AIDA

Advanced Infrastructure for Data Analysis

In-kind contribution INA-ITA-S23

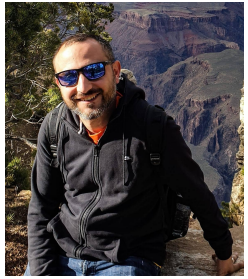
“Staff effort in support of Rubin commissioning: ML tools for instrumental monitoring and analysis”

LSST@Italy 2026 - Astronomical Observatory of Capodimonte, Napoli (Italy)
20th January, 2026

AIDA - Team

Dr. GIUSEPPE RICCIO, Researcher @ INAF, Astronomical Observatory of Capodimonte

In-Kind Contribution Leader, main designer and developer of AIDA, long experience on software engineering, astrophysical data analysis and machine learning.



Dr. STEFANO CAVUOTI, Senior Researcher @ INAF, Astronomical Observatory of Capodimonte

Co-developer of AIDA, long experience on astrophysical data analysis and machine learning.

Prof. Massimo Brescia, Associate Professor @ Dept. of Physics, University of Naples Federico II

Program leader for the Italian In-kind Program, previous contribution leader of ITA-INA-S23 in-kind, specialized in Data Science and Machine Learning solutions for Astrophysics.

Past Contributors

Dr. Silvia Pietroni, postdoc @ INAF, Astronomical Observatory of Capodimonte (Dec.2022 - March 2024), software developer

In-Kind Recipients

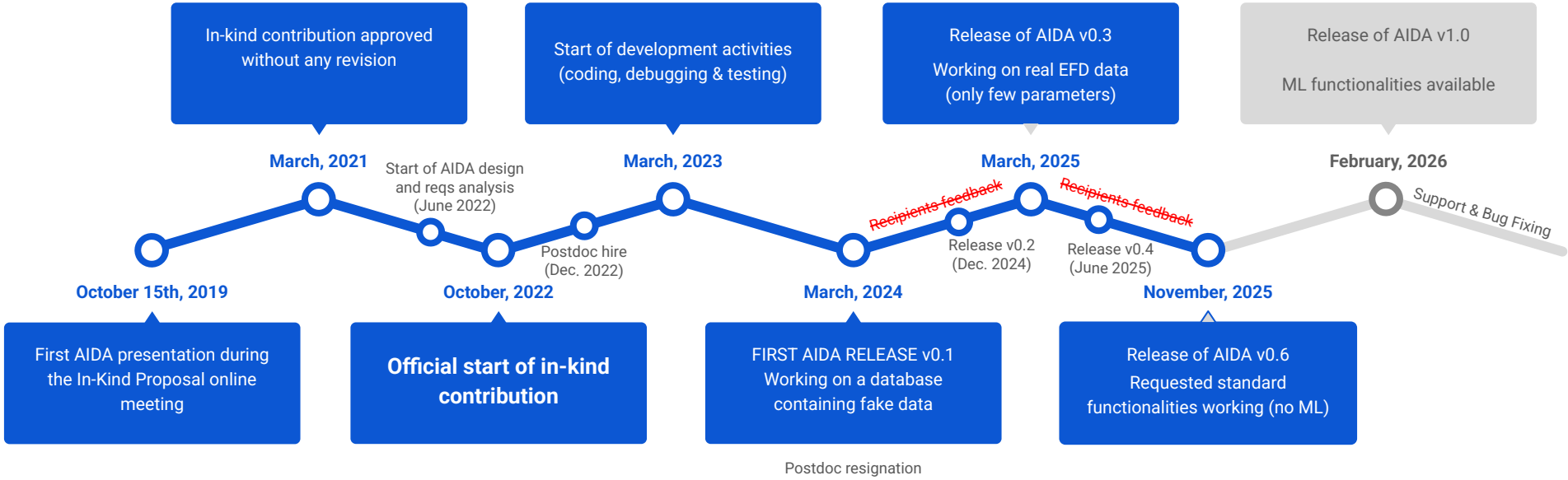
- LSST Commissioning
- Leanne Guy
- Andy W. Clements

Science Collaborations and Teams

- Rubin Commissioning Team
- Informatics and Statistics Science Collaboration (ISSC)

AIDA - In-kind contribution

"Directable software development effort in the general area of LSST focal plane instrument HouseKeeping/telemetry and science analysis, including machine/deep learning methods for prediction/classification tasks, as well as trend analysis and statistical characterization of focal plane instrument and science data/images systems."



