IVOA Interopability Workshop
2021-05-26

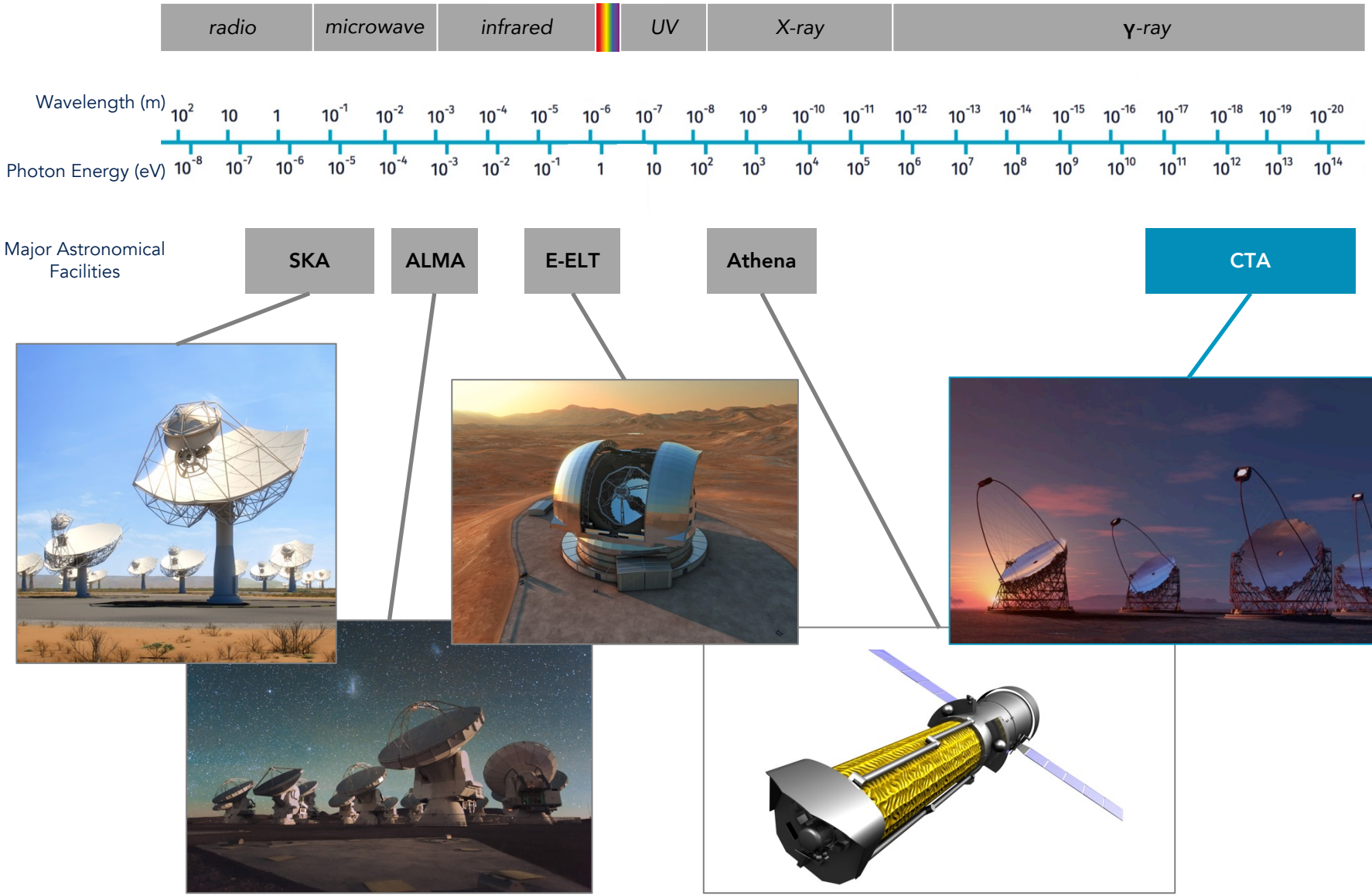
Matthias Fuessling,
CTA Observatory

OUTLINE

- CTA
- Data Models
- MM / MWL Context

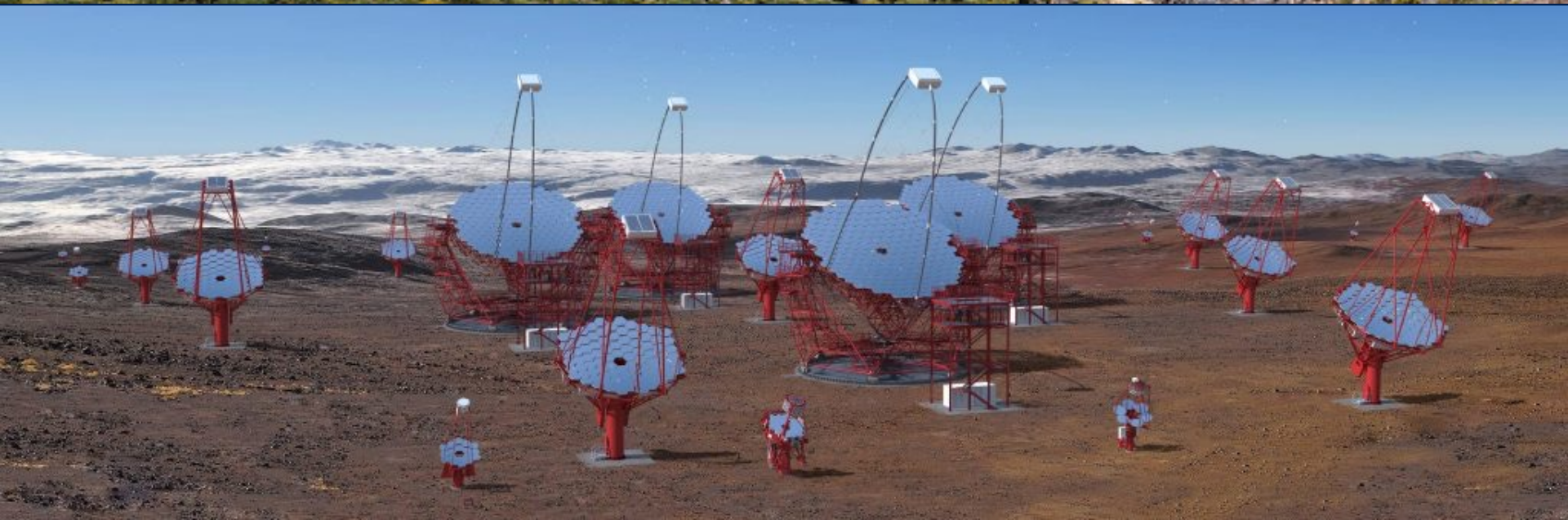
OUTLINE

