EPO and a Changing World
ASP Conference Series, Vol. 389, © *2008*
C. Garmany, M. Gibbs, and J. W. Moody

SpaceScience@Home: Authentic Research Projects that Use Citizen Scientists

Bryan J. H. Méndez

Space Sciences Laboratory, University of California at Berkeley, 7 Gauss Way, Berkeley, CA 94720-7450

Abstract. In recent years, several space science research projects have enlisted the help of large numbers of non-professional volunteers, "citizen scientists", to aid in performing tasks that are critical to a project, but require more person-time (or computing time) than a small professional research team can practically perform themselves. Examples of such projects include SETI@home, which uses time from volunteers computers to process radio-telescope observation looking for signals originating from extra-terrestrial intelligences; Clickworkers, which asks volunteers to review images of the surface of Mars to identify craters; Spacewatch, which used volunteers to review astronomical telescopic images of the sky to identify streaks made by possible Near Earth Asteroids; and Stardust@home, which asks volunteers to review "focus movies" taken of the Stardust interstellar dust aerogel collector to search for possible impacts from interstellar dust particles. We shall describe these and other similar projects and discuss lessons learned from carrying out such projects, including the educational opportunities they create.

1. Introduction

Some scientific endeavors require greater manpower (or computing power) than a team of professional scientists can practically accomplish. This is where the idea of "citizen science" is useful. Citizen science refers to a scientific project that makes use of a large group of volunteers to aid in performing critical project tasks. These volunteers may have no formal scientific training of any kind, but may still be able to aid in making basic observations, measurements, or computations. This is by no means a new concept. The Audubon Society has conducted the Christmas Bird Count (http://www.audubon.org/bird/cbc/index.html) since 1900 where birding enthusiasts across America help ornithologists take a census of early-winter bird populations.

The concept of the volunteers scientist is fundamental to the history of astronomy. For centuries amateur astronomers (those who do astronomy not as profession, but as a hobby) have been making critical contributions to the field. In the last century, their impact on the study of variable stars and minor planets has been staggering (http://www.aavso.org/aavso/membership/impact.shtml).

With the growth of the internet over the past two decades, ever more citizen science projects are taking shape. A great many of these projects are in the fields of zoology, meteorology, environmental science, and geology. For example, the United States Geological Survey (USGS) has been collecting online reports from citizens about the ground shaking following significant earthquakes (http://pasadena.wr.usgs.gov/shake/ca/). The business world has not let this idea go unnoticed either. Amazon.com has created a service called the Mechanical Turk (http://www.mturk.com/). This is an artificial, artificial intelligence service that pays "volunteers" for their efforts in everything from transcribing online videos to searching for Steve Fossett.

With the recent explosion of space science data now being collected world wide, citizen scientists can play a real and significant role in new discoveries. Projects using citizen scientists can not only help solve difficult problems space science, but also hold an enormous educational potential for those volunteers involved.

We will describe several recent citizen science projects in space science and discuss important lessons learned from them. We will also discuss ideas for future citizen science projects in space science.

2. SETI@HOME

SETI@HOME (http://setiathome.ssl.berkeley.edu/) from UC Berkeley's Space Sciences Laboratory is an early example of a citizen science project in space science. Radio SETI, the search for extraterrestrial intelligence, collects radio data from thousands of stars over a large frequency range looking for evidence of unnatural, narrow-band transmissions. Beginning in the 1990s, the project began generating enormous amount of data that proved too much for timely analysis by the resources of the project.

The team developed a program that runs on an internet-connected personal computer to download data from SETI, perform data reduction and basic analysis, then upload the results back to SETI. This program, SETI@HOME, runs in the background, only using the computer's free CPU cycles. Thus, a person can contribute to the project by volunteering their computer's idle time. The idea and program proved to be enormously popular; over five million people to date have subscribed to the SETI@HOME service. The success of SETI@HOME spawned the BOINC project (http://boinc.berkeley.edu/), which generalizes the SETI@HOME infrastructure for many different applications in massively distributed parallel computing.

The large number of subscribers to SETI@HOME demonstrates the strong interest in space science by the general public as well their strong desire to contribute to space science research. However, one drawback to the SETI@HOME approach is that it makes use of only the volunteers' computers, and does not engage the volunteers themselves in the process of research.

3. Clickworkers

At the beginning of this decade, the partnership of NASA Ames Research Center and the SETI Institute developed a citizen science project to identify cratering features on the surface of Mars. This project was called clickworkers (http://clickworkers.arc.nasa.gov/top).

In 2000, Clickworkers conducted a pilot test with over 800 volunteers using Martian surface imagery from the Viking orbiters. They compared the volunteers' work to that of professional planetary scientists and found a very tight match. Based on the success of the pilot test, Clickworkers began to make crater identifications in earnest in 2001 using Mars Global Surveyor imagery.

