(Introduction to) Getting Your Hands on Real Astronomy Data

L. M. Rebull
Caltech-IPAC/IRSA

11 May 2021

Context …

- A lot of times when I hear about "big data" kinds of problems, the data and often the things we learn from it doesn't take a ton of context to grasp, even if it takes sophisticated programs to pull/analyze the data.
  - Water usage in Toronto during the Stanley Cup…
- In the case of astronomy, there are some **big** hurdles to overcome in terms of:
  - Jargon both in terms of what is stored and what you could be looking for.
  - How you work with & interpret the data that are stored.
  - What needs to happen to raw data before you can do science.
- I'm assuming that y'all are very savvy about databases and coding, but not necessarily about the astronomy or how astronomy images work.
- In order to help you claw your way up the learning curve, we need to start a bit slowly, and maybe do some things on very small scales to learn about the basics before jumping whole hog into APIs, AI, ML.

I am SO NOT kidding about this. Even professional astronomers misuse the big catalogs all the way to journal articles! 😱

Outline

- Citizen Science – starting slowly
- FITS vs. JPG, GIF, etc.: why this matters
- Crash course in astronomical images
  - Color, artifacts, resolution
- Astronomical archives: IRSA
- *Key ideas have a box.*

Will put up a web page at the end which has a copy of this talk and links mentioned here.

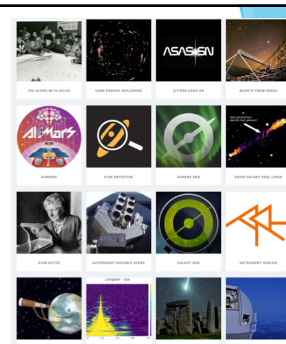
(Astronomical) Citizen Science

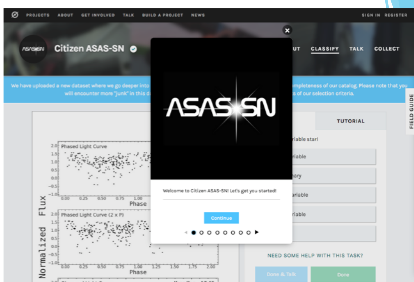
- Many, many programs out there.
- Most are :
  - Entirely web-based.
  - Require no special software (other than a web browser).
  - Require no deep astronomical knowledge.
  - (Yes, that means you're doing one at a time.)
- Some have lesson plans/labs/exercises.
- Some have hooks to go further when you are ready.

Many choices

- Zooniverse has a TON – go to the list of all Zooniverse projects, and scroll through the list!
- International Observe the Moon Night – moon watching
- Globe at Night, Great World-Wide Star Count – light pollution
- Citizen Sky – part of AAVSO, variable stars
- Cosmo Quest – Mars, Vesta, Mercury, and Moon crater mapping
- *And lots, lots more…*

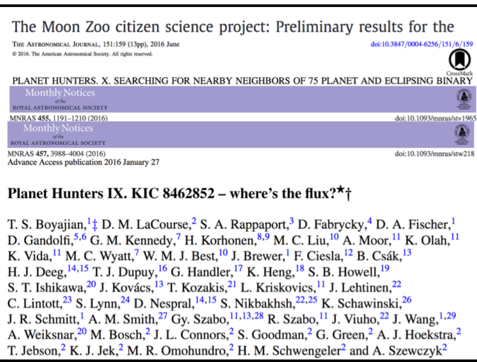
Zooniverse "menu"





Nice things

- Zooniverse (and others) have tutorials to help you get started.
- Usually the tasks they are asking you to do are not difficult – for a human – and can be easily explained via the tutorials.
- Kids can participate to the same degree as adults.
- ***You do not have to be perfect!*** Many people look at each item, and mistakes average out.
- Why are they going through this level of effort to get your help? Because the human eye is ***really*** good at identifying patterns, often way better than computers! (And ML just isn't there yet… **but you may be interested in these projects specifically for that aspect!**)
- You are contributing to real science!


The Moon Zoo citizen science project: Preliminary results for the

The Astronomical Journal, 151:159 (13pp), 2016 June doi:10.3847/0004-6256/151/6/159
© 2016. The American Astronomical Society. All rights reserved.

PLANET HUNTERS. X. SEARCHING FOR NEARBY NEIGHBORS OF 75 PLANET AND ECLIPSING BINARY

Monthly Notices of the Royal Astronomical Society
MNRAS 455, 1191–1210 (2016) doi:10.1093/mnras/stv1965

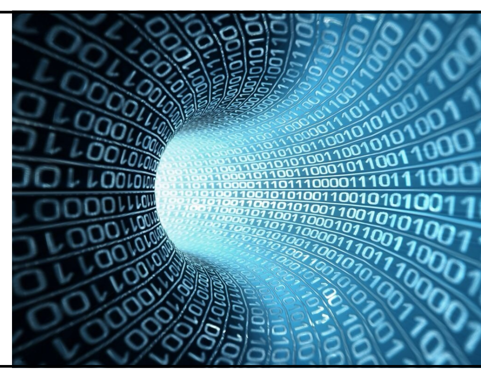
Monthly Notices of the Royal Astronomical Society
MNRAS 457, 3988–4004 (2016) doi:10.1093/mnras/stw218
Advance Access publication 2016 January 27

