# csirocollogo1.gifAstromining: Using Astronomy Datasets to develop Critical Inquiry Skills.

**Version 2.0, presented at CONASTA 60, Darwin, Tuesday 12 July 2011.**

**Abstract:** Astronomy is undergoing a revolution in the size and availability of real data, much of it freely and publicly available via the web. Rather than use simplified or fake data and examples, students can now access the same data as used by professional astronomers and use it to make their own discoveries. This workshop presents several schemes bringing real data to students including *SkyServer*, *GalaxyZoo* and *PULSE@Parkes*. Examples of how to incorporate this into a range of classroom activities, ranging from traditional printed handouts to online and interactive use are presented. Possibilities for both guided and open-ended investigations will be discussed.

## What is Authentic Astronomical Data?

This is data obtained from a variety of astronomical instruments, primarily telescopes but it can also include results from computer simulations in theoretical astrophysics and the like. Data can be from ground or space-based instruments and may include:

* Spectra
* Photometric readings
* Images
* Other information such as magnetic field strength, temperature and so on.

Data is not limited to optical/visible waveband of the electromagnetic spectrum and may vary from radiowavelengths to gamma rays. It may also include data from non-electromagnetic spectrum sources. This could include chemical composition of meteorites, particle energies for cosmic rays through to simulations of gravitational waves.

## Why use Authentic Data?

This is an important question. Why use authentic or real data instead of simulated or simplified data? Too often we use data that has been simplified, adapted or massaged to make it easier for students to plot, analyse and clearly show a trend. Whilst this is a valid approach, especially when introducing concepts it can give students a misleading idea of the realities of scientific investigations. The advent of the internet, web tools and massive databases in astronomy means that today there is a wealth of real data available for anyone to use. Using real data presents students with real challenges, data that is fuzzy, incomplete or in unusual formats. It also, however, provides great opportunities for students. They may be viewing data that no one else has ever seen. They can make discoveries, use a variety of tools used by professional scientists.

## Data Formats

Data can be available in a wide variety of formats. Some of these may be readily useable in software likely to be found in schools. At its simplest the data may a *.txt* file or *rtf* that can be used by word processors such as *Word™.* More generally it may be a *.txt*, *csv* or *ASCII* file that could be imported into commonly available spreadsheets such as *Excel™* or a database.

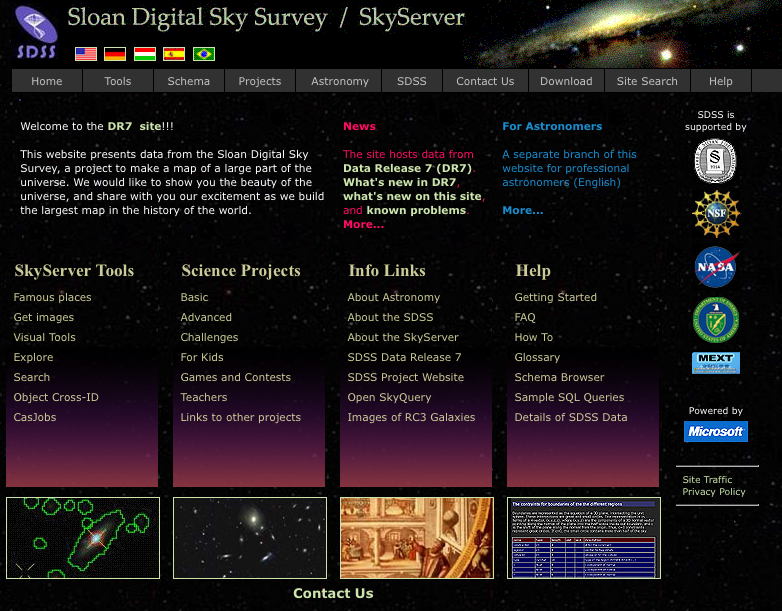
Images are a key component of much of astronomy. Whilst many spectacular public media release images such as those from the Hubble Space Telescope are freely available as JPEGS or high-resolution TIFF images these standard image formats are of limited scientific use. Fortunately there is a semi-standard format for raw astronomical images called *FITS* (Flexible Image Transport System). This uses metadata incorporating information such as filters, exposure times, coordinates and more in a header ASCII file that can be read by suitable programs. See the links section at the end of this paper for some places to start exploring astronomical image processing.

## Examples of Useful Educational Projects

There are many projects and sites utilising authentic data that are suitable for upper secondary students. Here are examples of some to indicate the variety and potential uses.