A new version of Clickworkers is now in beta testing using HiRISE imaging from the Mars Reconnaissance Orbiter. Volunteers may be able to request their own observations to be taken in this new version.

There are plans to extend Clickworkers from identifying craters in images of Mars to craters in images of the asteroids Vesta and Ceres which will be visited by the DAWN mission in 2011 and 2015, respectively.

Contact for Clickworkers: Virginia Gulick, SETI Institute, vgu-lick@mail.arc.nasa.gov

4. SPACEWATCH FMO Project

The SPACEWATCH program at the University of Arizona studies various kinds of small Solar System objects, such as asteroids, comets, and trans-Neptunian objects, and searches for objects that could be potential hazards for Earth (http://spacewatch.lpl.arizona.edu). In 2003, SPACEWATCH began a three-year, privately funded, citizen science program to search for Fast Moving Objects (FMOs; http://fmo.lpl.arizona.edu/FMO_home/index.cfm). FMOs move through the sky rapidly because they are close to Earth and are likely to be asteroids in a near Earth orbit. FMOs appear as streaks, or arcs in single epoch images. These streaks are difficult for software to differentiate from other possible features, such as cosmic ray strikes, background galaxies, or artificial satellites. People, however, demonstrate an inherent ability to easily distinguish between such features.

The SPACEWATCH FMO Project transmitted image data from Kitt Peak to Tucson in near real time and made that imagery available to the public on the same night it was collected via the internet, to allow for follow-up observations. It is essential to quickly follow up on possible Near Earth Objects (NEOs) as they can quickly be lost without accurate orbit determinations. Only fresh, unreviewed images were presented on-line.

SPACEWATCH recruited and trained hundreds of volunteers online, worldwide to examine image data. Volunteers were required to view an online training tutorial and pass a test before participating. Potential NEO candidates from volunteers were screened and evaluated and appropriate action was taken on the viable ones.

From 2003 to 2006, 52 volunteers spotted trails in images left by FMOs. Of those trails there were 43 discoveries of asteroids, 3 rediscoveries, 15 recoveries, 7 artificial satellites, and 8 unconsolidated discoveries.

There were some important lessons learned during the project. First, in order to calibrate the efficiency and reliability of the hundreds of volunteers, it would have been useful to inject fake candidates into the images sent out to users. Doing this would have allowed the SPACEWATCH team to prioritize the evaluation of FMO candidates. Candidates from volunteers with higher reliability could be reviewed first. Additionally, it would have been useful to have images reviewed by more than one volunteer, enhancing the efficiency of prioritizing candidates.

Additionally, there were problems with some volunteers who were not content to participate within the collaborative environment of the project. There were instances of volunteers hoarding images under multiple fake user names, and attempts to "discover" objects and send them to the Minor Planet Center as original observations, etc. Again, duplicate image delivery and the use of fake candidates could have helped here as a countermeasure. Other possible solutions might have been to work with specific groups of reliable people rather than the public at large. For example, classroom students could have been recruited to help and incentives given for their collegial participation.

Another important, if mundane, lesson learned was that adequate financial resources are needed to maintain such a project. For this project, the level of effort required to maintain the website, supervise, train, and correspond with volunteers was underestimated.

Contact for Spacewatch: Robert S. McMillan, University of Arizona, bob@lpl.arizona.edu

5. Stardust@home

In 2006, a new citizen science program was created at the UC Berkeley Space Sciences Laboratory. Inspired by SETI@HOME, Stardust@home (http://stardustathome.ssl.berkeley.edu) uses volunteers to help in the search for interstellar dust collected by NASA's Stardust mission.

NASA's Stardust mission launched in 1999 to collect dust from comet Wild 2 and return the samples to Earth. It was the first ever comet sample return mission. The Science objective of Stardust was to understand the materials and conditions that went into the formation of the Solar System.

Along the way to Wild 2, Stardust also collected Interstellar dust entering the Solar System from outside. Stardust used a two-sided, tennis-racquet-shaped collector with tiles of aerogel to collect star dust. The front side of the collector was used for collecting dust from Wild 2, while the back side of the collector was pointed into the incoming stream of interstellar dust twice during the flight to Wild 2. This was also the first-ever sample return mission from the Galaxy.

The density of interstellar dust is very low, and we expect only about 45 dust particles in total to have been captured in the 1,000 square cm collector. An automated microscope at the Johnson Space Center scans through the aerogel collector taking digital "focus movies." There will be over 1 million of these. Finding the interstellar dust particles will be like searching for 45 ants in a football field looking at one 5cm x 5cm square at a time.