Planet Hunters IX. KIC 8462852 – where's the flux?★†

T. S. Boyajian,[1‡] D. M. LaCourse,[2] S. A. Rappaport,[3] D. Fabrycky,[4] D. A. Fischer,[1] D. Gandolfi,[5,6] G. M. Kennedy,[7] H. Korhonen,[8,9] M. C. Liu,[10] A. Moor,[11] K. Olah,[11] K. Vida,[11] M. C. Wyatt,[7] W. M. J. Best,[10] J. Brewer,[1] F. Ciesla,[12] B. Csák,[13] H. J. Deeg,[14,15] T. J. Dupuy,[16] G. Handler,[17] K. Heng,[18] S. B. Howell,[19] S. T. Ishikawa,[20] J. Kovács,[13] T. Kozakis,[21] L. Kriskovics,[11] J. Lehtinen,[22] C. Lintott,[23] S. Lynn,[24] D. Nespral,[14,15] S. Nikbakhsh,[22,25] K. Schawinski,[26] J. R. Schmitt,[1] A. M. Smith,[27] Gy. Szabo,[11,13,28] R. Szabo,[11] J. Viuho,[22] J. Wang,[1,29] A. Weiksnar,[20] M. Bosch,[2] J. L. Connors,[2] S. Goodman,[2] G. Green,[2] A. J. Hoekstra,[2] T. Jebson,[2] K. J. Jek,[2] M. R. Omohundro,[2] H. M. Schwengeler[2] and A. Szewczyk[2]

If these citizen science opportunities are the only thing you remember from this talk, and especially if you really get into these, then I've done my job…

However… maybe that just whets your appetite and you want more, more, more!!



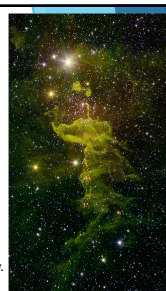
Then …

- For decades (centuries?), you just took your data home when you left the telescope.
  - Corollary: if you didn't have a telescope, no data.
- Early archives are where you went to:
  - Retrieve your own (unreduced) data.
  - **Maybe** you might look for other data from a particular source (at a particular location) from a particular data set.
  - For some data sets, download the whole thing and sort it out later…

…and Now

- Now, that PLUS:
  - Publicly funded assets mandated to make data accessible.
    - Corollary: if you don't have a telescope, you can still get data. (more equitable!)
  - Are there other data on this target from another date? Survey? Wavelength?
  - That someone else reduced and archived?
  - Other data on similar objects?
- Soon:
  - As data get bigger and bigger, won't be able to pull all data out of the archive and take home.
  - Can you get started analyzing before going home?
  - Mission evolving from "search-and-retrieve" to "do [some] analysis in situ."

Role of (Astronomy) Archives



- Ingest new data (and new reprocessing of old data).
- Maintain/serve vital repository of irreplaceable data:
  - Support for *observation* planning and *mission* planning.
  - Resource for original science.
  - Make high level science products available.
- Enable cutting-edge research:
  - Among many other things, that means APIs and Virtual Observatory.
- Science discoveries waiting in the archives that were never imagined or expected by the mission or even program PIs.

Astronomical Archives

- It's really true that there are **more data** than professional astronomers can hope to completely mine, with more coming in **all the time**.
- (Professional astronomers often pull off the low-hanging fruit, but that doesn't mean there's no other fruit on the tree!)
- **Any facility that comes from public funding is supposed to have a publically accessible archive. These are your data!**
- NASA is super good about this. NSF getting there. Europe is catching up. Other countries (e.g., China) not so much…
- NASA data alone will keep you busy for quite a while.
  - Astronomical data ←
  - Moon, planetary data
  - Earth observing data

Astronomical Archives cont'd

- Archives designed for the professional astronomy community.
- Archives usually designed to be easily accessible .. And if the archive team has done their job, **you** should be able to get into it without trouble!
- However, you do have to "reach across the barrier" to become familiar with some conventions, jargon, etc. It's not going to be packaged up for you with a nice little ribbon!

Concept #1

Don't use MSIE
(Internet Explorer!)
Or, for that matter,
Edge.

Ditch MSIE/Edge

- Don't use either as your default browser.
- Go get Firefox or Chrome.
- MSIE/Edge is buggy and has security holes.
- Most astronomers use Mac or Linux machines, so are not using MSIE/Edge.
- Most archives are operating more or less on a shoestring. It takes a lot of resources to make sophisticated web-based tools work on MSIE/Edge. With very few astronomers using MSIE/Edge, that's a lot of resources to pour into something with little reward. Thus, MSIE/Edge support is often dropped, **so some archives won't work with MSIE/Edge.**

Concept #2

Image "mechanics"

Concept #2a: Bit depth

- All digital images are "ones and zeroes" – your computer just knows how to handle those ones and zeroes.
- Images from your digital camera (JPGs, GIFs) are typically "8 bits deep."
- That means that there are only 2 to the power of 8 (2^8=) 256 discrete levels of information for each pixel (per color plane).
- That usually doesn't matter for you (unless you are a graphic designer or web developer…or an astronomer).
- Astronomical images have a much greater **dynamic range**; they are at least 16 or even 32 bits deep -- that is, there are at least 2 to the power of 16 (2^{16}) or 65,536 and possibly (2^{32}=) 4.3 billion possible discrete levels of data for each pixel.
- That kind of detail gets lost when you save an image as a JPG. *The computer is compressing 65,000 levels into 256.*

Think of bit depth like this…

Array of thimbles

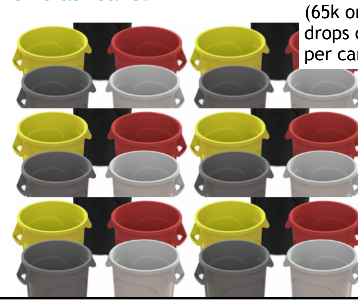
Can't get very much more than a dribble of water in each one. ("256 drops"!) Each overflows quickly! This is ok for images taken with your cell phone.

Teacups are bigger

Now can get more than a dribble of water in each one. But they still overflow easily!

Ooh! Trashcans are bigger still!

Array of trashcans!



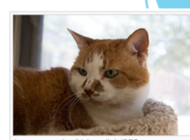
LOTS of water in each one! (65k or 4.3b drops of water per can!)

