DEAVI: Dynamic Evolution Added Value Interface

Deborah Baines,[1] Ignacio de la Calle,[1] Jose Maria Herrera-Fernandez[1], Aitor Ibarra[1], Jesus Salgado[1], Luis Valero-Martin[1]
[1]*Quasar Science Resources S. L., Edificio Ceudas, Ctra. de La Coruna Km 22.300, 28232, Las Rozas de Madrid, Madrid, Spain;*
deborahbaines@quasarsr.com

Abstract. We present DEAVI, an Added Value Interface (AVI) to manage and exploit data from the ESA missions Gaia and Herschel. AVIs are software packages that provide scientists with the mechanisms to submit their own code to be executed close to the ESA mission archives. GAIA AVIs are deployed at the Gaia Added Value Interface Platform (GAVIP), a Python-based platform designed and developed by ESA and hosted at the European Space Astronomy Centre (ESAC). The proposed AVI is part of the software package being developed by Quasar Science Resources for the StarFormMapper (SFM): A Gaia and Herschel Study of the Density Distribution and Evolution of Young Massive Star Clusters project, funded by the European Union under the Horizon 2020 programme.

1. Introduction

The European Space Agency (ESA) operates numerous missions, both operational and scientific. Two of ESA's scientific missions are Gaia[1] and Herschel[2]. The main objective of Gaia is to obtain a three-dimensional map of the Milky Way, i.e. the positions of the stars of our galaxy and their radial and positional velocity measurements. In the case of Herschel, one of its main goals was to study the formation and evolution of stars and galaxies and their interaction with the interstellar medium. Both missions can be considered a success given the number of discoveries they have made. However, combined, they can give a more complete pitcure of star formation. The StarFormMapper (SFM)[3] project is a European H2020 RIA project that proposes a combined study of the data of both missions covering all stages of star formation, from the formation of molecular cores to the dispersion of gas in young clusters. The SFM consortium is a collaboration between the Universities of Leeds (UK, Coordinator), Cardiff (UK) and Joseph Fourier, Grenoble (FR), as well as the Spanish company Quasar Science Resources, S.L. (QSR).

[1] http://sci.esa.int/gaia/

[2] http://sci.esa.int/herschel/

[3] https://starformmapper.org/

QSR[4] is a private company that provides consulting Software and System Engineering services for Research and Development projects. The team includes Computer System Analysts, Software and Data Archive Engineers and Scientists. In this project, QSR is developing the necessary software tools in order to handle the scientific algorithms for the analysis of the combined Gaia, Herschel and other data of young star clusters, including the visualisation of the results. These tools have been collected in the first version of the Dynamic Evolution Added Value Interface (DEAVI) software which is presented in this work.

2. Dynamic Evolution Added Value Interface

The DEAVI developed by QSR consists of a virtual infrastructure that can store and run algorithms, as well as visualise data in 2 and 3 dimensions. In addition, the possibility of data exchange between the client and the server has been incorporated through the implementation of a Simple Application Messaging Protocol (SAMP; Taylor et al. 2015) interface.

The infrastructure has been developed in a Docker container (Merkel 2014), which is a structure that wraps a software module containing all the needed elements to invoke and run the software. Docker containers can be invoked using input parameters and can be integrated in virtualised environments such as Amazon Web Services, Puppet, VMware, etc. Figure 1 shows a simple scheme of the virtual architecture where the main components are highlighted, and are described below.

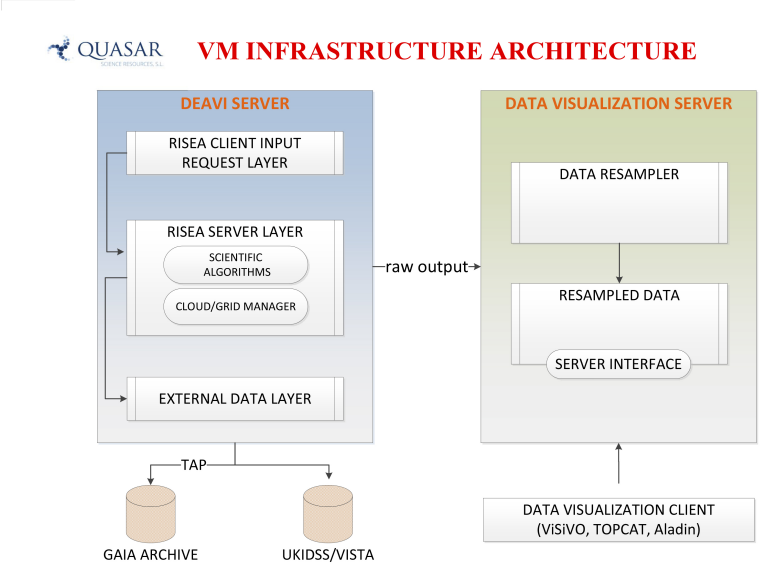

Figure 1. Virtual Server architecture design for the SFM project showing the different subsystems. RISEA server/client stands for Remote Interface to the Stellar Environment Algorithm, and is the main component of the system. TAP stands for Table Access Protocol to access archive data.