The interstellar dust particles themselves are not visible in the focus movies, only the tracks they leave in the aerogel. We had no knowledge of the condition of the aerogel until it returned, nor do we really know what the particle tracks look like. In order to use pattern recognition software to find tracks we would have to first teach the software to recognize particle tracks and differentiate them from other possible features. To do that we would need to find a dozen or so particles!

We need people to look through the movies. People show a remarkable ability to visually discriminate between various features on the surface of the aerogel and possible particle tracks. However, the task of manually searching through a million focus movies would be overwhelming for a small research group but not for an army of enthusiastic volunteers.

We place the focus movies online for volunteers to examine. We originally estimated that it will take 30,000 person-hours to complete the task (assuming four people view each focus movie). We have had over 24,000 people registered for Stardust@home (leading to hundreds of views per focus movie).

It will take several months to complete the scan of the collector with the automated microscope, that is the limiting time factor. The staff at Johnson Space Center is not able to scan the collector consistently due to other constraints of the Cosmic Dust Lab.

Volunteers use a Virtual Microscope (VM) to examine focus movies. The VM works directly within a web browser. No special software is needed, simply an internet connection and a fair amount of RAM. A simple online training session and test are required for volunteers to learn how to use the VM. Each focus movie is viewed by many people at random. Calibration movies (movies known to have no particle tracks, or movies containing simulated particle tracks) are placed into the data stream to gauge each volunteer's sensitivity (find tracks when they are there) and specificity (find no tracks when they are absent). Each user receives an overall score based on their responses to calibration movies, which make up about 20% of the movies seen through the VM. When viewed in the VM, each real focus movie receives a score based on the number of times it has been seen and weighted by the score of the volunteer who flags it as either containing a track or not. The movies are sorted into a prioritized list based on their scores, and the UC Berkeley Stardust@home team follows up on movies that receive a high enough score.

The first focus movies were made available online on August 1st, 2006. After completing more than 1/3rd of the collector, we have a few dozen candidates, all of which are much smaller than expected and very subtle. So far they appear very different from expectations based on simulated tracks made with test particles fired into aerogel. In August, 2007, Stardust@home began Phase 2 of the search, which revisits the prior focus movies at high magnification. When scanning of the rest of the collector resumes at the Cosmic Dust Lab at Johnson Space Center, a Phase 3 of the search will commence.

The volunteers for Stardust@home (who have named themselves "dusters") have been a remarkable group and are very dedicated. The Stardust@home website features a 'Community' section where dusters can communicate with each other and the Stardust@home team via a bulletin board. This has been an effective way of communicating with such a large number of people. Dusters often answer each other's questions and small group of moderators selected from the dusters help keep things in order on the board.

We provide different kinds of incentives to reward dusters for their efforts. The website features a list of the dusters with top scores. Dusters achieving various score milestones receive a certificate of participation, and we invite top scorers to visit our lab at Berkeley for a tour. Additionally, the first to find a track will have the privilege of naming the dust particle and will also be given coauthorship of scientific papers about the discovery.

To recruit volunteers we initially advertised the project via a press release during Stardust's return in January, 2006. Hundred of news articles around the world were written about the project and dozens of radio and television news spots carried the story as well. Our website collected the email addresses of more than 110,000 people who were interested in the project. Once we began the project more than 24,000 volunteers registered to participate. Our initial analysis of an online survey of 10% of the dusters shows that they are 20% Female and 80% Male, of ages ranging from 6 to 99, and from over 60 countries on 6 continents.

Contact for Stardust@home: Bryan Méndez, University of California at Berkeley, bmendez@ssl.berkeley.edu

6. Systemic

Systemic, which launched in November 2005 (http://www.oklo.org/), takes a somewhat different approach to citizen science. This project from the University of California at Santa Cruz has produced simulated extrasolar planetary systems. Participants download software to observe and analyze and characterize the simulated planetary systems. Their analyses are compared to the original simulations to test for possible biases. The Systemic collaboration also serves as a community center for people to find the latest news about extrasolar planetary research and for some amateur astronomers who are trying to make their own observations of extrasolar planets via the transit method.

7. GLOBE at Night

Most of the space science citizen science projects mentioned so far make use of their volunteers by having them perform analysis on data collected by professional scientists. The GLOBE at Night project (http://www.globe.gov/GaN/) does the opposite. In order to characterize light pollution conditions around the world, volunteers are asked to observe the night sky on specific dates and report counts of stars seen at their locations. The Great World Wide Star Count (http://www.starcount.org/index.html) was another such program carried out in collaboration with GLOBE at Night. Volunteers may participate at various levels depending on their interest. Participants range from avid amateur astronomers to school children. GLOBE at Night began in 2006 and so far in 2007 has recorded nearly 8,500 observations world wide.