Concept #2a: Bit depth

- "8 bits deep" = Images from your digital camera (JPGs, GIFs) = 256 (=2^8) drops can fit in each bucket (pixel).
- "16 or 32 bits deep" = Astronomical images = 65,536 (=2^{16}) drops can fit in each bucket, or even 4.3 billion (=2^{32}) drops can fit in each bucket.
- [AND, in your camera, you have one array of buckets for each color (red, green, blue), but we will get to that momentarily!]
- The analogy is imperfect, but does ok. Higher numbers of pixels in your camera (e.g., comparing cell phone to "real" camera) usually refers to how densely the buckets are packed together and how many total buckets there are. How deep the buckets are = bit depth.
- Key concept: astronomical images can hold more information than JPGs or GIFs.

Concept #2b: Compression

- JPGs are "lossy compressed" – they compress the data such that the file takes up less disk space, but it actually loses information!
- Someone worked hard to collect all those photons .. Don't lose information at the last step!
- (Loss-less compression is ok.)
- If you want to play with real astronomical images, you can't be using JPG, GIF, PNG…



Low compression (high quality) JPEG

High compression (low quality) JPEG

Concept #2c: FITS files!

- **FITS = Flexible Image Transport System**
- Most professional astronomers use this format. Not compressed, no loss of information at all.
- FITS images consist of a plain text header and the binary image.
- The header usually contains "metadata" – information about the data: coordinates of the image; maybe things about, e.g., target, telescope, astronomer(s), observation date/time, wavelength, even data reduction steps.
- The binary image can be one plane or many planes of images. It can also be a table of data (like a spectrum, or literally a table).

FITS, cont'd

- Need to have software that can read FITS.
- Image J, MaxIm DL, …
- Photoshop (not free) has a free plugin (FITS Liberator).
- Astronomers like free software.
  - IRSA tools I will show you shortly.
  - ds9 (Google that with "Harvard" to find it).
  - Python tools (look for Astropy).
  - …

Concept #3

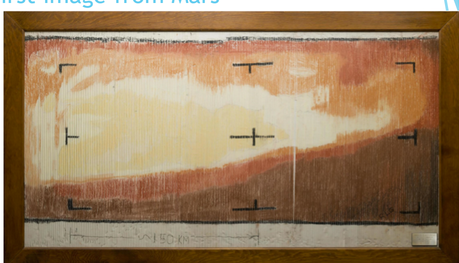
Color in images

Concept #3a: Color images

- Ok, so we have our array of 'buckets' that have collected 'water' (light). How will we display those water levels?
- How will we represent the numbers on our screen?
- (*historical detour…*)

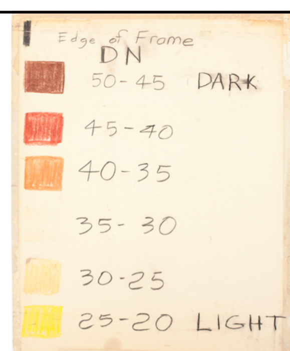
First image from Mars

Mariner 4, July 15, 1965



http://www.directedplay.com/first-tv-image-of-mars/

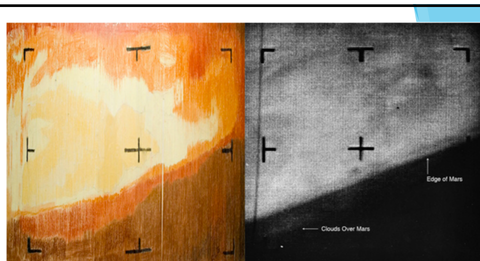
…print out the digits and color over them based upon how bright each pixel was.



- Brown/red color scheme, but didn't know Mars was red.
- Was looking for the colors that best represented a grey scale

since that was what they were going to get anyway.

- Could have been green or purple!



Hand-done Mars image (left) and actual Mariner data (right).

Color-by-number

0 1 2 3 4 5 6 7 8

Within an image, assign a number to a color.

What if have values more than 8?

Color-by-number

0-1 2-3 4-5 6-7 8-9 10-11 12-13 14-15 16-17

Now you can show data ranges from 0-17 (not just 0-8).

What color should 7.9 be?

Color-by-number

0-1 2-3 4-5 6-7 8-9 10-11 12-13 14-15 16-17

Now you have a more continuous spectrum of colors.

What if you have numbers >>17?

Color-by-number

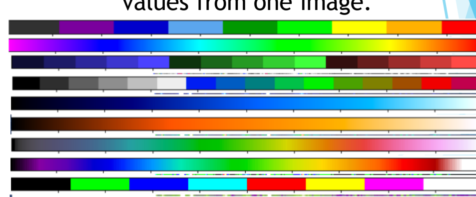
10^0 - 10^1 10^1 - 10^2 10^2 - 10^3 10^3 - 10^4 10^4 - 10^5 10^5 - 10^6 10^6 - 10^7 10^7 - 10^8 10^8 - 10^9

Now you can display a MUCH larger range of numbers!

But, you are still working with just one image. Values from that one image are displayed as a color.

Color Tables

Color tables describe the range of colors to which you are mapping the data values from one image.



Color Stretch

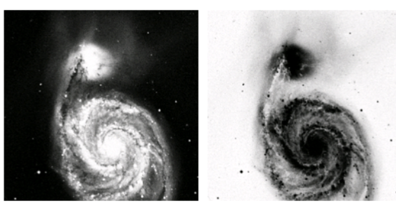
Color stretch describes how you translate the the range of data values to the range of colors from one image.

0	1	2	3	4	5	6	7	8
0-1	2-3	4-5	6-7	8-9	10-11	12-13	14-15	16-17
10^0 - 10^1	10^1 - 10^2	10^2 - 10^3	10^3 - 10^4	10^4 - 10^5	10^5 - 10^6	10^6 - 10^7	10^7 - 10^8	10^8 - 10^9

(I've demonstrated linear and logarithmic stretches here, but there are myriad others.)

