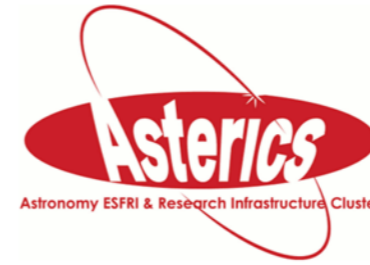# OPEN DATA AND TOOLS FOR GAMMA-RAY ASTRONOMY

Christoph Deil, MPIK Heidelberg
June 16, 2016

ASTERICS European Data Provider Forum and Training Event 2016, Heidelberg

This workshop will gather researchers and technicians on the topic of on-line publishing of astronomical data and services. It will offer an opportunity to identify common challenges and problems, to exchange solutions, and to share perspectives.
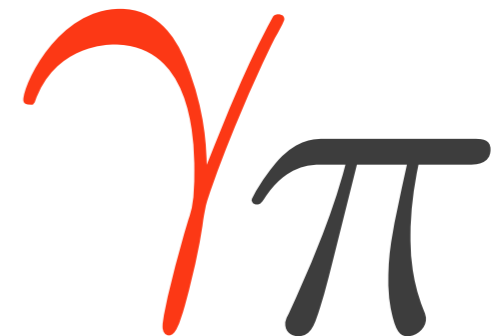
You

Thank you for this event! I'm learning a lot and hope to get some feedback!

Will try to give you an overview of some activities and challenges in gamma-ray astronomy.

Me

H.E.S.S.

cta
cherenkov telescope array

$\gamma\pi$

GAMMA-RAY ASTRONOMY

Space and ground telescopes

➤ Mathieu Servillat covered many things yesterday.

➤ I'll try to focus on some different aspects.

Fermi-LAT

H.E.S.S.

CTA

"

What is the status of data and tools in gamma-ray astronomy?

SPACE–BASED GAMMA–ASTRO

➤ Current mission: Fermi-LAT

➤ Data

   ➤ High-level fully open.
      NASA, HEASARC

   ➤ Event lists, IRF, spacecraft.
      FITS, mission-specific

➤ Software

   ➤ Science tools freely available

   ➤ Mostly mission specific (not
      much multi-mission re-use)

