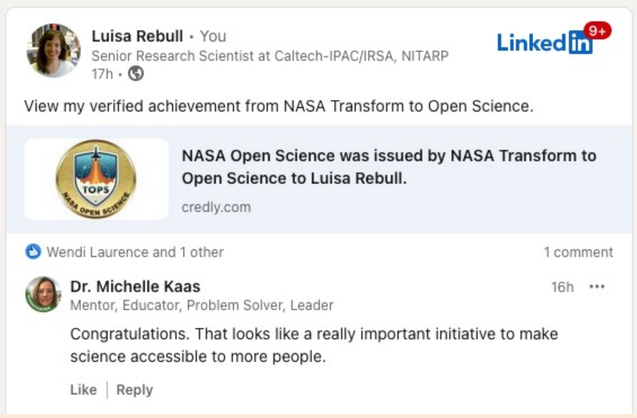
# Gaps in Understanding: Getting NASA Astrophysics Research Data Into the High School Classroom

http://nitarp.ipac.caltech.edu/

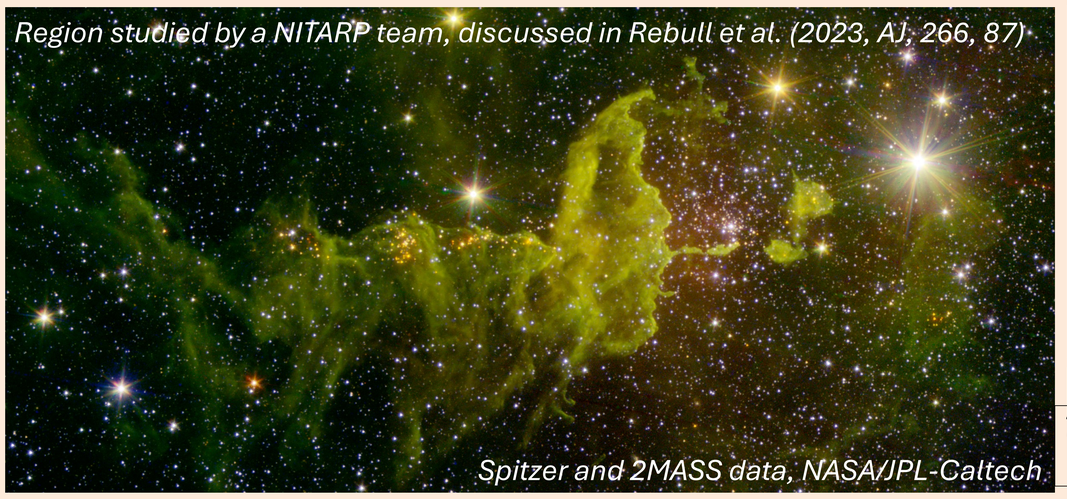
**L. M. Rebull (Caltech-IPAC/IRSA, NITARP)**

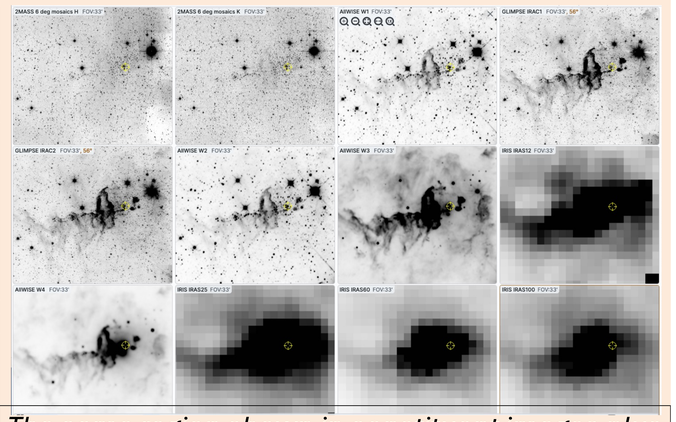



*← This is an unsolicited, unprompted comment from a childhood friend on my TOPS badge on LinkedIn. She has a PhD, but it's not in the sciences. Note "make **science accessible** to more people." Astronomers LOVE open data; it's already embedded in our culture. But to make research-grade open-access astronomy data truly **accessible to more people**, **you need scaffolding**. That's what this poster is about.*

**What is IRSA?** There is a tremendous wealth of data housed at NASA's Infrared Science Archive (IRSA). IRSA provides access to more than 700 billion astronomical measurements, including all-sky surveys in 24 bands. IRSA also provides tools that enable interactive exploration of the data. Approximately 15% of refereed astrophysics journal articles annually use data that ultimately come from IRSA.