One image

- One image can be represented as shades of grey. (or red pastels like Mars image!)
- Astronomical images in one filter are often shown in greyscale or reverse greyscale.
- This is the least distracting way to display details in the image!
- (Also uses less printer ink!)



Concept #3b: Color images

There is no such thing as a color image!

Whut?

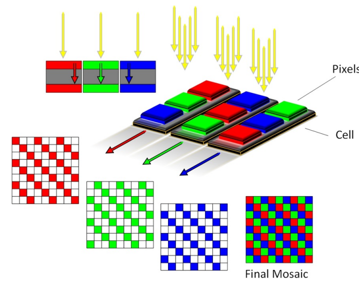
One image … to 3-color image!

- What if we took one image, and rather than scaling it to shades of black to white, did scales of black to red?
- Another image: black to green
- Another image: black to blue
- **Stack them up:** this is how you (your tv, your camera) construct a 3-color image!
- (faint in all 3=black; bright in all 3=white)

All color images ..

- **..are made up of 3 (sometimes 4) separate images.**
- Photographs: 3 emulsions (R,G,B)
  - (RGB=red, green, blue)
- Digital cameras/TVs: 3 sets of pixels (R,G,B)
- Digital systems sometimes have c,m,y,k
  - (cyan, magenta, yellow, "key"=black)
- Astronomical images have filters!

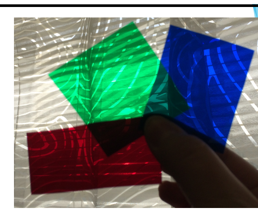
Demonstration of 3-color images from 1902



Astronomy

- You take an image in one filter….

Filters



- Kinda sorta like pieces of colored cellophane (here, theater gels).
- Only let in light over a narrow(er) range in wavelengths.
- It is not a coincidence that I picked red, green, and blue (R, G, B) here!
- Purpose: Make the light going through your telescope go through a filter before it hits your detector. Only record light that makes it through the filter.

Astronomy

- You take an image in one filter.
- You assign a color to an image from that filter.
- Repeat as necessary!
- Conventionally shortest wavelength is blue, longest wavelength is red.
- Assigning shades of color? Can do it ANY WAY YOU WANT.
- How you map the numbers to the colors matters for what details you bring out in the image.

"True color"

- Lots of astronomy public images describe things as 'true color' or 'false color'.
- What is "true color", really?
  - http://www.planetary.org/blogs/emily-lakdawalla/2015/12291508-two-epic-photos-of-earth.html
- "False color" makes it seem like we're hiding something.
  - But how do you make a "true color" IR image?
- **"Representative color"** is a better choice.

Common misconception

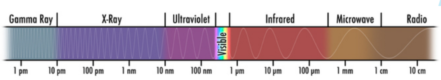
- Some people have the (incorrect) idea that astronomical images are obtained first, then broken into colors later.
- If you think of taking just one image and changing the color table, then yes, that works.
- But to make a color image, you need three images taken in different filters (like R, G, B).
- **You can't separate the photons after you record them.** You have to take 3 images in 3 filters, then combine them; you can't differentiate photons afterwards.

Concept #4

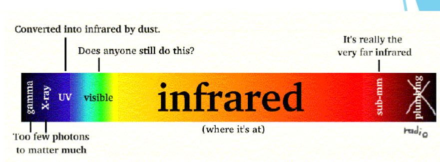
Astronomical data…

Concept #4a: multiple wavelengths!

- In astronomy, we're studying things we can never visit (or often even see from another angle).
- We have to take advantage of all the light we get from these things, even light that doesn't make it through the atmosphere to us!
- Lots of telescopes, data, **archives** from gamma rays through radio!



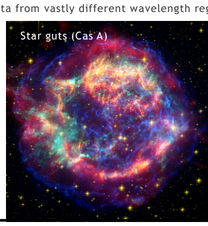
IPAC's Electromagnetic Spectrum

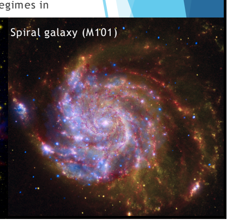


Slide from G. Rieke

Light beyond the visible

- Lots of images in wavelengths other than visible - so "true color" is black.
- Lots of images combining data from vastly different wavelength regimes in the different color planes.
- These combine Chandra (X-ray) data, Hubble (optical) data, & Spitzer (infrared) data.





Concept #4b: sky coverage, resolution

- Not every telescope looks at everything in the sky. (Some are surveys, many are not.)
  - In any given archive, your object may not have been observed. (yet?)
- Different telescopes may use different wavelengths of light (radio through gamma rays).
- Different telescopes have different pixel sizes, so a "pinpoint" of light may look like a pinpoint to one telescope but a much larger blob to another.
- (Also sometimes there are real astrophysical differences so what looks like a point source at one wavelength may legitimately be a blob at another wavelength.)

M16: Pillars of Creation

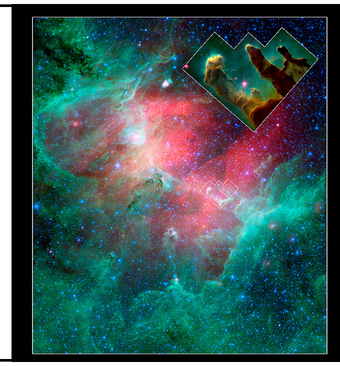
HST-WFPC2 (optical)

Hester & Scowen (Arizona State/ NASA), Nov 1995

RGB image!

