OPEN GAMMA-RAY ASTRO GAMMA ASTRO DATA FORMATS

Christoph Deil, MPIK Heidelberg
IACT DL3 meeting, Meudon, April 6, 2016

INTRODUCTION

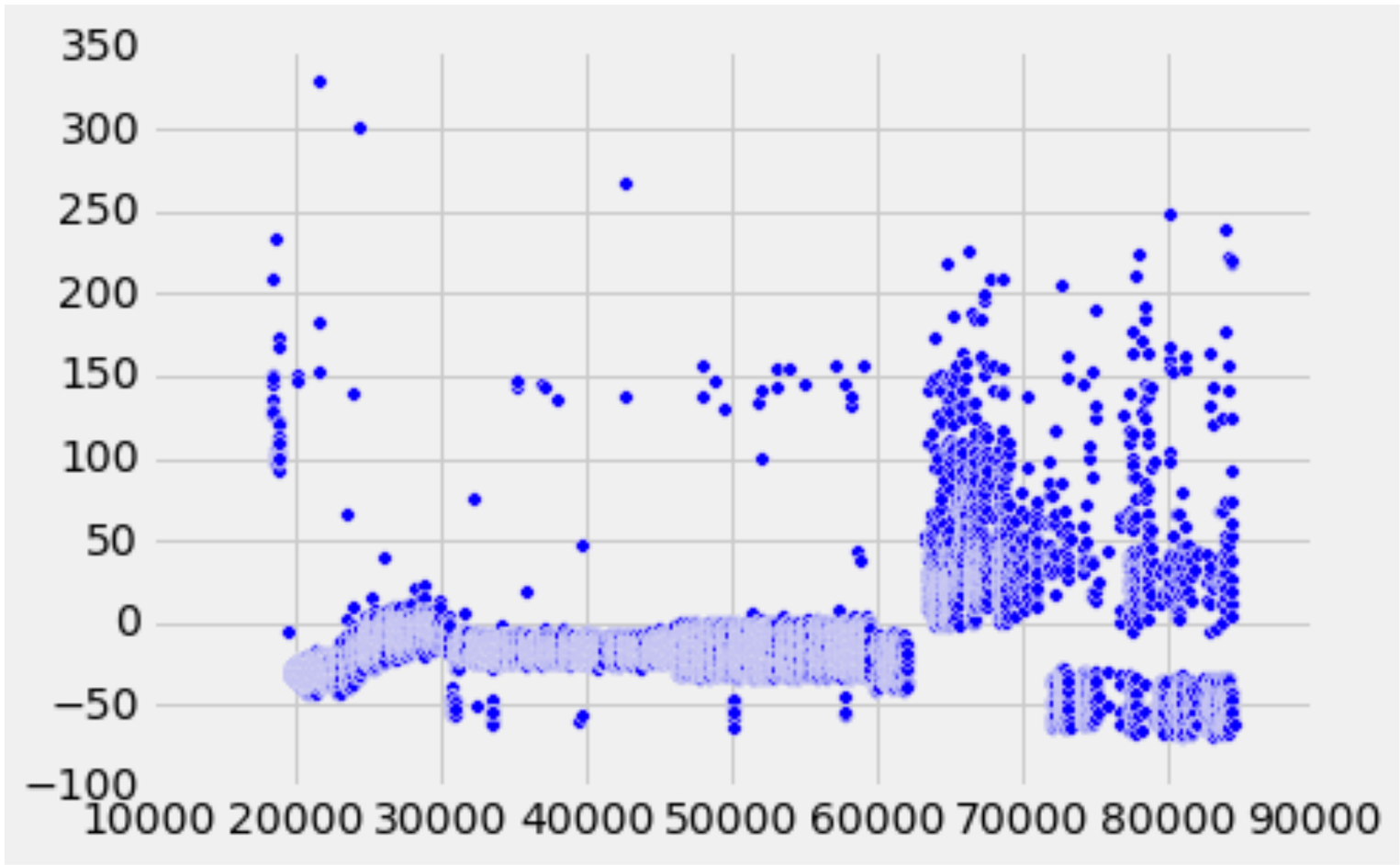
WHY?

- Why did we start an **open gamma-ray astro** effort and start writing a **gamma-astro-data-formats** public spec?
- Motivation 1: **multiple DL3 producers** in H.E.S.S.
  - For the past 5 years we (in H.E.S.S.) have been hoping for a DL3 data model and format specification from CTA.
  - This didn't happen.
    Or did it? I'm hoping to find out today!
  - In H.E.S.S. we have three chains that export DL3 to FITS. The communication effort to make the data consistent became too high, some discussions were going in circles because we didn't have a spec to refer to.