   ➤ Not open development.
      Very few scientists read,
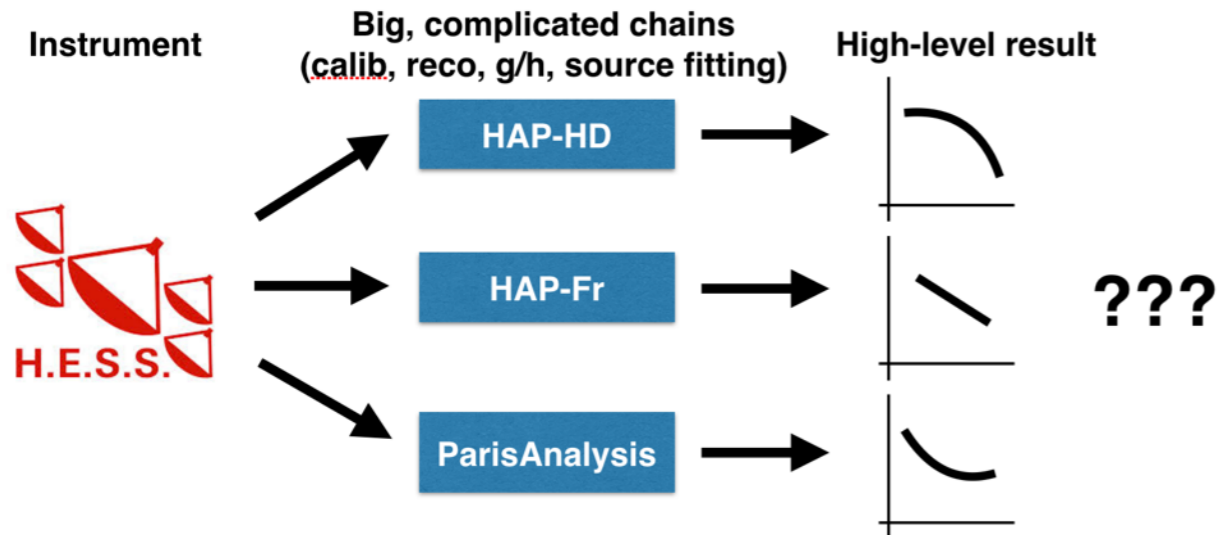   understand or extend the code.
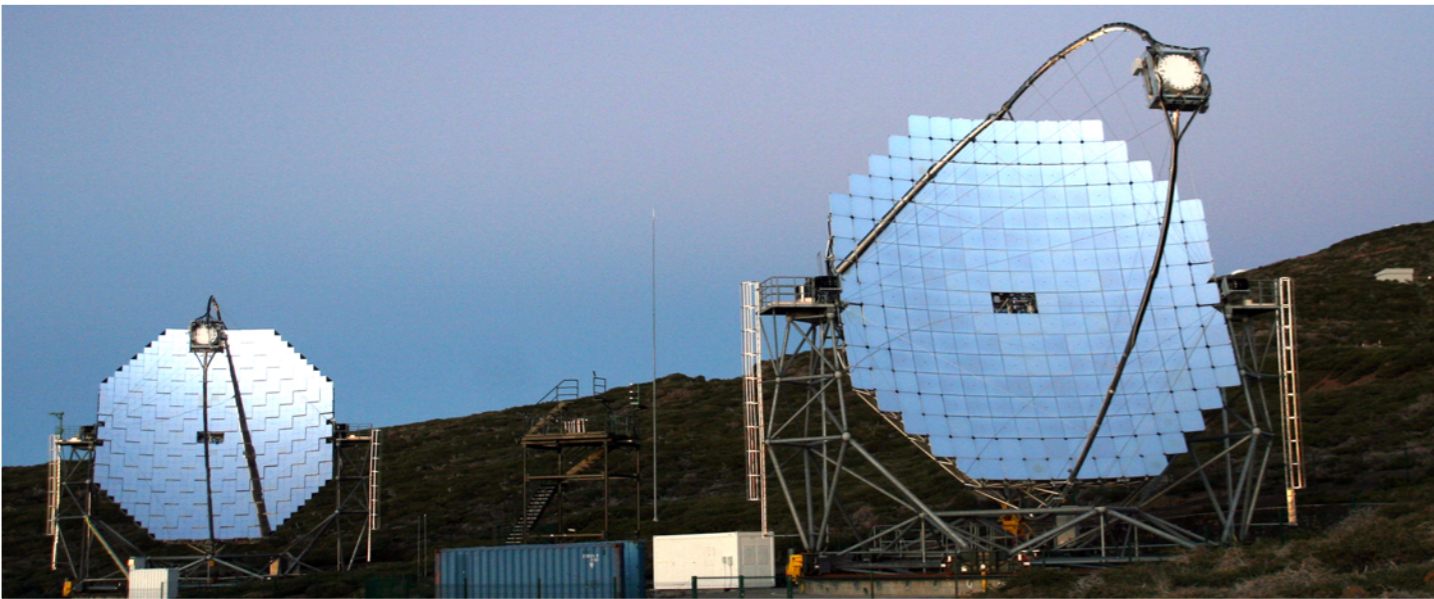
GROUND-BASED GAMMA-ASTRO

➤ A few "second generation" imaging atmospheric Cherenkov telescopes (IACTs) in operation for ~10 years.

➤ Built and operated by collaborations of ~100 people.

➤ So far in ground-based gamma-ray astronomy:

    ➤ All work (instrument, data, software, analysis, papers) done in the collaborations.

    ➤ Data is **not** open.

    ➤ Tools are **not** open.

Instrument → Big, complicated chains (calib, reco, g/h, source fitting) → High-level result

HAP-HD, HAP-Fr, ParisAnalysis → ???

GROUND-BASED GAMMA-ASTRO

➤ Current IACTs (like H.E.S.S.) use large proprietary C++/ROOT software chains.

➤ Data at all levels (even images and spectra at high level) consists of serialised C++ ROOT objects -> can only read it with software that wrote it.