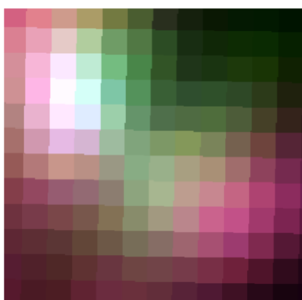
Image is ~1.2 arcmin square

Blue=[O III] (5007 A); Green= Ha (6560 A); Red=[S II] (6731 A)

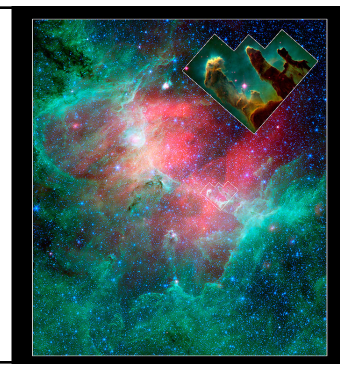


Spitzer : 4.5, 8, 24, and 70 microns *(cmyk image!)*

(Flagey et al. 2007)

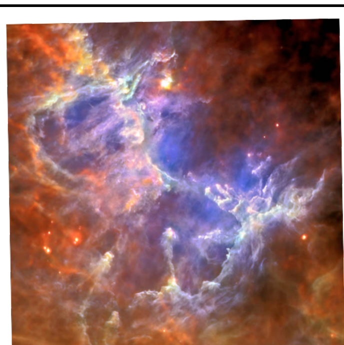


IRAS : 12, 25, 100 microns



Spitzer : 4.5, 8, 24, and 70 microns

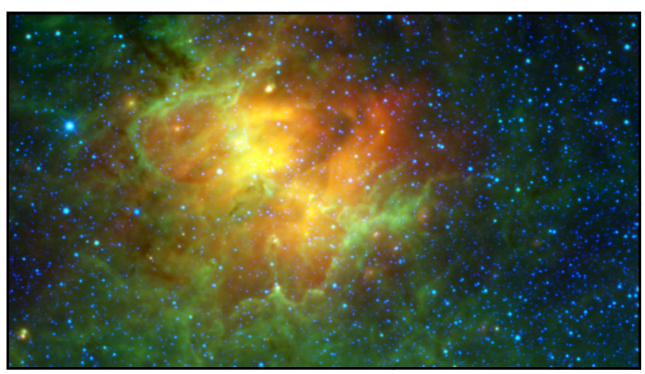
(Flagey et al. 2007)



Herschel : 70, 160, and 250 um.

Dust is ~10K (red) to ~40K (blue).

ESA/Herschel/PACS/SPIRE/Hill, Motte, HOBYS Key Programme Consortium



Concept #5 (last one!)

Artifacts!

Concept #5: Artifacts

- You're working with real data.
- Not prettied-up for public consumption.
- There will be artifacts – that is, stuff that isn't really part of the sky – in the image.
  - It's part of how the telescope responds to light, or a plane flying through the image, or even a bit of dust.
- Many big images are actually composed of lots of smaller ones, knit together ("mosaic").
  - You may be able to see these tile boundaries.

Lights here are not *really* gigantic fuzzy blobs with funny purple halo/shadows lurking nearby. This is just how this camera+detector responded to this lighting situation, where these lights are very bright. Same thing for astronomy: There are features that are just a result of how the telescope+instrument+detectors respond to a point source of light, and these features are more prominent when the source is very bright.

Artifacts

- Neil deGrasse Tyson on Colbert: **"Excuse me. Just because you don't understand what you're looking at doesn't mean it's aliens."**
- If you find something that looks weird, DO NOT assume it's really in the sky. It is much more likely to be saturation effects, scattered light, a plane, resolution issues, just how the telescope responds to bright light … Read all the documentation and understand the possible artifacts … educate yourself, try to convince a friend that it's not artifacts.
- (Don't immediately email the archive's helpdesk that you've found evidence of alien life that the government has been hiding …)

No giant space praying mantises .. Please …

Data acquisition and reduction

- Different telescopes work in entirely different ways to acquire data (even within the same wavelength regime), and the steps you take, the artifacts you have to deal with, the final product(s) you come up with is entirely customized to the telescope+instrument you use.
  - Same telescope, different instruments, totally different pipelines and final products.
- You need to take into account telescope+instrument peculiarities, but also turn individual images into quantitative, calibrated measures of brightness.
  - Detectors are linear over a regime; need to remove effects from telescope+instrument and determine how bright objects are on an absolute scale.
- Big facilities/surveys often (but not always) have established data reduction pipelines on which you can rely – but there too, you have to RTFM and learn what to believe (or not believe) in the data products.

OK, now … real archives!

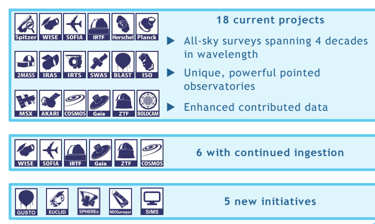
NASA Astronomical Archives

- IRSA (@Caltech) – Infrared & longer wavelengths
- Exoplanet Archive (@Caltech) – exoplanets
- NED (@Caltech) – galaxies
- MAST – Optical
- HEASARC – High energy
- ADS – Literature
- (Other US archives, plus Europe, Canada, …)

IRSA (where I work!)

- IRSA = NASA's Infrared Science Archive.
- Physically at Caltech in Pasadena, CA.
- Charter is to provide interface to all NASA infrared and sub-mm data sets (~1 μm to ~1 cm).
- Founded 1993, original home of IRAS data (1983).
- More than 12% of all refereed journal articles published today use data at least in part that came from IRSA.
- More than 5 PB of images and ~250 TB in databases. (1 PB=1024 TB=10^6 GB)
- All-sky coverage in 24 bands.

