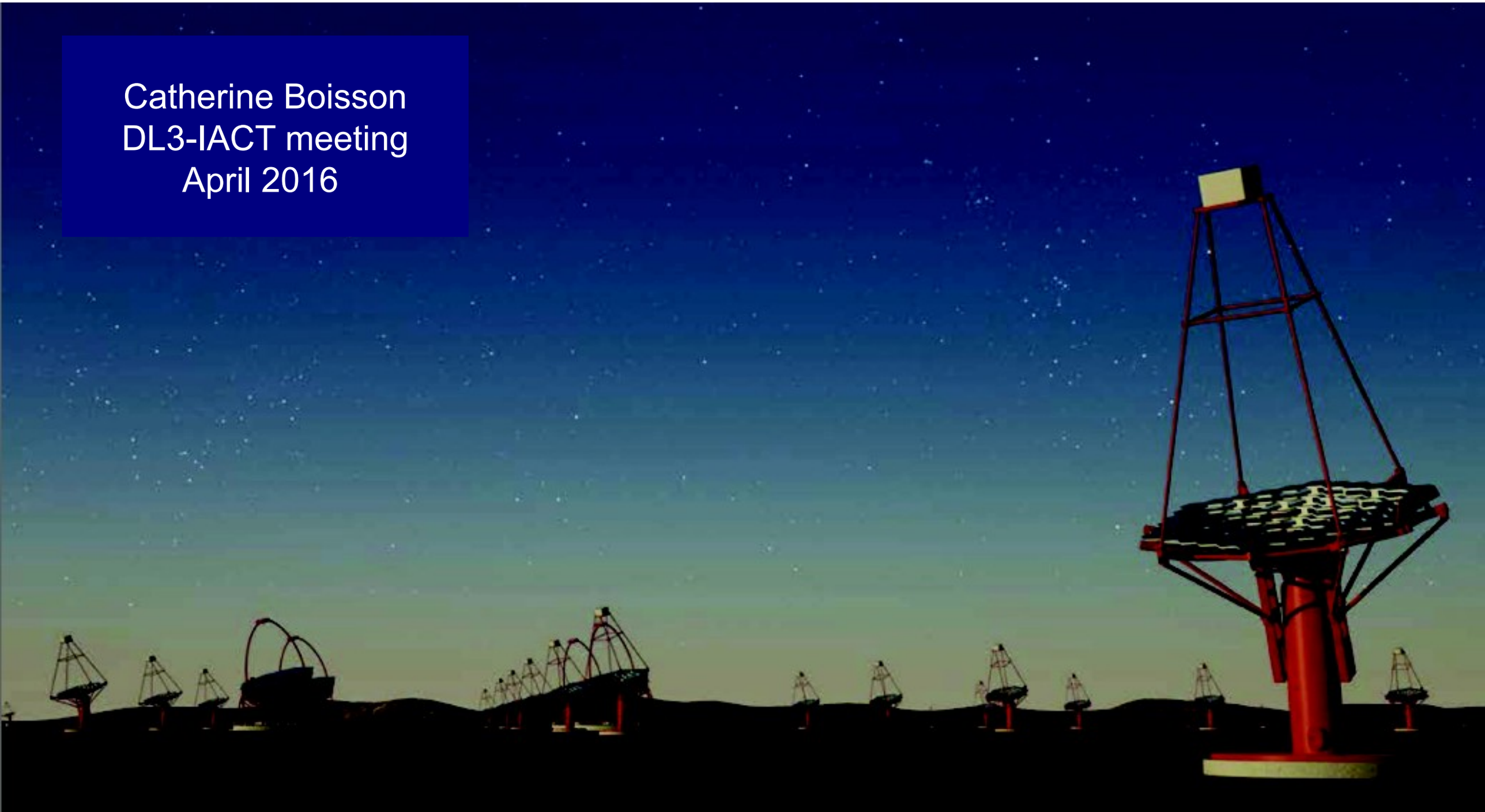
IACT Data to the End User

Catherine Boisson
DL3-IACT meeting
April 2016

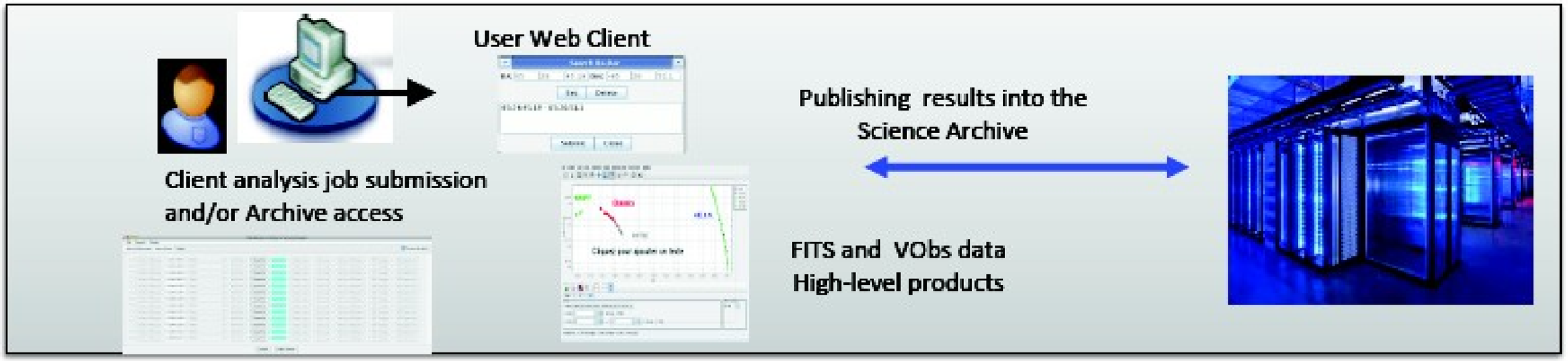


Challenges

Distant telescopes and a worldwide community:
reliable high-bandwidth intercontinental connection

Data rate, tens PB/year:
challenges for streaming, on-site processing, archive, multithread pipelines processing, and long term sustainability

Open access to Observatory data for a worldwide community:
"Scientific Analysis System" integrating together Data Centre, Archive, Software and e-infrastructures

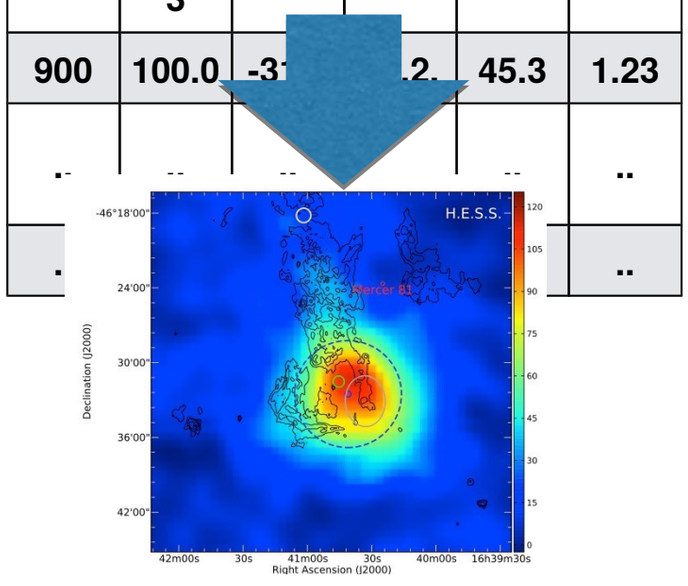


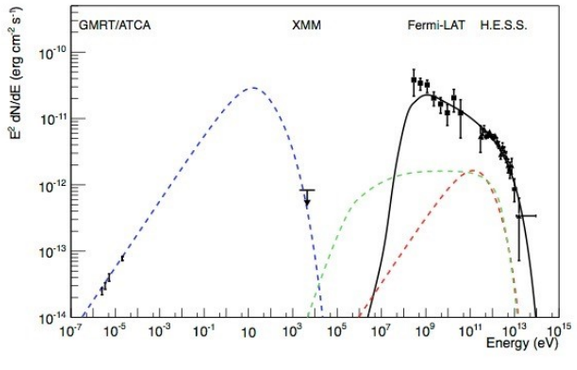
Data (level / name)		Content	Reduction
0	Raw	Telescope → disk	1
1	Calibrated	Physical quantities	1
2	Reconstructed	γ-ray shower parameters	10-1
3	Reduced	Selected for Physics	10-2
4	Science	Spectra, lightcurves, skymaps…	10-3
5	Observatory	Legacy observatory data (e.g. catalogue)	10-3 – 10-5