- CTA
- Data Models
- MM / MWL Context

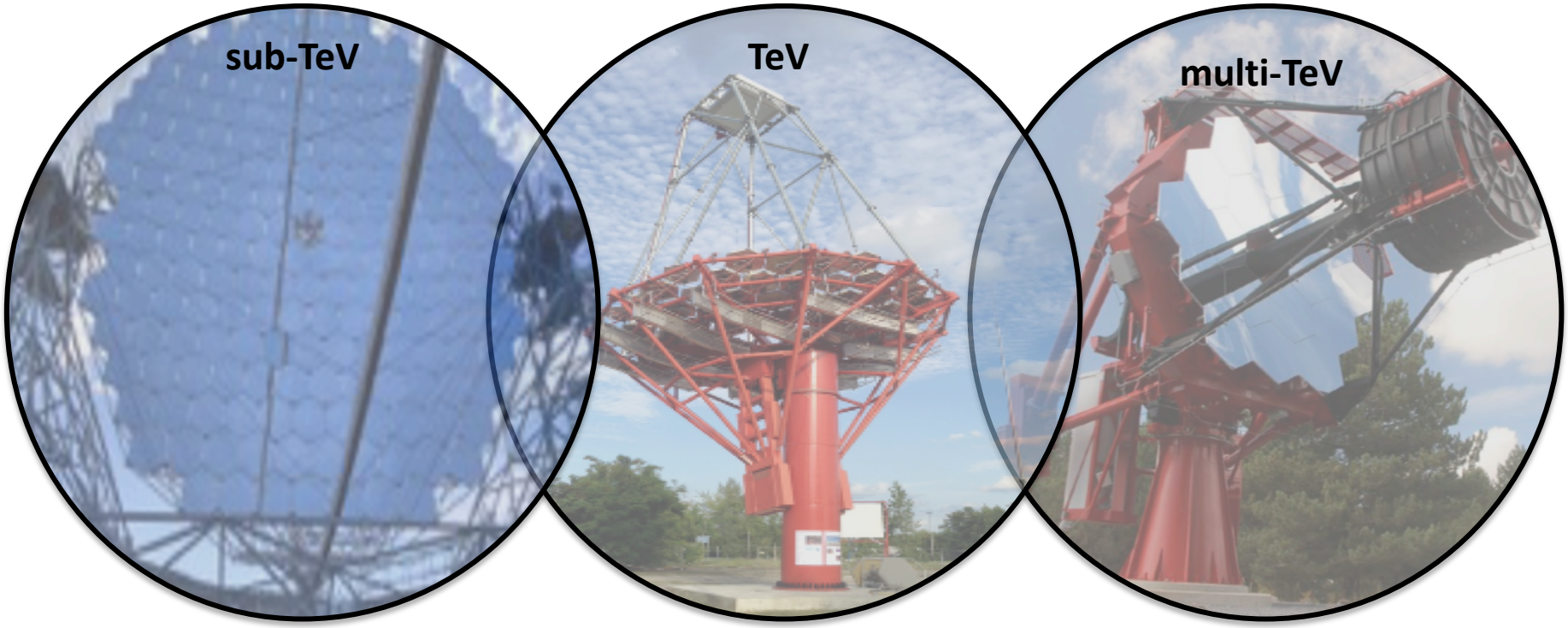

radio
microwave
infrared
UV
X-ray
γ-ray
Wavelength (m)
Photon Energy (eV)
Major Astronomical Facilities
SKA
ALMA
E-ELT
Athena
CTA

