

Gordon K. Squires¹ (squires@ipac.caltech.edu), Robert Hurt¹, Frank Summers², Carolyn Brinkworth¹, Michael Greene³

¹ Caltech/IPAC ²STScI ³JPL

E/PO Considerations for Data Processing

Background

Images and other multimedia resources are the primary product of astronomy in the eyes of the public. Likewise, they represent the most powerful hook for E/PO. Yet even with the latest and greatest web connectivity tools, access to such visual resources is at best haphazard.

Consider the realistic example of an educator searching for images of a specific galaxy for a brochure on astronomy. The idea would be to compare several different views of the galaxy from different observatories. How could this imagery be obtained?

One approach would be to go to a web search engine, type in something like – "M51 Galaxy" and look over the resulting pages of "hits". Even with image-specific engines the results are unsatisfactory. There will likely be dozens of copies of one or two of the most popular images that appear repeatedly on many websites. But usually it will be at low resolution, intended for on-screen viewing and useless for print or broadcast work. Moreover, finding the source site with the highest quality images and the original caption text may be difficult, as it would be buried amongst dozens of other superficially similar results.

To address the challenge described above, a new standard has been developed to associate key information within the headers of the images themselves. This is analogous to the metadata in fits headers of science images, but tailored to the needs of outreach imagery (jpgs, tiffs, etc). The Astronomy Visualization Metadata (AVM) standard includes key information such as WCS, observatory, instrument, passbands, exposure time, as well as derived descriptors such as common name, subject matter (using a standard taxonomy) and so on -- http://www.virtualastronomy.org/avm_metadata.php . A significant number of major observatories and facilities, including Hubble, Spitzer, Chandra, ESO, GALEX, WISE, LSST and NuSTAR have adopted the AVM standard and are embedding their outreach imagery accordingly.

Looking to the future of large data volumes, it is desirable to connect both the science and E/PO interface to the images through a common database. To achieve this, it will be necessary to build the data pipelines to write the science metadata in a format compatible with the international standards established for E/PO images. LSST is facing a similar challenge in this regard plans to include within the science data pipelines, database and interface the information required for E/PO purposes. A model that has this information recreated manually for a separate E/PO database or interface is not practical, given the enormous data volume produced.

We highly encourage the use of the new AVM metadata standard by future Astrophysics missions, and encourage the community to incorporate these standards as part of the science data pipeline.