- Low-level (calibration)
- Mid-level (reconstruction)
- High-level (user-analysis)

CTA Analysis Challenges

t	RA	Dec	az	alt	E
501	128.2	-37.1	321.1	45	0.112
600	130.23	-36.2	322.1	45.1	0.242
620	120.124	-33.33	312	45.7	0.434
640	121.1	-34.1	323.4	43.0	35.0
720	123.23	-28.1	322.1	43.5	0.401
900	100.0	-31	.2.	45.3	1.23
..	..	..	..	..	..





@ Kosack

What science data products do you get?

- **event-lists**: list of reconstructed showers with direction and energy (not necessarily gamma-rays!)
- **IRF tables**: instrumental response functions, matched as close as possible with observing conditions for each exposure (N-dimensional tables, larger than the event lists in size)
- **TECH tables**: atmosphere and data-quality measurements as a function of time during the observations, for use to make "good-time-interval" cuts, depending on science goal

What do you want to produce?

- Sky images, generally **data cubes**, e.g. f(RA,Dec,Energy):
  - total counts (gamma + background)
  - estimated gamma counts
  - flux
- Light-curves for a region of interest
- Spectra for a region of interest

More challenges

Large PSF + many sources = *confusion*

- need to model both CR background and diffuse gamma-ray background
  - *not so easy! Unlike at GeV energies, no good model for diffuse gammas > TeV…*
- many sources are non-point-like (extended, complex morphology)
  - may need to model many overlapping compact sources

Not sufficient to ask for spectrum from a region of interest…

- may contain too many components to be useful
- have to be clever

Science operations

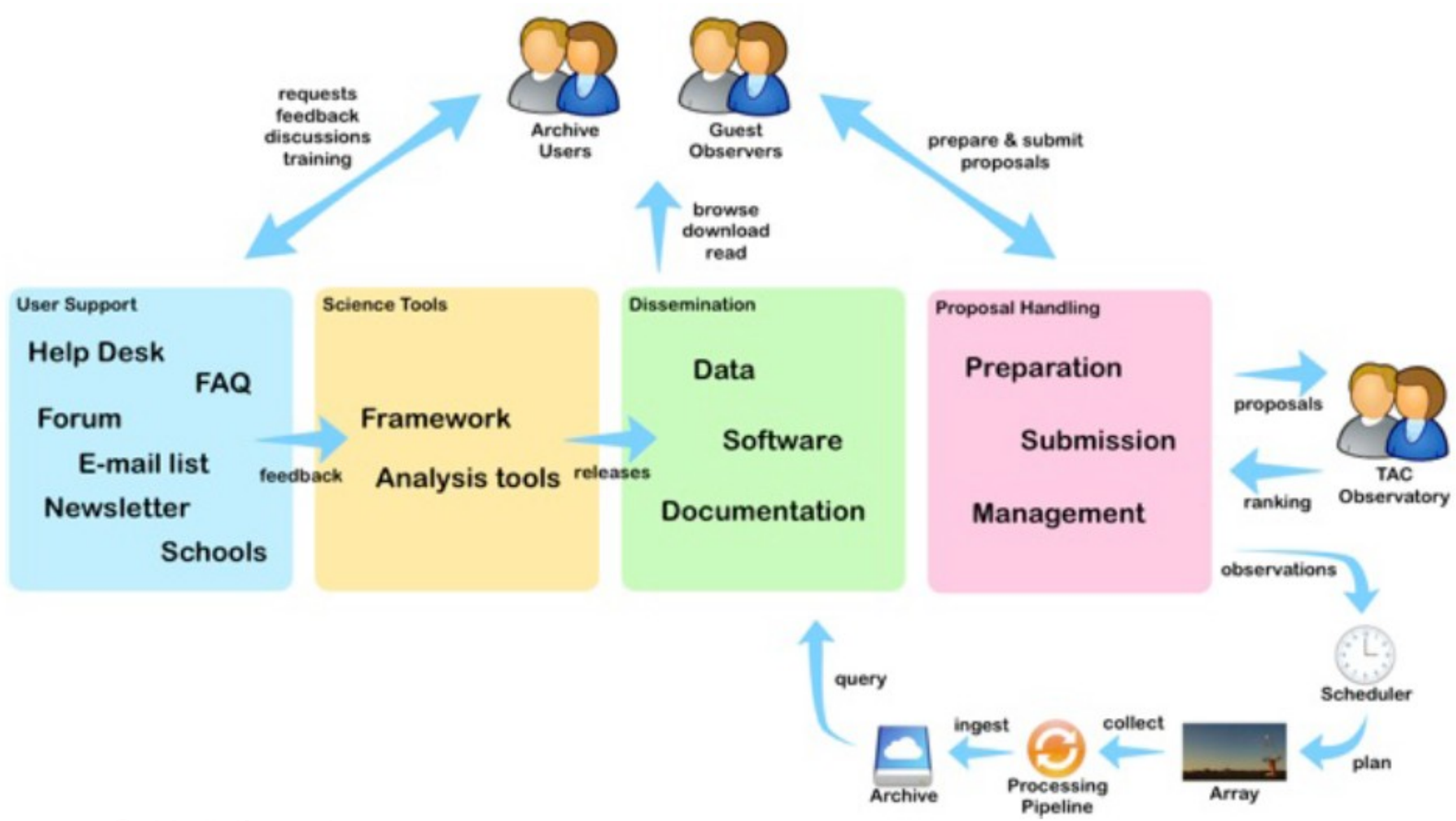
- **Data dissemination**
  - Allows Guest Observers and Archive Users to have access to data
    - Data rights / proprietary period granted to PI