CTA: THE CHERENKOV TELESCOPE ARRAY

Next generation ground-based gamma-ray observatory



SCIENCE CASES AND DESIGN



- Deepest sensitivity for short timescale phenomena

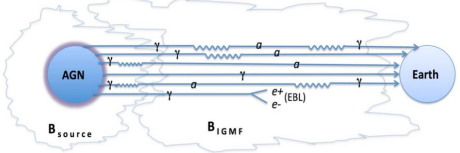
Time domain unexplored
→ cosmological sources

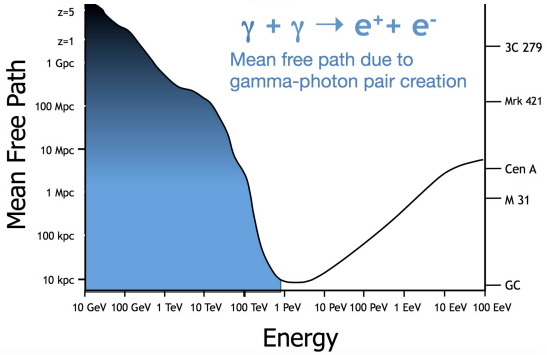
- deepest sensitivity ever
- arcmin angular resolution
- large FoV

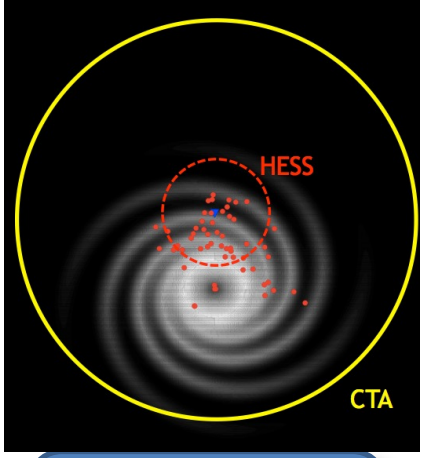
Surveys & precision studies

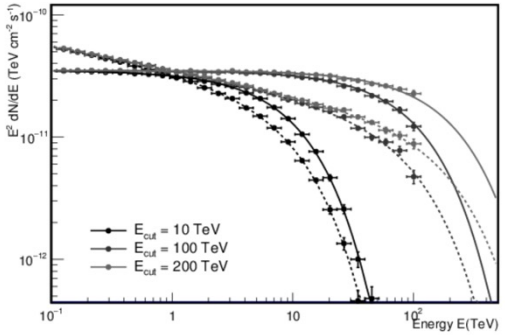
- Precision measurements in a still little explored energy range

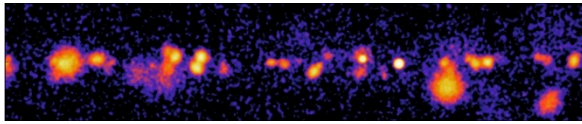
100 TeV range unexplored, precision studies

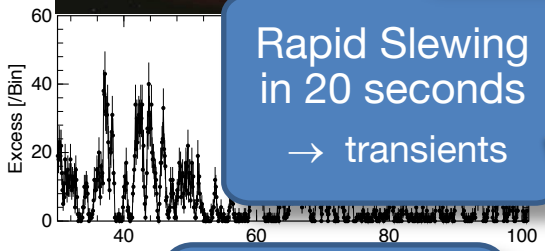


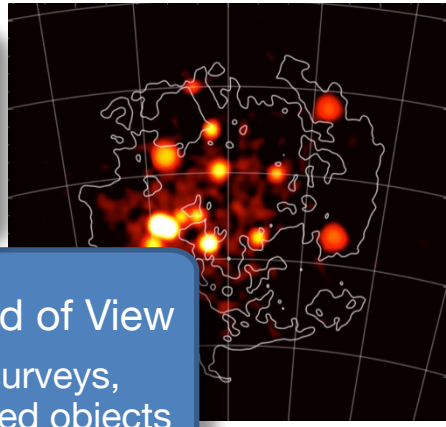












Energies down to 20 GeV
→ Cosmology++

10 x Sensitivity, Large Collection Area
→ all topics

Energies up to 300 TeV
→ Pevatrons

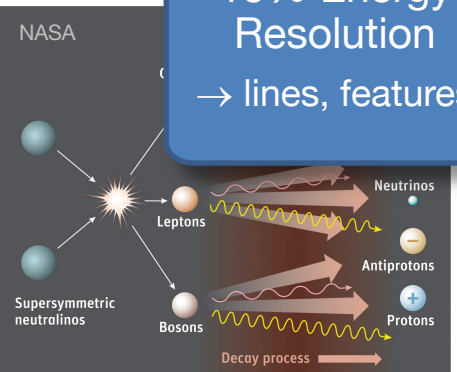
Rapid Slewing in 20 seconds
→ transients

