



Software Training with HSF and IRIS-HEP

New Perspectives - June 27, 2023

Guillermo Fidalgo

University of Puerto Rico at Mayagüez

What are HSF and IRIS-HEP?

- HEP Software Foundation
<https://hepsoftwarefoundation.org/>
- Institute for Research and Innovation in Software for High Energy Physics
<https://iris-hep.org/>



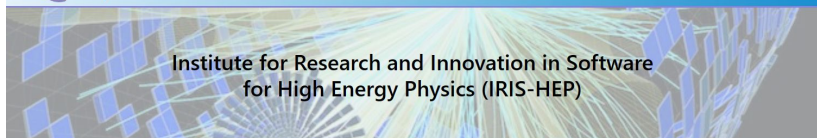
The HEP Software Foundation facilitates cooperation and common efforts in High Energy Physics software and computing internationally.

HSF Event Generator Tuning Workshop, 27-28 June 2023 (more info)

PyHEP.dev 2023 Developer's Workshop, 25-29 July 2023 (more info)

PyHEP 2023 "Python in HEP" (virtual) Workshop, 9-12 October 2023 (more info)

JuliaHEP 2023, 6-9 November 2023 (more info)



Computational and data science research to enable discoveries in fundamental physics

IRIS-HEP is a software institute funded by the National Science Foundation. It aims to develop the state-of-the-art software cyberinfrastructure required for the challenges of data intensive scientific research at the High Luminosity Large Hadron Collider (HL-LHC) at CERN, and other planned HEP experiments of the 2020's. These facilities are discovery machines which aim to understand the fundamental building blocks of nature and their interactions. [Full Overview](#)

News and Featured Stories:

Upcoming Events:

Jul 12-14, 2023 *Virtual*
HSF/IRIS-HEP Software Basics Training (Virtual)

Jul 17-21, 2023 *Princeton University*
CoDaS-HEP 2023 - Computational and Data Science Training for High Energy Physics

Jul 24-26, 2023 *Princeton University*
3rd MODE Workshop on Differential

Context

Experimental collaborations

- Bigger, spread over continents
- CMS and ATLAS ~ 8000 users, DUNE - 1200 users

Big, distributed computing resources, manpower

Detectors building, instrumentation and detector operations require expertise takes years of experience and involvement

Large data set volumes to process

Emerging technologies, novel techniques

Investment in organised training (hands-on)

- Mitigate some of the challenges
- Build future workforce
- Careers in HEP or other STEM areas

Organised Software Training is essential

- HSF and IRIS-HEP to the rescue!

Why not leave everything to the experiments?

- $O(10k)$ HEP people need training to tackle computing challenges
- General software skills matter as much as experiment-specific training
- Experiments share a lot of software topics
- Software topics are vast & evolving! (e.g., ML, Python, C++, Git, CI/CD)
- **We can cover more ground together** instead of reinventing the wheel...