AIDA - Goals

AIDA is a portable and modular web application, designed to provide an efficient and intuitive software infrastructure to support monitoring of data acquiring systems over time, diagnostics and both scientific and engineering data quality analysis, particularly suited for astronomical instruments

The AIDA design was mainly focused on the following goals:

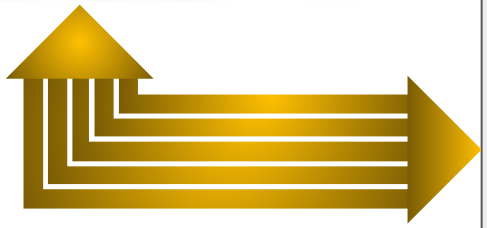
✔ **Daily mission operations**, for instruments health assessment and reporting:
- Monitoring and verifying the nominal quality of data;
- Production of user-defined reports (automatic and/or on-demand);

✔ **Short/Mid/Long term activities**, to investigate and report instruments behaviour over time:
- Analysis of instruments HouseKeeping/TeleMetry (HK/TM) trends;
- Analysis of instruments systematic effects;
- Detection and analysis of instrumental features, degradation and anomalies;

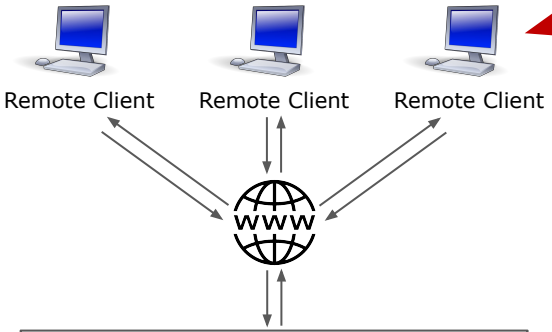
✔ **Data Analysis**:
- Basic and on-demand advanced statistics for data correlation and quality assessment;
- Machine/Deep Learning based classification/regression on local data.

AIDA, as in-kind contribution, has been designed and tailored for monitoring and analyzing **EFD (Engineering and Facilities Database)** LSST Commissioning data

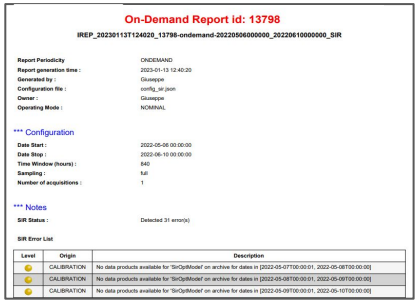
AIDA - Overall View



multi-thread/multi-process tasks both for data retrieval and analysis



AIDA - Main Features



On-Demand Report id: 13798

IREP_20230113T124020_13798-ondemand-20220506000000_20220610000000_SIR

Report Periodicity	ONDEMAND
Report generation time :	2023-01-13 12:40:20
Generated by :	Giuseppe
Configuration file :	config_sir.json
Owner :	Giuseppe
Operating Mode :	NOMINAL

*** Configuration

Date Start :	2022-05-06 00:00:00
Date Stop :	2022-06-10 00:00:00
Time Window (hours) :	840
Sampling :	full
Number of acquisitions :	1

*** Notes

SIR Status :	Detected 31 error(s)

SIR Error List

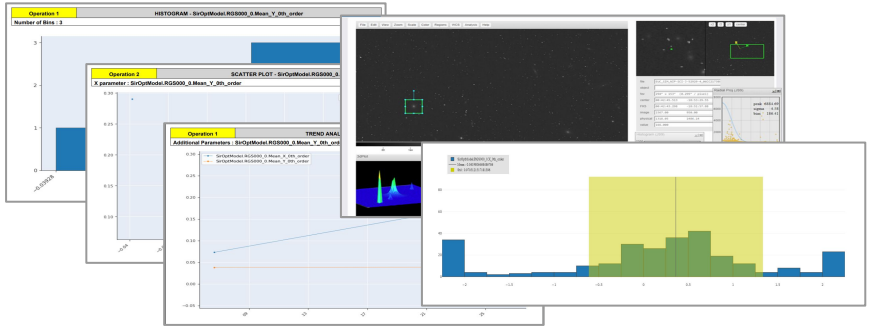
Level	Origin	Description
	CALIBRATION	No data products available for 'SirOptModel' on archive for dates in [2022-05-07T00:00:01, 2022-05-08T00:00:00]
	CALIBRATION	No data products available for 'SirOptModel' on archive for dates in [2022-05-08T00:00:01, 2022-05-09T00:00:00]
	CALIBRATION	No data products available for 'SirOptModel' on archive for dates in [2022-05-09T00:00:01, 2022-05-10T00:00:00]

Instrument monitoring, report generation and delivery

- ✓ periodic report generation on a user-defined parameters list and delivery to remote archive
- ✓ on demand customised report generation on a user selected parameter list, locally stored

Visualization/Exploration

- ✓ series of plots on user selected parameters/data products and ranges
- ✓ observed images (static view, dynamic windowing, statistical characterization)



Statistics

- ✓ standard (default) estimators
- ✓ special estimations (tables/images)
- ✓ statistical analysis on image pixels

