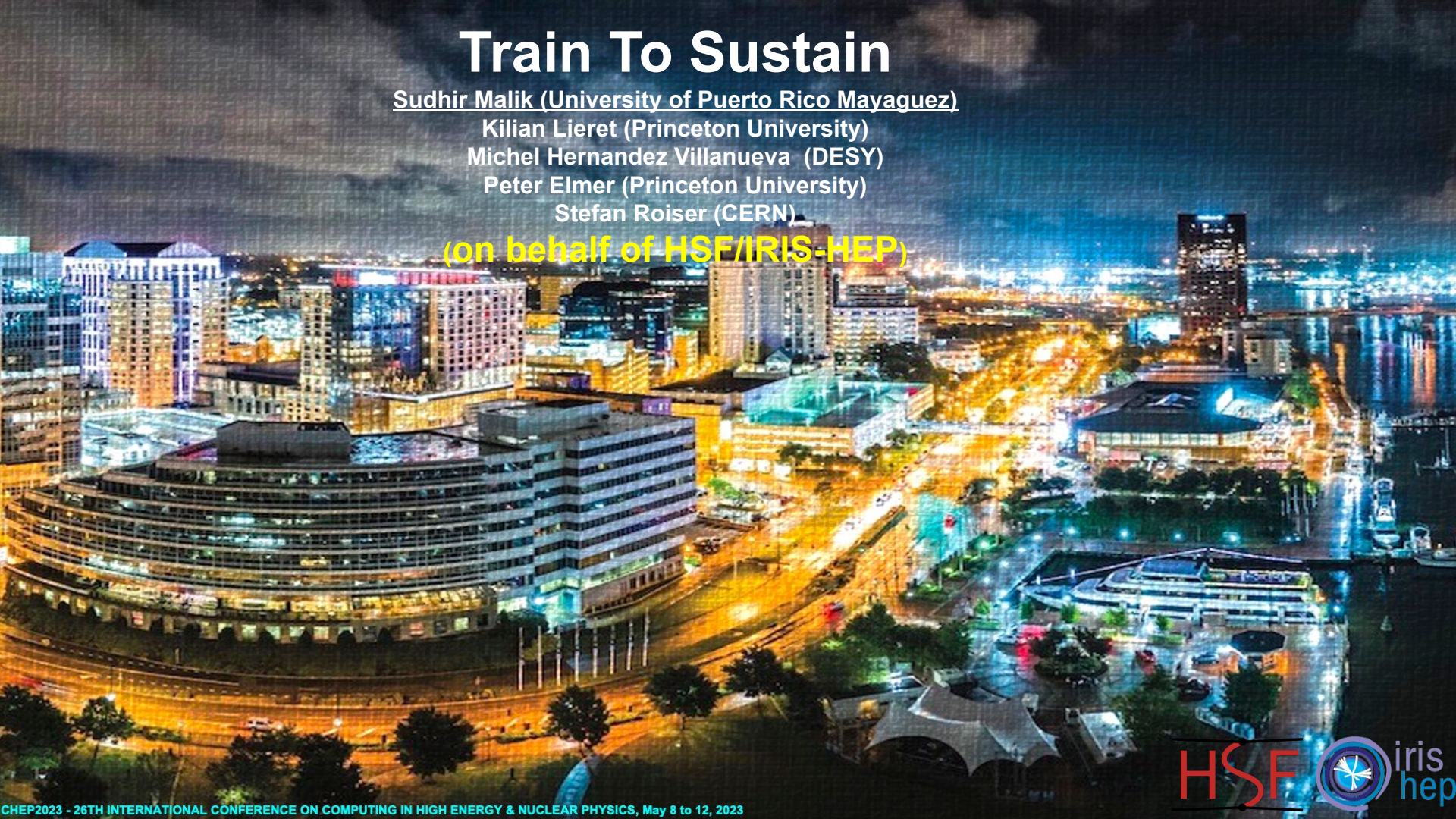
Train To Sustain

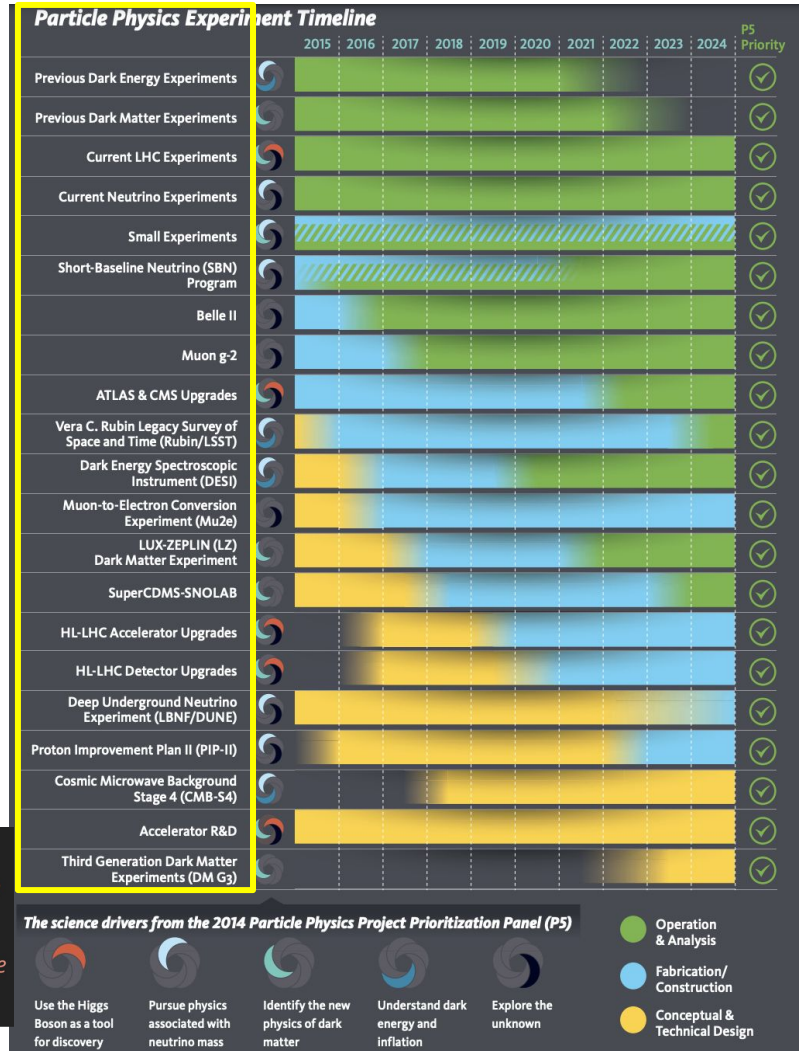
Sudhir Malik (University of Puerto Rico Mayaguez)
Kilian Lieret (Princeton University)
Michel Hernandez Villanueva (DESY)
Peter Elmer (Princeton University)
Stefan Roiser (CERN)
(on behalf of HSF/IRIS-HEP)



