Activities of UCM-GAE within the Task 3.2 D-GEX

J. Rosado, J.L. Contreras, M. López, F.J. Franco

Asterics
Astronomy ESFRI & Research Infrastructure Cluster

UNIVERSIDAD COMPLUTENSE MADRID

Participation in OBELICS

UCM staff in ASTERICS:

- José Luis Contreras (PI): 4.1 PM
  contrera@gae.ucm.es
- Francisco Javier Franco: 3.3 PM
  fjfranco@fis.ucm.es
- Marcos López: 3 PM
  marcos@gae.ucm.es
- Jaime Rosado: 18 PM
  jaime_ros@fis.ucm.es

Co-leader of WP3.2 D-GEX (Data Generation and information EXtraction)

Survey of data formats to be used in ESFRI projects and development of software solutions.

Participation in WP3.3 D-INT (Data systems INTgration)

Scaling-up existing databases and storage architectures beyond the Peta-scale level.

Members of CTA

Data model coordination, among other tasks.

Work Plan for D-GEX (ideal)

1. Data Format Survey

   Survey and comparison of standards used in ESFRI projects and related pathfinders.

2. Prototype Development

   Prototypes of Data Access Libraries (DAL) for selected formats, tests of libraries on CTA and evaluation of results.

3. Extension to other ESFRI Projects

   Proposal for Standard DAL (SDAL) to support ESFRI projects, training sessions and documentation.

4. Final Evaluation

   Development of SDAL library, implementation on ESFRI projects and final evaluation and lessons learnt reports.

More info at the pages of the group in the ASTERICS wiki (internal):

https://www.asterics2020.eu/dokuwiki/doku.php?id=intra:wp3:task3.2ucm

Data Format Survey

- Formal deliverable of the project (del. number D3.3).
- Revised by members of OBELICS and approved on 3rd May.

II. DELIVERY SLIP

	Name	Partner/WP	Date
Author(s)	Jaime Rosado and José Luis Contreras	UCM/WP3	21/04/2016
Amendments	Tammo Jan Dijkema, Tamas Gal, Dominique Boutigny	ASTRON/WP3 FAU/WP3 CNRS/WP3	26/04/2016
Reviewed by	Rob van der Meer	ASTRON/WP1	03/05/2016
Approved by	Rob van der Meer	ASTRON/WP1	03/05/2016

- Available at the internal wiki pages:

https://www.asterics2020.eu/dokuwiki/doku.php?id=intra:wp1:deliverables

- Working document made for this Survey also available at:

https://www.asterics2020.eu/dokuwiki/doku.php?id=intra:wp3:task3.2ucm

Contents of the Survey

- Experiments classified according to the type of data:
  - Image-based experiments: VIS and NIR telescopes (E-ELT, EUCLID, etc.).
  - Event-based experiments: gamma-ray, neutrino and cosmic-ray observatories (CTA, KM3NeT, Auger, etc.).
  - Signal-based experiments: radio interferometric arrays and GW detectors (SKA, LOFAR, Virgo, etc.).

- Possible synergies are identified:
  - Science-ready products from all types of experiments in formats compliant with the VO framework.
  - Common file formats and standards for data of experiments within the same category.

HDF5 for raw data

- HDF5 has advantages for management of large amount of data and is being establishing as a standard with many available tools and libraries.

- We plan to test it for CTA raw data and think it is suitable for other event-based experiments (not only):
  - Metadata stored along with data.
  - Well suited for hierarchical data, e.g.,
    array event → camera event → pixel signal.

- Would be a contribution to WP3.3 D-INT.

Standards for high-level data

- Data from event-based experiments for distribution to users are basically:
  - Event lists: type of particle, energy, direction...
  - IRFs: effective area, point spread function...
- Recent initiative to define a unified standard based on FITS for Imaging Atmospheric Cherenkov Telescopes: CTA, MAGIC, H.E.S.S., etc. (see talk of T. Hassan).