### SkyServer

*Skyserver* (<http://skyserver.sdss.org/> )is the educational interface for the Sloan Digital Sky Survey, SDSS. This survey ‘maps one-quarter of the entire sky in detail, determining the positions and absolute brightnesses of hundreds of millions of celestial objects. It will also measure the distances to more than a million galaxies and quasars’ (from Skyserver website). The Skyserver site accesses the same database as that used by professional astronomers but was designed from the outset to be used by non-professionals, primarily school students. It has an extensive range of tools, science projects, info links and help pages.

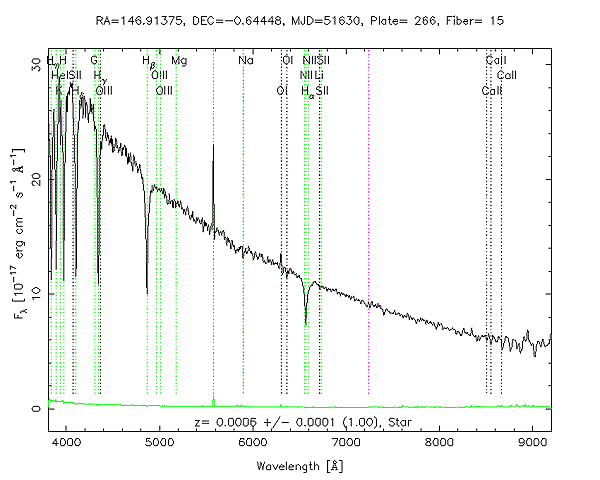


SkyServer homepage.

Projects are grouped into *basic*, *advanced* and *challenges* (which tend to be open-ended), each with detailed tutorials and extensive examples. The project section also includes simpler activities ‘For Kids’, games and puzzles and a detailed teacher guide. The Advanced Projects section includes the following topics:

* Hubble Diagram
* Color
* Spectral Types
* Hertzsprung-Russell Diagram
* Galaxies
* Sky Surveys
* Quasars
* Image Processing

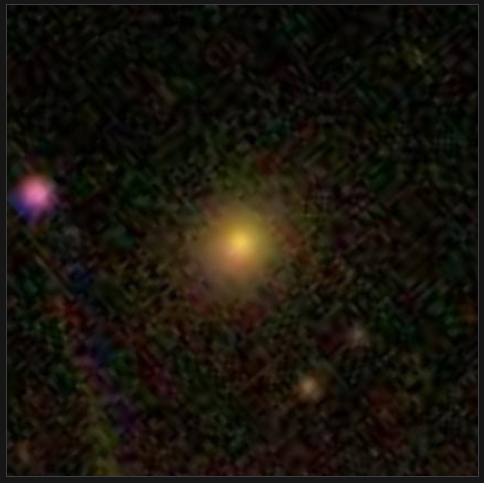
An example classroom activity on spectral classification using real data from SDSS is available at: <http://outreach.atnf.csiro.au/education/teachers/resources/>. Generating this was quite simple and quick using the tools on the site.



An example of a typical stellar spectrum from the SDSS database.

*Skyserver* is an excellent resource to inspire, educate and challenge high school students. The breadth and quality of the tutorials and other educational material provided allow you to cover much of an astronomy or astrophysics course focused on this site. Highly recommended.

### Galaxy Zoo

*Galaxy Zoo* (<http://galaxyzoo.org/> ) is perhaps the best-known and most successful example of what has become known as *Citizen Science*. Galaxy Zoo utilizes the vast SDS database (and now data from the Hubble Space Telescope) to tackle a specific problem of galaxy classification. Whilst not specifically aimed at formal education or high schools the project has many strengths that make it attractive for classroom use. It does not require any software to be installed as it runs over the web. If you wish to have students use it individually or in small groups they will have to register & sign-up after which they work through a short training tutorial that tests their understanding of the task. Feedback is provided at each stage. Once this is completed users are free to start classifying galaxies. The interface is clean, minimal and engaging. The process of classifying galaxies can be quite addictive. If you have access to the web and a data projector in your classroom you can run this as a whole-class activity, asking them to vote or justify their classifications. If you do not have online access in your room you can print out a selection of galaxies prior to the lesson. The beauty of *Galaxy Zoo* is its simplicity. The data is literally fuzzy and generates lots of discussion. There are currently Galaxy Zoo projects running; galaxy classification for Hubble data; <http://galaxyzoo.org/>, one on galaxy mergers; <http://mergers.galaxyzoo.org/> and one trying to identify supernovae; <http://supernova.galaxyzoo.org/>.

An example of a galaxy for classification in Galaxy Zoo. database.

Galaxy Zoo has over 400,000 followers/users worldwide. It has an active forum including an educators’ section where there are many excellent ideas and examples of how to incorporate it in classroom use.

The conceptual, pedagogical and software tools underpinning Galaxy Zoo now form the basis of a broader consortium, the *Zooniverse* <http://zooniverse.org> that seeks to house and encourage a broad range of projects.

