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/

2

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)	O(3k)	O(1k)
Undergrads	Ph.D. students	Postdocs

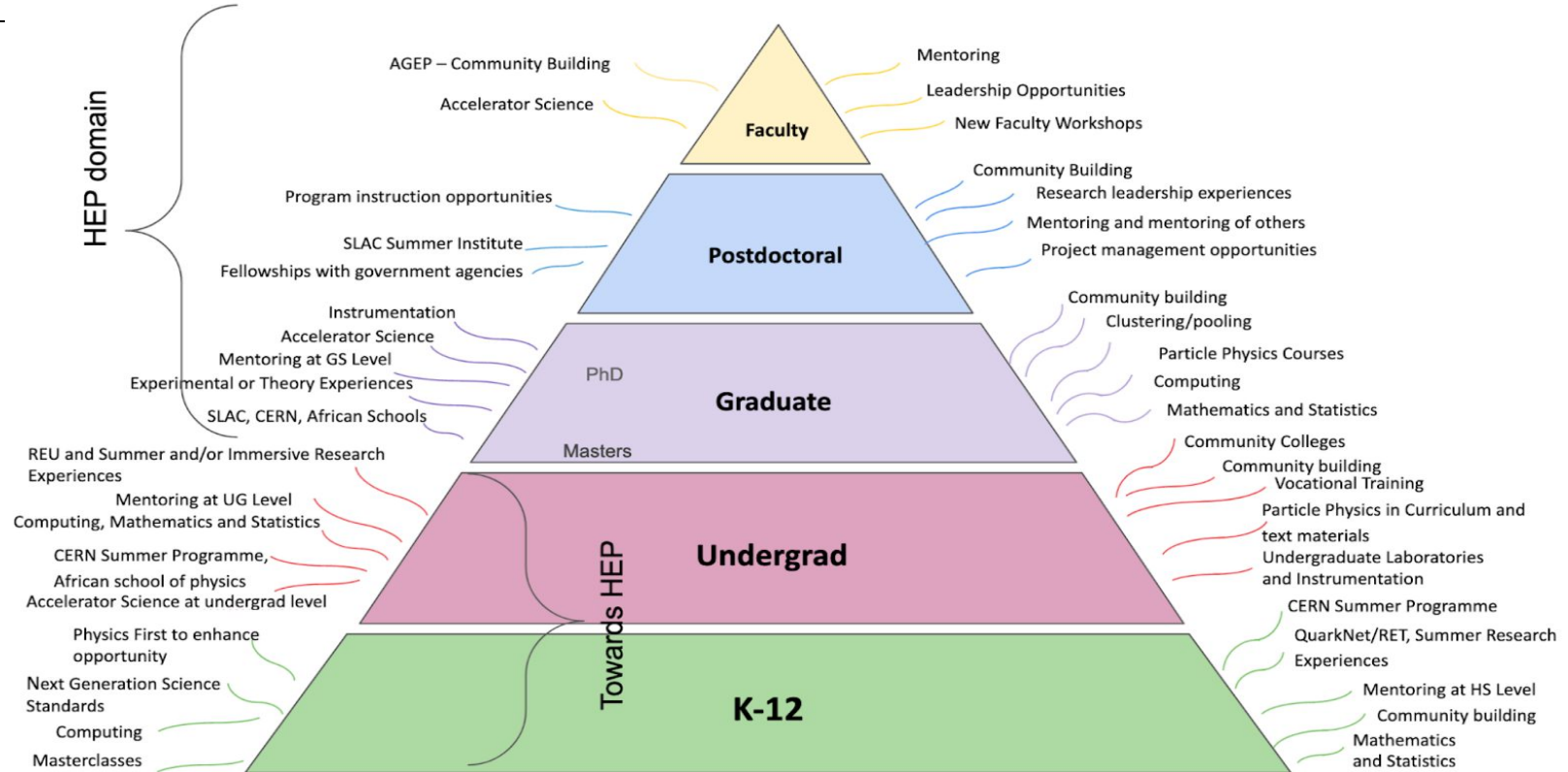Teach prerequisites & **Democratize science**