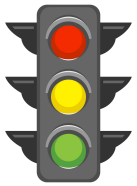
NISP-DARK.DET_12.DarkSubtraction_Diag.std	
Mean	103.97691889162952
Standard Deviation	55.174765923434988
Median	93.656876321575720
Min	26.925978812776677
Max	208.55603258218244
RMS	117.70919444504895
Variance	3044.2547947058424
Kurtosis	-1.4517662364055537
Skewness	0.13761575723724910
MAD	54.554519334551074
NMAD	80.881422289920050

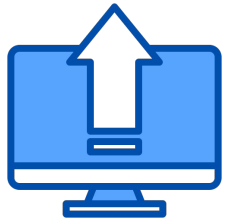
Machine Learning

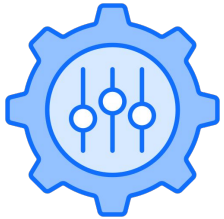
- ✓ Regression, classification and clustering experiments on available data

AIDA - Additional Features

- **Flagging System**
  - Users can associate a semaphore-like flag to each kind of experiment, to indicate its status and generate a related PDF report, summarizing its results and flags
- **Local Data Analysis**
  - Users can generate plots on data and visualize/analyze images uploaded from local machine
- **Users and System Management**
  - Online user registration to be confirmed by administrators
  - User password recovery system
  - Administrators can:
    - enable, disable or remove users
    - set Operating Mode (Nominal, Commissioning, Contingency)
    - enable/disable and configure systems to monitor
    - set web app configuration (SMTP server, number of processors to use…)
- **Logging System**
  - Each operation performed is logged into local DB
- **Easy Step-by-step Installation Procedure**
- **Customizable plots graphics**
- **Data backup system**

