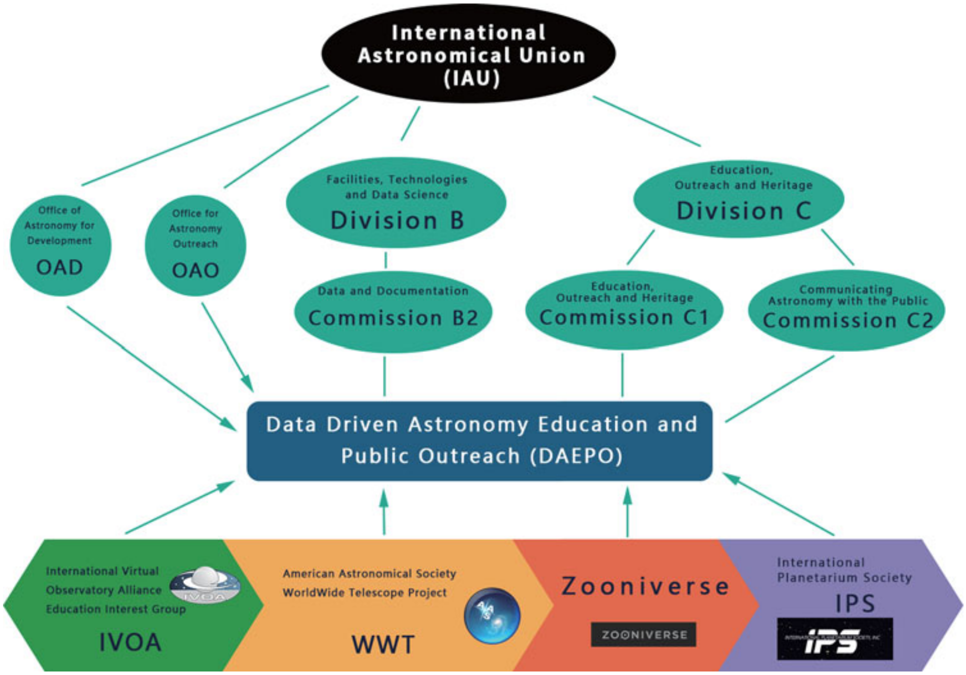

Figure 1. The organizational structure of IAU working group DAEPO.

With data related EPO projects become more and more active,people realized the importance of such projects to astronomy community and the development of astronomy EPO. In April 2017, Dr. Chenzhou Cui from National Astronomical Observatories, CAS (NAOC), PI of the Chinese Virtual Observatory, formally proposed the idea of DAEPO. He is the first one to use the phrase to summarize this kind of project. The International Astronomical Union (IAU) working group DAEPO established by him is hosted at the IAU Division B (Facilities, Technologies and Data Science) Commission B2 (Data and Documentation), and organized jointly with Commission C1 (Astronomy Education and Development), Commission C2 (Communicating Astronomy with the Public), Office of Astronomy for Development (OAD), Office for Astronomy Outreach (OAO) and several other non IAU communities, including IVOA Education Interest Group, American Astronomical Society Worldwide Telescope Advisory Board, Zooniverse Project (Fig. 1). It initially consisted of 16 members. All of them are professionals in related areas, including Dr. Beatriz Garcia (the chairman of IAU Commission C1), Dr. Pedro Russo (the chairman of IAU Commission C2), Dr. Sze-leung Cheung (the director of OAO), Dr. Kevin Govender (the director of OAD), Dr. Mark SubbaRao from IPS, Dr. Chris J. Lintott from Zooniverse Projct and Dr. Karen O'Flaherty from ESA. (daepo.china-vo.org)

According to the information on the official website of DAEPO working group, it has 3 major objectives. First, to act as a forum for people who interested in to discuss the value of the astronomical data in EPO, the advantages and benefits of data driven astronomy EPO, and the challenges it is facing. The working group also plan to provide guidelines, curriculum, data resources, tools, and e-infrastructure for DAEPO projects and activities. The third, it will provide best practices for reference. It is believed that with the promotion of this working group, the development of DAEPO will be more rapidly.

https://doi.org/10.1017/S1743921321000594 Published online by Cambridge University Press

In order to elaborate DAEPO projects more clearly, some rough definition of DAEPO summarized here. In a broad sense, all activities, projects, groups, organizations, concepts or platforms use astronomical data to carry out education and public outreach can be called DAEPO. In a narrow sense, DAEPO projects are specially designed interactive activities to use astronomical data as core, with the purpose of spreading astronomical knowledge. It usually has 5 elements: organizer or designer, astronomical data, platform or tool, astronomical knowledge, target group or participants. In the following chapters, best cases of DAEPO will be introduced according to the classification mentioned early in this paper. In addition to project introduction, topics including the scope of the audience, the effect of knowledge dissemination, whether astronomical data is the core (indispensable), if the participants make any scientific discoveries will be discussed.