CHEP2023 - 26TH INTERNATIONAL CONFERENCE ON COMPUTING IN HIGH ENERGY & NUCLEAR PHYSICS, May 8 to 12, 2023

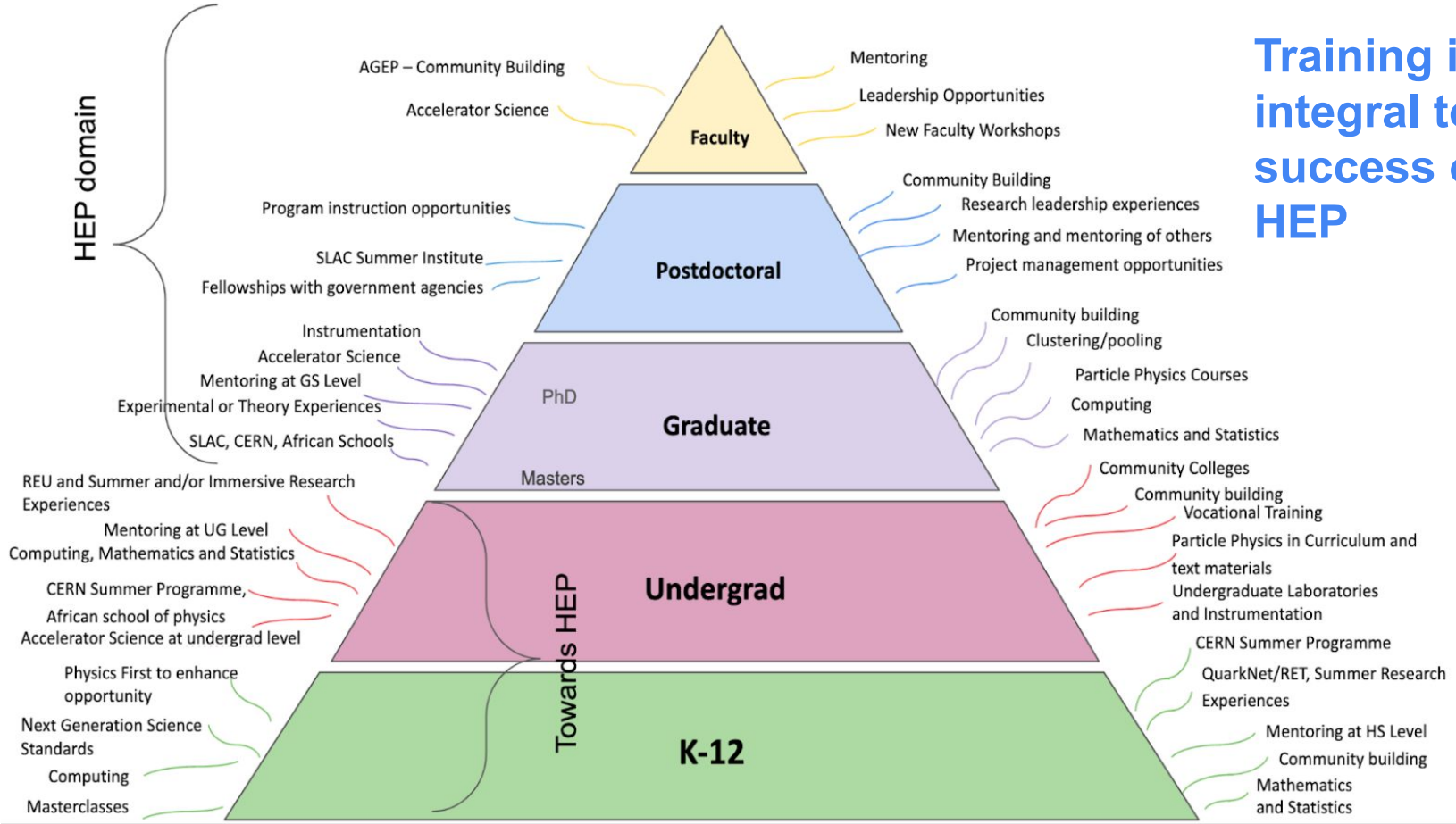
Context

- 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, disruptive changes (COVID, architecture, ideas)
- Investment in organised training (hands-on)
  - Mitigate some of the above challenges
  - Build future workforce
  - Careers in HEP or other STEM areas
- Organised Software Training is essential



- *Particle Physics Propels U.S. Progress*
- *Particle Physics is Discovery Science*
- *Particle Physics Builds STEM Leaders*
- *Particle Physicists Deliver Discovery Science*
- *Particle Physics and Quantum Information Science*
- *Particle Physics Progress and Priorities*
- *Particle Physics Makes a Difference in Your Life*
- *Particle Physicists Value Diversity*
- *Particle Physicists Advance Artificial Intelligence*
- *Particle Physics Institutions*