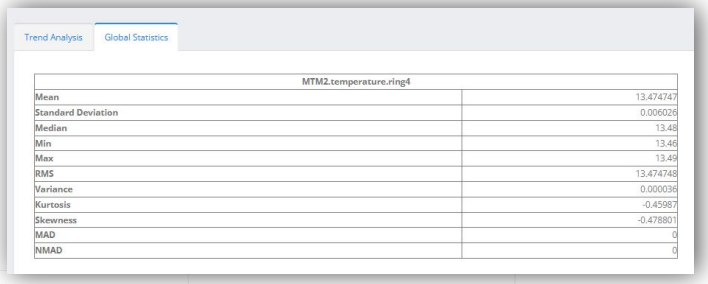


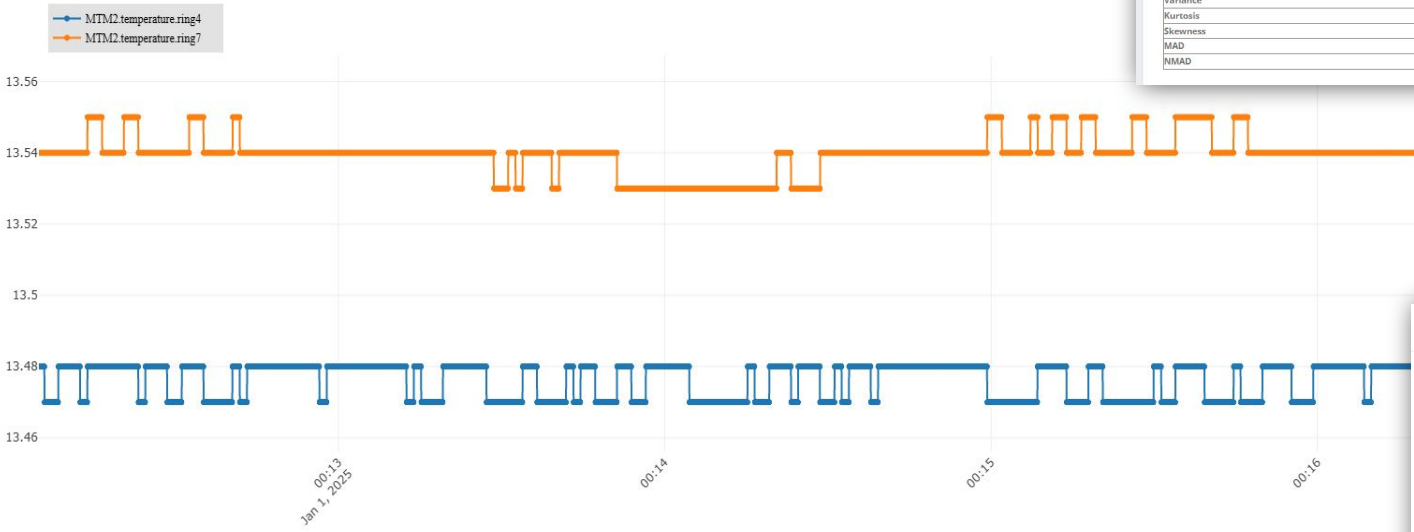


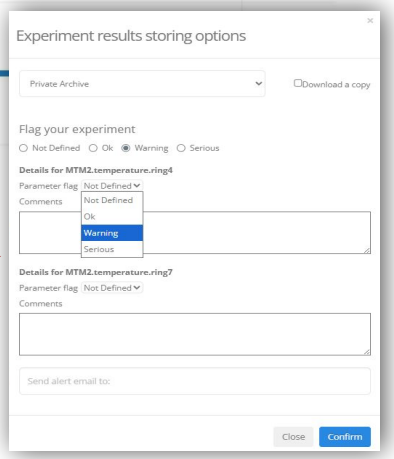


AIDA - Plots

All plots are interactive, navigable and customizable



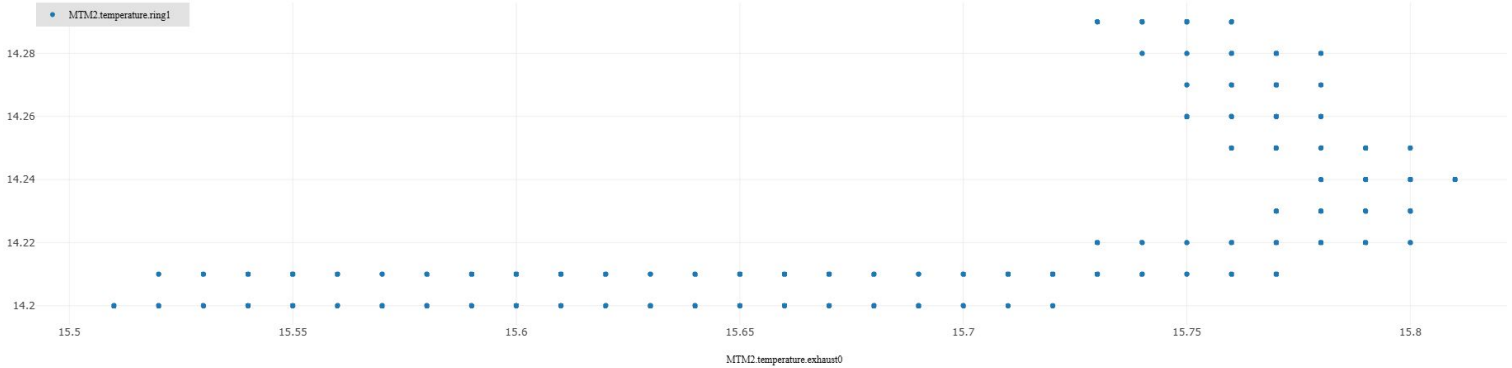




Flagging system example: form suitable to flag and dispatch any anomaly detected by the user during visual inspection of data.