      large FoV , segmented observations , serendipitous sources
    - Storage & distribution → **comprehensive metadata** + format that enables multi-wavelength analyses and cross-correlation with existing data bases → Virt Obs compliance
    - Science tools
    - User support

- **Science Tools**
  - Software package to derive astronomical images, spectra and light curves (data levels 4, 5 from 3)

- **User Support**
  - Documentation, Tutorials, Help Desk, Training

Observer Access



@ Knödleseder

IACT as an Observatory

Opening observation time to guest observers implies:

- Make data available to the PI for a *proprietary* period

- Offering access to the data in a transparent way, using astronomical standard formats

- Offering standard tools ('pipeline software') for analysis and documentation to process these data

Data from experiments are all stored in custom format which can only be interpreted with collaboration software.

Storage & distribution → format that enables multi-wavelength analyses and cross-correlation with existing data bases

→ *Virt Obs compliancy*

IACT as an Observatory

Deliver back data products

- Data systematically pipeline-processed to generate a range of scientific and auxiliary data products

- The minimal data product to provide to the Guest Observer :
  - list of events with the relevant quality informations,
  - required instrument response functions,
  - documentation,
  as well as science data (images, spectra, light curves...)

- The archive will be distributed through services

- Link to ARCHIVE – see M. Servillat's talk

Data policy

Products to deliver

- Data Level 3 and higher

  i.e. photon lists of gamma-ray candidate events, with the option of different selection parameters, including appropriate information on event quality

→ what we are concerned with this week

- Data Level 0

  upon request, subject to availability of computing resources;

  the Observatory is not obliged to provide software or user support for analysis of such data.

