

Advanced Scientific Programming in Python – Latin America

Summer School Proposal Executive Summary [Draft May 17, 2019]

Introduction

Scientific software development is no longer reserved for a reduced number of scientist working in highly specialized branches in the field of computer science. Powerful programming techniques are now available to the general scientific community and across many academic and non-academic institutions and fields, making *the ability to program* a highly desired and important skill to conduct state of the art research. However, there is a considerable lack of education regarding such programming practices and techniques in the community of scientists who are generally not properly trained for the creation and maintenance of scientific software. Very often scientists spend (waste) a significant amount time writing, maintaining and debugging sub-optimal self-written code, often reinventing the wheel, instead of being able to focus their efforts on their topics of research. The result is the creation of poorly written scientific code which is not efficient and becomes very limited and often obsolete very fast.

With the explicit intent of responding to these needs in the scientific community the *Advanced Scientific Programming in Python Summer School* was launched in Europe twelve years ago, providing the scientific community with a unique opportunity to learn techniques and tools to develop efficient and sustainable code with hands-on experience on the subset of these techniques that matter most, teaching them clean language design, extensibility of code, and good practices in scientific computing and data visualization. Each year, this self funded effort has attracted scientist who want to acquire software engineering skills to tackle new challenges in their areas of expertise for the benefit of the scientific community. The summer school grew quickly, and in recent years, the number of applications to the school has grown to be over 300 applications per year, for the 30 places typically available (with a 50/50 gender split). Its success and worldwide demand has lead to the creation of the *Advanced Scientific Programming in Python Summer School - Asia - Pacific*, and for now two years the school has been offered to a public of the Asian continent.

The purpose of this project is to bring the teachings and benefits of the school to a Latin American audience and its scientific community. The number of latin american scientist who are directly working with scientific software development has dramatically grown in the past few years without any effort of alleviating such problems. For this, we propose the creation of the *Advanced Scientific Programming in Python Summer School - Latin America* with the aim of bringing together scientists from all latin american countries in order to participate in the first ASPP in the American continent. During the 7-day intensive course, participants will explore and learn the fundamentals ideas of scientific code development from leading experts in the field. After the summer school the participants will be able to return to their home countries and benefit to the development of the scientific community of each nation. We aim to hold the first instance of the school during the month of July of 2020 in Mexico City due to its position as a leading developing country of latin america and due to its being home of the largest and most prestigious research institution in the Ibero-American world.

In line with the *open source* commitment of the scientific Python community, the complete material, lectures, exercises, solutions, and code produced will be available online under a Creative Commons license. The summer school will not charge fees, and we expect it to remain free from commercial interests. This opens up the possibility of using the material for the development and creation of new courses tailored to the specific needs of a specific target audience. Many social programs can be developed in order to attack the problem from an earlier stage, creating courses for various stages of education for example, high school students, college students, science teachers, etc.

The previous success of the other instances of the summer school is very highly related to the engaging and interactive environment that has been created by a group of highly capable and accessible scientists who have created a teaching methodology which fosters learning. The teaching and learning approach itself is largely inspired by the Agile technique from software engineering, which promotes flexible and responsive experience, and nurtures learning opportunities. In order to reproduce this environment, we aim to hold the school in a remote location in Mexico City in order to create a profound learning experience away from distractions and daily routines. Informal chats during shared meals and recreation time and during coding exercises will provide opportunities to all participants to form close professional and personal relationships, fostering future collaborations and openness within the next generation of Latin-American scientists.

Structure of the Summer School

The summer school Advanced Scientific Programming in Python - Latin America (ASPP - LA) will follow the structure and curriculum of the previously organized schools in Europe (2008-2019) and the Asia-Pacific region (2018 and 2019).

During the summer school, we will teach a selection of advanced programming techniques and best practices in software development (which are standard in the industry) to researchers in the natural and social sciences, engineering and mathematics. The school will consist of lectures and discussions along with practical hands-on exercises which will supplement the lectures and will let participants practice the techniques. All of the newly acquired skills will be tested in a real programming project: participants will team up to develop an entertaining scientific computer game.

We use the Python programming language for the entire course. Python works as a simple programming language for beginners, but more importantly, it also works great in scientific simulations and data analysis. These features make the language a very popular choice in almost any branch of science. We show how clean language design, ease of extensibility, and the great wealth of open source libraries for scientific computing and data visualization are driving Python to become a standard tool for the programming scientist.