AVAILABLE FOR ALL ANALYSIS TOOLS

AIDA - Plots



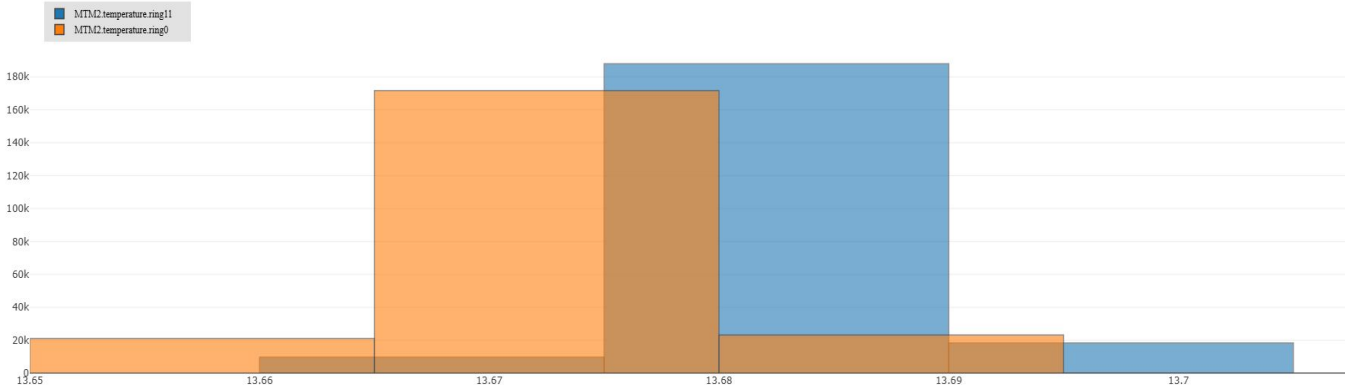
SCATTER PLOT

HISTOGRAM

Bins and counts can be generated from retrieved data

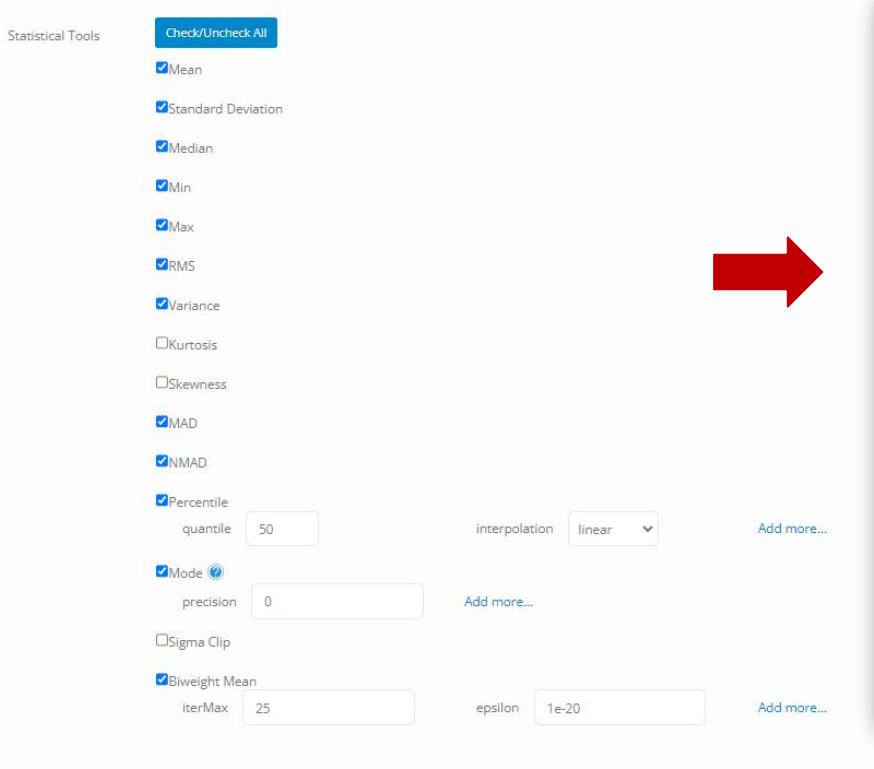
or

Directly retrieved and visualized from repository (if stored in)



AIDA - Statistics

An **expandable** collection of **base** (mean, median, RMS, σ, variance, min-max, MAD, NMAD, kurtosis, skewness) and **special** (mode, percentiles, biweight mean, σ-clipping) estimators

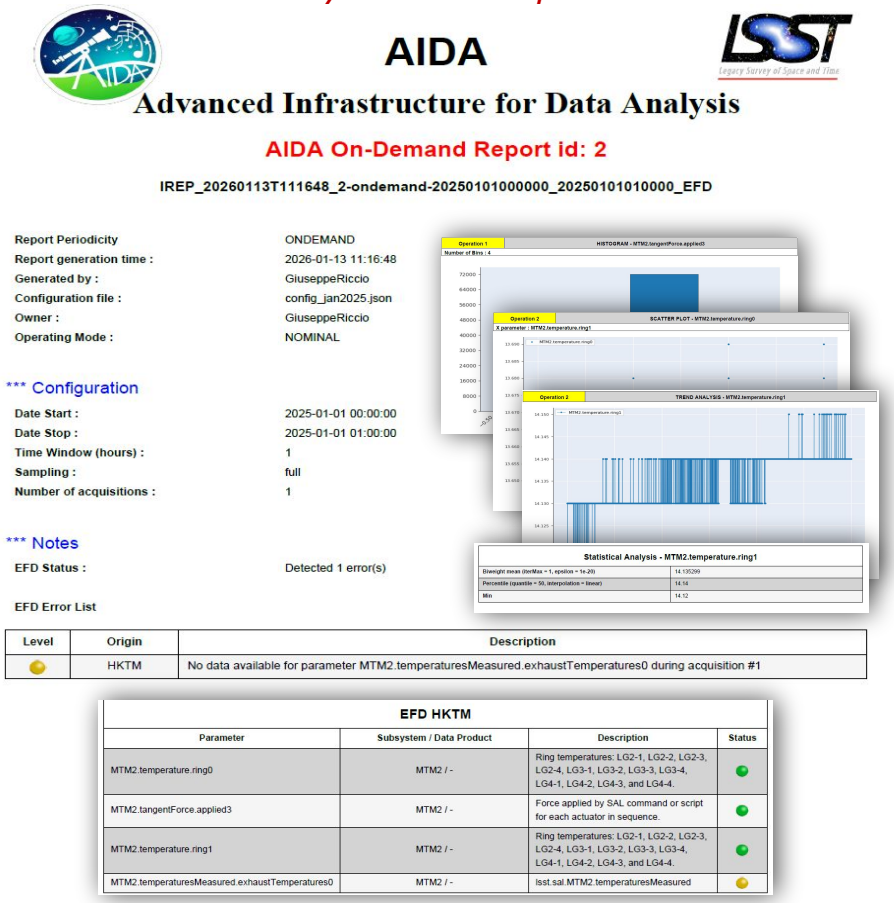


MTM2.temperature.ring0	
Mean	14.154394
StandardDeviation	0.032839
Median	14.15
Min	14.1
Max	14.21
RMS	14.154433
Variance	0.001078
MAD	0.03
NMAD	0.044477
Percentile quantile=50 interpolation=linear	14.15
Mode precision=0	Mode : 14.11 Counts : 11178
BiweightMean iterMax=25 epsilon=1e-20	14.154334