➤ No well-defined data levels and models.

➤ No data format for interchange between IACTs (or even chains within H.E.S.S.)
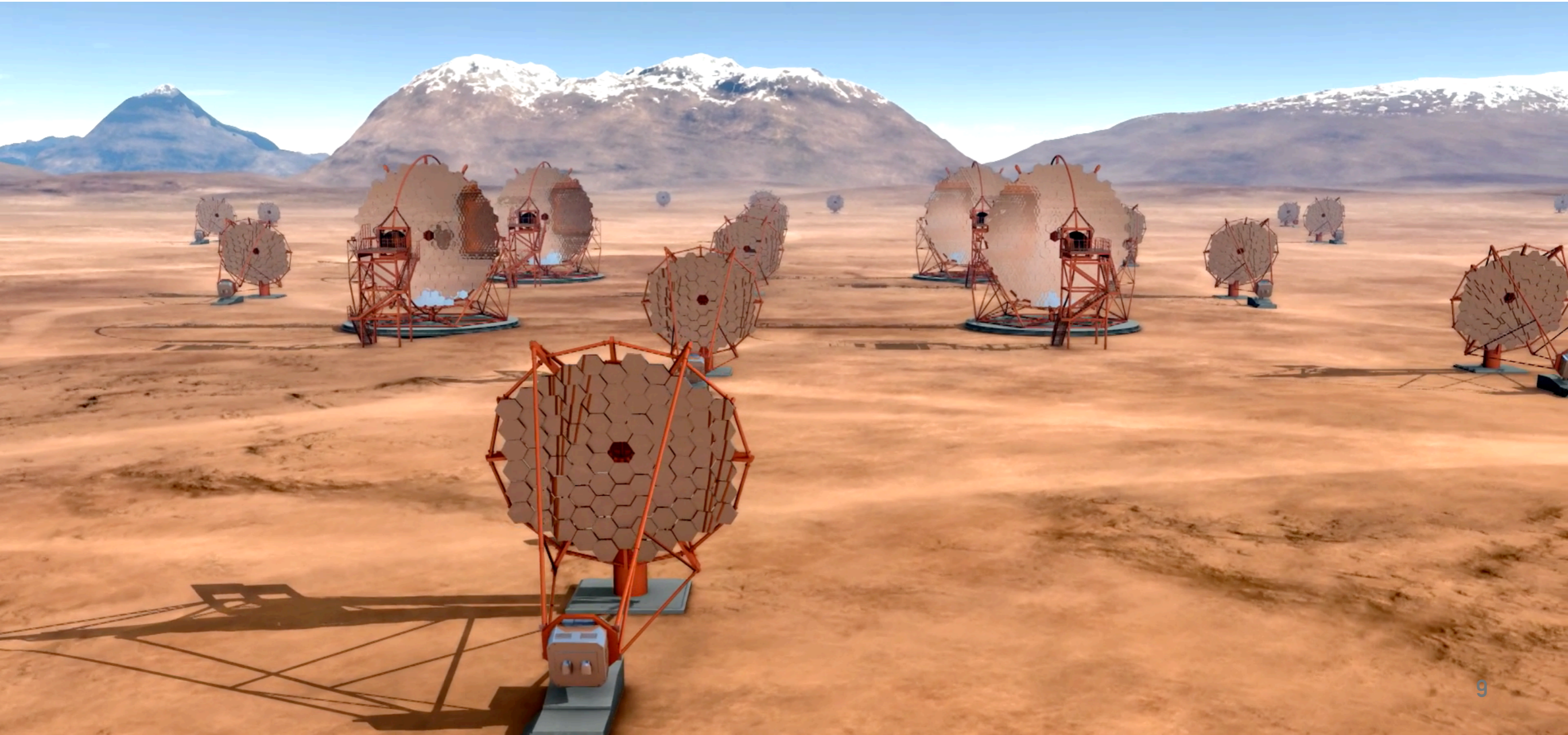
"

Towards open data and tools for
gamma-ray astronomy …

CHERENKOV TELESCOPE ARRAY (CTA)

An observatory.
Need open high-level data model and formats.
*Need open-source high-level tools.*
*Data will start flowing very soon — ASAP!*

CHALLENGES

➤ Organisation and sociological

  ➤ How to get people from existing IACTs and CTA to efficiently work together on open data and software and accelerate contributions and adoption?