$O(6k)$

Undergrads



Teach prerequisites &
Democratize science



$O(3k)$

Ph.D. students



Impart best practices



$O(1k)$

Postdocs

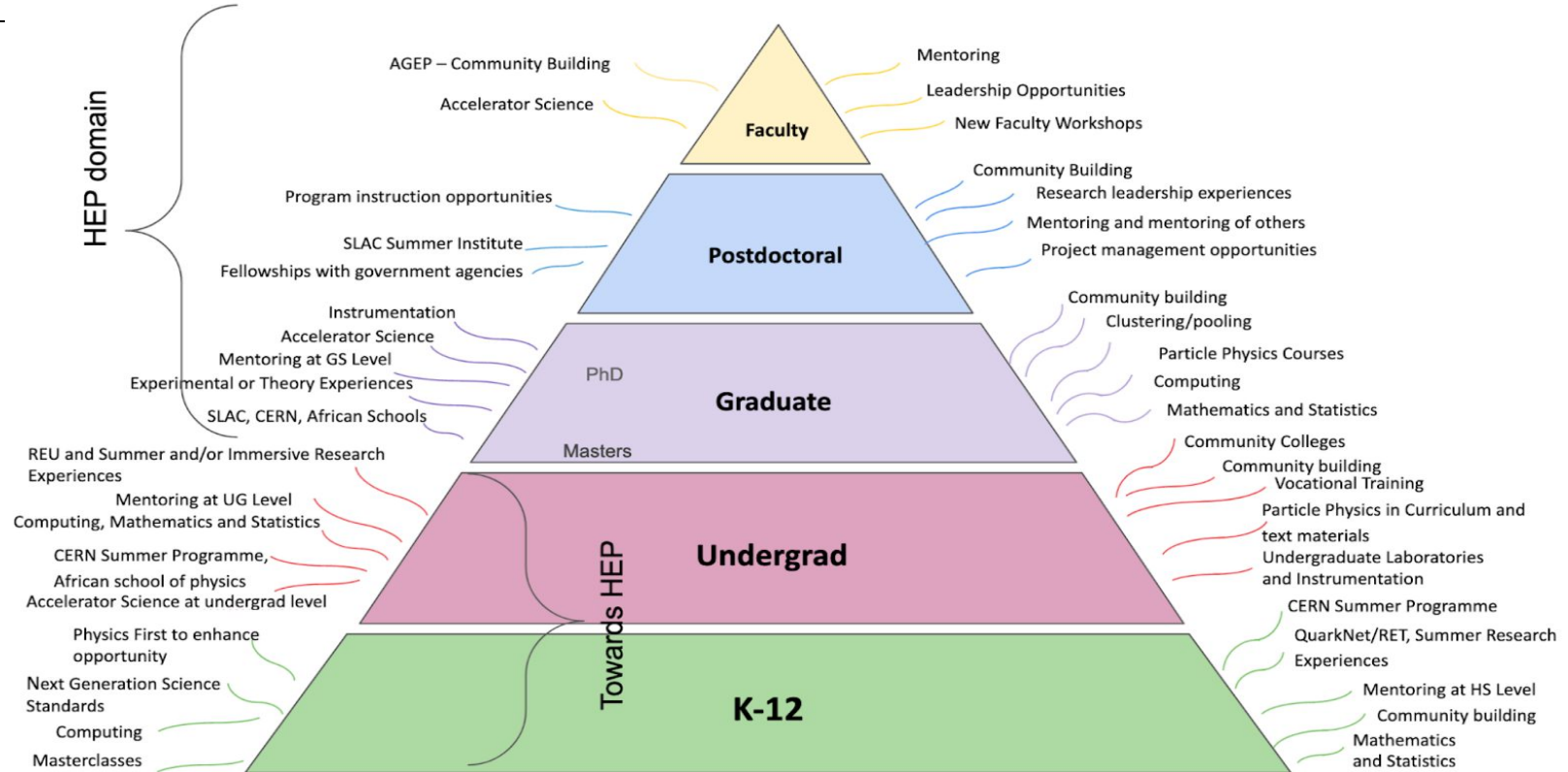


Put in touch with
most recent developments



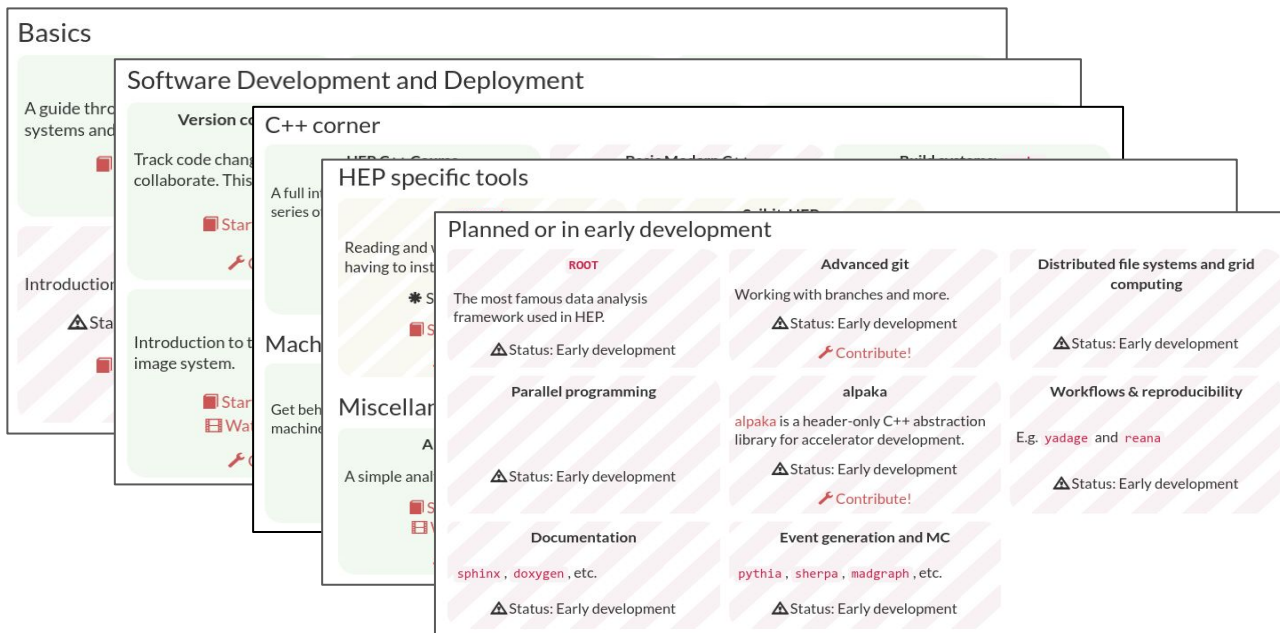
We need a **unified, scalable, and sustainable** software training framework

Workforce Pipeline: Training is integral to the success of HEP



Workshops and Training Material

Training Curriculum



GitHub

Search or jump to... Pull requests Issues Codespaces Marketplace Explore

HEP Software Foundation Training Material
Training and educational material for the high energy physics community.
41 followers https://hepsoftwarefoundation.org/... @HsfTraining hsf-coordination@googlegroups.com

Overview Repositories 42 Projects 5 Packages Teams 3 People 28 Settings

README.md

Welcome to HSF Training

Weekly meetings

October 2022

- 24 Oct Training WG Planning Meeting
- 17 Oct Training WG Planning Meeting
- 10 Oct Training WG Planning Meeting

Community pages

Our community

- Amber Roepe (@huer)
- Andrea Valassi
- Clemens Lange
- Dan Guest

Monthly hackathons

Containerization Training Hackathon

The big goal: Training in software and computing is an essential ingredient for the success of any HEP experiment. As most experiments have similar basic prerequisites we want to join our efforts and create an introductory software training course for new HEP newcomers from all experiments. This course is made of self-organizing training modules and should consist of software skills needed in the early to mid while training but practices for writing scientific software. We have started this work here and have completed and tested several of our modules to great success.

Slack

Recognition

Thanks goes to these wonderful people (emoji key) who contributed

Authors

- Meiriv Osn Evans
- amorenor
- Guillermo A. Fidalgo-Rodriguez
- Marchabot

Most training modules are website built from easy-to-read source files

HSF Home Code of Conduct Setup Episodes Extras License Improve this page Search...

Machine Learning on GPU

Machine Learning on GPU 0 - Setup

Watch later Share

Watch on YouTube

< Complete video walkthroughs!

This tutorial explores Machine Learning using GPU-enabled PyTorch for applications in high energy physics. It follows directly from the [Introduction to Machine Learning lesson](#) written by Meiriv Evans.

Prerequisites

- A Kaggle account. [Click here to create an account](#)
- Basic Python knowledge, e.g. through the [Software Carpentry Programming with Python lesson](#)
- Basic ML knowledge, e.g. through the [Introduction to Machine Learning lesson](#)

Lessons build on each other

Introduction

For physicists working on analysis in data-intensive fields such as particle physics, it's quite common these days to start developing new machine learning applications. But many machine learning applications run more efficiently on GPU.