AIDA - Reports

Document containing a collection of analysis performed on a list of systems and parameters

- ✓ 2 kind of reports:
  - ➢ **Periodic** report automatically generated at a chosen frequency
  - ➢ **On demand** customised report generation
- ✓ Configuration through reusable **JSON files**, created by the user
- ✓ All available plots/statistics can be performed on each available system/parameter
- ✓ 2 reports format : **XML** and **PDF**
- ✓ Reports are stored on AIDA DB, but they can ingested **to remote repositories**
- ✓ Some basic analysis on data (for example, over thresholding, notification of missing data on repository…)
- ✓ Intensive use of multi-threading/multi-process
- ✓ Data to analyze are temporary stored into HDF5 files



AIDA

Advanced Infrastructure for Data Analysis

AIDA On-Demand Report id: 2

IREP_20260113T111648_2-ondemand-20250101000000_20250101010000_EFD

Report Periodicity	ONDEMAND
Report generation time :	2026-01-13 11:16:48
Generated by :	GiuseppeRiccio
Configuration file :	config_jan2025.json
Owner :	GiuseppeRiccio
Operating Mode :	NOMINAL

*** Configuration

Date Start :	2025-01-01 00:00:00
Date Stop :	2025-01-01 01:00:00
Time Window (hours) :	1
Sampling :	full
Number of acquisitions :	1

*** Notes

EFD Status : Detected 1 error(s)

