Astronomical Data Analysis Software and Systems XXX
ASP Conference Series, Vol. 532
J. E. Ruiz, F. Pierfederici, and P. Teuben, eds.

AIDA, a Modular Web Application for Astronomical Data Analysis and Instrument Monitoring Services

Giuseppe Riccio, Stefano Cavuoti, and Massimo Brescia

INAF - Astronomical Observatory of Capodimonte, Salita Moiariello 16, I-80131 Napoli, Italy; giuseppe.riccio@inaf.it

Abstract. In the last decade, Astronomy has been the scene of the realization of panchromatic surveys, with sophisticated instruments acquiring a huge quantity of exceptional quality data. This poses the need to integrate advanced data-driven science methodologies for the automatic exploration of huge data archives, and the need for efficient short- and long-term monitoring and diagnostics systems. The goal is to keep the quality of the observations under control and to detect and circumscribe anomalies and malfunctions, facilitating rapid and effective corrections, ensuring correct maintenance of all components and the good health of scientific data over time. In particular, this requirement is crucial for space-borne observation systems, both in logistical and economic terms. AIDA (Advanced Infrastructure for Data Analysis) 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. Given its modular system prerogative, 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 techniques and methods. A specialized version of AIDA has been recently appointed as focal plane instrument operation diagnostics, analytics and monitoring service within the Science Ground Segment of the Euclid space mission.

1. Introduction

In the last decade, Astronomy has been the scene of an unprecedented scientific and technological evolution, culminating in recent confirmations of gravitational waves and in the observation of a black hole, giving life to the multi-messenger astrophysics, to underline the achievement of the possibility to observe and study the sky from multiple points of view. That period laid the foundations for the realization of panchromatic surveys, with ground-based and space-borne instruments, characterized by a wide field of view, combined with a very high spatial resolution, capable of acquiring a huge quantity of exceptional quality and deep data. Gaia (Gaia Collaboration et al. 2016), Euclid (Laureijs et al. 2010), the Legacy Survey of Space and Time (LSST) at Vera C. Rubin Observatory (Ivezić et al. 2019), the Extremely Large Telescope (ELT, Neichel et al. 2018), the James Webb Space Telescope (JWST, Gardner et al. 2006), the Square Kilometer Array (SKA, Dewdney et al. 2009) are just a few examples of observational projects, already active or in progress, capable of providing full volumes of data that can be considered as Big Data systems. Such tools pose two sets of problems: on the one hand, the huge amount of data poses the need to integrate advanced data-driven science methodologies for analysis and automatic exploration of data archives, exploiting resources and solutions proposed by Astroinformatics. On the other, the technological complexity of the focal plane instruments of these observational projects, highlights the need for efficient short- and long-term monitoring and diagnostics systems, with the possibility of keeping the quality of the observations under control, and of being able to reveal and circumscribe anomalies and malfunctions, facilitating rapid and

effective corrections and ensuring correct maintenance of the electro-mechanical, software and optical devices over time. In particular, this requirement is vital for spatial observation tools, both in logistical and economic terms.

AIDA (Advanced Infrastructure for Data Analysis) is a portable and modular web application, which was created with the aim at providing an efficient and intuitive software system to support monitoring over time, diagnostics and data analysis of a generic observation system, possibly composed of multiple focal plane instruments. Given its modular system prerogative, it is possible to extend its functionality, by integrating and customizing HouseKeeping (HK), TeleMetry (TM) monitoring and diagnostics systems, as well as scientific data analysis solutions, which can be based on Machine/Deep Learning (MDL) and data mining techniques and methods.

2. AIDA Overview

AIDA is primarily a portable, multi-task and multi-user web application, specialized for tabular/image data analysis and monitoring, equipped with a local data repository and organized as a series of software modules specialized on several multi-process and multi-thread tasks, which require a high-speed and reliable network connection with data archives. It can be executed via web browser on a wide number of devices, including smartphones and tablets, since the front-end has a responsive design based on HTML5[1] (HyperText Markup Language), CSS3[2] (Cascading Style Sheets), Bootstrap[3] 5 and Javascript-based[4] scripting, thus ensuring a high level of compliance with the most common web platforms. The back-end environment, responsible for local configuration, communication with data repositories and analysis is based on PHP7[5] (PHP: Hypertext Preprocessor) and Python 3[6]: PHP is used to dynamically populate the graphical user interface with instrument data (for instance, to list the available systems in drop-down menus for plots configuration), while Python is dedicated to data retrieval and analysis modules. Finally, AIDA uses AJAX[7] (Asynchronous JavaScript and XML) technology to interface the front-end web client with the back-end server.

The internal architecture of AIDA is based on a local MySQL[8] database (DB) system, which includes various kinds of information either persistent or volatile. In order to reduce the size of the local hard drive system, all data remotely collected during the working time and used to perform analysis/monitoring tasks, are deleted at the end of any involved process and/or after the user logout. Only the data products generated by the use of the application is persistently stored and maintained in the local archive.