3. Application of data products in EPO

Let's go back to the "pillars of creation" we mentioned at the beginning. This kind of picture generated by astronomical data is very attractive and may be used in many EPO scenarios like physics classes, popular science articles, movies, museums, etc. Although strictly speaking these activates are not DAEPO, but they are the most widely used method. They cultivate the essential environment for all DAEPO projects. In most cases, this kind of activates is composed by data products from professional astronomer and communication activities. The content of data products is relatively simple, could be a picture, a short video or a piece of sound. These activities have the widest audience and flexible forms compared with other EPO activities. The audience of such activities usually passively accept information, and the dissemination efficiency of knowledge usually depends on the communication means used.

3.1. *Data visualization*

Vision is the most important feeling of human beings. At least 80 percent of the external information of a people is obtained by vision. Naturally, the results of data visualization provided by large astronomical telescopes become a major part of astronomy EPO activities. These astronomical data visualization results are mostly spectacular, bizarre, beautiful and beyond imagination. They embody the mystery of the universe and bring new stimulation and attraction to audience.

The famous Hubble Space Telescope of NASA launched in 1990 generates about 10TB of new data every year. As one of the most influential space-based optical telescope, it provides the world hundreds of pictures of planets, stars, galaxies, nebulae, asteroids, deep space objects, etc. These beautiful and magnficent pictures always spread quickly and attract public attention. After properly combining popular culture and using some imagination, these data products tend to be more influential. The image of Pluto released by NASA's New Horizons mission rapidly spread the whole world, because a heart-shaped area on it was noticed. Besides, the phrase "Pluto's heart" is mysterious and interesting enough to attract attention.

Another noteworthy instance is the first-ever black hole picture realeased by the Event Horizon Telescope (EHT) collaboration. Under normal circumstances, black holes are invisible because their gravity is so extrem, even light can not escape. But for many people with little background knowledge, it is too abstract to understand. This visualization generated by around 5PB data make people understand black hole more intuitively. The topic about this picture was viewed 860 million times on the Chinese microblog platform called Weibo. (Figure 2 UP)

https://doi.org/10.1017/S1743921321000594 Published online by Cambridge University Press

Figure 2. (UP) Screenshot of Weibo topic of the first-ever black hole picture. (DOWN) Screenshot of Weibo topic of welcome package from University of Chinese Academy of Science.

3.2. *Data sonification*

Similar to data visualization, data sonification technology turns data to sound according to a certain internal logic. Some astronomers began to study it to help visually impaired people. And the meaningful products of the study can be also used in the field of EPO. For example, Jeff Cooke and his colleague from the Deeper, Wider, Faster (DWF) program build data sonification tools to help detect and study transient events in the Universe (Andreoni & Cooke 2019). CosMonic Projct developed by Ruben García-Benito uses data sonification to show astronomical theories. In 2020, the University of Chinese Academy of Science sent all freshmen an welcome package which including data sonification of 15 pulsar signals captured by FAST. The topic was very popular on Weibo and was viewed more than 110 million times in a short time. (Figure 2 DOWN)

3.3. *Means of communication*

Both data visualization and sonification products could be very simple and isolated. The audience with no research background in corresponding area can hardly interpret and learn anything from them. In this case, the astronomer, educator or media reporter who use these products need to provide external information to help people understand the knowledge behind the data. The final audience range and communication effect largely depends on these external information and the means of communication which closely related to the development of mass communication of human society. From the presses and publications, to science exhibitions, lectures, classes, summer camps, MOOCs, internet platforms, different communication way affect the efficiency of communication. But in general, the threshold for audience of these activities is very low. As long as the participants are interested in, they can be part of the process. And because audience usually passively accept information, the communication efficiency is not very optimistic.