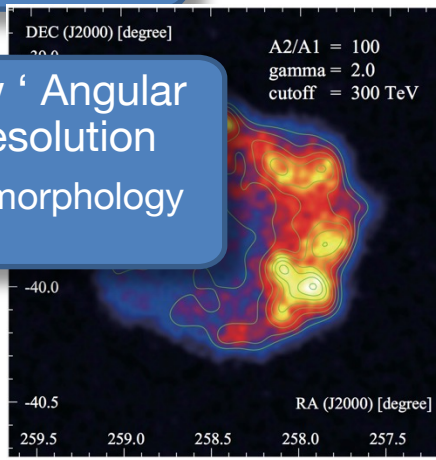
8° Field of View
→ surveys, extended objects

cta
cherenkov telescope array

10% Energy Resolution
→ lines, features

Few ' Angular Resolution
→ morphology





https://doi.org/10.1142/10986

Contents

Chapters and corresponding authors:

CTA SITES: TELESCOPE ARRAY SITES, HEADQUARTERS, SCIENCE DATA MANAGEMENT CENTER

CTA AS AN OBSERVATORY

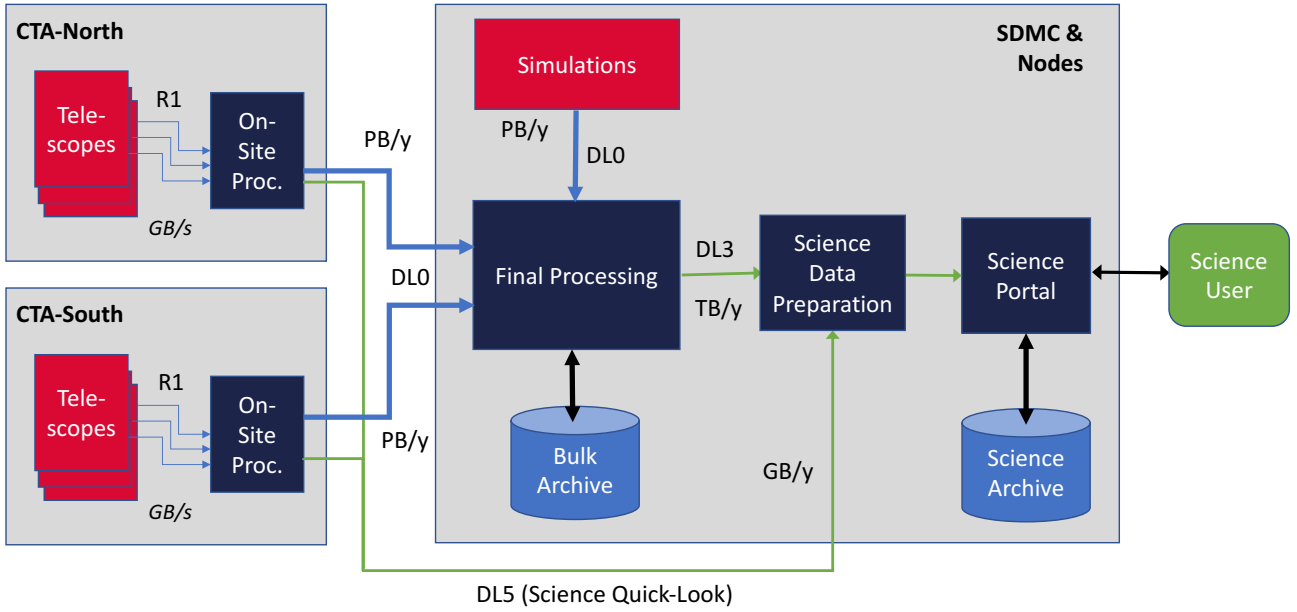
- A Guest Observer Facility
  - For the **first time** in this waveband
    - Existing instruments are run as experiments
  - Annual cycles, TAC ranking, long-term schedule
  - Proposal preparation support, tracking, helpdesk +
  - Public science data archive
    - After proprietary period
- Two Telescope Arrays – one Observatory
  - Inter-site coordination
  - Uniform approach to science operations
- “Data” is the final product of CTAO
  - Users will receive their data fully calibrated in FITS format, and be provided analysis tools
- Open Observatory
  - Follow FAIR principles
  - Support open data and open science
  - Adhere to astronomical standards
  - Support VO

