

Analysis Preservation Training

Containerization + CI/CD

Jan 16 - 20, 2023

<p>Instructors (in the recordings):</p> <ul style="list-style-type: none">• Docker: Brendan Regnery• Singularity: Marco Mambelli, Kilian Lieret Wouter Deconinck, Michel Villanueva, Aman Desai• GitHub: Emery Nibigira• GitLab: Giordon Stark	<p>Mentors:</p> <ul style="list-style-type: none">• Aman Desai• Aman Goel• Emery Nibigira• Giordon Stark• Guillermo Fidalgo• Jaydip Singh• Kilian Lieret• Marco Mambelli• Matt Snyder• Michel Villanueva• Wouter Deconinck	<p>Local organising committee:</p> <ul style="list-style-type: none">• Sudhir Malik• Michel Villanueva• Wouter Deconinck• Kilian Lieret
--	---	---



**If you aren't recording this on Zoom,
enable captioning and start
the recording ...**

(just a reminder)

Everyone is Welcome

- You are physicists working in international collaborations. All of you should know this page:
 - [The CERN code of conduct](#)
- Built on a set of core CERN values →
- Taken together, provide the basis for respect: respect for others, respect for the organization and respect for its mission.
- We encourage a culture of openness where all contributors feel free to engage in the discussion.



What is Analysis Preservation?

- Methodology intended to ensure that materials produced in the analysis process
 - Datasets
 - Software
 - Documentationare reusable and understandable in the future.
- **Data is expensive. Experiments are unique.**
 - Preserving the knowledge around them is a must.
- Analysis preservation is work-in-progress in most of the experiments.
 - Many technical challenges involved.
 - We will not focus in a particular preservation scenario.
 - Instead, we will review the tools useful for preservation.

What are we learning this week?

- We will take a quick tour learning the basic functionality of tools popular in analysis preservation.

- **Containerization technologies**

- Docker

- Singularity/Apptainer

- **Continuous Integration/Deployment (CI/CD)**

- GitLab pipes

- GitHub actions



Monday

Welcome

Kickoff/Orientation
[15 CET]

Preservation@CMS
[15:30 CET]

Help with Setup
[16-17 CET]

Tuesday to Thursday
Work on your own, when you want

Watch and work through tutorials:
[Indico Agenda](#)

Friday
Hands-on sessions

Block 1:
[10-12 CET]

Block 2:
[13-15 CET]

Block 3:
[17-19 CET]

Block 4:
[21-23 CET]

Monday

Welcome

Tuesday to Thursday

Work on your own, when you want

Friday

Hands-on sessions

Kickoff/Orientation
[15 CET]

Preservation@CMS
[15:30 CET]

Help with Setup
[16-17 CET]

Get on the same page with logistics and debug **initial setups/installations.**

Watch and work through tutorials:
[Indico Agenda](#)

Block 1:
[10-12 CET]

Block 2:
[13-15 CET]

Block 3:
[17-19 CET]

Block 4:
[21-23 CET]

Monday

Welcome

Tuesday to Thursday

Work on your own, when you want

Friday

Hands-on sessions

Kickoff/Orientation
[15 CET]

Preservation@CMS
[15:30 CET]

Help with Setup
[16-17 CET]

Watch and work through tutorials:
[Indico Agenda](#)

Block 1:
[10-12 CET]

Block 2:
[13-15 CET]

Block 3:
[17-19 CET]

Block 4:
[21-23 CET]

Work through all of the content here and learn/work at your own pace with **our virtual support on Slack**.

Channel: [#2301-analysis-preservation](#)



Monday
Welcome

Tuesday to Thursday
Work on your own, when you want

Friday
Hands-on sessions

Kickoff/Orientation
[15 CET]

Preservation@CMS
[15:30 CET]

Help with Setup
[16-17 CET]

Watch and work through tutorials:
[Indico Agenda](#)

Sign up for mentoring sessions

Deadline: Wed. 4 pm (CERN), 10 am (ET), 11pm (Peking) We will assign you to one of the sessions afterward

Join the room indicated for your specific hands-on session.