NEED FOR A SPEC EXAMPLE 1 — RUN START TIME

- What exactly is “Run start time”?
  You would think it’s easy / obvious, but there are subtle differences how it can be defined and computed, as well as data quality issues.

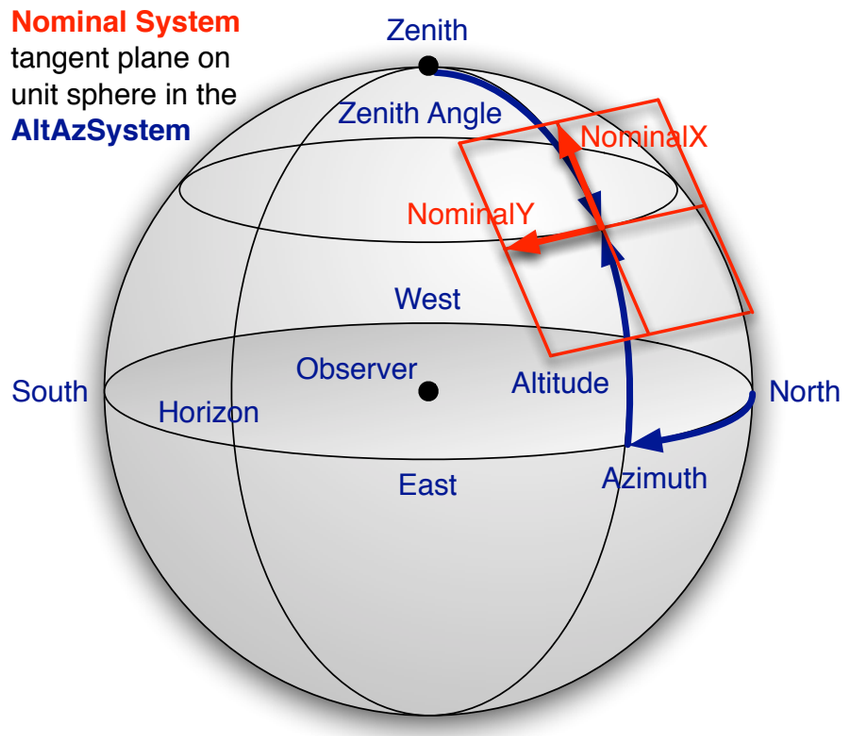
Run start time difference between two HESS chains (usually ~10 sec off!???)

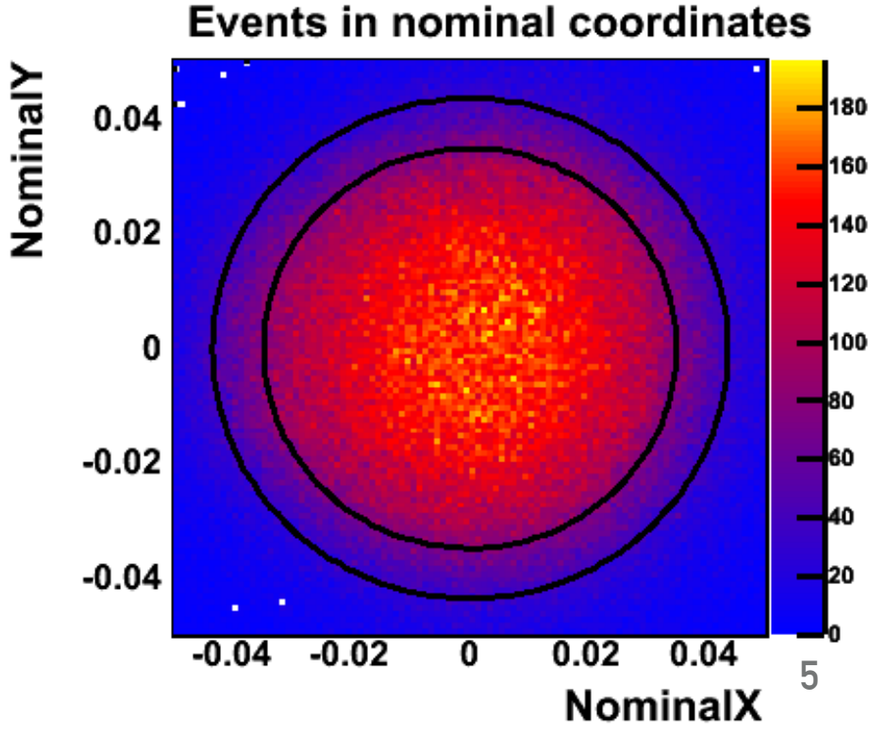


Run number = Observation ID

NEED FOR A SPEC EXAMPLE 2 — FIELD OF VIEW COORDINATES

- Event lists have “field of view coordinates” DETX and DETY.
- These are used for background modeling at the moment.
- No exact definition written down
  >> H.E.S.S. exporters and open-source science tools inconsistent for years
  - FOV frame aligned with ALTAZ or RADEC frame?
  - Direction of X and Y?
  - TAN projection applied?
  - With or without refraction?