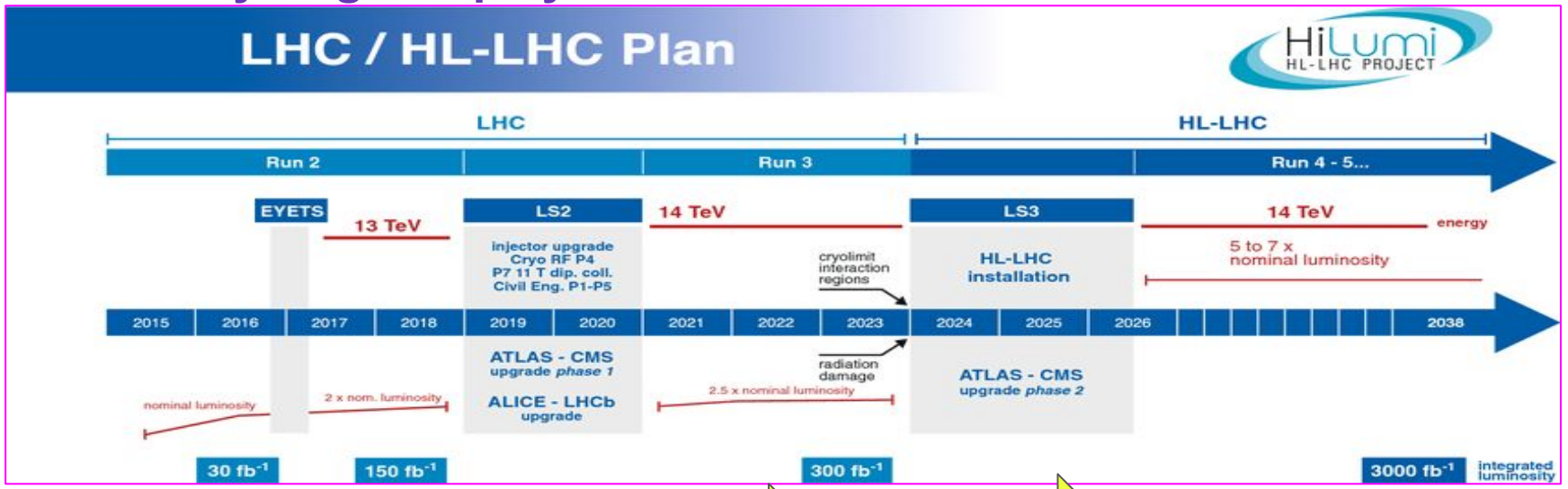
Workforce Pipeline



Training is integral to the success of HEP

Training Efforts

Several synergistic players and drivers



FIRST-HEP

HSF-India

IRIS-HEP institute (Inst. for Research and Innovation in Software in HEP)

SIDIS - Software Institute for Data Sciences

SWIFT-HEP - SoftWare and InFrastructure Technology for High Energy Physics

HEP Software Foundation

Paradigm

- No standard curricula for HEP students exists
- Not all HEP students can attend university-offered software courses
- HEP students in many cases don't receive any programming training
- Students trained as physicists but asked to be data analysts

Democratize science by making software prerequisites accessible to everyone

Experiments need **Cyberinfrastructure professionals** and lifelong learners

We need a unified, scalable, and sustainable software training framework powered by the entire community

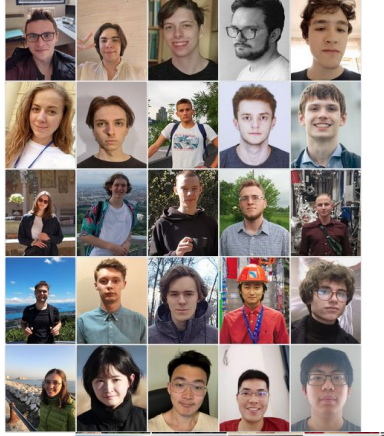
HSF HEP Software Foundation

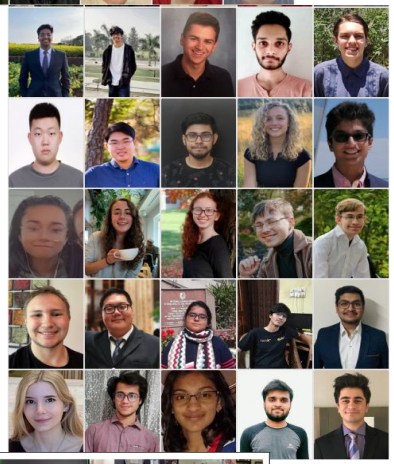
HSF/IRIS-HEP is leading training efforts