The success of *Galaxy Zoo* has generated other astronomy-related citizen science projects. As of July 2011 these include:

*Icehunters.org* <http://www.icehunters.org/>is the newest project where you can try and identify Kuiper Be;t Objects in the search for the final traget for the New Horizons mission.

*The Milky Way Project* <http://www.milkywayproject.org/>where users identify bubles of nebulosity in the Milky Way.

*Moon Zoo* <http://www.moonzoo.org/> which uses high-resolution imagery of the lunar surface from the Lunar Reconnaissance Orbiter to study cratering on the Moon. This would suit students in upper Primary and lower Secondary.

*Planethunters.org* <http://www.planethunters.org/> uses data from the Kepler space telescope to try and identify possible exoplanets.

*Solar Stormwatch* <http://solarstormwatch.com/> that asks participants to view videos of solar flares from a pair of telescopes in space, the STEREO mission, to try and identify solar storms. This is also a great project for classroom use and worth exploring in more detail.

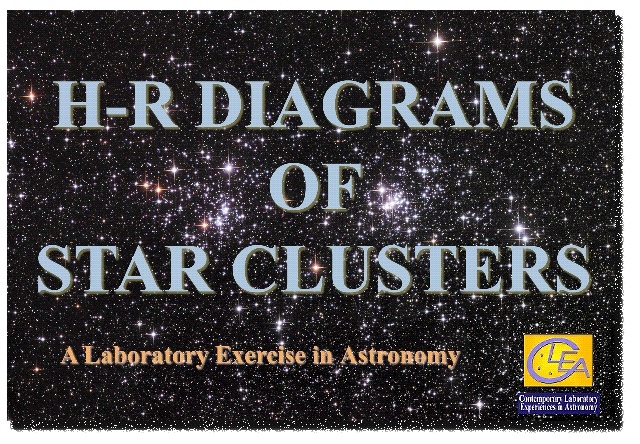
### Project CLEA (Contemporary Learning Experiences in Astronomy)

*Project CLEA* <http://www3.gettysburg.edu/~marschal/clea/CLEAhome.html> differs in implementation and intent to Galaxy Zoo as it is based on a suite of free astronomy exercises that you can download and install on PCs (only one is Mac compatible at present). The activities cover a wide range of astronomical topics and challenges but many/most incorporate real data. They allow students to simulate the tasks of observing, gathering then analysing data. Aimed at college-level student in the US they nonetheless are easily implemented at the senior high school level in Australia.

Perhaps the most exciting and encompassing one is their new suite, VIREO:

VIREO, the Virtual Educational Observatory, is a simulated observatory which can access a huge database of astronomical information, both through a set of dedicated catalogs and via on-line databases. It provides a set of optical and infrared telescopes of various sizes along with a radio telescope. Auxiliary equipment includes CCD and infrared imagers (which produce FITS files), an aperture photometer, a single-slit spectrometer and a multi-object fiber-fed spectrometer, and multiple tunable radio receivers. Analysis tools are provided for astrometry, photometry, spectrum analysis, and a variety of other purposes.

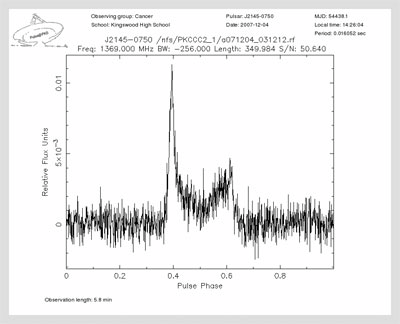
(from the Project CLEA website).

It includes many of their earlier activities that are each still available as separate packages. Many of these directly fit into requirements of senior physics astronomy or astrophysics topics. These include:

* Photoelectric Photometry of the Pleiades
* HR Diagrams of clusters
* Spectral Classification of Stars
* The Hubble Redshift-Distance Relation
* The Revolution of the Moons of Jupiter

There are also another 11 activities currently available for download. Each comes with pre and post-diagnostic tests, teacher/software manual and student manual/workbook. The materials and resources are excellent and worth exploring in detail.

### PULSE@Parkes

*PULSE@Parkes* (*PUL*sar *S*tudent *E*xploration online at Parkes) is a project operated by the Australia Telescope National Facility, part of CSIRO Astronomy and Space Science, <http://outreach.atnf.csiro.au/education/pulseatparkes/>. The project allows high school students to observe pulsars directly via remote control of the iconic 64-m Parkes radio telescope in NSW. Each observation is typically two to five minutes in length, generating a data file that can then be analysed by students using online data tools. At present we provide one two-hour observing slot to a school per month. All of the data obtained since the project’s inception in late 2007 is freely available online in the project data archive. A typical pulse profile from an observed pulsar is shown at right. Sometimes radio frequency interference means that the pulsar signal is swamped, resulting in much messier plots than this one. On other occasions, some of the selected project pulsars are “off”, displaying no noticeable signal. Students have to make decisions based on the visual feedback from the plots during their real time observations as to whether to keep observing or whether they should switch to a different pulsar.

