CTA Data in Australia's ASVO?
(March 2018)
G.Rowell (for CTA-Australia)

---

1 Overview

The Cherenkov Telescope Array (CTA) is the next generation TeV gamma-ray observatory currently under construction. Its data products from the full (north + south) arrays will be publicly available after some proprietary time (likely 1-2 years post-observation). This therefore presents the first opportunity to integrate astrophysical information from the highest energy photons so far observed with data from radio to X-ray telescopes already integrated into the online repositories such as virtual observatories (VOs).

Here we outline the case for CTA's high level data products to be a part of Australia's All-Sky Virtual Observatory (ASVO), and discuss potential levels of CTA data that may be appropriate.

CTA data can easily be incorporated as part of 'multi-colour' studies of steady/extended sources or of transient/variable sources via light curves. The scientific benefit of such broad-band and multi-messenger studies is these days beyond doubt and the large world-wide effort to construct new wide-field instruments from radio to X-ray astronomy is a testament to this. CTA's order of magnitude better sensitivity than current TeV gamma-ray telescopes brings TeV astronomy truly into the mainstream. In Australia with its many wide-field radio telescopes (and large-scale optical surveys), the ASVO could be a prime site to link with data from ASKAP, MWA, UTMOST and eventually, the SKA.

Within CTA there are significant efforts to ensure the various data levels conform to VO and user community standards, with the aim of connecting CTA data into other established VOs (e.g. GAVO [1]). The high energy astrophysics community has considerable experience in VO and archive tools with NASAs HEASARC [2] being a driving force in this area for many years.

2 CTA Data Levels

CTA utilises the ground-based stereoscopic Cherenkov imaging technique to reconstruct the direction and energy of gamma-rays in the >30 GeV band, and up to energies beyond 100 TeV (limited by photon statistics).

Table 1. summarises the anticipated data levels (DLs) that CTA will produce.

DATA (EVT)	SHORT NAME	DESCRIPTION	ARCHIVE BRANCH
Level 0 (DL0)	RAW	Data from DAQ written to disk	RDA
Level 1 (DL1)	CALIBRATED	Physical quantities measured in the camera: photons, arrival times etc.	RDA
Level 2 (DL2)	RECONSTRUCTED	Reconstructed shower parameters (energy, direction, particle ID	RDA
Level 3 (DL3)	REDUCED	Set of selected (e.g., gamma-ray) events	SDA
Level 4 (DL4)	SCIENCE	High level data products (spectra, skymaps, lightcurves)	SDA
Level 5 (DL5)	OBSERVATORY	Final products of the Observatory such as CTA survey sky maps or source catalog.	**HDA (VO-compliant)**

Table 0: Data levels (DLs) for CTA (from CTA Data Archive description 23/03/14)

The volume of each DL is expected to be ~12 PB/yr for DL1 down to ~0.7 PB/yr for DL5. Besides the real data from CTA telescopes, the detection technique requires detailed knowledge of the atmospheric conditions during data taking. This could be on a daily basis or even down to time-scales of hours or minutes in the case of triggered observations of transients. Thus, instrument response functions (IRFs) will need to accompany DLs of level 3 or lower in the case where event reconstruction is required. The data volume of such IRFs is however much lower than their respective DL counterparts (a factor or 10 or so).

3 CTA Data Levels in ASVO

In the first instance DL4 & DL5 data will be the most desirable to incorporate into the ASVO as they will conform to standard formats such as FITs images, light curves & VOtables. Images can be in units appropriate for immediate astrophysical interpretation such as erg/cm^2/s in various energy bands.

Medium level DL3 CTA data products would also be a ideal way to enable the wider community to perform their own analysis of public CTA data tuned to their specific requirements. DL3 data comprise the events passing the standard criteria for their 'gamma-ray likeness' (and so de-selecting the cosmic-ray background events).

Using DL3 data as input, publicly available software pipelines [3] can be used to produce DL4 and DL5 level products with user-selected spatial regions and energy bands as well as up-to-date IRFs if needed. Such pipelines are based on the codes used in X-ray and GeV gamma-ray astronomy (e.g. Fermi-LAT). They are available now and are being constantly improved by the CTA Consortium using mock data generated by Monte-Carlo simulations (the so-called CTA 'data challenge'). In principle, these pipelines might be installed on ASVO (or some other cloud) so as to run them virtually, freeing up the user to concentrate only on high level decisions (e.g. energy bands, regions of interest, time windows of interest).

Data at the DL2 level (individual telescope event parameters) could also be considered for the ASVO to enable the wider public to test new methods of lower-level event reconstruction. The distribution policy for CTA DL2 data and lower is still under development but is likely to be restricted to CTA Consortium members and/or countries. However, after several years such data may be more widely distributed. For example in the next few months HESS is about to release parts of its lower level data DL2 products (from 2004 observations) along with accompanying analysis pipelines as a test case for this scenario. Additionally, the FACT telescope [4] (small 4m TeV gamma-ray telescope a La Palma) monitors many variable and flaring active galactic nuclei (AGN) and from late 2017 it provided its low level data and analysis tools online.

4 TeV Gamma-Ray Data in the ASVO – 2018/2019

The forthcoming release of the HESS Galactic Plane Survey (HGPS) [5] will be the first example of such high level (DL4/DL5) products surveying the TeV gamma-ray Milky Way. This information will be very useful to the radio community as ASKAP and MWA survey data become widely available.

Over the coming year, we suggest to consider incorporating the HESS HGPS products into the ASVO as a trial run. This will be an ideal preparation for CTA's real data products which should start to flow from ~2021.

For CTA we suggest investigation of the requirements needed to host its DL4 and DL5 data (~PB/yr), with possible extensions to DL2 and DL3 data in the longer term. AAL's investments in this area might in be used to assist with data management, their merging with data products from other telescopes in effective ways, and persons experienced with the analysis tools required to convert DL2/DL3 to DL4/DL5 products. These latter aspects are important to ensure the ASVO offers a unique service that aligns with Australia's astronomical interests and strengths.

5 References

[1] GAVO German Astrophysical Virtual Observatory - http://www.g-vo.org/pmwiki/Main/HomePage

[2] NASA HEASARC - https://heasarc.gsfc.nasa.gov/

[3] ctools - http://cta.irap.omp.eu/ctools/

[4] FACT - https://www.fact-project.org/

[5] Abdalla etal (HESS Collab.) 2018 “The H.E.S.S. Galactic plane survey” A&A in press