Block 1:
[10-12 CET]

Block 2:
[13-15 CET]

Block 3:
[17-19 CET]

Block 4:
[21-23 CET]

If you haven't done yet...

[1] Join the Slack channel: [invite](#)

If you have troubles to join, let us know now.

[2] Follow the setup pages

- Docker ([setup](#))
- Singularity/Apptainer ([setup](#))
- CI/CD with Github ([here](#)) or GitLab ([here](#)).

[3] Take a look at the [Analysis example with CMS open data](#).

[4] [Sign up for mentoring sessions](#)

Deadline: [Wed. 4 pm \(CERN\), 10 am \(ET\), 11pm \(Peking\)](#)

We will assign you to one of the sessions on Wednesday afternoon (ET)



**Meet your
mentors!**



Aman Desai

My research:

Searches for New Physics Beyond the Standard Model of Particle Physics

My expertise is:

Physics analysis, Machine learning techniques

A problem I'm grappling with:

Symbolic calculation with Python

I've got my eyes on:

Statistical analysis in Particle Physics, Monte Carlo Simulations

I want to know more about:

SUSY, Extra Dimension theory, Detector Simulation



Emery Nibigira

Research Associate, U. of Tennessee

My research: *Search for exotic long-lived particles using CMS detector and Tracking detector operation and performance*

My expertise is: High Energy Physics (HEP) Data analysis, Tracking detector, Docker, Continuous Integration and Deployment (CI/CD).

A software and computing problem I'm grappling with: Analysis code performance, Enhance sensitivity for the discovery of Beyond Standard Model particles.

I've got my eyes on: New tools in High Energy Physics!

I want to know more about: How to improve time management at work.



Guillermo Fidalgo Rodríguez

Masters Student in Physics
University of Puerto Rico - Mayagüez
gfidalgo@cern.ch

My research:

Analysis on Exotic Physics: Emerging Jets.
I do Machine Learning for DQM Studies for the CMS Tracker.

My expertise is:

Using python for ML Studies

A problem I'm grappling with:

Designing a new HLT Trigger for analysis

I've got my eyes on:

A Framework laptop for replacing my old pc.

I want to know more about:

C++, Arduino

How to efficiently teach a dog new tricks





Jaydip Singh

Ph.D. Student at University of Lucknow, India , working with DUNE, SAND and INO - collaboration.

My research:

Reconstruction of neutrino energy to do Neutrino Astronomy, WIMP detection and Neutrino-Nucleus Interactions physics.



My expertise is:

Developing algorithm for particle energy reconstruction at LArTPC.

A problem I'm grappling with:

Muon Energy reconstruction and neutrino astronomy with the DUNE far detector.

I've got my eyes on:

AI algorithms for the reconstruction of particle interaction information.

I want to know more about:

Application of AI algorithm for the particle detection and new physics.





Kilian Lieret

*Postdoc at Princeton University,
IRIS-HEP, CMS experiment*

My research: Charged particle tracking with graph neural networks ([github](#))

My expertise is: Scientific python stack, graph neural networks pytorch geometric, hyperparameter optimization

A software and computing problem I'm grappling with: clustering in learnt latent spaces

I've got my eyes on: probabilistic programming frameworks and applications

I want to know more about: applications of graph neural networks in physics and beyond





Marco Mambelli

*Software Developer at Fermilab
in Batavia, IL*

My research:

I work on workflows and distributed computing system.

In particular the GlideinWMS and HEPCloud projects that are used to run all analyses and simulations for CMS, and most Fermilab experiments.

