Summer School: Advanced Scientific Programming in Python - Latin America

Project Proposal [Draft May 27, 2019]

México City, México

July 2020

Scope

La escuela de verano Advanced Scientific Programming in Python (ASPP) - Latinoamérica seguirá la estructura y curriculum de las escuelas previamente organizadas en Europa (2008-2019) y la región de Asia-Pacífico (2018, 2019). Esta exitosa escuela de verano ha mostrado una necesidad por enseñar a científicos los principios de *computo científico e ingeniería de sistemas* en el lenguaje de programación Python. El objetivo de éste intenso programa es enseñar técnicas avanzadas de programación, así como las mejores practicas en el desarrollo de software que son estándares para realizar investigación en ciencias naturales y sociales, ingeniería y matemáticas.

Las herramientas y conceptos introducidos en este taller ayudará a jóvenes investigadores a usar menos tiempo escribiendo, resolviendo problemas del código y entendiendo programas de computación. Adicionalmente, esperamos que después del curso, participantes serán capaces de contribuir a la comunidad científica de código abierto. Con tan solo una semana de clases, reconocemos que estos ambiciosos objetivos unicamente serán parcialmente cubiertos. Sin embargo, los participantes serán expuestos a una amplia cantidad de recursos y referencias, que en caso de ser estudiados y practicados, entrenarán a la nueva generación de futuros científicos capaces de escribir mejor códigos y también hacerlo re-usable y fácil de entender en caso de que alguien quiera reproducir sus resultados.

Éste curso consistirá en clases y discusiones con una componente practica The course will consist of lectures and discussions along with practical hands-on exercises which will supplement the lectures and will let participants practice the techniques. The participants are assumed to have prior knowledge (at the level of a competent user) on some high-level programming language used in scientific computing (for example, C, C++, Java, Fortran, MATLAB, etc.) and

a basic knowledge of Python. The lectures will move at a rapid, though pedagogical pace. In general, the material covered will be more than the average amount that a participant can digest in a single week. Therefore, it will be mandatory that the participants familiarize themselves with some (online) prerequisite material that we will provide. Moreover, we encourage participants to continue with further study and implementation of the techniques after the course.

The course will not be an introductory course to the Python programming language, though questions on basic features of the language will be refreshed as we go along the lectures if necessary. It will also not be a specialized course on software engineering or computer science. Rather, the lectures will be focused on the needs of scientists who need to optimize their software work-flow. Correspondingly, there will be a strong hands-on component, where participants are encouraged and expected to participate, discuss and proactively learn not only from the faculty but also from other participants.

[c1]Merge this section

Vision

The Advanced Scientific Programming in Python has been a student initiative with the goal of better understanding how to make scientific source code reproducible, acurate and reusable that many scientists use along their career. The school will bring together scientists from all over Latin-America for the first ASPP in the American continent. During the 7-day intensive course, students will explore and learn the fundamentals ideas of scientific code development from leading experts in the field. The school will be held in a remote location near Mexico City to create a deep, focused learning experience away from distractions and daily routines. Informal chats during shared meals, recreation time and during coding excercises 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.

Need

There is a considerable education lack in scientific programming courses. Across many academic and non-academic institutions, programming has become a key skill to conduct research. However, in many cases our interaction with codes ranges from simply copy-paste code found online to developing and running new modules-packages. Whilst many of us are familiar with the basics of programming, less well understand the full capabilities and standards used to facilitate openness and sharability of code. Additionally, the codes we commonly use do not aliviate the generation of bugs which leads to use trial-and-error or rely on

---

[c1] Josue

advice from others. Therefore, a better understanding on coding practices as well as the underpins on Python, will provide the required tools to increase the participants efficiency and code robustness while programming.

Organizing Committee

- Carlos Echeverría Serur, PhD Student at the Institut für Mathematik of the Technische Universität Berlin (Alumnus ASPP-Europe 2018).

- Luis García Ramos, PhD Student at the Institut für Mathematik of the Technische Universität Berlin (Alumnus ASPP-Europe 2018).