iris hep

Software Training

- Software training hub for new researchers in
  - High Energy Physics
  - Related communities - Nuclear, Neutrino, Astro,Theory
- Skills are essential
  - To produce high-quality and sustainable software needed to do the research, solve future challenges
- Thousands of users in the community
  - Sustainability (challenging) is the centerpiece of its approach
- The training modules are
  - Open source - GitHub, Slack, Websites, Indico, youtube videos
  - Enable technical continuity, collaboration and nurture the sense to develop software that is reproducible and reusable
- Made huge input impact to Snowmass 2021 process on Community Engagement Efforts
- Training Scientists, Postdocs, Graduate Students, Undergrads
- Broader Impacts - Training High School Teachers, diversity
- Pivotal Role in making training integral of HEP future





Scalability and Sustainability

GitHub

Search or jump to... ests 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 (she/her)

Monthly hackathons

HSF iris hep

✨ Containerization Training Hackathon ✨

🚀 The big goal!

we do a little hacking

Slack

Recognition

Authors

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

Meirin Oan Evans

amorenobr

Guillermo A. Fidalgo-Rodríguez

Marchela04

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 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 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

Publications and Visibility

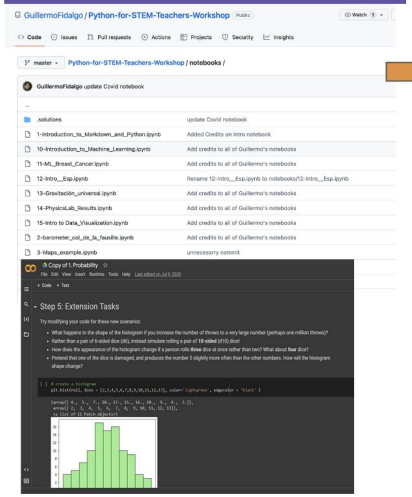
Training Talks & Papers

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 trainin
- Breaks barriers and enables diversity