➤ Technical

  ➤ What is a good IACT high-level data model?

  ➤ Which existing standards and recommendations (FITS, OGIP, VO) should we adopt?

  ➤ Where to invent new ones?
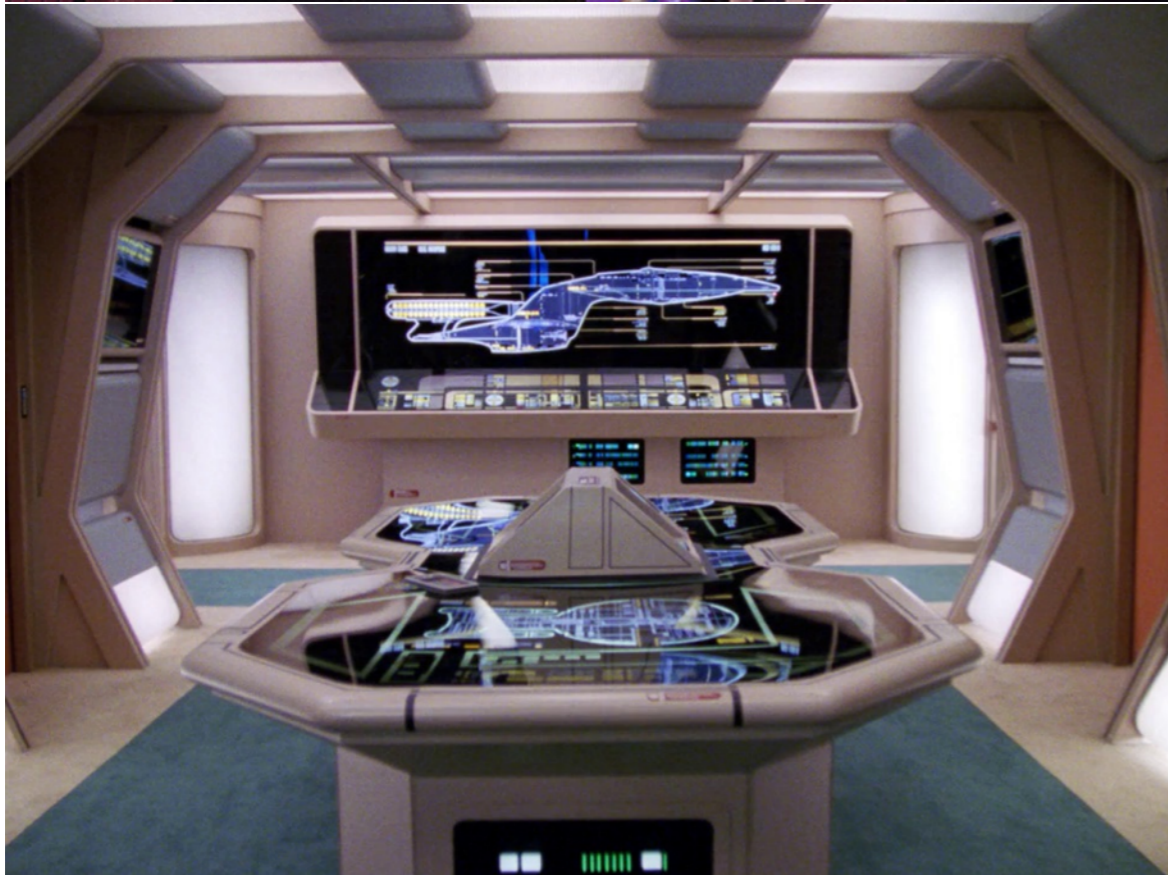
THE KEY — A COMMON IACT DATA LEVEL 3 MODEL AND FORMAT