Rich data sets = enormous science potential

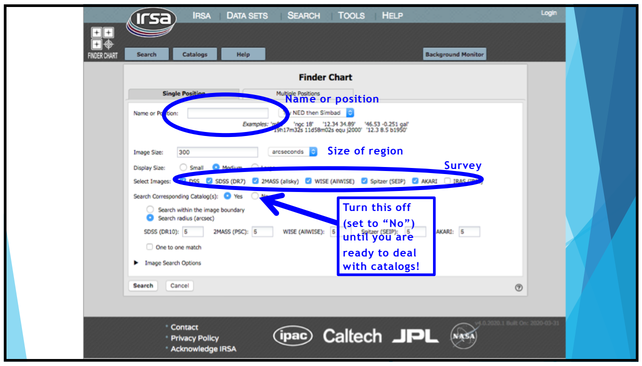


Firefly Tools

- Many IRSA tools have a similar look-and-feel.
- That software is called "Firefly" and is open source (see GitHub).
- Interactive tables, plots, images, overlays.
- (There are also APIs to interact with Firefly as well as the archive itself.)

Getting started

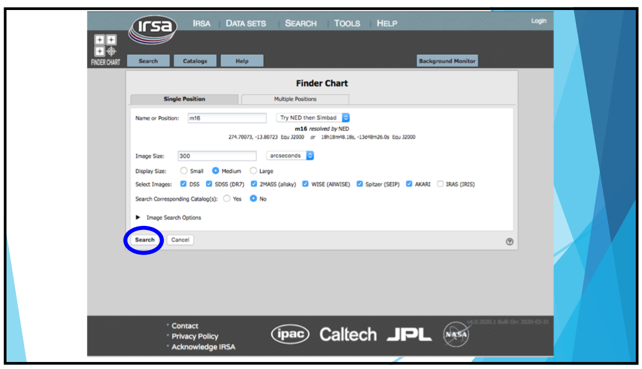
- Going to point you towards Finder Chart because it gives you the same chunk of sky in several wavelengths from several different surveys all at once.
  - http://irsa.ipac.caltech.edu/applications/finderchart/
- (There are lots of video tutorials in the IRSA YouTube feed that explain how to use many of the archives with a Firefly interface.)

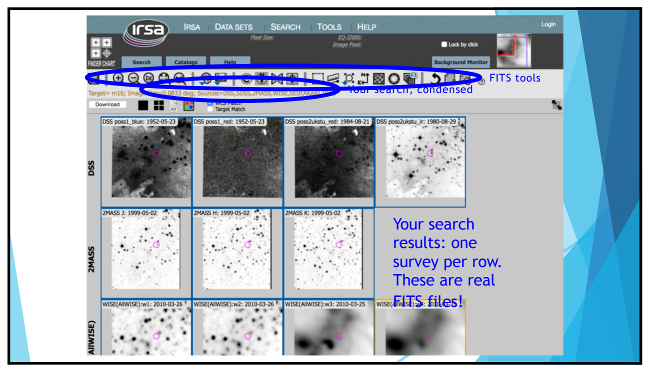


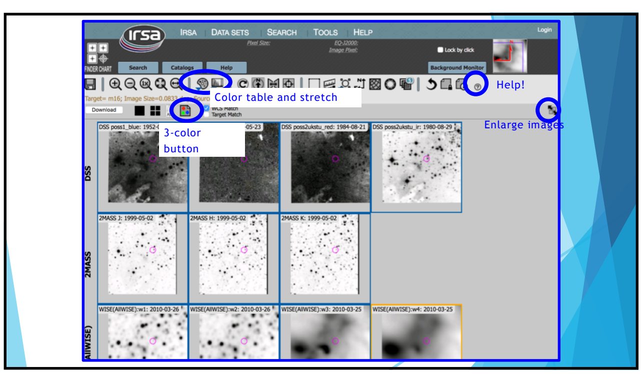
Bands

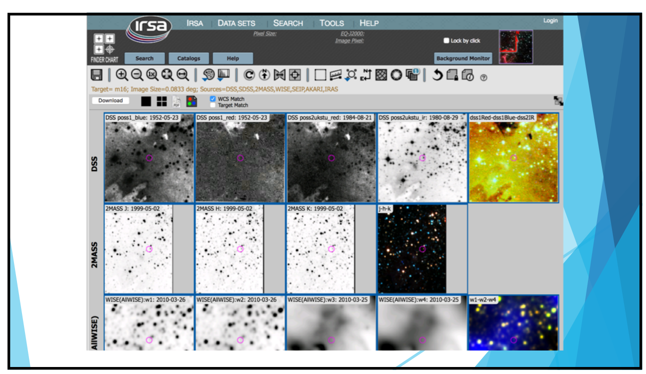
- *Optical*: DSS (Palomar Observatory Sky Survey): blue, red, NIR
- *Optical*: SDSS (Sloan Digital Sky Survey): *ugriz* (blue to NIR)
- *Infrared*: 2MASS (2-Micron All Sky Survey): *JHK* (NIR, 1-2 microns)
- *Infrared*: WISE (Widefield Infrared Survey Explorer): 3.5, 4.6, 12, 22 microns
- *Infrared*: SEIP (Spitzer Enhanced Imaging Products): 3.6, 4.6, 5.8, 8, 24 microns
- *Infrared*: AKARI (Japanese satellite): 60, 90, 140, 160 microns
- *Infrared*: IRAS (Infrared Astronomy Satellite): 12, 25, 60, 100 microns

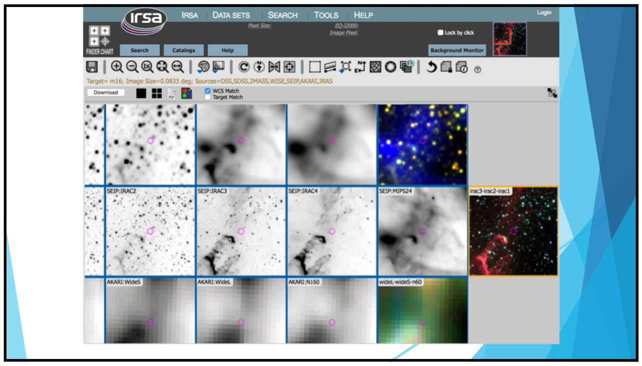
- Can also load in a FITS file from other IRSA holdings, off your disk, or off the web!