Contact for GLOBE at Night: Connie Walker, NOAO, cwalker@noao.edu

8. Galaxy Zoo

A new space science project for the citizen scientists out there has entered the scene in 2007 in the form of Galaxy Zoo (http://www.galaxyzoo.org/). The goal of the project is to obtain morphological classifications for a million galaxies observed in the Sloan Digital Sky Survey (SDSS-II). Morphologies help astronomers learn more about galaxy evolution and large scale structure in the Universe. Images of galaxies are taken for the Sloan Digital Sky Survey by a robotic telescope in New Mexico. The data are automatically reduced by software. The software also automatically searches the images for galaxies. The software can make many basic measurements of the galaxies, including size and brightness. But one property for which software is inadequate is that of determining morphological classification (e.g. spiral, elliptical, irregular, etc.). Here is another case where people are needed because they are better at pattern recognition in images than computer software.

Volunteers are presented with a tutorial and quiz to teach them how to classify galaxies and to recognize erroneous galaxy detections. Then, after registering, volunteers are presented with images from SDSS-II and asked to classify the galaxies as spirals (clockwise spiral, anticlockwise spiral, or edge-on), elliptical, merger, or star/don't know. Images are seen by multiple people for maximum reliability.

Volunteers are able to communicate with each other and the Galaxy Zoo team via an electronic bulletin board. They also may review galaxies they have classified and print images of them. Since the galaxies are initially identified by automated software, the volunteer viewing an image may be the first human being to ever lay eyes on that particular galaxy, a rather romantic notion.

As of this writing there are approximately 110,000 users of Galaxy Zoo who have completed 30 million classifications.

The contact for Galaxy Zoo: Jordan Raddick, John Hopkins University, raddick@pha.jhu.edu

9. Future Projects

Many possible citizen scientists projects in space science are being planned for the future. There are an enormous number of possibilities, especially with new all sky surveys being completed in the next several years (e.g.: LSST, WISE) and ever growing easy access to astronomical data via robotic telescope networks for education and the Virtual Observatories (NVO: http://www.us-vo.org/, and IVOA: http://www.ivoa.net/). We describe one possible project as an example, Asteroids@home.

9.1. Asteroids@home

The Wide-field Infrared Survey Explorer (WISE, http://wise.ssl.berkeley.edu) of NASA will launch in late 2009 and detect many possible Near Earth Objects (NEOs) that will require an enormous effort to confirm. In most cases, NEOs are asteroids in near Earth orbits, possibly in Earth crossing orbits, which raises the possibility of future collisions. WISE will have the advantage of being able to discover NEOs that are too faint to be seen by Earth-based searches, because asteroids are brighter in infrared than in visible light. In addition, WISE will survey portions of the sky closer to the Sun than Earth-based telescopes typically survey.

Candidate detections in the WISE data will be made by automatic software, but the thousands of anticipated candidates will need to be individually confirmed by a trained human eye to find the expected several hundred NEOs that can be discovered by WISE. Candidate NEOs will be identified by comparing sky exposures from one orbit to the next and looking for objects that appear in one frame but at a different location in the next frame. Automated software is effective at finding differences but not in correctly identifying them as moving objects. Software is often confused by artifacts such as cosmic rays, diffraction spikes, and objects at the detection threshold. People have shown a remarkable ability to quickly distinguish real NEOs from other objects.

WISE EPO at UC Berkeley will build on the successful model of Stardust@home to create an online project for WISE called Asteroids@home allowing the public to aid in the search for NEOs in the WISE data. Image data of candidate NEO detections will be served up to Asteroids@home volunteers via the Web, so that they can determine if the candidate is likely a real NEO or some other object. Training tutorials will be put in place to train the volunteers on what to look for in WISE data, and volunteers would need to take and pass a test to be able to participate in the search. We will calibrate each volunteer's responses by randomly adding calibration data into the mix of candidate images. The calibration images would be images that we know a priori are indeed NEOs or are not. NEO candidate detections would be ranked by a score determined from the number of volunteers viewing them weighted by the individual calibrations of those volunteers.

10. Conclusions

Citizen scientists are willing and able to join in the process of discovery and are an invaluable resource for many kinds of projects that, until recently, seemed impractical. For those planning to carry out a new citizen science project it is important to consider some key ingredients: provide ample background information for the volunteers; train volunteers with tutorials and consider requiring a test for participation to insure volunteers do indeed take the tutorials; provide incentives and rewards for volunteers to participate; calibrate volunteer responses whenever possible; promote the project via the media and/or through other projects themselves; provide as much feedback to the volunteers as possible; be sure to budget for the resources needed, because online citizen science projects require significant personnel time to design and maintain websites and software and to communicate with volunteers; people genuinely wish to participate in a real science project but are not interested in gimmicks created solely for the sake of outreach, so be sure that the project is an authentic task. Finally, if you have a project that could benefit from a large number of volunteers who do not need specialized knowledge, do it.