- Josué Martínez Moreno, PhD Student at Research School of Earth Sciences of the Australian National University, Canberra (Alumnus ASPP-APAC 2019).

- Ricardo Méndez Fragoso, Professor at Department of Physics, School of Sciences of the National Autonomous University of Mexico, (Invited to ASPP-Europe 2019).

- Florencia Noriega, Software Engineering Lecturer at CODE University of Applied Sciences, Berlin (Alumna ASPP-Europe 2016).

- Tiziano Zito, from the Institut für Psychologie of the Humboldt Universität zu Berlin (Co-Founder of ASPP-Europe).

Faculty (Preliminary List)

- Tiziano Zito, Institut für Psychologie, Humboldt Universität zu Berlin.

- Rike-Benjamin Schuppner, Institute for Theoretical Biology, Humboldt-Universität zu Berlin, Germany.

- Stéfan van der Walt, Researcher, Leader of the Software Working Group, Berkeley Institute for Data Science, University of California at Berkeley, USA.

- Josué Martínez Moreno, PhD Student at Research School of Earth Sciences of the Australian National University, Canberra, Australia.

- Florencia Noriega, Lecturer in Software Engineering at CODE University of Applied Sciences, Berlin, Germany.

- Olmo Zavala,

- 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.

- Fernando Pérez, Assistant Profesor, Berkeley institute of Data Science, University of California at Berkeley, USA.

- Ricardo Méndez Fragoso, Professor of Physics, Faculty of Sciences of the National Autonomous University of Mexico.

- [c1]Who else may be interested? Ideally we should invite people from the US, and America Latina

Tutors (Preliminary List)

- Carlos Echeverría Serur.

- Luis García Ramos.

- Josué Martínez Moreno.

- Ricardo Méndez Fragoso.

- Florencia Noriega.

Note: The budget will fund travel, accommodation and *Per Diem* expenses of lecturers and organizers. Typically, organizers also serve as lecturers. The budget may not allow for funding of organizers who do not actively participate in the school as lecturers. Additional budget will be use to support accommodation or travel expenses of international and local students.

Participants

Participation is free, i.e. no fee is charged, however participates should take care of travel, living, and accommodation expenses by themselves. We may be able to provide a small travel allowance if resources are sufficient (see "Funding" below).

Admission Requirements

These are the requirements for participating in ASPP - Latin America:

- Applicants should be research students, Masters, PhD students, Postdocs or Research Scientists in the natural or social sciences, engineering or mathematics, at an institution in South, Central or North America. Priority will be given to applicants from Latin America and applicants from underrepresented groups.

[c1] *Josue: Text added.*

- Applicants are absolutely required to know a high-level programming language used in scientific computing (for example, C, C++, Java, Fortran, MATLAB, etc.) and the basics of Python to get the most out of this Summer School.

- Applicants should provide proof of their English proficiency by providing a letter of motivation and a curriculum vitae of the applicant written in the English language. Applicants are not required to provide results from a proficiency test.

Lectures

These lectures will build from feedback obtained during the previous ASPP-Europe and ASPP-Asia-Pacific schools (topics include potential speakers).

- Version control (GitHub) -

- Tidy data analysis and visualization - Ricardo Méndez Fragoso [c1]

- Testing and debugging scientific code - Luis García Ramos

- Advanced NumPy - Carlos Echeverría

- Organizing, documenting, distributing scientific code and continuous integration -

- Advanced scientific Python: context managers and generators -

- Writing parallel applications in Python -

- Profiling and speeding up scientific code with Cython and numba -

Schedule

Monday XX July 2020			
Time	Topic	Speaker	Tutors
9:00-9:30	Introduction		
9:30-11:00	Git & Github		
11:00-11:30	Coffee Break		
11:30-13:00	Git & Github		
13:00-14:00	Lunch Break		
14:00-15:00	Git & Github		
15:00-16:00	Data visualization	Ricardo Méndez	
16:00-16:30	Coffee Break		
16:30-18:00	Data visualization	Ricardo Méndez	

