Asterics
Astronomy ESFRI & Research Infrastructure Cluster

# ASTERICS - H2020 - 653477

# Third main thematic training event

## ASTERICS GA DELIVERABLE: D3.12

| | |
|---|---|
| Document identifier: | ASTERICS-D3.12-final.docx |
| Date: | **19 July 2018** |
| Work Package: | **WP3 OBELICS** |
| Lead Partner: | **LAPP** |
| Document Status: | **Report** |
| Dissemination level: | Public |
| Document Link: | www.asterics2020.eu/documents/ASTERICS-D3.12.pdf |

Abstract

This report gives an overview of the third main thematic event organized by WP3-OBELICS entitled "Second ASTERICS-OBELICS International School on Advanced software programming for astrophysics and astroparticle physics". This event was the deliverable D3.12 under the grant agreement. This report discusses the overall outcome and feedback collected from students, tutors and organizers that will be useful to organize similar event in future.

# I. COPYRIGHT NOTICE



# II. DELIVERY SLIP

| | Name | Partner/WP | Date |
|---|---|---|---|
| From | Jayesh Wagh | CNRS-LAPP | 23/07/2018 |
| Author(s) | Jayesh Wagh | CNRS-LAPP | 23/07/2018 |
| Reviewed by | G. Cimo' | ASTRON / JIVE | 31/07/2018 |
| Approved by | R. van der Meer | ASTRON | 06/08/2018 |

# III. DOCUMENT LOG

| Issue | Date | Comment | Author/Partner |
|---|---|---|---|
| 1 | 25/07/2018 | First draft | Jayesh Wagh, LAPP |
| 2 | 06/08/2018 | Second draft | Vincent Poireau, LAPP |

# IV. APPLICATION AREA

This document is a formal deliverable for the GA of the project, applicable to all members of the ASTERICS project, beneficiaries and third parties, as well as its collaborating projects.

# V. PROJECT SUMMARY

ASTERICS (Astronomy ESFRI & Research Infrastructure Cluster) aims to address the cross-cutting synergies and common challenges shared by the various Astronomy ESFRI facilities (SKA, CTA, KM3Net & ELT). It brings together for the first time, the astronomy, astrophysics and particle astrophysics communities, in addition to other related research infrastructures. The major objectives of ASTERICS are to support and accelerate the implementation of the ESFRI telescopes, to enhance their performance beyond the current state-of-the-art, and to see them interoperate as an integrated, multi-wavelength and multi-messenger facility. An important focal point is the management, processing and scientific exploitation of the huge datasets the ESFRI facilities will generate. ASTERICS will seek solutions to these problems outside of the traditional channels by directly engaging and collaborating with industry and specialised SMEs. The various ESFRI pathfinders and precursors will present the perfect proving ground for new methodologies and prototype systems. In addition, ASTERICS will enable astronomers from across the member states to have broad access to the reduced data products of the ESFRI telescopes via a seamless interface to the Virtual Observatory framework. This will massively increase the scientific impact of the telescopes, and greatly encourage use (and re-use) of the data in new and novel ways, typically not foreseen in the original proposals. By demonstrating cross-facility synchronicity, and by harmonising various policy aspects, ASTERICS will realise a distributed and interoperable approach that ushers in a new multi-messenger era for astronomy. Through an active dissemination programme, including direct engagement with all relevant stakeholders, and via the development of citizen scientist mass participation experiments, ASTERICS has the ambition to be a flagship for the scientific, industrial and societal impact ESFRI projects can deliver.

# VI. EXECUTIVE SUMMARY

The second ASTERICS-OBELICS International School on Advanced software programming for astrophysics and astroparticle physics was organized from 3-8 June 2018 in Annecy, France. This school was organized following the second ASTERICS-OBELICS workshop, where the need to introduce machine learning techniques in the field of astronomy and astrophysics was discussed to address the big data challenges. This school was organised and hosted by the CNRS LAPP (Laboratoire d'Annecy de Physique de Particules) laboratory. This five-day training event brought together over 80 (60 students, 13 tutors and 15 organizers) PhD students, postdocs, senior researchers from the domain of astrophysics and astroparticle physics from renowned research institutes in Europe as well as Africa. The school provided theoretical and hands-on training to acquire efficient and fast computer programming techniques, as well as introduced participants to Machine Learning. Thanks to the tutors, the school successfully addressed the programming needs of the astrophysics community.