The aim of this lesson is to:

- demonstrate how to move an existing machine learning model onto a GPU
- discuss some of the common issues that come up when using machine learning applications on GPUs

The skills we'll focus on:

1. Understanding a bit about GPUs
2. Using Python & PyTorch to discover what kind of GPU is available to you
3. Moving a machine learning model onto the GPU
4. Comparing the performance of the machine learning model between the CPU and the GPU

Enough verbosity for self-study

Visibility - Publications, Talks, Conferences

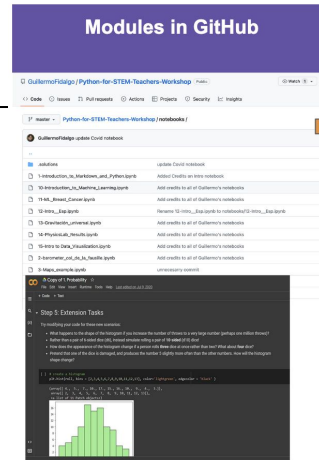
- Our efforts are visible
- Invited to conferences
- Produced publications

Training Talks & Papers

Date	Type	Title	Note
	talk	HSF / IRIS-HEP Training Activities (Coordinated Ecosystems Workshop)	
2022-10-12	talk	Training Challenge (IRIS-HEP retreat)	
2022-09-12	talk	Teaching Python the Sustainable Way: Lessons Learned at HSF Training (pyHEP 22)	
2022-09-05	talk	Sustainable Software Training Delivery at the HEP Software Foundation	
2021-02-28	paper	Software Training in HEP	Published in CSBS
2021-06-29	talk	Software Training and Sustainable HEP	Video available
2021-05-21	talk	Software Training in HEP	Video available
2020-11-19	talk	Community building	Video available
2020-11-19	talk	HSF Training: Making "that thing my postdoc taught me once" available for everyone	Video available
2018-07-08	paper	HEP Software Foundation Community White Paper Working Group - Training, Staffing and Careers	

Broader Impacts

- Software awareness and skill development among high school students via teachers
- Developed Software module
- Coding Camps
- Relation with community of teachers to expand and sustain our efforts
- Access to wider community of teachers to get software training
- Breaks barriers and enables diversity



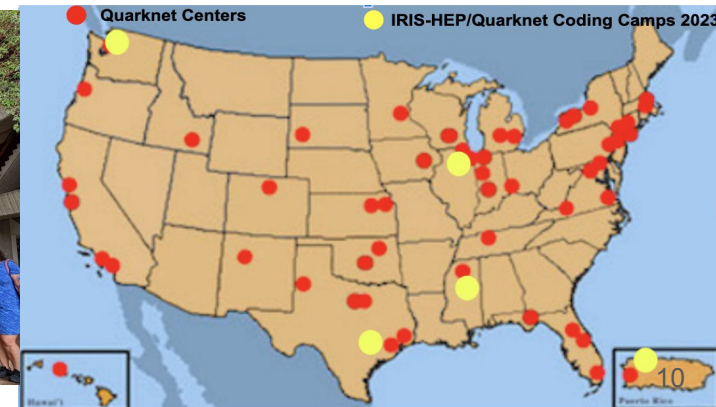
“student hat” Engage, explore, cexplain

- Teachers work in groups
- Running Python code
- Using pre-Jupyter notebook
- Review basic coding
- Use CMS data



“teacher hat” Elaborate and Evaluate

- Teachers develop implementation plans for their own classroom
- Writing Jupyter notebook
- Adapt and apply appropriately to their classroom



Why you should contribute?

Give back to the community!

Great way to start with open source!

You want to have a high impact and advance HEP?

Develop new skills and gain teaching experience

Develop networks with people from diverse communities and collaborations!

Training might be your most effective choice!

Instill best practices to your students

Get recognition for your contributions

You don't need to be an expert to contribute. Start out as a student, and you can be an educator now!

How to contribute?

Join our Slack workspace

https://join.slack.com/t/hsftraining/shared_invite/zt-1g6ga2pgg-ybmNW55kIgs81Kadqw4g

Join our mailing list to find out about training events

<https://groups.google.com/g/hsf-training-wg>

Attend weekly working group meetings

Monday 16:00 CET weekly

<https://indico.cern.ch/category/10294/>

Contribute to training materials

Pull requests to training materials are enthusiastically invited!

<https://github.com/hsf-training>



Thank you!