As web application it does not require any client-side installation and can be fully executed via any web browser through a network connection. In order to preserve the safety of remote data, used by the application, the access to AIDA is reserved to registered users, after having been authenticated by the AIDA administrators, through a standard login/password mechanism. Only administrators have the rights to add/remove active users and to control all the operations performed with AIDA. But any authorized user has full access to all AIDA services, including

---

[1] `https://www.w3.org/TR/2012/CR-html5-20121217`

[2] `https://www.w3.org/Style/CSS`

[3] `https://getbootstrap.com`

[4] `https://www.ecma-international.org/publications/standards/Ecma-262.htm`

[5] `https://www.php.net`

[6] `https://www.python.org/download/releases/3.0`

[7] `https://www.scriptol.fr/ajax/ajax_adaptive_path.pdf`

[8] `https://www.mysql.com`

the possibility to download and analyze all instrument HKTM and Scientific TeleMetry (STM) data, according to the instruments and data archive permission policy assigned to AIDA by the related project managers.

AIDA exposes four main functionalities:

- *generation of data analysis/monitoring reports* (periodic or on user demand), pre-configured offline or through interactive dedicated forms, by collecting the results of a given number of operations for short/mid/long-term data monitoring/analysis;
- *visualization/exploration of tabular/image type data*, by a series of dynamically navigable plots or a built-in image explorer module;
- *statistical analysis of data*, through standard or advanced estimators;
- *machine/deep learning analysis*, by means of a wide set of classification, regression, clustering and dimensional reduction models and built-in modules for pre- and post-processing of tabular/image type data.

In order to be easily tailored to a specific instrument/project and to be extended with new graphical or analytics tools, AIDA has been designed as a modular system, based on Object-Oriented Programming, easily adaptable to any specific data archive system. The formats adopted for internal data products representation are JSON[9] (JavaScript Object Notation), for configuration files and XML[10] (eXtensible Markup Language) for plots and reports:

- available plots and statistics are defined as classes/functions linked to a specific table in the AIDA DB. To add a new operation, it is sufficient to implement the related class/function and register it within the local DB;
- a JSON configuration file is associated to every system monitored by AIDA, including the list of parameters to handle and the connection protocol with the remote archive. To add a new system, it is sufficient to generate and register its own configuration file in the local DB, and to implement the specific class and methods to interface AIDA with system data repository;
- each generated report is locally stored in both PDF and XML formats. Since, as said above, all data downloaded for analysis tasks are deleted from the local DB at the end of the user session/experiment, the XML version is particularly suitable to permit any future re-generation of same analysis/plots output without repeating the time/computing expensive download of input data from the remote archive and the related processing.

Most of AIDA services and functionalities are based on an intensive use of multi-processing and multi-threading mechanisms. Furthermore, several MDL models are developed to exploit the parallel programming paradigm through embedded modules interfacing with the GPUs (Graphical Processing Units) available on the AIDA hosting machine. It is hence evident that the AIDA computing performances strongly depend on the specific architecture of the hosting server. The statistical functions, standard and advanced types, exposed by AIDA are based on an already validated built-in library, dedicated to the scientific analysis of data related to the Science Ground Segment of the ESA Euclid Space Mission (Brescia et al. 2019). In particular, a customized version of AIDA, for instance IODA (Instrument Operation Data Analysis), has been recently officially appointed as the Instrument Operation monitoring and analysis tool for the Euclid mission.

Acknowledgments. The authors acknowledge financial contributions from the agreement *ASI/INAF 2018-23-HH.0, Euclid ESA mission - Phase D*

[9] `https://www.json.org`

[10] `https://www.w3.org/XML`

References

Brescia, M., et al. 2019, in ADASS XXVI, edited by M. Molinaro, K. Shortridge, & F. Pasian, vol. 521 of ASP Conf. Ser., 228

Dewdney, P. E., Hall, P. J., Schilizzi, R. T., & Lazio, T. J. L. W. 2009, IEEE Proceedings, 97, 1482

Gaia Collaboration, et al. 2016, A&A, 595, A1. 1609.04153

Gardner, J. P., et al. 2006, Space Sci.Rev., 123, 485. astro-ph/0606175

Ivezić, Ž., et al. 2019, ApJ, 873, 111. 0805.2366

Laureijs, R. J., Duvet, L., Escudero Sanz, I., Gondoin, P., Lumb, D. H., Oosterbroek, T., & Saavedra Criado, G. 2010, in Space Telescopes and Instrumentation 2010: Optical, Infrared, and Millimeter Wave, vol. 7731 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 77311H

Neichel, B., Mouillet, D., Gendron, E., Correia, C., Sauvage, J. F., & Fusco, T. 2018, in SF2A-2018: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, Di. 1812.06639