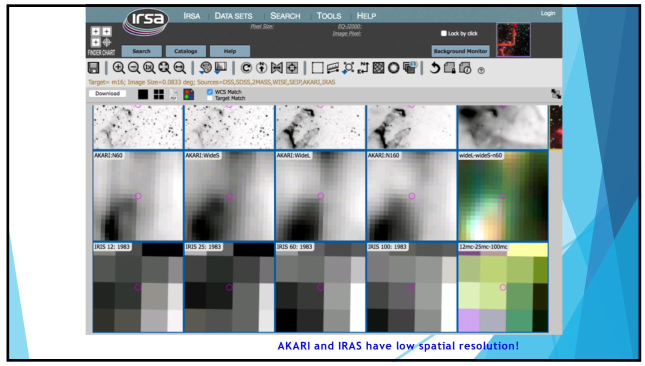


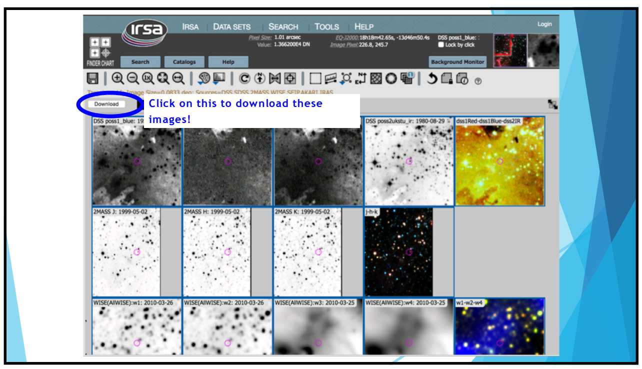


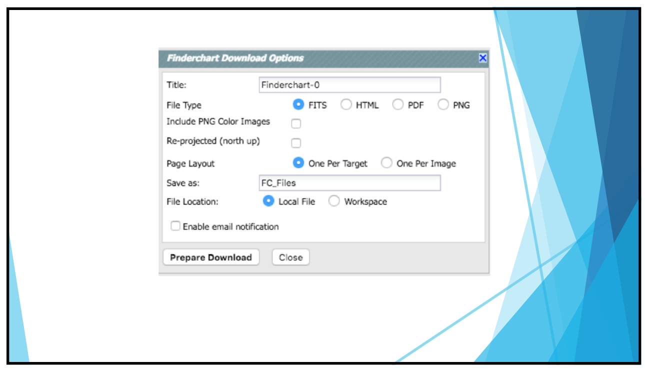












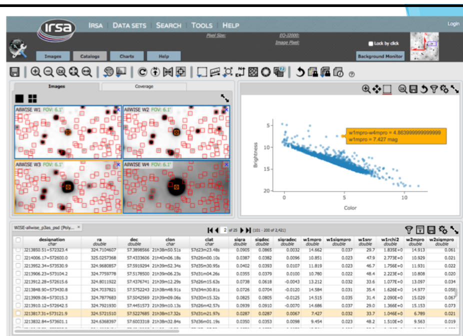
Challenges to explore

- In the 1780s, Charles Messier was a comet hunter who started a list of "not comets."
- The list is ~100 objects of all sorts (star clusters, galaxies, dead star guts, baby star nurseries).
- These objects are bright & famous, and so most telescopes (even pointed ones) have observed them.
- Investigate the Messier Catalog online; find a Messier object of each broad type (nebula, globular cluster, galaxy); see what it looks like in the infrared (or other wavelengths from other archives). You will need to request different sizes of images!
- Which types of Messier objects look different and which look the same across these wavelengths? Why?

Next steps

- Beyond qualitatively looking at images, there are a lot of quantitative measures of objects online.
- We measure brightnesses in some seriously funky units (beyond what I can cover here), but you can pull and explore those catalogs using Firefly tools. You can make plots that are interactive. (→)
- You can measure brightness of an object as a function of time, and we see lots of variations. There are a lot of light curves online too. THIS is where there is about to be an explosion of data and is where a TON of astronomers are trying to make ML work to pick out the interesting sources. (→)
- Spectroscopy is kind of like using a prism – you break up the light into a rainbow. From the bright/dark pattern in the rainbow, you can figure out (among other things) what kinds of atoms/molecules are in the object emitting the light. Lots of spectra online too. (→)

This is a bright, red baby star.



Next steps

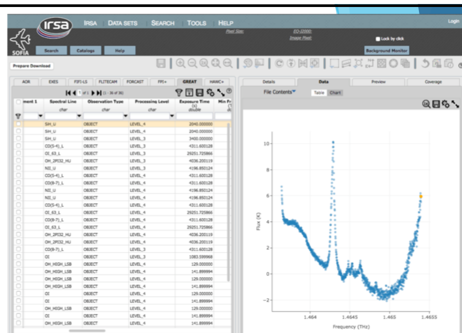
- Beyond qualitatively looking at images, there are a lot of quantitative measures of objects online.
- We measure brightnesses in some seriously funky units (beyond what I can cover here), but you can pull and explore those catalogs using Firefly tools. You can make plots that are interactive. (→)
- You can measure brightness of an object as a function of time, and we see lots of variations. There are a lot of light curves online too. THIS is where there is about to be an explosion of data and is where a TON of astronomers are trying to make ML work to pick out the interesting sources. (→)
- Spectroscopy is kind of like using a prism – you break up the light into a rainbow. From the bright/dark pattern in the rainbow, you can figure out (among other things) what kinds of atoms/molecules are in the object emitting the light. Lots of spectra online too. (→)

This is a pulsating star – it pulsates once every 0.53239 d.