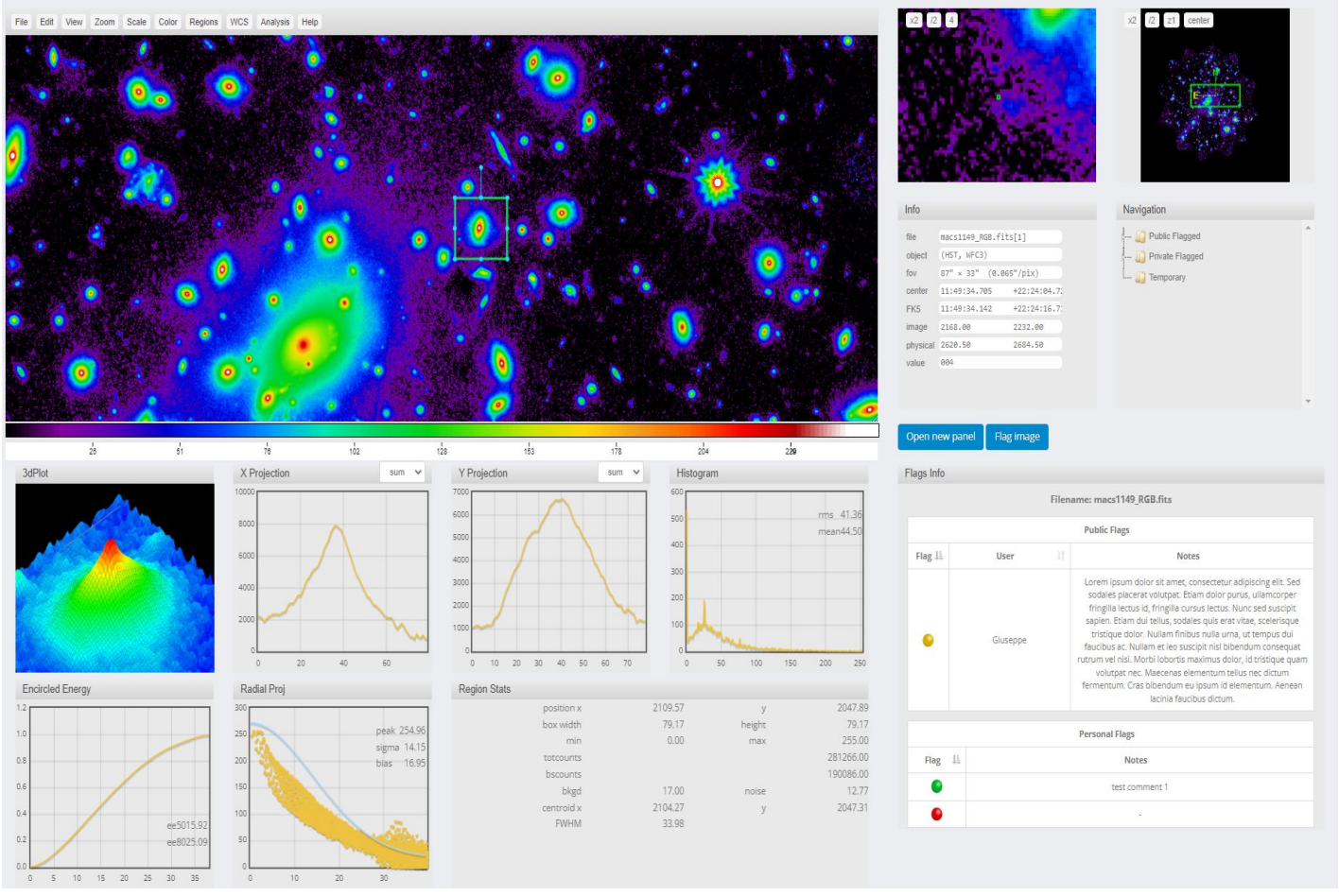
EFD Error List

Level	Origin	Description
	HKTM	No data available for parameter MTM2.temperaturesMeasured.exhaustTemperatures0 during acquisition #1

EFD HKTM

Parameter	Subsystem / Data Product	Description	Status
MTM2.temperature.ring0	MTM2 / -	Ring temperatures: LG2-1, LG2-2, LG2-3, LG2-4, LG3-1, LG3-2, LG3-3, LG3-4, LG4-1, LG4-2, LG4-3, and LG4-4.	
MTM2.tangentForce.applied3	MTM2 / -	Force applied by SAL command or script for each actuator in sequence.	
MTM2.temperature.ring1	MTM2 / -	Ring temperatures: LG2-1, LG2-2, LG2-3, LG2-4, LG3-1, LG3-2, LG3-3, LG3-4, LG4-1, LG4-2, LG4-3, and LG4-4.	
MTM2.temperaturesMeasured.exhaustTemperatures0	MTM2 / -	lsst.sal.MTM2.temperaturesMeasured	

AIDA - Image Visualization

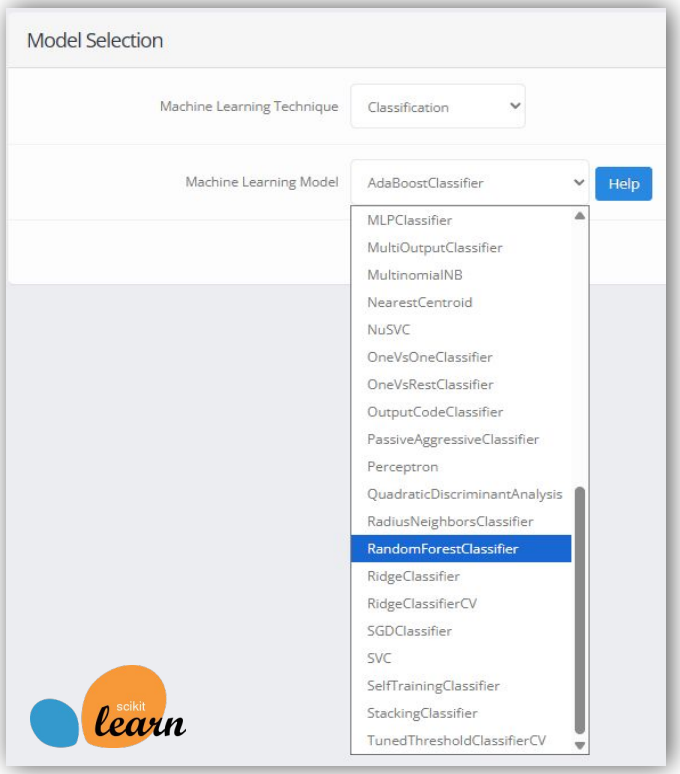


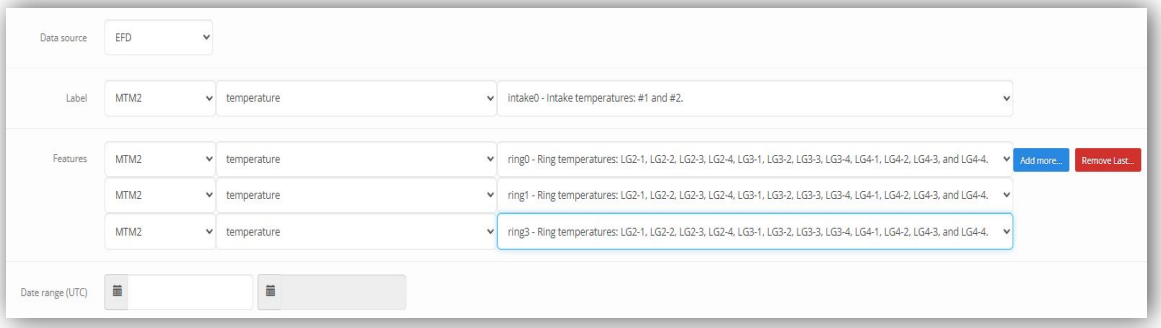
AIDA is able to support the direct display and dynamical navigation of **FITS/PNG/JPG images**, multiple frame buffers, region cursor manipulation, thumbnail extraction, many scale algorithms and colour maps.