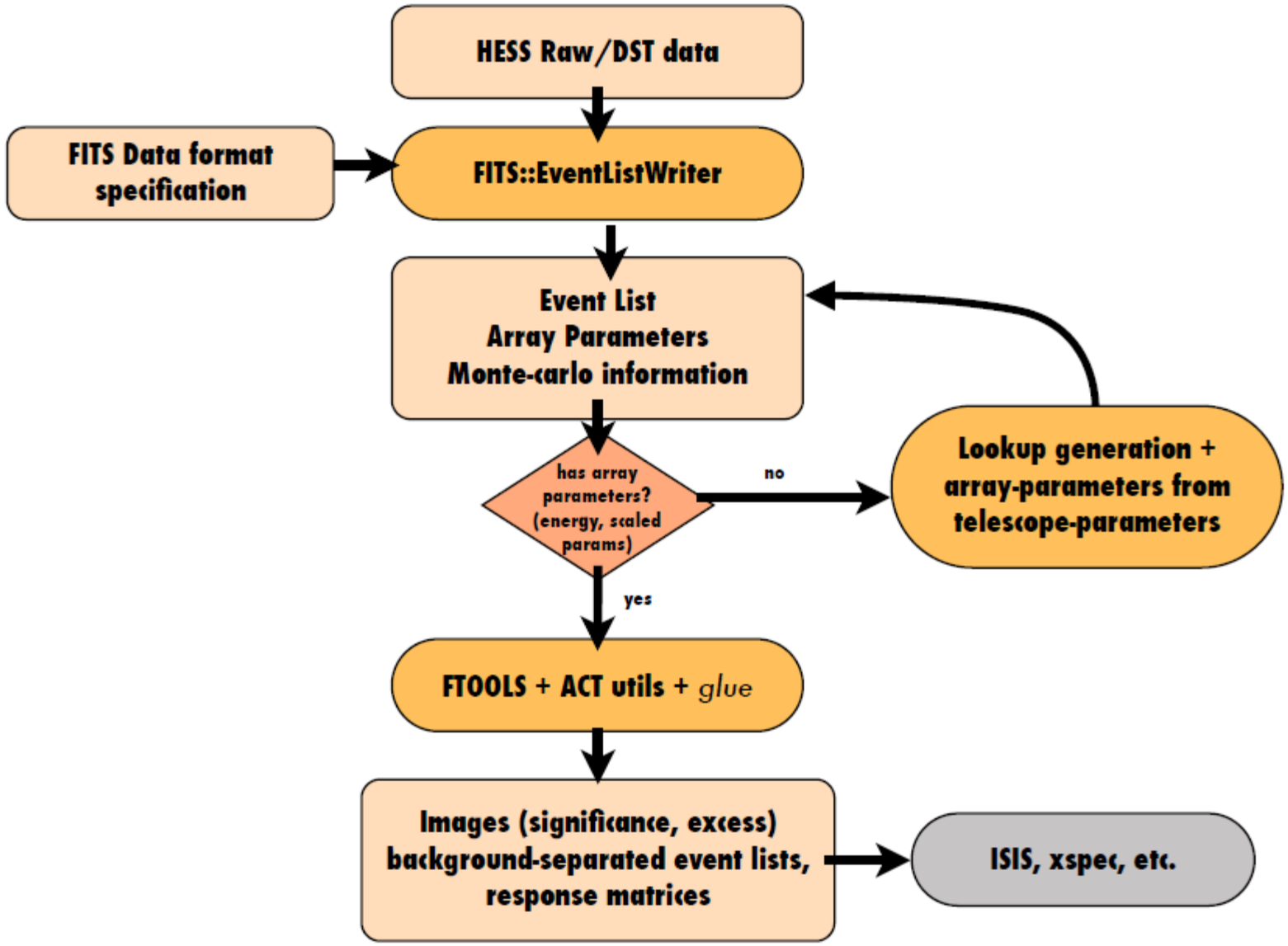
IACT as an Observatory

Deliver high-level analysis software

- Distribution of software tools to Guest Observers requires central software releases, up-to-date documentation and extensive user support

→ only one analysis chain or different approaches?