Next steps

- Beyond qualitatively looking at images, there are a lot of quantitative measures of objects online.
- We measure brightnesses in some seriously funky units (beyond what I can cover here), but you can pull and explore those catalogs using Firefly tools. You can make plots that are interactive. (→)
- You can measure brightness of an object as a function of time, and we see lots of variations. There are a lot of light curves online too. THIS is where there is about to be an explosion of data and is where a TON of astronomers are trying to make ML work to pick out the interesting sources. (→)
- Spectroscopy is kind of like using a prism – you break up the light into a rainbow. From the bright/dark pattern in the rainbow, you can figure out (among other things) what kinds of atoms/molecules are in the object emitting the light. Lots of spectra online too. (→)

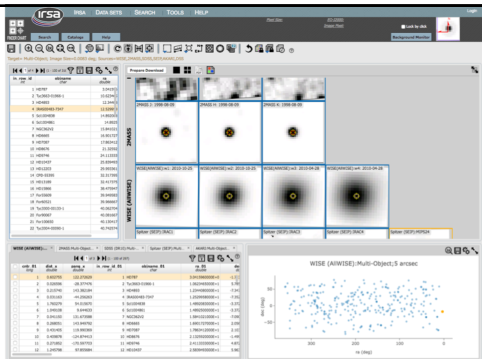
This is a spectrum (plotted against frequency rather than wavelength) of a star-forming region, taken by SOFIA, NASA's flying infrared observatory. (Yes, really!)



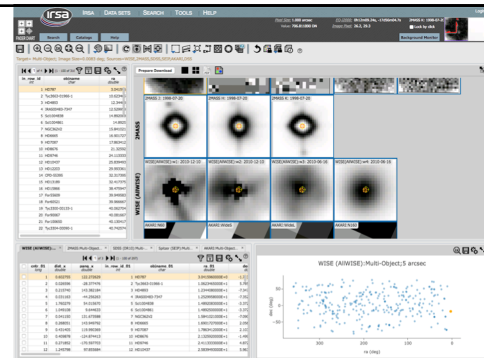
Cautions

- The catalogs that are available for large data sets are usually done in a "lights out" pipeline.
- That means that you have to RTFM, pay attention to the error flags that are provided with every measurement in every catalog.
  - Not everything there is a detection.
  - Not everything listed as a detection might even be a detection!
- Even professional astronomers get fooled, and poor work makes it into journal articles!!

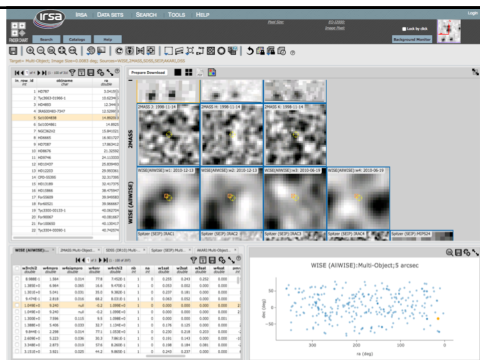
Normal point source:
Well-centered on target, bright, clearly detected with a good entry in the catalog.



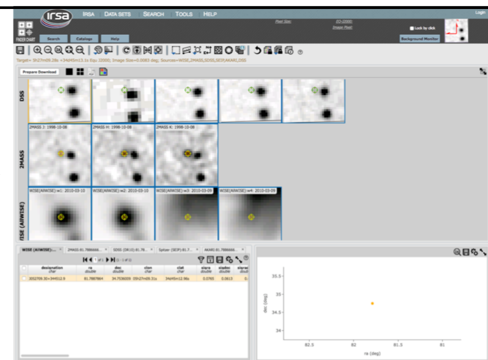
Saturated point source:
Catalog tells you this isn't good (except for W4).



Saturated point source:
Catalog tells you this isn't good and indeed our requested position is offset from the match.



Catalog tells you this is a good detection, but when you use the quantitative measures in context with other things known about this source, the W3W4 measurements cannot be correct…and indeed when you look at the images, they're not convincing.



Next steps

- If you're interested in learning more about basic astronomy, I recommend the free, online textbook at OpenStax.
- If you're interested in learning more about the strange units astronomers use to measure the brightnesses of objects, I have a collection of movies in my YouTube feed talking about filters, magnitudes, and colors.
- If you're interested in software to work with astronomical data, look for Astropy.
  - PyVO – Python code for accessing Virtual Observatory (VO). (I've linked an overview video to the links page I'll share at the end.)
  - astroML is Python module for astronomy ML/data mining.
  - (But you still need to be an educated consumer about the data!!!)

More interested in the planets?

- There is a HUGE, vibrant community of "amateurs" who play with NASA images of planets (Cassini, Juno, Mars rovers/spacecraft, …)
- Images are posted to these archives as soon as possible after downloading from the spacecraft; often you are looking at them at the same time the scientists are.
- Emily Lakdawalla (at the Planetary Society, http://www.planetary.org/) is a **huge supporter of/advocate for** this community, and has *many* tutorials on how to get access to and work with those images.

Summary

- A lot of data already out there. A lot available via web interfaces.
- You can do citizen science and really help out!
- You can poke around in the archives and get real data but you do need to know some basics:
  - FITS files not PNG, GIF, JPG.
  - You can change the scaling and color table to whatever you want to bring out whatever details you want.
  - Explore the web interface first – you can't break anything, and you can learn some interesting things!
  - Weirdness in the images/catalogs = probably artifacts, not aliens ☺
  - Use the data wisely; pay attention to, e.g., error flags.
- Other archives too, and there is a rich community of people working with images from other planets in our SS.

For more information

This takes you to a web page that has links to all the stuff I talked about (plus a copy of the talk).

http://web.ipac.caltech.edu/staff/rebull/outr/datalinks.html

Do you know a high school educator who wants to get involved in real astronomy research? Ask me about NITARP, or go to http://nitarp.ipac.caltech.edu