Modules in GitHub

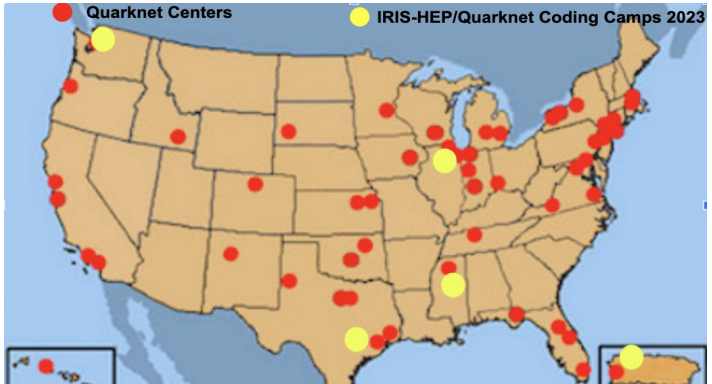


"student hat"
Engage, explore, cexplain

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

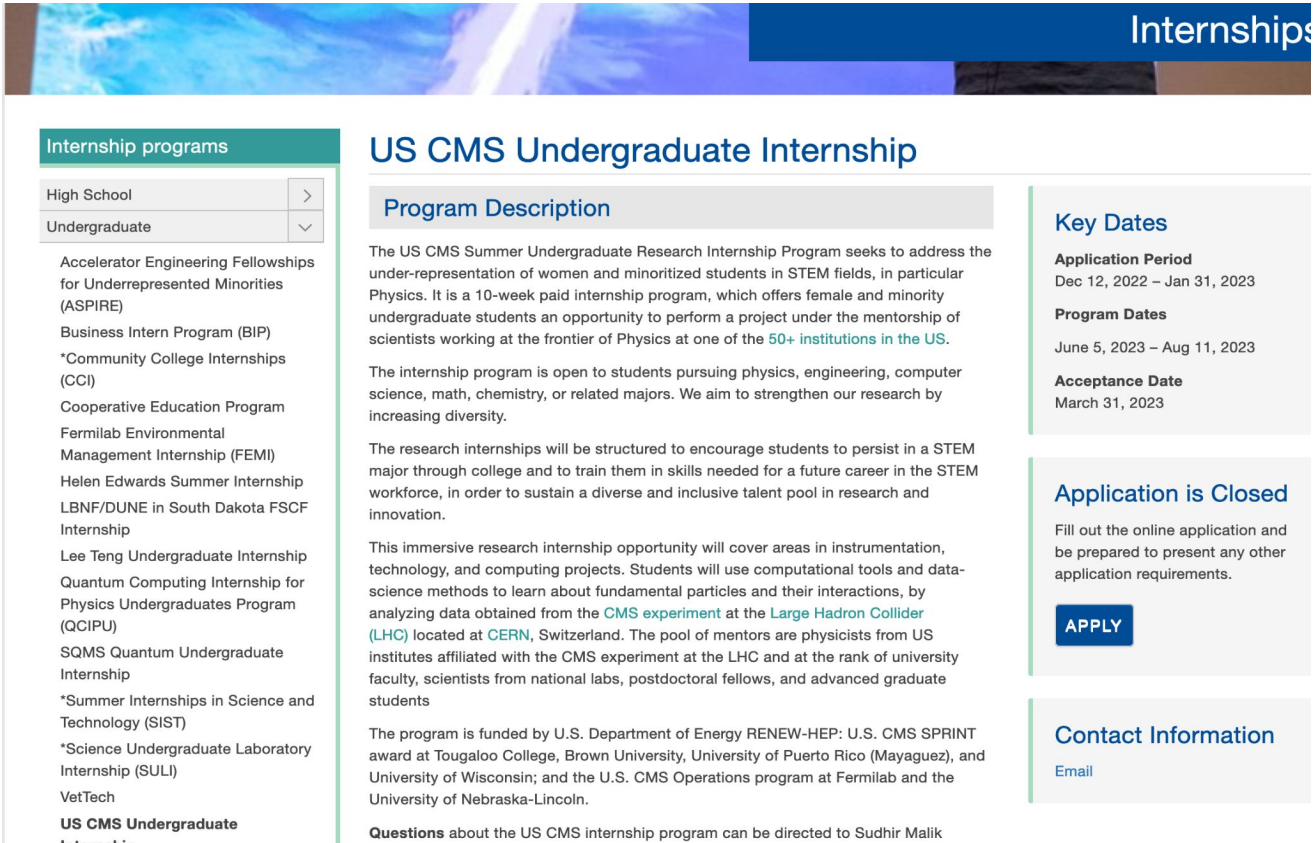
"teacher hat"
Elaborate and Evaluate

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



USCMS Internship Program

- Gives minoritized, MSI and HBU students opportunity for HEP tools
- Software Training Curriculum provided by HSF/IRIS-HEP



Please visit related talks (at CHEP2023)

Training and on-boarding initiatives in HEP	Allison Reinsvold Hall (US Naval Academy)	8 May 2023, 11:45 15m Marriott Ballroom I (Norfolk Waterside Marriott) Oral Track 8 - Collaboration, Reinterpretation, Outreach and Education Track 8 - Collaboration, Reinterpretation, Outreach and Education
Train to Sustain (This Talk)	Sudhir Malik (University of Puerto Rico Mayaguez)	8 May 2023, 12:15 Chesapeake Meeting Room (Norfolk Waterside Marriott) Oral Track 5 - Sustainable and Collaborative Software Engineering Track 5 - Sustainable and Collaborative Software Engineering
Building a Global HEP Software Training Community	Kilian Lieret (Princeton University)	9 May 2023, 17:30 Marriott Ballroom I (Norfolk Waterside Marriott) Oral Track 8 - Collaboration, Reinterpretation, Outreach and Education Track 8 - Collaboration, Reinterpretation, Outreach and Education
Software Training Outreach In HEP	Cordero, Danelix (CROEM High School, Mayaguez, PR)	11 May 2023, 11:45 Marriott Ballroom I (Norfolk Waterside Marriott) Oral Track 8 - Collaboration, Reinterpretation, Outreach and Education Track 8 - Collaboration, Reinterpretation, Outreach and Education

Thank you to the organisers for the opportunity to give this talk

Thank you to all contributors to HSF/IRIS-HEP Training

https://hepsoftwarefoundation.org/training/community.html