WHY?

- Why did we start an **open gamma-ray astro** effort and start writing a **gamma-astro-data-formats** public spec?
- Motivation 2: science tool interoperability and multi-mission analysis
  - At the "Python in gamma-ray astronomy" workshop, November 16 — 20, 2015, MPIK Heidelberg http://gammapy.github.io/PyGamma15/ we talked about writing down specs for DL4 and DL5 for IACTs as well as for Fermi-LAT to improve science tool interoperability and make multi-mission analysis easier.
- E.g. how to store fitted model results, flux points, light curves, observation specification lists, … in FITS, JSON, YAML or XML
- This hasn't happened yet, people are prototyping serialisation formats, but no-one is working on a spec yet.

EXAMPLES FOR DL4 FORMATS I'D LIKE TO SEE IN THE SPEC

- Victor Zabalza is using a table with certain keys for flux points for his SED fitting package, see the Naima data format spec page.
  - If we write down simple short specs defining keys and their semantic, we improve tool interoperability.
    (For some things, VO standards exist, so there we just have to implement and promote them in our codes.)
- Toby Burnett has developed the pointlike data format (sparse HEALPIX) for efficient binned Fermi-LAT data analysis many years ago, that we now want to try out for IACT analysis.
  - Such things are also in scope for the **gamma-astro-data-formats** spec, it's not just about IACT DL3.

ORGANISATION, HOW TO CONTRIBUTE

WHAT IS THE "OPEN GAMMA-RAY ASTRO" EFFORT?

- Just started in the past 6 months.
- Jeremy Perkins (Fermi-LAT, NASA) set up a mailing list:
  https://lists.nasa.gov/mailman/listinfo/open-gamma-ray-astro
  (just a handful of posts so far)
- I set up a Github organisation:
  https://github.com/open-gamma-ray-astro
- At the moment there are two repositories:
  - https://github.com/open-gamma-ray-astro/gamma-astro-data-formats
    "Data formats for gamma-ray astronomy" — open spec attempt
    Latest version: http://gamma-astro-data-formats.readthedocs.org/
  - https://github.com/open-gamma-ray-astro/2016-04_IACT_DL3_Meeting
    This meeting — attempt to bring people from current IACTs and CTA together to work on IACT DL3 and get feedback / buy-in / contributions for the open spec.

HOW TO CONTRIBUTE

- You can contribute by
  - File an issue on Github requesting some correction or addition (or removal) from the spec.
  - Contribute edits via a pull request!
  - Sending feedback to the mailing list (good for important, high-level things)
- Who has the commit bit?
  - At the moment I'm giving it to anyone that contributes.
  - I've never had a bad experience with this collaboration model. The issue is rather get people to contribute to specs at all.
- Stable versions (e.g. on Zenodo) and if / how this could lead to content for official CTA specs to be discussed at this workshop for the first time.