My expertise is:

Distributed scientific computing, coding and system engineering

A problem I'm grappling with:

Efficient use of GPUs on supercomputers

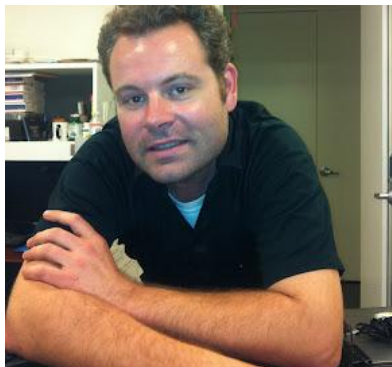
I've got my eyes on:

New tools to ease collaboration

I want to know more about:

Quantum computing





Matt Snyder

Senior Technology Analyst, Brookhaven National Laboratory

My expertise is:

observability (elastic, grafana); git

A problem I'm grappling with:

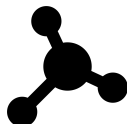
docker-compose and podman-compose reconciliation

I've got my eyes on:

prometheus, loki

I want to know more about:

singularity, gitlab ci/cd





Michel Villanueva

*Research Fellow, DESY
Working in tau lepton physics
and Distributed Computing at Belle II*

My research: *Precision measurements
with tau leptons*

My expertise is: Data analysis in distributed computing environments.

A software and computing problem I'm grappling with: Scalability of the Belle II analysis workflow in the high-luminosity scenario.

I've got my eyes on: Sustainable operation of the grid. Training newcomers and get fresh ideas!

I want to know more about: Columnar analysis, its capabilities and potential use cases in our experiment.





Wouter Deconinck

Associate Professor, University of Manitoba, Canada

My research:

Electroweak precision measurements:

- ep (proton's weak charge, $\sin^2\theta_w$, Electron-Ion Collider),
- ee (electron's weak charge, MOLLER),
- neutron beta decay (beam lifetime at NIST)

My expertise is:

Parity-violating electron scattering, user-centered design, agile management, system integration

A problem I'm grappling with:

How to develop a future-proof software environment for a large experimental collaboration

I've got my eyes on:

Acceleration on heterogeneous systems

I want to know more about:

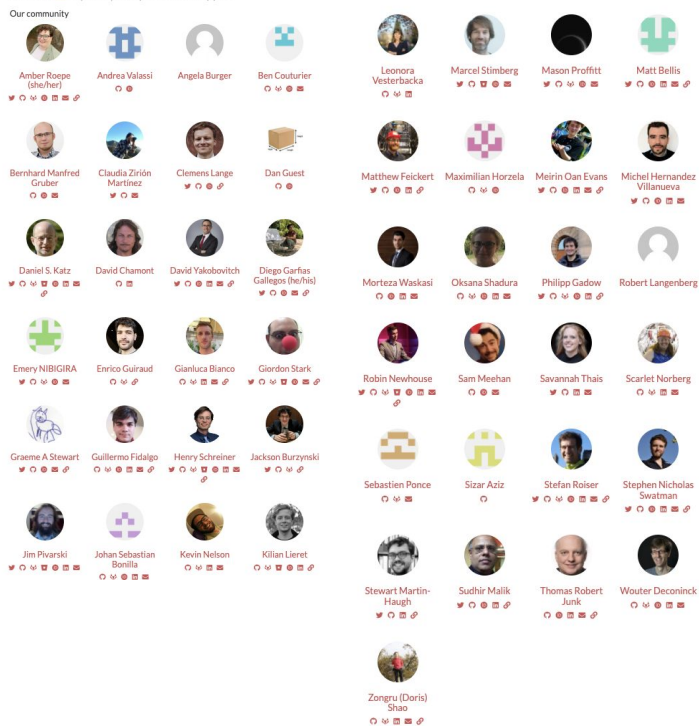
Multi-architecture container best practices



HSE Educator/Mentor community

<https://hepsoftwarefoundation.org/training/educators.html>

Here is the list. Section for a special thank you for everyone who made our workshop possible.



- **Join our hackathons**
 - **In 1 or more topics**
- **Join our community**
- **Become training Educator**
 - **a mentor, facilitator, instructor**
- **Open a PR for any of our modules**

Stay Tuned! Follow us on Twitter

<https://twitter.com/HSFTraining>

- **Training events**
- **New material**
- **Sharing job offers!**



HSF
"ftware Four"

... 🔔 Following

HSF Training
@HSFTraining

Software training group of [@hepsoftfound](#). Teach researchers to create sustainable software with sustainable, scalable training efforts!

📍 Interwebs 🔗 hepsoftwarefoundation.org/training
📅 Joined September 2022

Group picture!