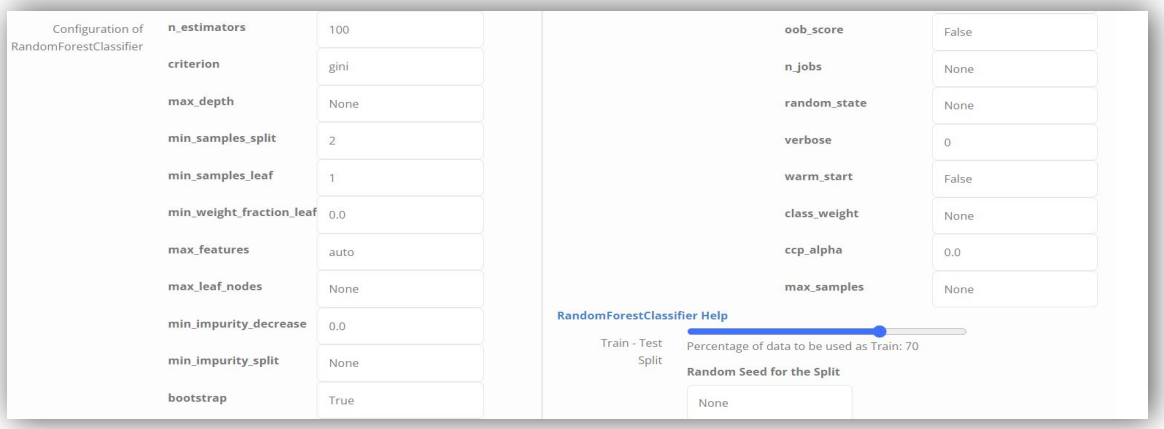
Based on JS9 Library :
https://js9.si.edu/

AIDA - Machine Learning tools

AIDA includes **more than 100 classification, clustering and regression models** based on **Machine Learning** to apply on available tabular data, useful in this case to identify operating anomalies or correlations between instrumental information







Code Name : Flexibility

AIDA has been designed as a **modular system**, based on **Object-Oriented Programming** and specific information on DB, so it is possible to **extend its functionalities**, by integrating and customizing monitoring and diagnostics systems, as well as scientific data analysis solutions, including machine/deep learning and data mining methods

- **Available plots and statistics are defined as classes/functions** linked to a specific table in DB. To add a new operation, it is sufficient to implement the related class/function and add it to the local DB;
- **A JSON configuration file is associated to every system monitored by AIDA**. It includes info about the instrument and connection to the related data and metadata archives;
- **Every system has a dedicated class** which implements methods for interfacing AIDA with the data repository. To add a new system, it is sufficient to create its own configuration file, implement the related class and methods, and fill DB with required information.

AIDA, even if designed for monitoring and analyzing EFD (Engineering and Facilities Database) commissioning data, can be easily suited **for any kind of survey and engineering data**

Conclusions

- ✔ The **AIDA web application** has been designed to provide an **efficient and intuitive software infrastructure** to support **monitoring** of data acquisition systems over time, **diagnostics** and both scientific and engineering **data quality analysis**, in particular for astronomical instruments
- ✔ It provides **a number of tools** for data analysis & system diagnostics, including plot generation, statistics, image viewer and Machine Learning module for classification, regression and clustering of data retrieved from official databases
- ✔ AIDA is **very easy to use** and does not require any installation on client machine, being based on Docker technology
- ✔ Being designed as a modular system, **it is possible to integrate and customize** monitoring and diagnostics systems, as well as scientific data analysis solutions (for example, a specific version of AIDA is the **official monitoring** and analysis tool for the **ESA Euclid space mission**)

CRITICAL ISSUES

AIDA has been designed for monitoring and analyzing **EFD** data during the **Vera Rubin/LSST Commissioning phase** but the development experienced a slowdown due to very few feedback from the recipients, difficulties in recruiting and early resignation of the recruited person

AIDA Code Repository

https://github.com/pepric/aida

Conferences & Publications:

Astronomical Data Analysis Software and Systems XXX (ADASS XXX). ASP Conference Series, Vol. 532, Proceedings of a virtual conference held 8-12 November 2020. Edited by Jose Enrique Ruiz, Francesco Pierfedereci, and Peter Teuben. San Francisco: Astronomical Society of the Pacific, 2022., p.459, https://aspbooks.org/custom/publications/paper/532-0459.html

Astroinformatics 2023, VIDEOMemorie of the Italian Astronomical Society, Astroinformatics 2023, Vol. 4, id. 45: https://www.memsait.it/videomemorie/volume-4-2023/VIDEOMEM_4.2023.45.mp4

109° Congresso Nazionale SIF (2023), 11-15 September 2023, https://2023.congresso.sif.it/talk/366

Challenges and Innovations in Computational Astrophysics (ChaICA) II, 18-21 November 2020, presentation

AIDA paper to submit to Astronomy & Computing **in preparation**

THANK YOU