OUTLINE

- CTA
- Data Models
- MM / MWL Context

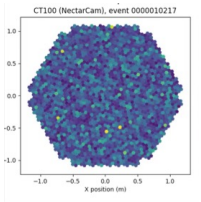
DATA FLOW & ANALYSIS CATEGORIES

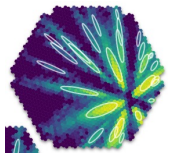
- Data Processing at different steps
  - Timescales of near real-time and next-day for quicklook and science alerts (on-site processing), months for final products (final processing)
  - Involves strict data quality assurance and verification of data products
- Strong data reduction along the processing steps
  - From PB/y (at raw data level) to GB/y (high-level science data)
- Open access through Science Portal
  - access to science archive, to science analysis tools
  - Exploration of quick-look data products



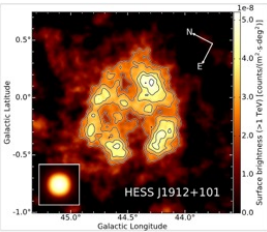
DATA LEVELS

Data Level	Short Name	Definition
R0	Raw Internal	On-site streamed raw data, not normally preserved long-term in this form. R0 content and format is internal to each device / controllable system, such as raw data transmitted from the physical device / system to its respective server in the on-site Data Centre.
R1	Raw Common	On-site stream raw data meeting common standards, transmitted on-site from a Camera or other on-site system to the OES. This is the first level of data seen by the OES, that will typically need some pre-processing from the R0 data format. Exceptionally, some R1 data may be stored for engineering purposes.
DL0	Raw Archived	All archival data from the data acquisition hardware/software, transmitted from the OES to the DPPS. This is the lowest level of data that are intended for long-term storage in the bulk archive. This includes both camera event data and technical data from other sub-systems, such as non-camera devices or software.
DL1	Processed	Processed DL0 data that may include telescope-level (TEL) data and parameters derived from them. Typical contents include calibrated image charge, Hillas parameters, and a usable telescope pattern. DL1 data is not normally stored long-term.
DL2	Recon-structed	Reconstructed shower parameters such as energy, direction, particle ID, and related signal discrimination parameters. Does not include telescope-level (TEL) information. For each event this information may be repeated for multiple reconstruction and discrimination methods. DL2 data is not normally stored long-term.
DL3	Reduced	Sets of selected events with a single final set of reconstruction and discrimination parameters, along with associated instrumental response characterizations and any technical data needed for science analysis.
DL4	Binned	Data product produced by binning of DL3 data, including data cubes and maps which are suitable for combination/summation to produce DL5 products.
DL5	Science	Data product produced by combination of DL4 products an extraction target specific region(s) of interest. Includes for example light-curves and spectra, along with associated data such as source models and fit results.
DL6	High-Level	High-level or legacy observatory data, such as survey maps and source catalogues.





event_id	RA	DEC	E
1	23,3	-40,1	0,01
2	24,6	-40,5	20,0
3	23,5	-41,12	0,45
4	21,3	-38,2	1,03



DATA LEVELS

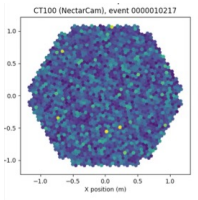
Data Level	Short Name	Definition
R0	Raw Internal	On-site to each respect… …nternal …n to its
R1	Raw Common	On-site system proces… …-site …rposes.
DL0	Raw Archived	All arch the low camera… …. This is …h …ware.
DL1	Processed	Proces… Typical data is … …n. DL1
DL2	Recon-structed	Recons parame repeate… …ination …be …ng-term.
DL3	Reduced	Sets of associa… …g with …lysis.
DL4	Binned	Data pr combin… …for
DL5	Science	Data pr Include results. …terest. …and fit
DL6	High-Level	High-le…

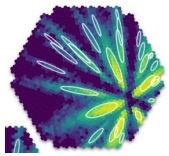
Image/Data Cube
Per Telescope

Mixed Tel/Event (Temporary)

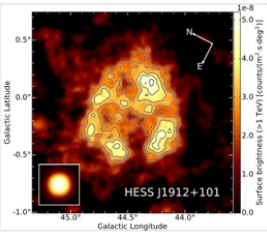
Event List

Binned





event_id	RA	DEC	E
1	23,3	-40,1	0,01
2	24,6	-40,5	20,0
3	23,5	-41,12	0,45
4	21,3	-38,2	1,03