[4] http://www.quasarsr.com/

2.1. DEAVI server

The DEAVI server has three different layers:

- The first layer is the Remote Interface to the Stellar Environment Algorithm (RISEA) Interface. This is a client interface to access the scientific algorithms, either in production or in development, as designed by the scientific team within the consortium. The Client allows adding, modifying and implementing physical conditions, input parameters and output data.

- The second layer is the RISEA Server which runs on the virtualised infrastructure and is used to handle and inject different algorithms.

- The third layer is the Data Access Interface. It consists of a set of data access mechanisms allowing access to the Gaia and Herschel data, as well as to the auxiliary data used by the algorithms. Herschel data is available for download using the Archive Interoperability interface (HAIO) from the ESA Science Archives. Access to the Gaia catalogue is available using the IVOA Tabular Access Protocol (TAP; Dowler et al. 2011) interface implementation.

2.2. Data visualisation server

The Data visualisation server is the other important part of the virtual architecture. It consists of two components: a Data Resampler that resamples the data on the server side to allow client visualisation, and a Data Visualisation Client that allows the exploration of the results in 3D. Thanks to this part of the software, scientists can visualise data in the client and interact with them using the Bokeh, Astropy and D3.js libraries. As an example, figure 2 shows a simple algorithm that obtains, from the Gaia catalogue, the positions of a number of stars in a cluster and plots them together with arrows, representing their proper motions. In this prototype, the graphical user interface implements the following sections:

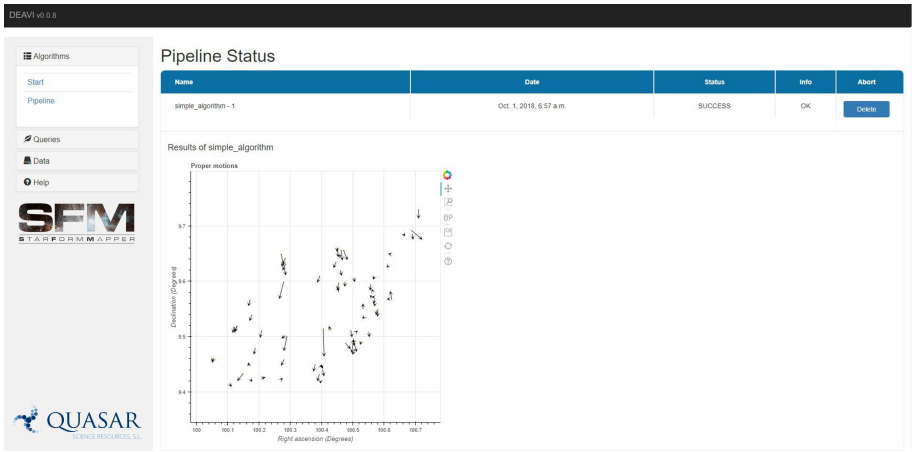

Figure 2. Example of a simple algorithm in DEAVI: star positions with their proper motions. The libraries used to create the plot are Astropy and Bokeh.

- Algorithms. This section is dedicated to the execution of the algorithms developed by the scientific team of the consortium. Within this section it is possible to choose and execute an algorithm, visualise its status and the results.
- Queries. To speed up the process of making queries to the scientific archives a friendly interface has been implemented to perform queries on GAIA and Herschel data. The results of these queries can then be fed as input to the available algorithms.
- Data. This section is where the different stored user data is displayed.
- Help. This last section is dedicated to providing the necessary information to the user about the use and functionalities of DEAVI.

3. Deployment of DEAVI in GAVIP

The concept of exploitation platforms has gained prominence over the past few years. The main purpose of these platforms is to offer the possibility of executing algorithms close to the data when the data to be explored is massive. An example of these platforms is the Gaia Added Value Interface Platform (GAVIP; Vagg et al. 2016). GAVIP is a Python-based platform that allows the global scientific community to run scientific code. It is installed at the European Space Astronomy Centre (ESAC) where the data of the missions Gaia and Herschel are stored. Taking into account both factors, we have decided to deploy our software in GAVIP. This adds two important advantages: a) allows data processing without moving the mission data through the network; b) more computing power thanks to the use of the ESAC infrastructure (RAM, HDD, etc.).

4. Conclusions

The StarFormMapper is a project funded by the Horizon 2020 program of the European Union for the study of massive star and star cluster formation. SFM combines data from two of ESA's major space missions, Gaia and Herschel. Quasar Science Resources has developed a value-added interface capable of working simultaneously with both data sets. This interface allows scientists to add new algorithms and visualise original data and results in a friendly and intuitive way. The first version of the software has been deployed on the GAVIP platform at ESAC. Since the data of the Gaia and Herschel missions are also at ESAC, the movement of data through the network is avoided.

Acknowledgments. This work was supported by the SFM project which receives funding from the European Union's Horizon 2020 Research and Innovation Action (RIA) programme under Grant Agreement No 687528.

References

Dowler, P., Rixon, G., & Tody, D. 2011, ArXiv e-prints. 1110.0497
Merkel, D. 2014, Linux J.
Taylor, M. B., Boch, T., & Taylor, J. 2015, ArXiv e-prints. 1501.01139
Vagg, D., O'Callaghan, D., McBreen, S., Hanlon, L., Lynn, D., & O'Mullane, W. 2016, ArXiv e-prints. 1605.09287