Impart best practices

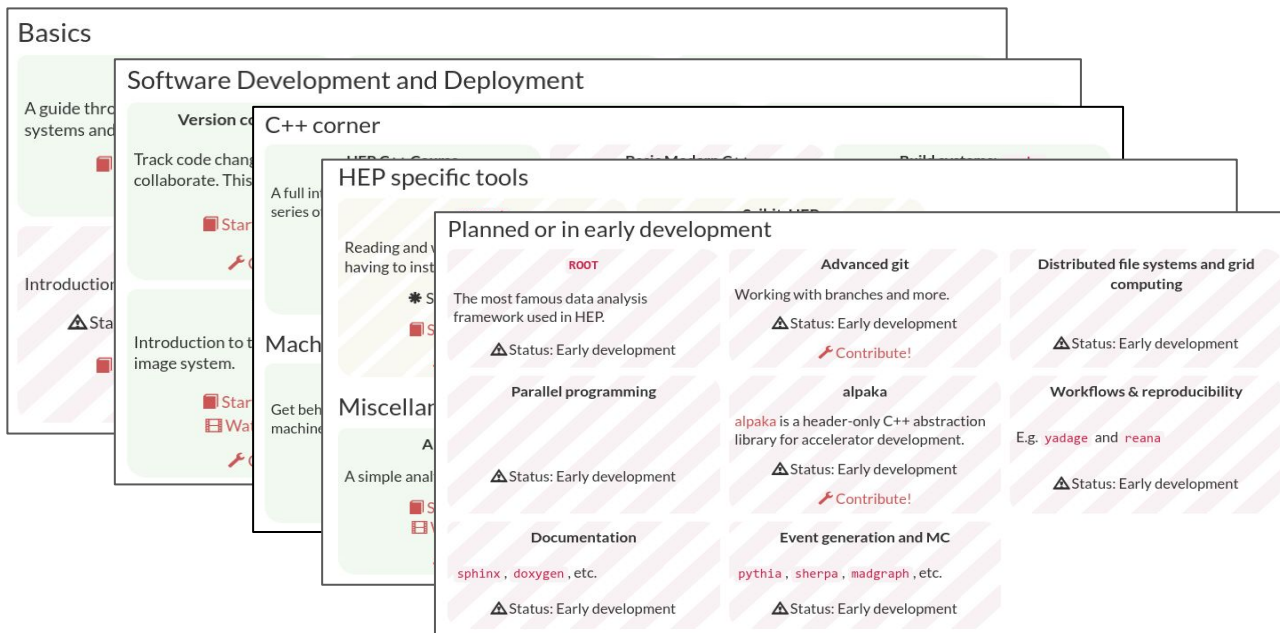Put in touch with **most recent developments**

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

4

Workforce Pipeline: Training is integral to the success of HEP

Workshops and Training Material

Training Curriculum

GitHub

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 🤗

Community pages
Our community

Weekly meetings
October 2022
- 24 Oct — Training WG Planning Meeting
- 17 Oct — Training WG Planning Meeting
- 10 Oct — Training WG Planning Meeting

Monthly hackathons
✨ Containerization Training Hackathon ✨
🎯 The big goal!
we do a little hacking

Slack

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

---

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

< 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 Meirin 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

8

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
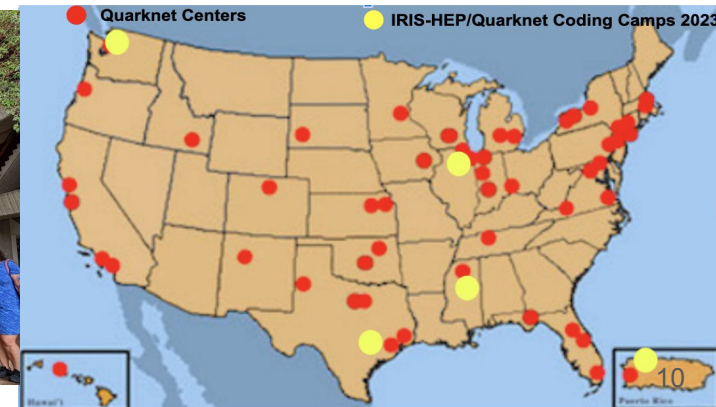- Breaks barriers and enables diversity
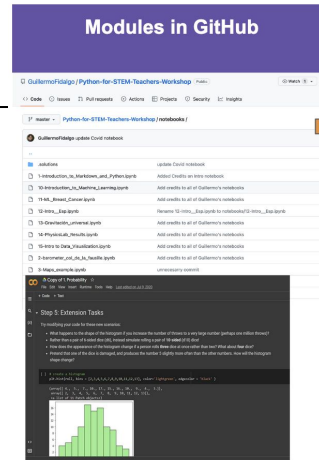
Why you should contribute?

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-ybmNW55kJgsgJ81Kadqw4g

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!

@hepsoftfound   @hsf-training   hepsoftwarefoundation.org

we do a little hacking