Effected by the development of big data and Internet technology, people are immersed in all kinds of visual stimulation. Simple images of the universe in recent years were not able to attracted public attention as long as before. Stories and interesting details need to be connected with data products to create some social topics. In addition to pictures and sounds, video generated by data visualization may also be helpful.

4. Interactive data platforms

Interactive data platform usually is a data-intensive online platform with various of EPO functions and data built-in. Specially designed and developed by scientists and engineers, these platforms allow participants to interact with astronomical data flexibly and freely. Users can learn, review and apply scientific knowledge on platforms actively according to their own conditions. The emergence of these platforms are inseparable from the development of Big Data, Internet, cloud computing and information technology. It

https://doi.org/10.1017/S1743921321000594 Published online by Cambridge University Press

Figure 3. WWT classroom of Lingyuan No.2 High School in China.

improves the efficiency of knowledge transfer but also raises the threshold of participation. To participate, people must be able to access specific hardware devices and have some basic computer skills.

4.1. *Data visualization platforms*

For those who have sustained insterest in astronomy knowledge and a certain self-learning ability, a platform centralizes massive astronomical data and related information can provide more possibilities. Participants can log in at anytime and anywhere to browse the data themselves or to use the platform in an EPO activity. For example, the astronomy software Stellarium first released in 2007 is an platform embedded visualization data and rendered realistic skies. It provides convenient way for users to get the star distribution at any time and any place. It is used by amateur astronomer all over the world and been translated to nearly 100 different languages (stellarium.org). Another example is ESASky developed by European Space Astronomy Centre (ESAC) Science Data Centre (sky.esa.int). It is a science-driven discovery portal allows user to simply access astronomical data through browser (Merín *et al.* 2017). Although the main purpose of ESASky is to visualize the metadata of ESAC archives and external partners, its explore mode provides an excellent opportunity for public to access real astronomical data. Currently, the platform has Spanish, English and Chinese virsions.

As the first designed astrophysical data exploration platform in the era of Big Data, WorldWide Telescope (WWT) is a tool included real observational data and simulation data for multi-purpose, like demonstration of scientific research results and interactive STEAM education (Rosenfield *et al.* 2018). It has a web portal and also a PC client which is able to cache data for offline use. With the world wide telescope concept orginated in 2001 (Szalay & Gray 2001), the platform with same name was launched by Microsoft Research in 2008 and open-sourced in 2015. The management of it was transferred to American Astronomical Society (AAS).

In China, WWT debuted in 2008. Since then,the Chinese Virtual Observatory team keeps introducing it to teachers and educators in China and published articles, books to help them use astronomical data to orgnize EPO activities. Because of the flexible forms and functions of WWT platform, it can adapt to different groups of audience with different types of EPO activities. But the threshold for organizers is high. Till the end of 2020, totally 11 WWT teacher training workshops, 2 online sessions were held in China and four national WWT tour contests have been organized. Over 2,000 science educators and students participated directly in these programs. 8 classrooms and 16 planetariums equipped with WWT were already put into use or under construction (Fig. 3). In November 2019, the download times of the China-VO version of WWT reached 40,000.

https://doi.org/10.1017/S1743921321000594 Published online by Cambridge University Press

4.2. *Crowd Computing*

Originally designed and developed to support project SETI@home, Berkeley Open Infrastructure for Network Computing (BOINC) is an open-source middleware system for volunteer computing and grid computing. Now over 30 projects from diverse research areas are running on this platform, including astrophysics program like Asteroids@home, Einstein@home, Universe@Home, etc.† Benefit from the successful of SETI@home, volunteers from all over the world activily participate in these projects and help scientists with data analysis. Although this platform is successful in applying data processing, the participants are hard to learn anything from it because they only contribute their computer resources. Some of the projects have no detailed background, and the research related are too professional for general public to understood.

5. Scientific projects with public participation

Some DAEPO projects designed to let the public participate in the process of data acquire,classify and analyse. These projects usually proposed by astronomer in one particular field with a clear scientific goal. The organizer may set a requirement for participants and provide planty of background information. All people involed are consciously participating in the research process. According to the form of organization, there are two types of DAEPO projects fall into this category.

5.1. *School cooperation projects*

Astronomer from astronomical telescope or observatory can provide a certain observation time and data for EPO activities. They are usually familiar with telescopes' scientific objectives and are able to separate part of the research work to design some simple tasks for young students. In order to ensure the effect and for the convenience of management, some projects cooperate with interested and eligible teachers and schools. NSO mentioned before is one of the example. Although it is now open to users all over the world to apply observation time of Liverpool Telescope, cooperate teachers and students have higher authority.