# VII. Table of contents

# 1. Introduction

During the 2nd ASTERICS-OBELICS workshop in 2017 (described in D3.7), the need and applications of machine learning in astrophysics and astronomy were extensively discussed. Traditional methods to model and analyse large volume of complex datasets or big data generated from astronomy ESFRI projects have to be replaced with more advanced and powerful techniques such as Machine Learning. Machine Learning is yet to be adopted by the astronomy community. Hence it is important to introduce and train the researchers on these new techniques and understand their advantages over traditional methods.

Following this feedback, the third thematic WP3 training event was dedicated to advanced software programming for astrophysics and astroparticle physics with focus on machine learning. Many researchers from astronomy ESFRI projects were consulted to identify the list of exercises for theory and hands-on sessions to be included in the school programme.

Considering the feedback from the first international school in 2017, C++ and Python were selected as the programming languages for the hands-on sessions. However two months before the event, the school tutors for C++ modules had to withdraw their participation and hence the school was then focused on Python only. Registered participants were informed immediately about this change in the programme.

The school announcement was widely disseminated to H2020-ASTERICS members, ESFRI projects as well as European astronomical societies in the month of February 2018 and March 2018. Task 3.1 MAUD members along with the LAPP staff managed the organizational arrangements for the school participants. As for the registrations, we attained maximum number of registration requests by the registration deadline.

The school was organized and hosted by the Laboratoire d'Annecy de Physique de Particules in Annecy, a renowned touristic location in France. This five-day training event brought together over over 80 (60 students, 13 tutors and 15 organizers) PhD students, postdocs, senior researchers from the domain of astrophysics and astroparticle physics from renowned research institutes in Europe as well as Africa.

The format of the school included theory sessions as well as hands-on sessions so that participants with diverse background could obtain basic theoretical knowledge to participate in the hands-on sessions.

Post-event feedback from the participants has convinced us about the success of the event in terms of organization as well as scientific programme. There is still some room for further improvement in the scientific programme and organization and we hope that this report will

be useful not only for us for other members of the astronomy and astrophysics community organize similar event in the future.

# 2. Organizational Approach

We used feedback from last year's participants as well as a survey form sent to some of the representatives of ESFRI project within OBELICS and to the tutors of the first edition of the school. These inputs were used to further narrow down the potential scientific modules and define the duration of the school. While optimizing the scientific modules, simultaneously a local organization committee comprising of staff members from LAPP and OBELICS members at LAPP was formed to advance the organizational arrangements to host 80 participants.

## 2.1 Venue & Duration

The event was organized by CNRS-LAPP from 3-8 June 2018 in Annecy, a picturesque, and usually sunny, location in the month of June. The first week of June was selected as most of the accommodation arrangements options were already booked for the rest of the month due to other events happening in the city.

## 2.2 Dissemination

For dissemination purpose, we relied on email communications as well as social media dissemination. The first email communication was sent out in the month of February 2018. We requested ASTERICS members to disseminate the school announcement within their respective institutions, ESFRI projects as well as contacts in the national astronomical societies. As we attained maximum number of participants/students (60), we believe that the dissemination was efficient.

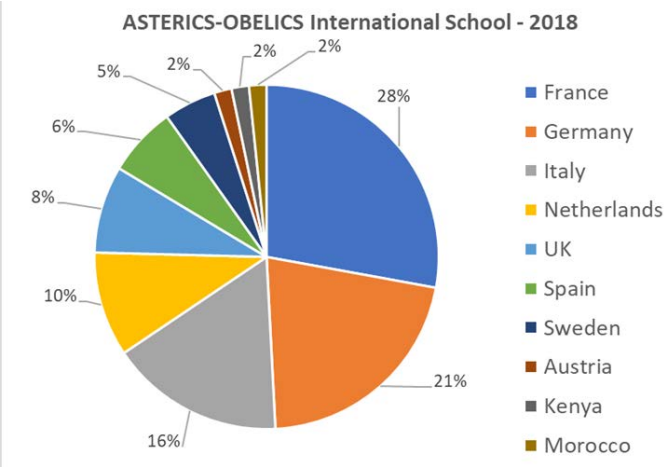


*Figure 1: Registrants for the 2nd ASTERICS-OBELICS International School, Annecy, June 2018*

## 2.3 Registration

We opened the school registration in the Month of February 2018. Interested PhD students and researchers were able to confirm their participation through online registration and payment of €150 of registration fee per participant on CNRS- Azure Colloque portal. The comprehensive registration fee covered all the following charges during the school

- Course fees (all modules): 3-9 June 2018.
- Registration kits.
- Shared accommodation (6 nights) at Centre Jean 23 guesthouse.
- Lunch: 3-9 June 2018.
- Coffee breaks: 3-9 June 2018.
- Social Events

We provided scholarships in the form of full registration fee waiver to 11 meritorious applicants from Europe and Africa. These applicants were selected by the OBELICS members at LAPP based upon academic excellence, need for financial support, importance of school modules in their respective research fields and quality of CV as well as letter of motivation.