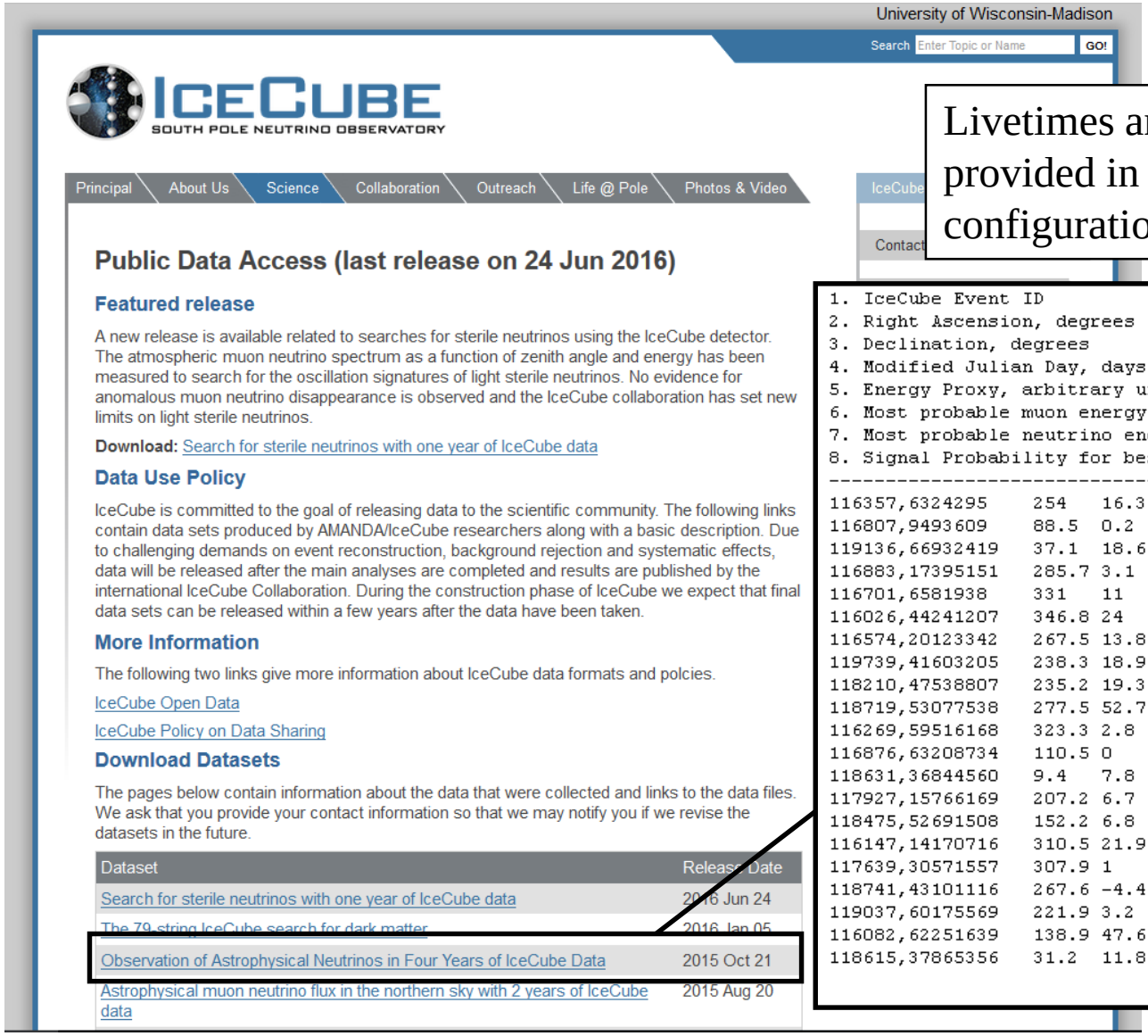
  https://gamma-astro-data-formats.readthedocs.io/
- The UCM group has proposed to extend this standard for other gamma-ray observatories (e.g., HAWC) as well as for neutrino and cosmic-ray observatories (e.g., KM3NeT).