Another example is the Pulsar Search Collaboratory (PSC) formed in 2007 to help scientists analyze more than 30TB data acquired by Green Bank Telescope.† Stared with local interested school and students, PSC has expanded nationwide (Williamson *et al.* 2019). Students have to attend an online training to participate and have actually dicovered new pulsars. Similar to it, PULSE@Parkes organized by Commonwealth Scientific and Industrial Research Organisation (CSIRO) Australia Telescope National Facility (ATNF) uses the observational time and data from Parkes radio telescope to study pulsars (Hollow *et al.* 2008).

5.2. *Citizen science*

The examples from above show there are some simple and repetitive tasks in astronomy research can be completed by students and the general public with certain training. Some astronomical discoveries in history have also proved this point. After some modification, projects like NSO and PSC can be easily extended to the public.

Galaxy Zoo is a DAEPO citizen science project originally designed for all the people around world. At the begining, it use the data set Data Release 7 (DR7) from Sloan Digital Sky Survey (SDSS) to let participants assist in the morphological classification of a large numbers of galaxies (Lintott *et al.* 2008). All galaxies are classified into one

† https://boinc.berkeley.edu/projects.php, Project list of BOINC
† http://pulsarsearchcollaboratory.com/home/about/, About PSC projects

https://doi.org/10.1017/S1743921321000594 Published online by Cambridge University Press

of six categories – elliptical, clockwise spiral, anticlockwise spiral, edge-on , star/don't know, or merger†. Later, more data set was added in, including optical Hubble Space Telescope surveys, CANDELS survey, MaNGA, UKIDSS, etc. Within a decade, more than 1.4 million galaxies was classified (Masters & Galaxy Zoo Team 2020).

After Galaxy Zoo, Zooniverse launched dozens of similar citizen science projects. Many of them can be called DAEPO projects. For example, the one cooperated with NASA using Wide-field Infrared Survey Explorer (WISE) mission data to find Planet Nine. In most of these projects, participants only have to classify or tag the data. More than a million people around the world participated projects on this platform. Over 180 papers were published in space area including 67 from Galaxy Zoo series of projects‡. A bizarre object called Hanny's Object near the spiral galaxy IC2497 is one of the most famous discoveries of the project (Lintott *et al.* 2009).

Popular Supernova Project (PSP) designed and created by Xingming Observatory and China-VO pushed a little further for the data used in this project is from an amateur observatory. Launched in 2015, over 20,000 user registered. They can start to browse the image after they pass the basic online test. Totally 24 supernovae and extragalactic novae were discovered. The youngest discoverer only 10-year-old. (psp.china-vo.org)

5.3. *Amateur observation*

Some astronomy EPO projects call on amateur astronomers to submit their observation results to help form valuable astronomical research materials. Like International Meteor Organization (IMO) collects meteor observational records from people around the world. The International Occultation Timing Association (IOTA) teaches amateur observe and time occultations to help the discovery of new double stars and other astronomical research. There are also projects like JUNOCAM from NASA allow participants to upload telescopic image of Jupiter captured by themselves to help NASA plan future mission.

6. Interdisciplinary study and activities

With the increasing professionalism of DAEPO projects, the threshold for participation is rising and the number of participants will relatively reduce. But they can provide better experience and more meaningful achievement for participants. If keep increasing the requirement of participants, for example, they have to be a well trained programmer or engineer, an interdisciplinary study activities can be organized. In 2013, Galaxy Zoo Challenge held by Zooniverse and Kaggle have data scientists write algorithm to classify the morphologies of galaxies. Over 326 teams participated§.

In China, China-VO team held an astronomical data mining contest with Alibaba Cloud to classify selected spectrum data from LAMOST. All participants can use the resource provided by Alibaba Cloud and apply machine learning method to analysis data¶. Nearly 1,000 people participated. China-VO organized 2 AI competitions base on PSP data in 2019 with FUTURELAB and Kaggle. 483 teams from universities and colleges participated in these contests.