SPEC WRITING

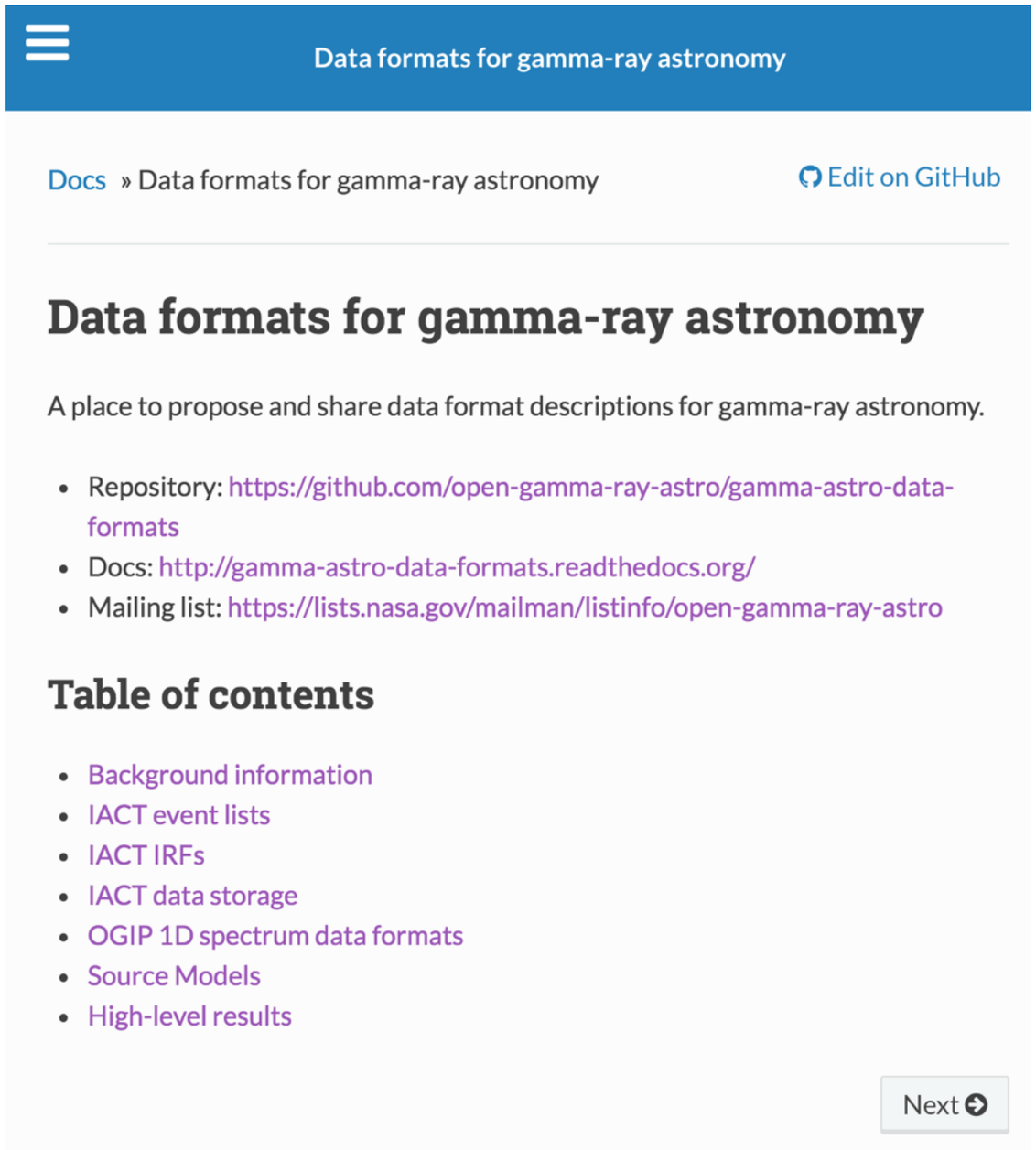
DEMO HOW TO CONTRIBUTE

- We want **YOU** to take the step and contribute to the gamma-astro-data-formats spec now, at this meeting and also after.
- So let's do a live demo how to contribute:
  - File an issue and discuss something
  - Edit files via the web interface (no git skills needed)
  - Edit files locally (git skills needed)
- If you have any questions about git or Github or RST and Sphinx, talk to me this week or email me anytime!

CAVEATS ABOUT GITHUB

- Github issues and pull requests are great for small and medium-sized changes (e.g. add and review one page in the spec) that take a day or week.
- It does **not** work well for very complex or controversial discussions.
- E.g. the discussion on IRF association to EVENTS via OBS_ID or GTI got completely stuck (see https://github.com/open-gamma-ray-astro/gamma-astro-data-formats/issues/20).
- For such things f2f meetings are best … that's why we're here!

IACT DL3 SPEC



LOOK AT CURRENT SPEC

- Instead of copy & pasting the content of the spec on slides here, let's look at the current spec directly: http://gamma-astro-data-formats.readthedocs.org/
  - Background info
  - IACT event lists
  - IACT IRFs
  - IACT data storage

NEXT STEPS FOR THE SPEC

- There seems to be agreement that IRF to EVENT association should be done via **GTIs** (not OBS_ID as we do now). This is a fundamental change and we need to define how exactly it should work (e.g. introduce GTI_ID?)
- Add **EVENT_TYPE** to represent event classes (use Fermi-LAT format?) and work out how IRF association works if there are multiple event types for a given observation / GTI.
- Precise definitions for **FOV coordinates** and time.
- Many possible improvements and extensions, small and large. Let's organise sessions for tomorrow to work on those later today.

SUMMARY

- We have recently started on “open-gamma-ray-astro” effort and a “gamma-astro-data-formats” document.
- So far describes what we do for IACT DL3 in H.E.S.S. and Gammalib / Gammapy.
- We think this could evolve into a community-developed IACT DL3 spec that works for all current IACTs and CTA.
- This will only work if there’s adoption and contributions from people outside H.E.S.S. and if this is a joint, not parallel effort with CTA.
- We would like to complete and improve the current IACT DL3 spec here, at this workshop, i.e. discuss and then edit the documents together this week.
- Thoughts?