A clear interface between low-level and high-level analysis

IACTs

**Common data level 3 (DL3)
data model and FITS format**

**Open source
science tools**
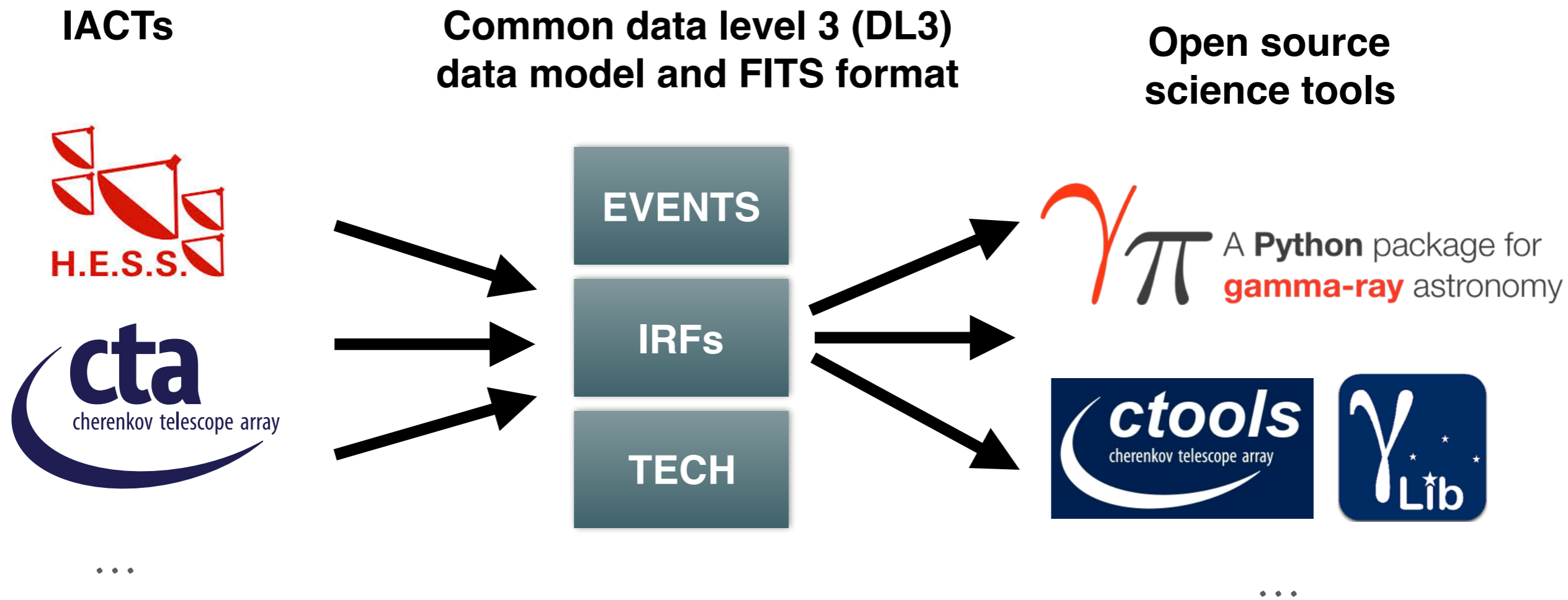
*Existing IACTs export their data and instrument response functions (IRFs) to FITS
(mostly H.E.S.S. so far, but there is interest from other IACTs)*
CTA and other new IACTs start producing DL3 in the right format from the start.

SPEC WRITING

Data formats for gamma-ray astronomy

Docs » Data formats for gamma-ray astronomy   ⚙ Edit on GitHub

Data formats for gamma-ray astronomy

A place to propose and share data format descriptions for gamma-ray astronomy.

- Repository: https://github.com/open-gamma-ray-astro/gamma-astro-data-formats
- Docs: https://gamma-astro-data-formats.readthedocs.io/
- Mailing list: https://lists.nasa.gov/mailman/listinfo/open-gamma-ray-astro

Table of contents

- General
- IACT event lists
- IACT IRFs
- IACT data storage
- OGIP 1D spectrum data formats
- High-level results

Next ➡

IACT DL3 OPEN SPEC

➤ Started in late 2015.

➤ First f2f meeting in Meudon in April 2016 (17 participants, all major IACTs present).