# 3. Scientific Programme

The school programme can be found on the school website:

https://indico.in2p3.fr/event/16864/page/1782-scientific-programme

It was based on following five modules.

- Efficient code design
- project management
- Python libraries for science
- Python libraries for astrophysics and astronomy
- Machine learning & statistics: connecting scientific data with probabilistic models.

Python was the language of reference during the school. All the five modules included a combination of theory and hands-on sessions. Following the feedback from last year, we tried to improve on duration for hands-on sessions and limit the duration of theory sessions as much as possible. To facilitate more tutor-student interaction, the participants were divided in two groups for hands-on sessions. In addition, we also merged theory and hands-on sessions together providing more flexibility to the tutors during the school. All the presentations and course material were made publicly available on the school website and it is free to download. In addition, an online chat channels on slack were opened in case the students wished to discuss certain exercises with other students or tutors.

**Module 1: efficient code design**

- Jupyter notebook, Python IDE
- Good code practice
- Python docstring
- Unit, functional, and integration testing

**Module 2: project management**

- Git
- Continuous integration
- Class, module & package in Python
- Debugging & profiling

**Module 3: Python libraries for science**

- Numpy & Pandas
- SciPy

- Matplotlib

**Module 4: Python library for astrophysics & astronomy**

- PyVO
- Astropy

**Module 5: Machine learning & statistics: connecting scientific data with probabilistic models**

- Probabilistic foundations of machine learning
- Practical introduction to unsupervised and supervised learning
- Use of scikit-learn and tensorflow in scientific applications

On the third day of the school, Tamas Gal from ECAP delivered a plenary lecture on **JULIA language.**

On the last day of the school a keynote lecture by Robert Lupton was organized. This keynote was entitled as **"Scientific Software: Art, Engineering, or Science? "**

## 3.1 List of tutors

In addition to identifying school modules, the OBELICS members also contributed to suggestions of the tutors for the school from their projects. Full list of tutors with their brief bio is available on https://indico.in2p3.fr/event/16864/page/1846-list-of-tutors

## 3.2 Course content

In order for tutors to develop the course material, it was important for them to understand the level of understanding of registered participants. We had sent out a survey to school participants to understand their knowledge level for all the modules and submodules. A summary of these inputs was shared with tutors. All the tutors were consulted while finalizing the theory and hands-on sessions well in advance before the school.

## 3.3 Infrastructure

All the participants had their own laptop. A list of recommendations on tools to be installed on the laptop was provided to the participants well in advance. LAPP had ensured a strong wifi network along with eduroam network. Slack chat rooms were set up for participants to interact and discuss during the school.

# 4. Feedback

After the event, an online questionnaire was sent to the school participants to have their feedback on the scientific programme as well as the organization of the school. We received inputs from 31 participants, a summary of these inputs is presented below.

## 4.1 Event Organization

Almost all of the participants were satisfied with the event organization, accommodation arrangements as well as the social events.

## 4.2 Expectations of the participants

Despite of having participants from diverse background as well as research projects, the school managed to fulfil the expectations of almost of all the participants. Nonetheless, couple of participants mentioned to provide more time for hands-on exercises.