[c1] *Josue*: Perhaps change the structure and remove Tidy data?s

Tuesday XX July 2020			
Time	Topic	Speaker	Tutors
9:00-9:30	Testing & debugging		
9:30-11:00	Testing & debugging		
11:00-11:30	Coffee Break		
11:30-13:00	Testing & debugging		
13:00-14:00	Lunch Break		
14:00-15:00	Advanced Numpy		
15:00-16:00	Advanced Numpy		
16:00-16:30	Coffee Break		
16:30-18:00			

Wendsday XX July 2020			
Time	Topic	Speaker	Tutors
9:00-11:00	Organizing, documenting, distributing & CI		
11:00-11:30	Coffee Break		
11:30-13:00			
13:00-XX:XX	Social activity		

Thursday XX July 2020			
Time	Topic	Speaker	Tutors
9:00-10:30	Parallel Python		
10:30-11:00	Coffee Break		
11:30-13:00	Parallel Python - Pycuda		
13:00-14:00	Lunch Break		
14:00-14:30	Project Introduction		
14:30-16:00	Programming Project		
16:00-16:30	Coffee Break		
16:30-18:00	Programming Project		

[c1] *Josue*: I would like to add some Scipy, what do you think?
[c2] *Josue*: Do we want a social activity?

6

Friday XX July 2020			
Time	Topic	Speaker	Tutors
9:00-10:30	Profiling, Cython & numba		
10:30-11:00	Coffee Break		
11:30-13:00	Profiling, Cython & numba		
13:00-14:00	Lunch Break		
14:00-16:00	Programming Project		
16:00-16:30	Coffee Break		
16:30-18:00	Programming Project		

Saturday XX July 2020			
Time	Topic	Speaker	Tutors
9:00-10:30	Programming Project		
10:30-11:00	Coffee Break		
11:30-13:00	Programming Project		
13:00-14:00	Lunch Break		
14:00-16:00	Programming Project		
16:00	Repository Freeze		
16:00-16:30	Coffee Break		
16:30-16:45	How to contribute to ASPP		
16:45-17:00	Programming project - Learning report		
17:00-18:00	Friendly Tournament		
18:00-XX:XX	Final Social Event		

c1 c2

Estimated Costs

Presupuesto dentro de 150000 y 200000 pesos

The budget will fund travel, accommodation and daily expenses of lecturers and organizers. Organizers can also serve as lecturers or tutors since the budget may not allow for funding of organizers who do not actively participate in the

---

[c1] *Josue*: I've removed 1 hour from the lunch break so people can go earlier home, do you think we should use that hour or shouldn't?

[c2] *Luis and Carlos*: We think they should have two hours of lunch break (to be discussed)

school as lecturers. Additional resources will be used to support accommodation or travel expenses of international and local students.

Estimated costs per Faculty Member (should be covered by the organizers)

- Flight (from Europe or US): Approx. 800USD

- Accommodation in CDMX: Approx. 400USD

- Local Transport: Approx. 100USD

- Food for the entire duration of the conference: Approx. 200USD

Estimated costs of Organization (sholuld be covered by the organizers)

- Computer Rental (40 computers): Approx. ???USD

- Venue Rental: Approx. ???USD

- Social Activity (Trip to Teotihuacán): Approx. ???USD

- Airport-Venue Transport: Approx. ???USD

Estimated average costs per Participant (should be covered by the participants)

- Flight (From a Latin American Country): Approx. 500USD

- Accommodation in CDMX: Approx. 300USD

- Local Transport: Approx. 50USD

- Food for the entire duration of the conference: Approx. 150USD

Tentative Partner Organizations (Sponsors/Funding)

- Secretaría de Ciencia, Tecnología e Innovación de la CDMX.

- Faculty of Sciences (FCIENCIAS-UNAM).

- Instituto de Investigaciones en Matemáticas Aplicadas (IIMAS-UNAM)

- Consejo Nacional de Ciencia y Tecnología (CONACYT)

- Instituto Latinoamericano de la Comunicación Educativa (ILCE)

- Instituto Politécnico Nacional (IPN)

Policies

Inclusiveness

We are striving hard to obtain a group of participants which is international and gender-balanced.

Behavior

See `https://numfocus.org/code-of-conduct`Code of conduct