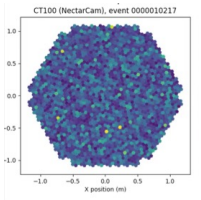


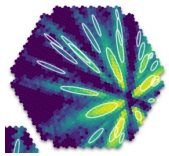
DATA LEVELS

Data Level	Short Name	Definition
R0		
R1		
DL0		
DL1		
DL2		
DL3		
DL4		
DL5		
DL6		

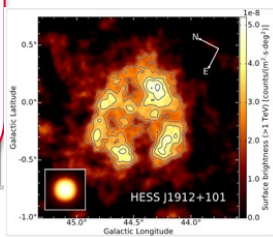
Internal to the Observatory

Available to the Science User

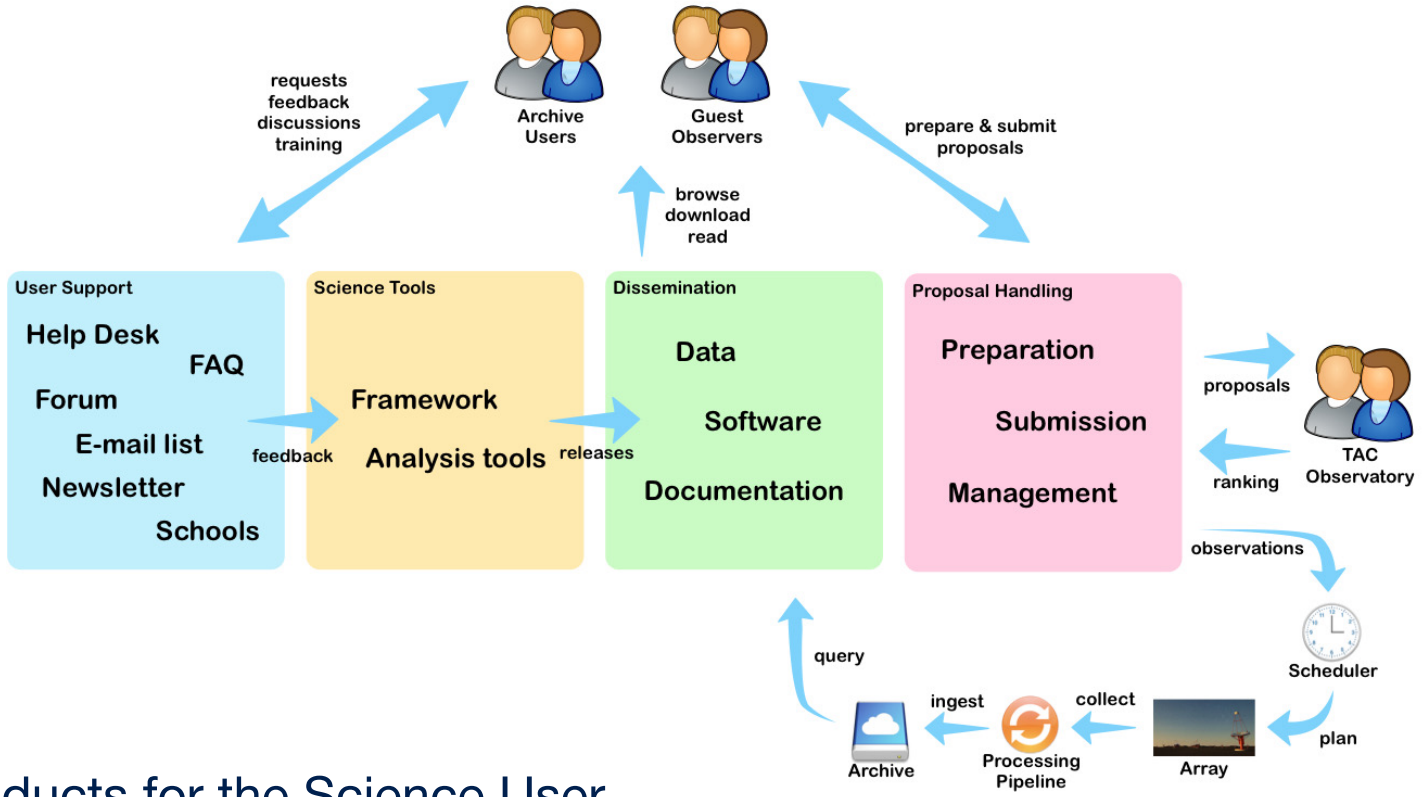




event_id	RA	DEC	E
1	23,3	-40,1	0,01
2	24,6	-40,5	20,0
3	23,5	-41,12	0,45
4	21,3	-38,2	1,03



SCIENCE USER PERSPECTIVE



- Products for the Science User
  - Photon (candidate) event list data (FITS) (DL3) with corresponding instrument response functions, background model
  - SkyMaps, Lightcurves, Spectra (DL5), Catalogues (DL6)
  - Science analysis tool suite, supporting documentation