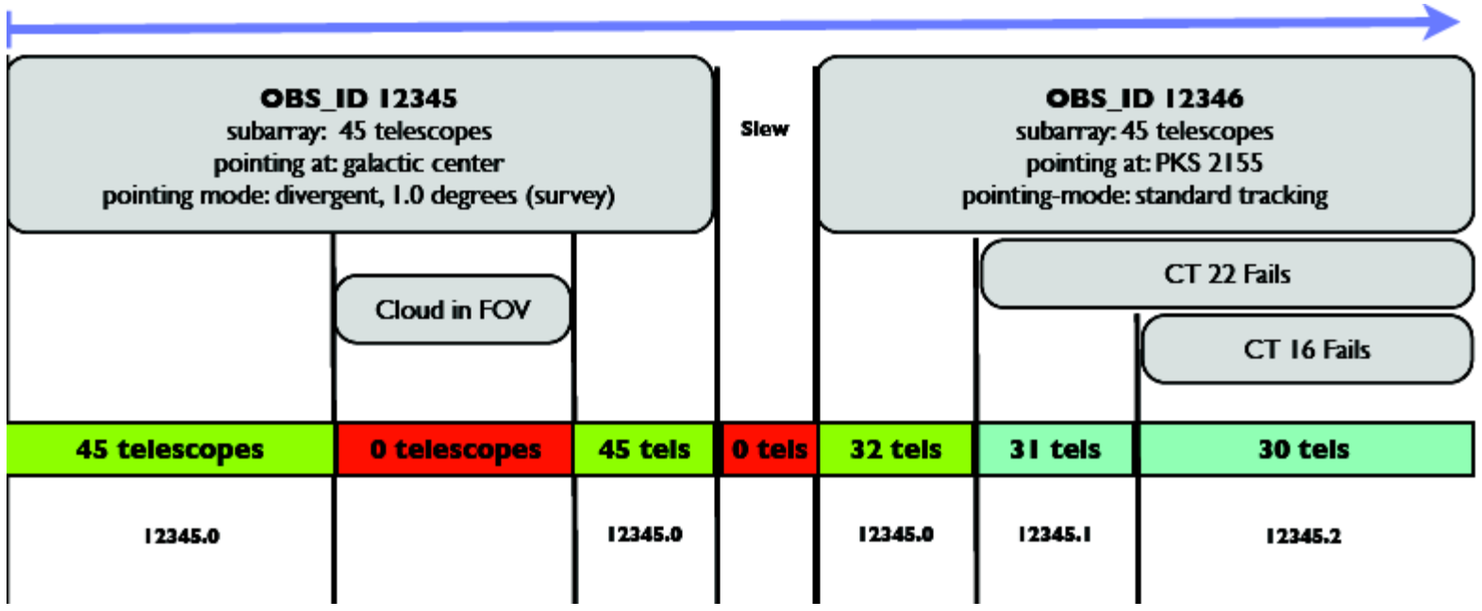
- The software environment should allow to:
  - access data from a Science Archive
  - perform interactive data reduction from event list to publishable products
  - visualize and manipulate data
  - perform science analysis
  - get access to documentation and help


HESS Raw/DST data
FITS Data format specification
FITS::EventListWriter
Event List
Array Parameters
Monte-carlo information
has array parameters? (energy, scaled params)
no
Lookup generation + array-parameters from telescope-parameters
yes
FTOOLS + ACT utils + glue
Images (significance, excess) background-separated event lists, response matrices
ISIS, xspec, etc.

Questions still pending

How should an observation be defined?

- affects how response tables, etc, are indexed and retrieved
- how do we store each possibility?



Questions still pending

GTIs to select time regions where data should be used, we need two levels of quality selection:

- GTI: boolean, tells when data should be used or not
- RTI: (response time interval) to tell you when array changed, requiring new response info
  - per telescope
  - lists telescopes that are “good” /participating for a time interval

STILL VALID ?

This meeting

Update event, IRF and TECH specs for IACT DATA
Have a definition for TECH3

Open DL3 data format specs and data challenges to a larger community than CTA-internal → development with that in mind.

Should work for all existing experiments

ASTERICS - IACT

To be published compliant with VO :
- event lists
- spectra, light curves, Images
- source catalogs

Data challenges proposed in DADI WP4 and OBELICS WP3 are coherent with some of the key issues explored in the CTA Data Management project.

Opportunity for CTA to secure the Virtual Observatory developments for the future “data dissemination challenges” of the CTA Observatory.

True also for running Cherenkov “pathfinders”: H.E.S.S. and MAGIC