Motivations of the proposal

- Open data promoted by EU and goal of ASTERICS.
- Crosscheck analyses by the scientific community.
- Profit from common (open) tools.
- Multi-messenger analyses.
- Already happening: event sets are already being made available (web, electronic journals).

Example for IceCube data

IceCube Public Data Access: https://icecube.wisc.edu/science/data



Livetimes and effective areas also provided in HDF5 for different detector configurations and particle types.

```
1. IceCube Event ID
2. Right Ascension, degrees
3. Declination, degrees
4. Modified Julian Day, days
5. Energy Proxy, arbitrary units
6. Most probable muon energy for best-fit, TeV
7. Most probable neutrino energy for best-fit, TeV
8. Signal Probability for best-fit
-------------------------------------------------------------------------
116357,6324295    254   16.3  55421.5     289916      755   1693  0.96
116807,9493609    88.5  0.2   55497.3     199981      604   880   0.83
119136,66932419   37.1  18.6  55911.3     157871      397   713   0.88
116883,17395151   285.7 3.1   55513.6     147002      422   709   0.8
116701,6581938    331   11    55478.4     140113      317   466   0.81
116026,44241207   346.8 24    55355.5     139728      339   442   0.86
116574,20123342   267.5 13.8  55464.9     131950      302   400   0.82
119739,41603205   238.3 18.9  55987.8     130382      326   394   0.85
118210,47538807   235.2 19.3  55702.8     106898      252   393   0.77
118719,53077538   277.5 52.7  55829.3     91292       156   198   0.66
116269,59516168   323.3 2.8   55405.5     66558       134   193   0.41
116876,63208734   110.5 0     55512.6     64650       147   206   0.45
118631,36844560   9.4   7.8   55806.1     63994       139   179   0.5
117927,15766169   207.2 6.7   55642       60582       125   185   0.45
118475,52691508   152.2 6.8   55768.5     56734       124   156   0.45
116147,14170716   310.5 21.9  55387.5     53805       112   178   0.55
117639,30571557   307.9 1     55589.6     53542       116   184   0.33
118741,43101116   267.6 -4.4  55834.4     51756       116   191   0.56
119037,60175569   221.9 3.2   55896.9     51631       109   158   0.58
116082,62251639   138.9 47.6  55370.7     51112       109   189   0.37
118615,37865356   31.2  11.8  55803       50376       109   190   0.49
```

Other activities on data model within CTA

- Definition of contents of raw data and metadata.
- Testbed for the different raw data formats proposed so far.
- Array configuration database.
- Modeling of data production and analysis.
- Use cases for continuous observation (slewing).

Many interactions with OBELICS tasks →
How to make all this compatible with OBELICS: deliverables and products ?

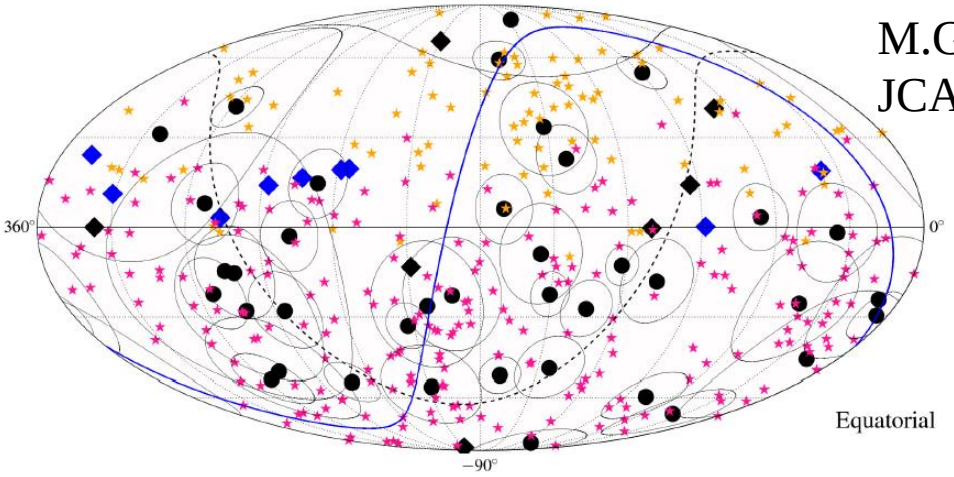
THANKS !