➤ Detailed work: on Github Important things: monthly telcons

➤ Has some buy-in from some the main IACT DL3 producers and consumers.

➤ Plan:

   ➤ a first stable version on Zenodo

   ➤ more f2f meetings

   ➤ a board and formal process

IACT DL3 DATA MODEL AND FORMATS

➤ At the moment it's all FITS and we're simply agreeing on the format and semantics of header keywords and data content

➤ Example for data model and format issues we're discussing:

  ➤ Abstractions and organisation:
  "What is an observation?" (support slew observations?)
  "Where to store livetime information?" (EVENTS, GTI, TECH)

  ➤ Data links:
  "How to associate events and instrument response?"
  (FITS grouping convention? Header keys? Index files?)

  ➤ Flexibility vs simplicity:
  "Fix physical units or allow both MeV and TeV?"
  "Fix time scale to TT or allow others from the FITS standard?"

H.E.S.S. PUBLIC TEST DATA RELEASE 1

....

➤ Idea

  ➤ A test data release to help the open spec and tools dev

  ➤ Not a science data release.

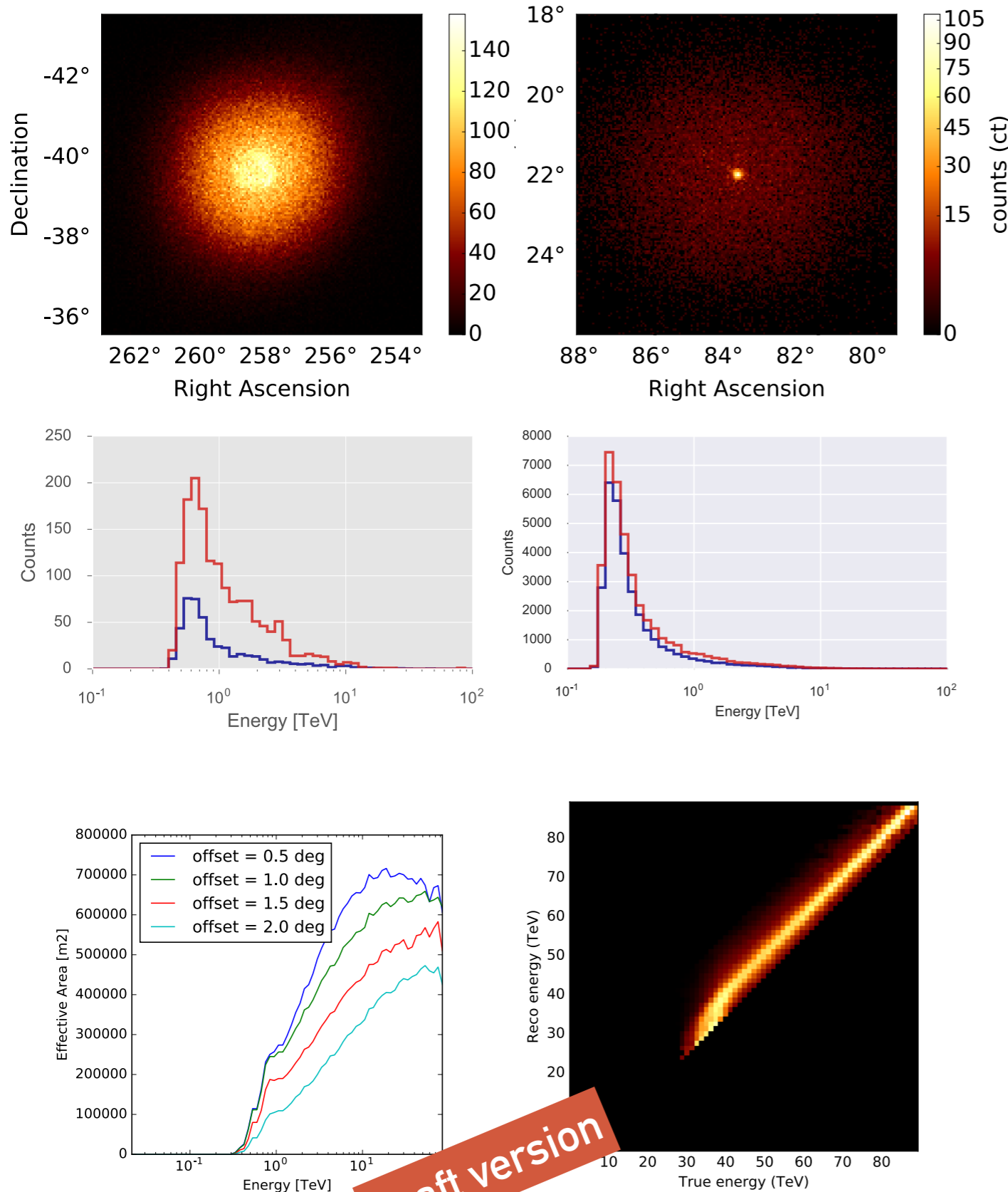➤ Very small sub-set of HESS 1 data:

  ➤ two point sources

  ➤ two extended sources

  ➤ a variable source

➤ Status: in preparation. Plan to release this summer.

➤ First time real ground-based gamma-ray DL3 data (events and IRFs) are openly available.

H.E.S.S. PUBLIC TEST DATA RELEASE 1 — TERMS OF USE

```
H.E.S.S. DL3 public test data release 1 (HESS DL3 DR1)
------------------------------------------------------

H.E.S.S. collaboration, 2016

The data and documentation is publicly released by the H.E.S.S. collaboration
as a contribution to the ongoing efforts to define a common open format for
data level 3 of imaging atmospheric Cherenkov telescopes (IACTs) and IACT
open-source science tool development, to enlarge the community involved in
IACT data analysis.

No scientific publications may be derived from the data. Using the data
for new claims about the astrophysical sources is not permitted.

When using this data, please include the following attribution:

    This work made use of data from the H.E.S.S. DL3 public test
    data release 1 (HESS DL3 DR1, H.E.S.S. collaboration, 2016).

Alternatively, use the following shorter version, e.g. for presentations:

    HESS DL3 DR1, H.E.S.S. collaboration

These terms of use must be included in all copies in full or part of the data.

For information on context, aims, use and contacts, as well as a description
of the dataset, see the hess_dl3_dr1.pdf document.
```

Draft version

H.E.S.S. PUBLIC TEST DATA RELEASE 1 — RELEASE PLAN

➤ Tarball and PDF on Zenodo.
(PDF additionally on ArXiV)

➤ Zenodo provides us for free with:

  ➤ DOI

  ➤ Archiving

➤ Reference open spec v1.0

➤ Add "terms of use" saying that
use for scientific publications is
not allowed.

DL3 public test data release 1

Draft version

OPEN-SOURCE TOOLS

➤ Two major science tool packages started with focus on IACTs.

➤ Gammapy

➤ Python package

➤ Build on Numpy, Scipy, Astropy

➤ Gammalib / ctools

➤ C++, Python wrapper, FTOOLs

➤ No dependencies except CFITSIO

➤ Both:

➤ Open source, open development

➤ Use and contribute to DL3 spec

SUMMARY

➤ Ground-based gamma-ray astronomy has a history of working in collaborations.

➤ In the past decade, the TeV source count has gone from ~10 to ~100 … change from "experiments" to "telescopes".

➤ CTA, the next-generation IACT, will be an observatory with open data and open-source tools.

➤ This presentation covered:
   - open IACT DL3 spec effort
   - planned HESS test data release
   - open-source science tools

MAKE GIFS AT GIFSOUP.COM

THOUGHTS?

➤ Most (all?) people in this room have thought about and been involved in similar efforts for a long time.

➤ I'm very interested to get some feedback now or later on what we're doing well and badly!

  ➤ How to make collaborative open spec and software development work well and sustainable for the coming decade?

  ➤ Comments on IACT DL3 data model and format?
  (see http://gamma-astro-data-formats.readthedocs.io/)

  ➤ Comments on how we plan to do the HESS public test data release? (Zenodo, terms of use, documentation)

  ➤ Which specific other standards or tools could be of interest for us?