DL3 DATA PRODUCTS

Output: Science Data

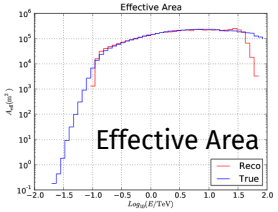
Event-List

event_id	RA	DEC	E	class	type	n_tels	...
1	23,3	-40,1	0,01			5	
2	24,6	-40,5	20,0			34	
3	23,5	-41,12	0,45			3	
4	21,3	-38,2	1,03			4	

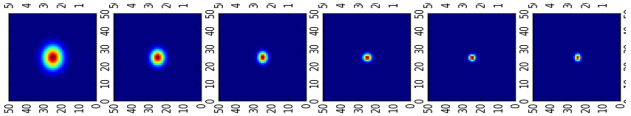
Technical Tables (for sub-GTIs)

TIME	Transparency	Temperature	Trigger Rate
580234.34	0.8	32	12034
580234.35	0.94	32	13023
580234.36	0.70	33	12532

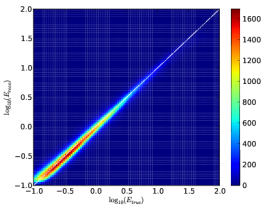
Instrumental Responses:

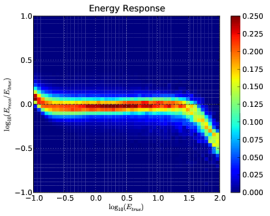


PSF



Energy Migration

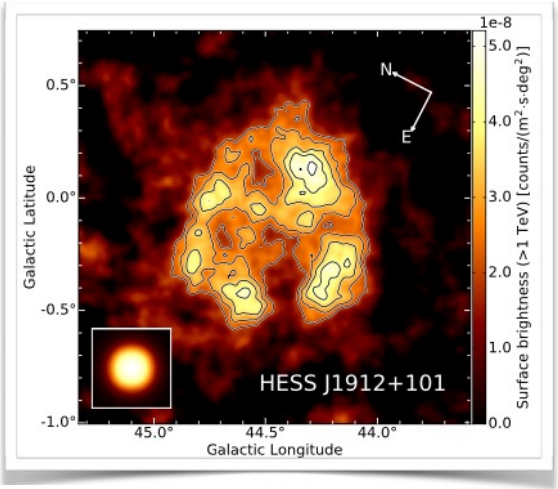


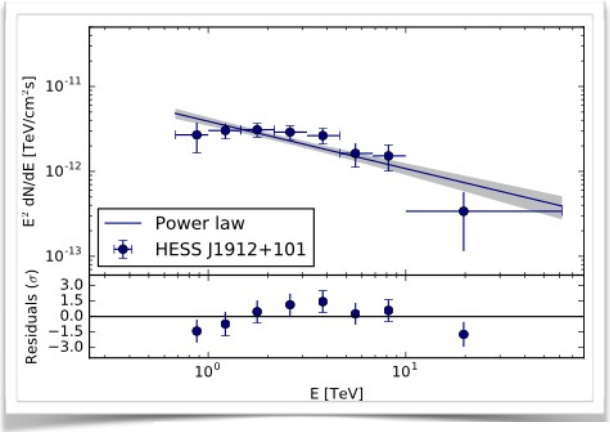


(note these are not CTA responses, just examples form HESS)

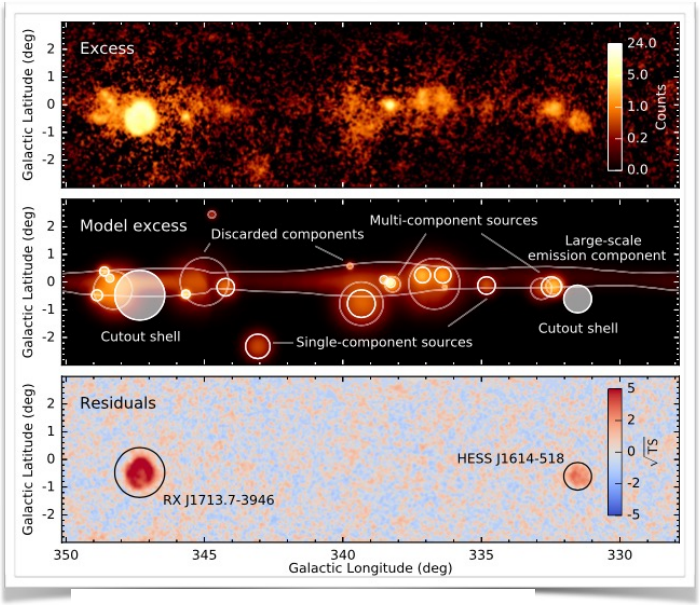
DL5 AND DL6 DATA PRODUCTS

- DL5 Data Products
  - Sky Maps
  - Spectra
  - Light Curves
  - Time Series
  - Phaseograms





