

Programa:Taller Open Data Science

26 y 27 de Septiembre 2019

Laboratorios de la Escuela Profesional de Ing. de Computación e Informática -
Facultad de Ciencias Físicas y Matemáticas -Universidad Nacional Pedro Ruiz Gallo

Talleristas

Arturo Sánchez

Ingeniero informático, PhD en Física Aplicada, especializado en ciencia de datos. Investigador científico del CERN (Organización Europea para la Investigación Nuclear) de Suiza. Administrador informático y co-fundador del Sistema ATLAS, experimento de investigación y monitoreo de los datos del acelerador de partículas. Entrenador y desarrollador de proyectos de código abierto (<https://github.com/artfisica>) en análisis de ciencias, educación, Big Data y Open Science. Representante Latinoamericano de la Red Global de Creative Commons - Venezuela.

Iskya García

Master en Ciencias Físicas. Investigadora científica y coordinadora de la iniciativa CEVALE2VE, especializada en análisis de datos científicos. Experiencia como docente en Ciencias Físicas en la Universidad Federal de Rio de Janeiro, Brasil, ha dictado talleres como instructora del proyecto ATLAS del CERN en diversas universidades de Argentina, Brasil y Venezuela. Su investigación se basa en la astrofísica y estudio del universo.

Organizaciones

CERN: La Organización Europea de Investigación Nuclear, cuenta con el laboratorio más grande de física de partículas, el propósito de su trabajo es ayudar a descubrir de qué está hecho el universo y cómo funciona. <https://home.cern/>

Physics Without Frontiers (PWF): Es una iniciativa que trabaja para inspirar, capacitar y motivar a estudiantes de física y matemáticas de todo el mundo, enfocándose en los países rezagados del acceso a la ciencia y tecnología, con el propósito de construir la próxima generación de científicos. <https://www.ictp.it/physics-without-frontiers.aspx>

CEVALE2VE: Es una iniciativa latinoamericana que busca fortalecer la divulgación didáctica, la formación académica y la investigación en el áreas de física y ciencias de la computación. Su propósito es construir redes de colaboración entre instituciones latinoamericanas para fomentar la investigación científica en los estudiantes. <http://www.cevale2ve.org/es/inicio/>

Proyectos Invitados

Experimento ATLAS (A Toroidal LHC ApparatuS, Aparato Toroidal del LHC): Es el moderno acelerador de partículas del CERN en Suiza. Su tamaño es de 46 metros de largo y 25 de diámetro, y pesa unas 7.000 toneladas. En el proyecto están implicados unos 3.000 científicos e ingenieros de más de 175 instituciones pertenecientes a 38 países diferentes. <https://atlas.cern>

Open Data ATLAS: Es el repositorio de datos abiertos del experimento ATLAS. El proyecto es una guía para aprender a visualizar estos datos, e incluso proporciona software de código abierto para que pueda hacer sus propios descubrimientos y proyectos. <http://opendata.atlas.cern/> <https://youtu.be/hPaOfikJ1M4>

Programa

El presente programa es tentativo, la revisión de los temas según el horario dependerá del desarrollo del taller, la participación y nivel de experiencia de los asistentes.

Jueves 26 de Septiembre 2019

9:30 am

Registro

Validación de registros y bienvenida

10:00 am - 1:00 pm

Introducción a la programación computacional

Sistemas Operativos Linux

Interfaz de Línea de Comandos o CLI (Scripting)

Principios de BASH

Principios de Python

Principios de C++

2:00 pm - 3:00 pm

Principios en manipulación de datos

Pre-procesamiento de Datos (Data Cleaning)

Estructura de Datos y formatos

Curación de Datos (Bookkeeping and Storage)

3:00 pm - 5:00 pm

Computación en la Nube (Cloud Computing)

Definiciones y Diagramas sobre la Cloud

Leyendo Datos (Data Reading)

Transformando Datos (Data Dealing)

Desarrollo de Software Distribuido y Versionado: GIT

Ejemplos de pipelines: creación de análisis de datos

5:30 pm - 6:00 pm

Cierre

Dudas, consultas y cierre del primer día de taller.

Viernes 27 de Septiembre 2019

8:30 am

Registro

Validación de registros y bienvenida

9:00 am - 10:00 am

Review del Día 1

Revisión de pauta, dudas y consultas

10:00 am - 1:00 pm

Desarrollo colaborativo (Data Analysis & Software)

Desarrollo de Código (Bookkeeping and Storage)

Principios de optimización de pipelines

Principios de Estadística aplicada al Análisis de Datos

Ejemplos de Análisis en Física de Altas Energías

Visualización de Datos y Publicación de Resultados

2:00 pm - 5:00 pm

Visualización de Datos y Publicación de Resultados

Ejemplos de Análisis de Imágenes

Publicación y Reproducibilidad

Licencias, Citaciones y Créditos

5:00 pm - 5:30 pm

Cierre

Cierre del Taller y Despedida