Open high-level data formats and software for gamma-ray astronomy

Christoph Deil, Jeremy Perkins, Catherine Boisson, Johannes King, Peter Eger, Michael Mayer, Matthew Wood, Victor Zabalza, Jürgen Knödlseder, Tarek Hassan, Lars Mohrmann, Alexander Ziegler, Bruno Khelifi, Daniela Dorner, Gernot Maier, Giovanna Pedaletti, Jaime Rosado, José Luis Contreras, Julien Lefaucheur, Kai Brügge, Mathieu Servillat, Régis Terrier, Roland Walter, Saverio Lombardi

Abstract

In gamma-ray astronomy, a variety of data formats and proprietary software exist, often developed for one specific mission or experiment. Especially for ground-based imaging atmospheric Cherenkov telescopes (IACTs), data and software have been so far mostly private to the collaborations operating the telescopes.

However, there is a general movement in science towards open data and software and the next big IACT array, the Cherenkov Telescope Array (CTA), will be operated as an open observatory.

We have created a Github organisation where we are developing high-level data format specifications. A public mailing list was set up and a first face-to-face meeting on the IACT high-level data model and formats took place in April 2016.

The hope is that this open multi-mission effort will help to accelerate the development of open data formats and open-source software for gamma-ray astronomy, leading to synergies in the development of analysis codes and eventually better scientific results (reproducible, multi-mission).

Resources

- Mailing list for announcements and high-level discussions (75 members, including people from all major gamma-ray collaborations): lists.nasa.gov/mailman/listinfo/open-gamma-ray-astro
- Github issues and pull requests are used for detailed discussions: github.com/open-gamma-ray-astro
- Data format specifications in HTML and PDF format, including example files: gamma-astro-data-formats.readthedocs.io/

Notes

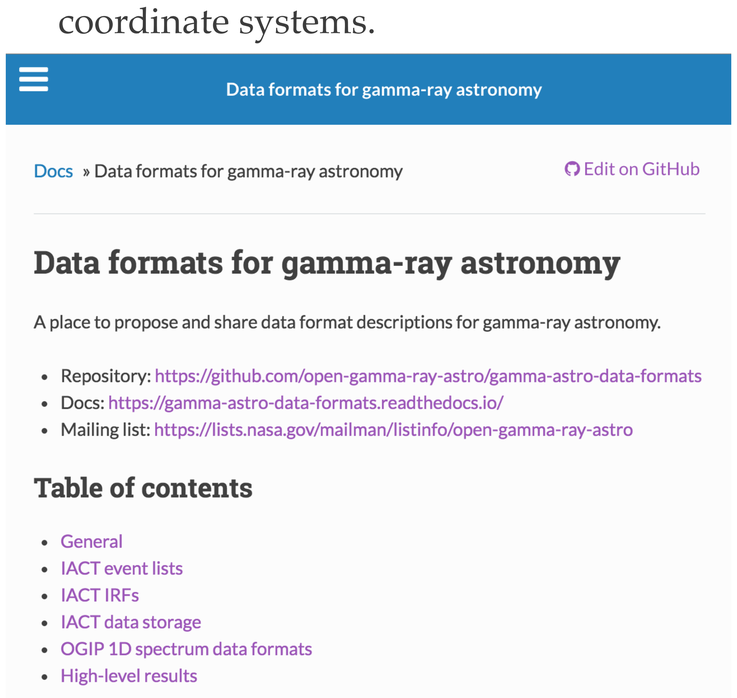
- The specification of a given format defines the names and semantics of data and header fields (mostly FITS is used).
- The spec should be easy to understand and useful. We include example files and some explanations, in addition to the detailed spec for a given format.
- The scope is high-level data, starting with event lists and instrument response functions (IRFs), what is called “data level 3” (DL3) in CTA.
- First stable release (archived on Zenodo, with a DOI) coming soon.

Contribute

- Use the existing formats.
  Give feedback!
  Propose additions and changes.
- Join the mailing list.
  Send an email with an idea or proposal.
- Create a Github account.
  File an issue with a correction or make a pull request proposing an addition.
- No formal approval process in place yet.
  This is a very recent effort.

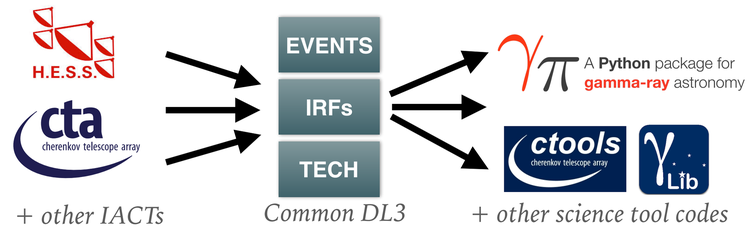
Content

- So far, the focus has been mainly on the definition of the IACT DL3 formats (see info in right column).
- Some work on higher-level formats for spectra and lightcurves, as well as HEALPIX images and cubes ongoing.
- A “general” section defining common things like details about time scales or coordinate systems.

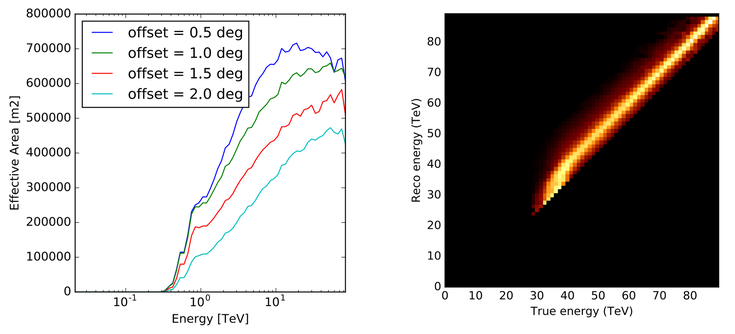


Example: IACT DL3

- Imaging atmospheric Cherenkov telescopes (IACTs) data level 3 (DL3): what is released to end users for analysis with science tool: EVENT, IRF, TECH
- First effort to develop an open common data model and format for IACT DL3.
- Prototyping by existing IACTs (H.E.S.S., MAGIC, VERITAS) and science tools (Gammapy, Gammalib/ctools) ongoing.
- Many points under discussion, e.g.
  — Observation modes, time intervals
  — EVENT and IRF linking
  — Pointing and livetime info
  — IRF axes and IRF validity ranges
  — Field of view coordinates



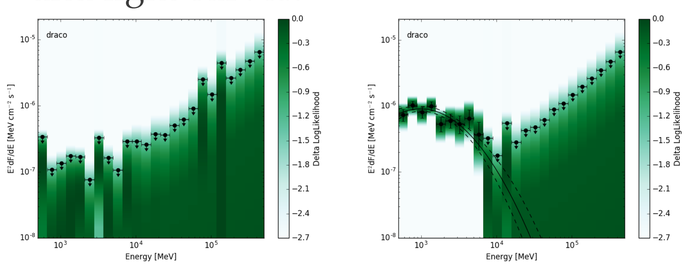
+ other IACTs *Common DL3* *+ other science tool codes*



Examples for IACT instrument response functions (IRFs): effective area (left) and energy dispersion (right)

Example: Likelihood SED

- A proposed format to store spectral analysis results for exchange and publication.
- Developed first in Fermipy and used for Fermi-LAT, now being adopted by other instruments and codes.
- “Flux points” and upper limits, but also a format for full likelihood profiles.
- Discussion on other high-level formats is ongoing, e.g. for global spectral models and light curves.



Examples for SEDs and likelihood profiles: a weak source (left) and a strong source (right)

Poster presented at Gamma 2016, Heidelberg in July 2016