BACKUP SLIDES

Use case

"Search for correlations between the arrival directions of IceCube neutrino events and ultrahigh-energy cosmic rays detected by the Pierre Auger Observatory and Telescope Array".

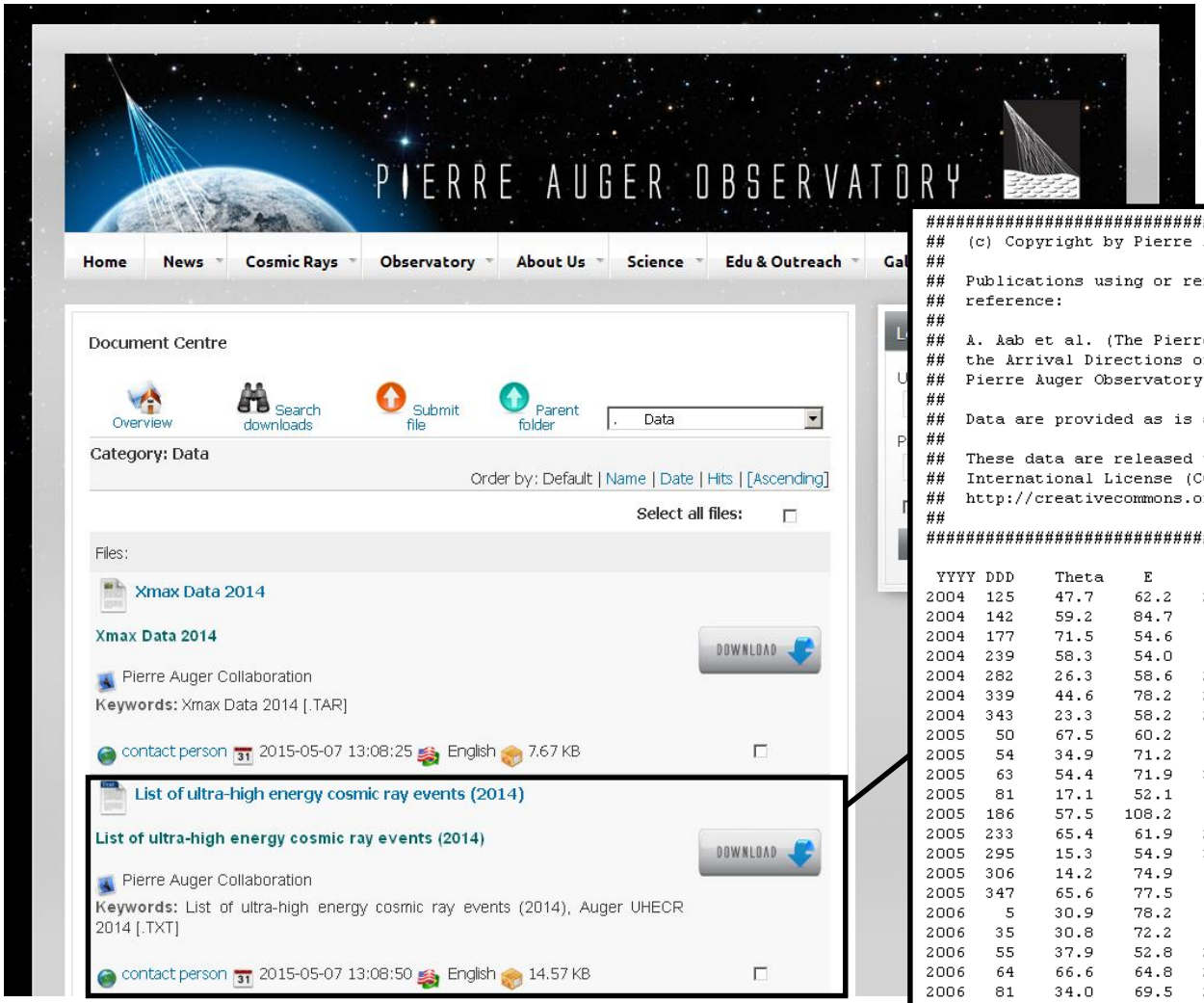
- HE event lists for selected observation periods.
- Angular resolutions: average values for Auger and TA, while event-dependent for IceCube.
- Model for an energy dependent Gaussian deflection of cosmic rays.
- Average relative exposures of Auger and TA as a function of declination.