Target Audience

The summer school is targeted at master or phd students, as well as post-docs and research scientists from Latin American countries. Competence in Python or in another language such as Java, C/C++, MATLAB, or Mathematica is required, and basic knowledge of Python is assumed. The course is planned to allow for an attendance of 30 international students (50/50 gender split), who fund their own travel and stay for a week. Importantly, participation in the school is free, i.e. no fee is charged, however participants should take care of travel, living, and

accommodation expenses by themselves. We aim to provide travel allowance for a number of participants if resources are sufficient (see section “Funding” below).

Preliminary Program

These lectures will build from educational materials given during the previous ASPP-Europe and ASPP-Asia-Pacific schools. The contents of the lectures, the exercises and the programming project are developed by all faculty members together in order to ensure a coherent set of materials, which gets further refined by integrating the student feedback collected by a survey at the end of each instance of the school.

- Day 1: Best Programming Practices - Best Practices for Scientific Computing - Version control with Git and how to contribute to open source with GitHub - Object Oriented Programming and Design Patterns.
- Day 2: Software Carpentry - Test-driven development, unit testing and quality assurance - Debugging, profiling and benchmarking techniques - Advanced Python 1: idioms, useful built-in data structures, generators.
- Day 3: Scientific Tools for Python – Advanced NumPy – The Quest for Speed 1: Interfacing to C with Cython – Programming in teams.
- Day 4: The Quest for Speed 2: Writing parallel applications in Python – Programming project.
- Day 5: Efficient Memory Management – Advanced Python 2: decorators and context managers – Programming project.
- Day 6: Practical Software Development – Programming project – The Scientific Game Tournament.

Preliminary Faculty

The number of speakers is approximately 10, and they also act as tutors for the students during the course. All speakers are computational scientists or software developers with a PhD in relevant disciplines. Speakers do not receive any honorarium, and take part in the summer school on their spare time:

- Tiziano Zito, Institute of Psychology, Humboldt Universität zu Berlin, Germany.
- Rike-Benjamin Schuppner, Institute for Theoretical Biology, Humboldt- Universität zu Berlin, Germany.
- Stéfan van der Walt, Berkeley Institute for Data Science, University of California at Berkeley, USA.
- Madicken Munk, Postdoctoral Researcher at the Data Exploration Lab, National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, USA.
- Kathryn D. Huff, Assistant Professor, Department of Nuclear, Plasma, and Radiological Engineering, University of Illinois at Urbana-Champaign, USA.
- Ricardo Méndez Fragoso, Professor of Physics, Faculty of Sciences of the National Autonomous University of Mexico, Mexico.

Organizing Committee

- Carlos Echeverría Serur, Scientific Assistant and PhD Candidate at the Institut für Mathematik of the Technische Universität Berlin, Germany (Alumnus ASPP-Europe 2018).
- Luis García Ramos, PhD Candidate at the Institut für Mathematik of the Technische Universität Berlin, Germany (Alumnus ASPP 2018).
- Josué Martínez Moreno, PhD Candidate at Research School of Earth Sciences of the Australian National University, Canberra, Australia (Alumnus ASPP-APAC 2019).
- Ricardo Méndez Fragoso, Professor at Department of Physics, School of Sciences of the National Autonomous University of Mexico, Mexico (Invited to ASPP 2019).
- Florencia Noriega, Software Engineering Lecturer at CODE University of Applied Sciences, Berlin, Germany (Alumna ASPP 2016).
- Tiziano Zito, Institute of Psychology, Humboldt Universität zu Berlin, Germany (Co-Founder of ASPP).

Funding

The total budget of the event varies in range between 10,000 and 20,000 USD, depending on how much we are able to obtain as donations from sponsors and/or partner organizations (facilities, lecture rooms, computer laptops, accommodation for faculty and students).

The budget will be used to fund travel, accommodation and daily expenses of faculty members and organizers. It will also cover the rent of 15 laptops (1,500 USD), organize catering for breaks (400 USD) and lunch for the speakers (700 USD), as well as 2 social events (2,000 USD). Students pay their own travel and accommodation costs. In order to open the school to a population of students from developing Latin American countries who do not have access to funding, we would like to offer financial support to fund travel or accommodation expenses for a number of international and local students (12,000 USD)

Tentative Partner Organizations (Sponsors/Co-Organizers)

- Secretaría de Ciencia, Tecnología e Innovación de la CDMX (SECTEI).
- Facultad de Ciencias, UNAM (FCIENCIAS-UNAM).
- Sociedad Mexicana de Física (SMF)
- Instituto de Investigaciones en Matemáticas Aplicadas (IIMAS-UNAM)
- Consejo Nacional de Ciencia y Tecnología (CONACYT)
- Instituto Latinoamericano de la Comunicación Educativa (ILCE)
- Instituto Politecnico Nacional (IPN)

Additional Information

More information and teaching materials from previous instances of the school can be found at the following website: <http://python.g-node.org>