At present, due to the high resource requirements, such DAEPO contests are rare. However, this may be an important direction of DAEPO in the future. In these activities, the audience are professionals from other fields. They will not only learn experience and

† https://data.galaxyzoo.org/, Data of Galaxy Zoo

‡ https://www.zooniverse.org/about, About Zooniverse

§ https://www.kaggle.com/c/galaxy-zoo-the-galaxy-challenge, The information of Galaxy Zoo Challenge, Dec. 20, 2013

¶ https://tianchi.aliyun.com/competition/entrance/231646/introduction/, Tianchi astronomical data mining competition, Apr. 15, 2018

https://doi.org/10.1017/S1743921321000594 Published online by Cambridge University Press

knowledge in astronomy, but also use astronomical data as research resources in their own fields. These DAEPO projects will help promote mutual exchanges and the progress in various scientific research areas. This is in line with the current development trend of astronomical research and the vision of IAU including OAD, OAO and the Office of Astronomy for Education (OAE). It will help further use of astronomy as a tool for development to benefit society. Currently, IVOA is working with OAD on utilization of astronomical data and related technologies.

7. Prospective development

Because of the diversity and flexibility of DAEPO, it is hard to predict the future development of it. But the trends can be discussed. With more and more people interested in DAEPO activates, professional, educator, students and public all begin to organize or participate in. For scientists, future research in some astronomy areas like time-domain astronomy may need the participation of amateur astronomers. Astronomy data scientists need to provide more available dataset and standards for DAEPO. For organizers, goals are to increase the scope of audience, reduce the threshold of participation and maintain scientific content. In order to achieve these goals, they need to work closely with scientists to carry out a delicate and appropriate design. For educator, suitable DAEPO projects can enrich teaching content and practice STEAM education. Combined with DEAPO projects and activities, some special tools and places like astronomical museum, interactive planetarium will be more common. In the future, The appear of more new media and technologies like virtual reality (VR), naked-eye 3D will bring more possibilities to DEAPO.

8. Acknowledgement

This paper is supported by National Natural Science Foundation of China (NSFC)(11803055), the Joint Research Fund in Astronomy (U1731125, U1731243, U1931132) under cooperative agreement between the NSFC and Chinese Academy of Sciences (CAS), the 13th Five-year Informatization Plan of CAS (No. XXH-13514, XXH13503-03-107). Thanks the teams of China National Astronomical Data Center (NADC), China-VO, and NAOC - Alibaba Cloud Astronomical Big Data Joint Research Center.

References

Gaia Collaboration, Brown, A. G. A., Vallenari, A., *et al.* 2020, arXiv:2012.01533

Qian, L., Pan, Z., Li, D., *et al.* 2019, Science China Physics, Mechanics, and Astronomy, 62, 959508. doi:10.1007/s11433-018-9354-y

Masters, K. L. & Galaxy Zoo Team 2020, Galactic Dynamics in the Era of Large Surveys, 353, 205. doi:10.1017/S1743921319008615

Cui, C., He, B., Yu, C., *et al.* 2017, arXiv:1701.05641

Rosenfield, P., Fay, J., Gilchrist, R. K., *et al.* 2018, apjs, 236, 22. doi:10.3847/1538-4365/aab776

Giordano, F., Racero, E., Norman, H., *et al.* 2018, Astronomy and Computing, 24, 97. doi:10.1016/j.ascom.2018.05.002

Andreoni, I. & Cooke, J. 2019, Southern Horizons in Time-Domain Astronomy, 339, 135. doi:10.1017/S1743921318002399

Merín, B., Giordano, F., Norman, H., *et al.* 2017, arXiv:1712.04114

Szalay, A. & Gray, J. 2001, Science, 293, 2037. doi:10.1126/science.293.5537.2037

Williamson, K., McLaughlin, M., Stewart, J., *et al.* 2019, The Physics Teacher, 57, 156. doi:10.1119/1.5092473

Hollow, R., Hobbs, G., Champion, D. J., *et al.* 2008, Preparing for the 2009 International Year of Astronomy: A Hands-On Symposium, 400, 190

Lintott, C. J., Schawinski, K., Slosar, A., *et al.* 2008, MNRAS, 389, 1179. doi:10.1111/j.1365-2966.2008.13689.x

Masters, K. L. & Galaxy Zoo Team 2020, Galactic Dynamics in the Era of Large Surveys, 353, 205. doi:10.1017/S1743921319008615

Lintott, C. J., Schawinski, K., Keel, W., *et al.* 2009, MNRAS, 399, 129. doi:10.1111/j.1365-2966.2009.15299.x