M.G. Aartsen et al., JCAP01 (2016) 037

Open data sets from Auger

Auger Document Center: https://www.auger.org/index.php/document-centre




```
################################################################################
##  (c) Copyright by Pierre Auger Collaboration 2014                          ##
##                                                                            ##
##  Publications using or referring to these data should cite the following   ##
##  reference:                                                                ##
##                                                                            ##
##  A. Aab et al. (The Pierre Auger Collaboration), "Searches for Anisotropies in  ##
##  the Arrival Directions of the Highest Energy Cosmic Rays Detected by the   ##
##  Pierre Auger Observatory", arXiv:1411.6111                                ##
##                                                                            ##
##  Data are provided as is and might be subject to future changes.           ##
##                                                                            ##
##  These data are released under the Creative Commons Attribution-ShareAlike 4.0  ##
##  International License (CC BY-SA 4.0). For more details see                ##
##  http://creativecommons.org/licenses/by-sa/4.0/                            ##
##                                                                            ##
################################################################################

 YYYY DDD    Theta     E       RA      dec      Glon     Glat
2004  125    47.7     62.2    267.2    -11.4     15.5      8.4
2004  142    59.2     84.7    199.7    -34.9    -50.8     27.7
2004  177    71.5     54.6     12.7    -56.6    -56.9    -60.5
2004  239    58.3     54.0     32.7    -85.0    -59.1    -31.8
2004  282    26.3     58.6    208.1    -60.1    -49.5      1.9
2004  339    44.6     78.2    268.4    -61.0    -27.6    -16.9
2004  343    23.3     58.2    224.7    -44.0    -34.1     13.1
2005   50    67.5     60.2     29.0    -14.0    174.9    -70.0
2005   54    34.9     71.2     17.5    -37.8    -76.0    -78.6
2005   63    54.4     71.9    331.2     -1.3     58.7    -42.4
2005   81    17.1     52.1    199.1    -48.5    -52.8     14.1
2005  186    57.5    108.2     45.6     -1.7    179.5    -49.6
2005  233    65.4     61.9    278.4     -1.3     29.7      3.4
2005  295    15.3     54.9    333.0    -38.1      4.4    -55.0
2005  306    14.2     74.9    114.8    -42.8   -103.9    -10.0
2005  347    65.6     77.5     18.3     29.2    128.6    -33.4
2006    5    30.9     78.2     18.9     -4.7    138.3    -66.8
2006   35    30.8     72.2     53.6     -7.8   -165.9    -46.9
2006   55    37.9     52.8    267.6    -60.6    -27.5    -16.4
2006   64    66.6     64.8    275.2    -57.2    -22.6    -18.6
2006   81    34.0     69.5    201.1    -55.3    -52.3      7.3
2006  100    33.7     54.7     28.8    -16.4   -179.9    -71.8
2006  118    57.3     56.3    322.5     -2.0     51.6    -35.6
2006  126    65.2     82.0    299.0     19.4     57.6     -4.7
```

Parameters in Event Lists

IACT open format	IceCube	Auger
EVENT_ID [integer]	EVENT_ID [float]	
TIME [s]	TIME [day]	YYYY DDD
RA & DEC	RA & DEC	RA & DEC
ALT [elevation]		THETA [zenith angle]
GLON & GLAT		GLON & GLAT
ENERGY [TeV]	ENERGY PROXY [arb. units]	ENERGY [EeV]
	MUON ENERGY [TeV]	
	NEUTRINO ENERGY [TeV]	
EVENT TYPE [event quality]	SIGNAL PROBABILITY	
TELESCOPE		
OBSERVER		
OBS_MODE		
…		