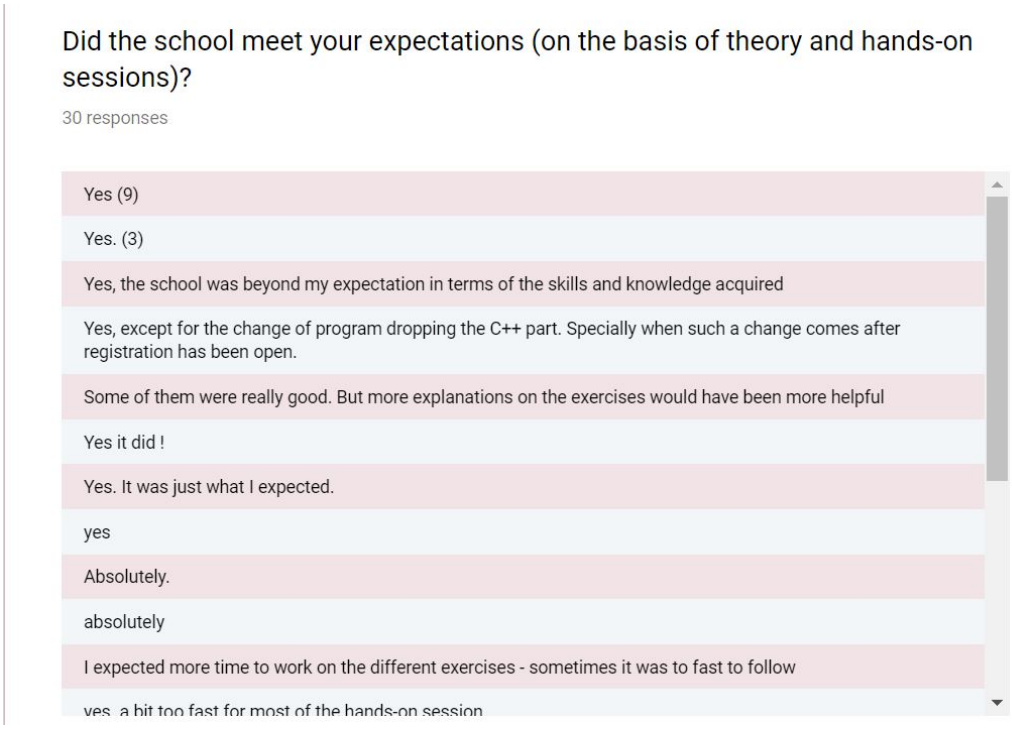


*Figure 2 Feedback from school participants regarding their expectations from the school*

## 4.3 Course Content

Vast majority of the participants informed that the training event has upgraded their programming knowledge and skills.