H.E.S.S. A&A 612, A8 (2018)



H.E.S.S. A&A 612, A1 (2018)

- DL6 Data Products
  - Catalogues

DATA FORMATS FOR GAMMA-RAY ASTRONOMY

- CTA aims to have an open and interoperable data model and format that is VO compliant
  - Note: we treat data model and format separately
- CTA currently working towards finalising definition of CTA data model for all data levels from R1 to DL3 and DL5
- For DL3, based on Open Gamma-ray Astro Data Format (OGADF)
  - https://gamma-astro-data-formats.readthedocs.io/en/latest/
  - Community-driven work by members from current IACTs (H.E.S.S., VERITAS, MAGIC) and CTA mainly in 2015-2018
  - used in current IACTs, that already lead to harmonisation
- CTA will evolve the OGADF further to meet its science needs
  - Revisions are currently ongoing for DL3, aim to have an updated version towards end of this year

DATA FORMATS FOR GAMMA-RAY ASTRONOMY IN IVOA

- Question:
  - What is needed to make the CTA data model and format interoperable, FAIR, and VO compliant?
  - What is needed for the standardisation* of a data format for the gamma-ray astronomy? Potentially within the IVOA?
  - Which parts map to already existing standards, which parts are new?
- Remarks:
  - H.E.S.S. already released parts of their data via the VO and it works
    - https://www.mpi-hd.mpg.de/hfm/HESS/pages/dl3-dr1/
  - There is potential interest on common data formats in gamma-ray astronomy beyond current IACTs:
    - e.g. HAWC, SWGO, ..
  - There is potential interest on common data formats from other wavelengths:
    - e.g. KM3NeT (neutrinos)
  - At the least, Interoperability is needed to support (joint) MM/MWL science cases
  - Discussions are currently on-going as part of the OGADF open forum, within the ESCAPE framework and/or as bilateral discussions between observatories

* There is still some confusion in our community on ‘Standardisation vs. commonalities vs. interoperability’ and when a data format becomes ‘a standard’

FURTHER TOPICS RELATED TO DATA PROCESSING

- Large-scale data processing to bring the raw data (DL0) to reduced and calibrated science data products (DL3) is a service task by the Observatory
  - Includes all steps from calibrations, simulations, event reconstruction, data volume reduction, …
- However, there are clearly benefits from commonalities in the framework of the VO
  - Provenance for data products and activities (long-standing work in IVOA by M. Servillat and C. Boisson, provenance in DIRAC (M. Sanguillon)
  - workflow management (e.g. common workflow language, FAIR, metadata)
  - Machine-learning
- These topics are currently discussed within ESCAPE framework

OUTLINE

- CTA
- Data Models
- MM / MWL Context

MM / MWL CONTEXT

- Support for MWL / MM physics is one of the key aspects of CTA in up-coming era of time domain astronomy
  - Follow-up on external Science Alerts or issuing of CTA-detected Science Alerts possible on O (1min) timescale
  - Coordinated campaigns for multiwavelength observations
  - Support for ToO programs spanning across multiple observatories
- Ongoing discussions between observatories on how to best organize operations, software and data, and in the community on their needs
  - ASTERICS, ESCAPE
  - CTA consortium Kavli-IAU Transients 2020 Workshop

https://arxiv.org/abs/2007.05546

International Coordination of Multi-Messenger Transient Observations in the 2020s and Beyond
Kavli-IAU White Paper

REMARKS ON SCIENCE ALERTS

- Discussions started within on enhancements of VO events (or similar format / supporting infrastructure)
  - User-contributed or curated information
    - esp. important for upcoming era of community brokers or user-contributed alert correlation and filtering
  - Provenance information
    - Enrich alerts with information on data quality, reconstruction or detection algorithms, multiple science cases, alert correlation, …
  - Allow for feedback loops (observatory-enhanced alerts)
    - Real-time negotiations to support “Facility ABC will observe if facility XYZ observes too simulat”
    - Re-inserting the events with additional follow-up information?
    - Notification that observations will start?

SCHEDULES AND VISIBILITY

- CTA aims to be able to
  - exchange observation schedules to support MM / MWL
  - Inform on visibility
  - Support multi-observatory scheduling (TBC)
- CTA aims to use appropriate standards and protocols from VO
- Detailed discussions or implementation not yet started with the relevant groups in VO

SUMMARY

- CTA will work as an open Observatory
  - User services and support in the core
- Data Challenges
  - Several PB/y raw data to be handled and processed to science-ready data products for the science users
  - High-quality science data products, software tools and services
- CTA aims to be compliant with VO standards and protocols and connect CTA to VO

CTA has interest in various topics related to VO, in particular:

- Data models (DL3 and higher)
- Provenance (all topics)
- Support for multi-messenger approach (VO events, schedules)