The first module for the project uses the concept of *dispersion measure* to determine approximate distances to pulsars, which are then plotted on a plane projection of the Milky Way. This module runs on an ATNF server over the web and is available at: <http://pulseatparkes.atnf.csiro.au/distance/> Anyone can access this so students not involved in observing with the telescope can still analyse all the data.

The Single Pulse module: <http://pulseatparkes.atnf.csiro.au/single_pulse/>allows user to determine the period of any observed pulsar using an iterative technique.

Future modules are planned and under development so check back to the project page regularly for announcements.

All pulsar data collected by CASS radio telescopes, including that from PULSE@Parkes (P595) are freely available and searchable from the new *CSIRO Data Access Portal*: <http://datanet.csiro.au/dap/>. This access the data now stored at *ANDS*, the Australian National Data Service: <http://www.ands.org.au/>.

## More Information and Links:

* *Australia Telescope Outreach and Education* website: <http://outreach.atnf.csiro.au/>
* An online reference for the links in this presentation may be found at:  
  <http://outreach.atnf.csiro.au/education/teachers/resources/authenticdata.html>

# *The ESA/ESO/NASA Photoshop FITS Liberator*: <http://www.spacetelescope.org/projects/fits_liberator/index.html> was initally a plug-in for Adobe Photoshop that allowed users to use the raw image files from the Hubble Space Telescope data archive to produce their own images but is now a free, stand-alone package.

* FITS explanation on Wikipedia: <http://en.wikipedia.org/wiki/FITS>
* *Galaxy Zoo Hubble*: <http://galaxyzoo.org/>
* *Galaxy Zoo The Hunt for Supernovae*: <http://supernova.galaxyzoo.org/>
* *Galaxy Zoo Understanding Cosmic Mergers*: <http://mergers.galaxyzoo.org/>
* *Icehunters.org*: <http://www.icehunters.org/> is a Zooniverse project to detect KBOs for finl target for the New Horizons mission.
* *Image Processing Resources for Astronomy Teachers*: <http://www.phy.duke.edu/~kolena/imagepro.html> also includes a list of astronomy image processing curriculum activities.
* *IRIS*, a free, powerful astronomical image processor (Windows): <http://www.astrosurf.com/buil/us/iris/iris.htm>
* *Jayanne English’s Visualization Information:* <http://www.physics.umanitoba.ca/~english/viz2009/>has some excellent tutorials and links for astronomical image processing.
* The Milky Way Project: <http://www.milkywayproject.org/>
* *Moon Zoo:* <http://www.moonzoo.org/>
* *Old Weather:* <http://www.oldweather.org/>
* *Planethunters.org:* <http://www.planethunters.org/>
* *Project CLEA*: <http://www3.gettysburg.edu/~marschal/clea/CLEAhome.html>
* *PULSE@Parkes*: <http://outreach.atnf.csiro.au/education/pulseatparkes/>
  + *Module One for PULSE@Parkes* (distance determination):  
    <http://pulseatparkes.atnf.csiro.au/distance/>
  + Module Two: Single Pulse used for determining pulsar periods:  
    <http://pulseatparkes.atnf.csiro.au/single_pulse/>
* *Solar Stormwatch*: <http://solarstormwatch.com/>
* *SkyServer* (Sloan Digital Sky Survey educational interface): <http://skyserver.sdss.org/>
* SkyVirtual Observatory: <http://skyview.gsfc.nasa.gov/> is a Virtual Observatory on the Net generating images of any part of the sky at wavelengths in all regimes from Radio to Gamma-Ray.
* Stellar Classification Online Public Exploration (SCOPE): <http://scope.pari.edu/> is a citizen science project in which participants classify spectra of stars.
* The Network for Citizen Science Projects and resources (Science for Citizens): <http://scienceforcitizens.net/> allows you to search for relevant citizen science projects in a wide range of disciplines.
* *Zooinverse*: <http://zooniverse.org/>

If you wish to discuss use of authentic data with your students, find out more about these projects or get involved in PULSE@Parkes please contact:

**Robert Hollow**

Education Officer & PULSE@Parkes Project Lead

CSIRO Astronomy and Space Science

Phone: +61 (0)2 9372 4247

Email: [robert.hollow@csiro.au](mailto:robert.hollow@csiro.au)

This document is available online at: <http://outreach.atnf.csiro.au/education/teachers/resources/> in Word Doc and PDF formats.