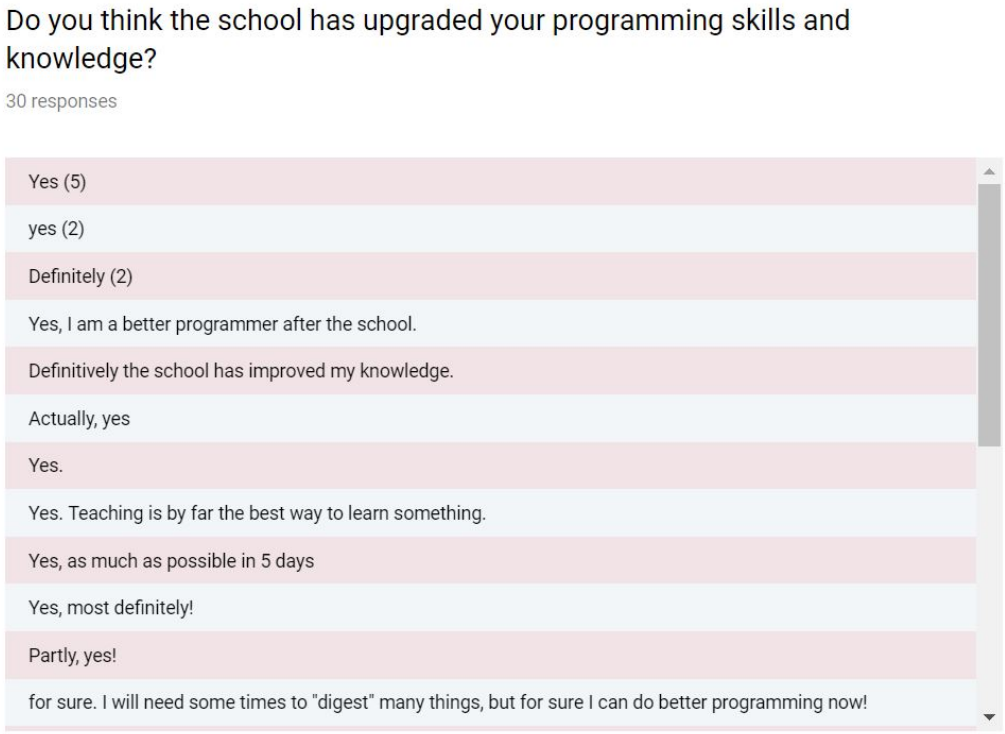


*Figure 3 Feedback on overall programming knowledge upgrade*

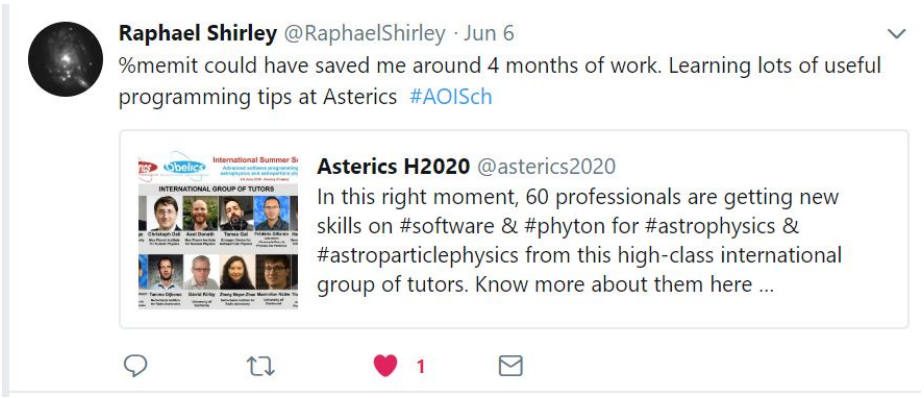


*Figure 4 Tweet from one of the participant on programming tips at OBELICS school*

The video interviews of the participants will be released on https://www.asterics2020.eu/ sharing their experiences during the school. We have summarised the feedback on each module below:

- **Efficient code design:** Most of the participants thoroughly enjoyed the course content. A couple of critical feedback included more time allocation to hands-on session.

- **Python libraries:** This module of the school was very well received by the participants and they found it very useful with some really important tips and tricks for their day to day programming. One of the critical feedback included need for PyVO hands-on session.

- **Machine Learning:** Despite of being a vast subject to handle within limited time, the participants were quite happy for being introduced to machine learning and to have gained very useful insights and tips in this cutting edge field. The critical feedback for this module was about having more hands-on sessions as most of the participants were beginner to machine learning.

# 5. Self-Evaluation

Overall, we believe the training event maintained its successful legacy and reputation. Once again we managed to address programming needs of the wider community of astrophysics and astroparticle physics. Concerning gender balance, about 20% of the participants were women. The students also widely shared their experience during the school on Social Media such as Twitter and Facebook. Event also received visibility on Elsevier[1] as well as EOSCpilot[2] website.

Based on the feedbacks from the participants during and after the school, we prepared a SWOT analysis.

**Strengths:** The scientific content of the school for the theory as well as hands-on sessions on machine learning and python programming designed by our tutors was extremely well adapted to the needs of school participants. Four social events featuring guided tour of Annecy, regional delicacies buffet, a dinner cruise and welcome reception made sure participants have ample time to socialize and network amongst themselves.

**Weaknesses:** As per the feedback from the some of the participants, it was observed that more time for hands-on sessions was necessary for machine learning.

**Opportunities:** It was a challenging task to organize 5 days event with theory and hands-on session on machine learning, python libraries for astronomy and astrophysics and efficient code design. Following the first edition of the school, we found that more time should be dedicated for hands-on sessions. Balance between theory and hands-on session duration was quite critical. Theory sessions were absolutely necessary as prerequisites for hands-on sessions because we had participants with different levels of understanding in programming. Thanks to our tutors we managed to achieve a good balance between theory sessions and hands-on sessions for all the modules though there is still room for improvement. Some participants also suggested to include C++ as one of the module in the school programme. We also learnt that there is a well-defined interest in the community for python programming in astronomy and an entire school dedicated to python programming will certainly serve the community needs.

**Threats:** The only potential threat we identified was adjusting time allocation for hands-on exercises as some participant found allocated time fairly sufficient whereas it not enough for others.

[1] https://www.journals.elsevier.com/astronomy-and-computing/news/the-second-asterics-obelics-international-school

[2] https://www.eoscpilot.eu/news/cnrs-lapp-france-sets-scene-annual-asterics-obelics-international-summer-school-4-8-june-2018.


*Figure 5 Group photo of participants of 2nd ASTERICS-OBELICS international school, CNRS-LAPP Annecy, 3-9 June 2018*