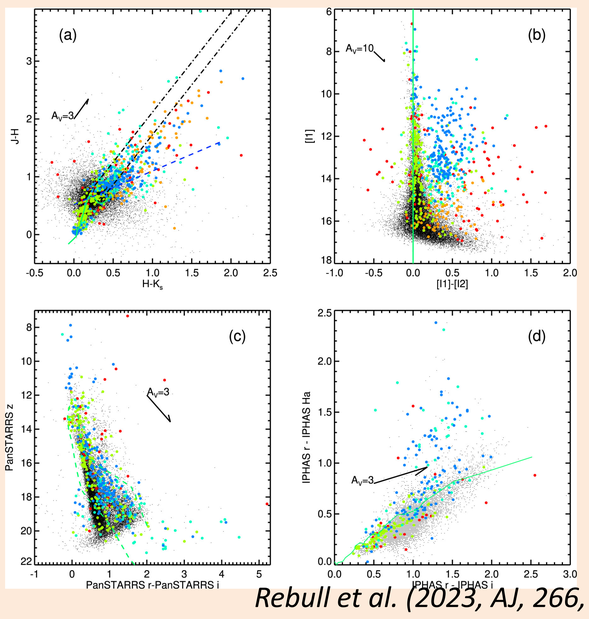
**What is NITARP?** NITARP, the NASA/IPAC Teacher Archive Research Program, has been running (in one form or another) since 2005. NITARP partners small groups of largely high school educators with a research astronomer for a year-long authentic research project. NITARP projects use data and tools from IRSA, the same data and tools used by professional astronomers. The teams experience the entire research process, from writing a proposal, to doing the research, to presenting the results at an American Astronomical Society (AAS) meeting. The program runs from January through January. Applications are available annually in May and are due in September. The educators' experiences color their teaching for years to come, influencing thousands of students per teacher.




*The same region shown in constituent images plus more of various wavelengths and resolutions.*

There is a tremendous amount of ready-to-use data in NASA Astrophysics archives, data that are ready for science right now, delivered to the archives by individual investigators and mission teams. **Using and understanding the data requires substantial baseline knowledge**, more knowledge than the average citizen scientist possesses. **The simplest data are images** (↑) but even there, users need to understand at minimum that each image is 16 or even 32 bits deep (as opposed to 8 bits deep like a jpeg), that each image recorded photons over a relatively narrow range of wavelengths (over a filter bandpass), that each constituent image can be represented as greyscale, that you can control how those recorded photons get mapped to shades of grey (or other colors, like "paint by number"), that three (or four) images can be combined into a color image, that you can combine images that are obtained from wavelengths outside the optical, and that the spatial resolution of the images varies as a function of both wavelength and telescope. If you master this content, and have software that can understand FITS files, you can probably play around with the images on your own.




*Rebull et al. (2023, AJ, 266, 87)*

**Catalogs are more complicated than images, because astronomers like strange units**. Because the ancient Greeks made a particular decision 2000 years ago (seriously), we have inherited some strange units. In order to use these catalogs, you must learn about and become fluent in this magnitude system. We make plots like these four (←), which are hard to understand if you're new at this. Here, the small black dots are background or foreground objects, and the larger colored dots are candidate young stars (from the region in the image above); blue & green are ones we are pretty sure are young stars, and red/orange are ones we are pretty sure are not young stars. Solid and dashed lines are where stars of various sorts appear, and they tell you things about these distributions; in general, the colored dots shouldn't be too similar to the black dots. Importantly, note plots like (b) and (c) – the numbers on the y-axis increase going *down* so that bright objects are at the top. Larger numbers actually mean fainter, because of our strange units. **In order to understand (or even create) these plots, you need help from an astronomer**. If you are working on your own, this is a *lot* harder than playing with images!

Astronomers deliver lots of ready-to-use, open data products to the archives (including IRSA), and that is AWESOME because it means that there is a lot of science waiting to be done right now in the archives. This is great for everyone (e.g., Scire+22 shows that delivered ready-to-use products like these increase the scientific output of an observatory by a factor of 2!). This is good particularly for novice astronomers of any sort, including high school educators/students. It is hard for a true novice working without support to just metaphorically walk in and use these data. In order for anyone to actually do science with these data, there are a lot of barriers to entry such as data formats, units, and even vocabulary, to say nothing of astrophysics. (There is also documentation specific to each data set that explains, e.g., data quality flags, but that is a learning curve that *all* users of that data must master, professionals and amateurs alike.) **In order to support novice users of astronomy data**, whether they are citizen scientists, amateur astronomers, or professional scientists from elsewhere at NASA, there needs to be some sort of scaffolding to help these novices up the learning curve. **NITARP has worked hard to provide such scaffolding for our educators and their students**. Since we work with our participants for a year, we have the luxury of extended interactions with them to cover astronomical images and color stretches/color tables; filters, magnitudes, and colors; color-color and color-magnitude diagrams; spectral energy distributions; and how to pull all of these things together to do astronomical research. This kind of basic scaffolding should be provided for users of NASA data if you want them to really understand & use the data.