# News from the HSF and "international direction of travel"
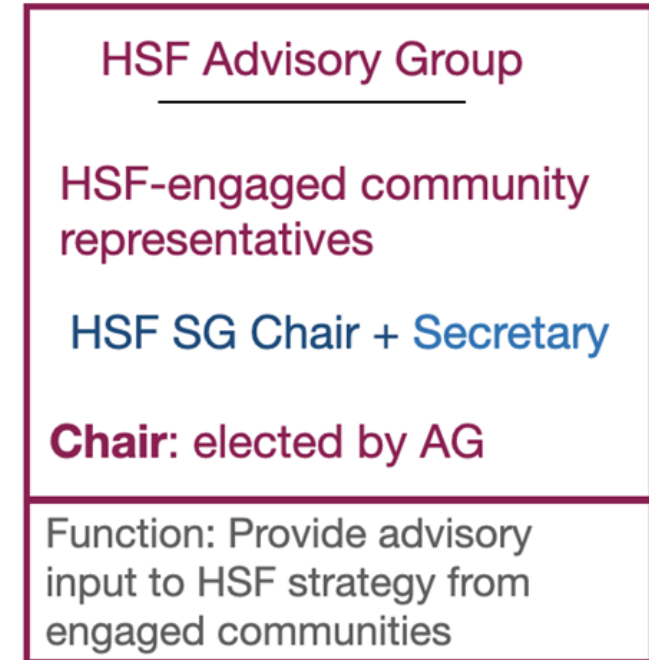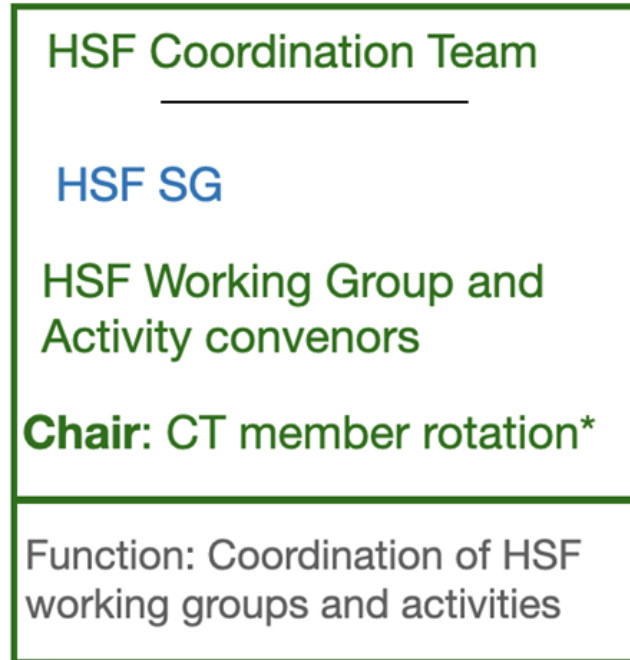
**Eduardo Rodrigues**
**University of Liverpool**

**SWIFT-HEP Workshop, Warwick University, 11th-12th November 2024**

# Note:

❑ **Largely an executive summary of recent reports on CERN SFT and HSF / community activities to the WLCG Management Board and the LHCC**

❑ **And summary from HSF own work / discussions**

- Proposals post the feedback at the [May DESY WLCG/HSF workshop](#):
  - Implementations followed over Summer and Autumn …



**HSF Coordination Team**

HSF SG

HSF Working Group and Activity convenors

**Chair:** CT member rotation*

Function: Coordination of HSF working groups and activities

*There is no standing chair for the HSF Coordination Team meetings which are focused on running the day to day work of the HSF

**HSF Steering Group**

**Chair:** elected by SG

Function: Decision-making body responsible for HSF strategy

**HSF Advisory Group**

HSF-engaged community representatives

HSF SG Chair + Secretary

**Chair:** elected by AG

Function: Provide advisory input to HSF strategy from engaged communities

# HSF evolution/organisation: Steering Group

- SG got formed over Summer and met several times

- **SG members get roles assigned**, in particular to follow at least one of the Activity Areas and discuss plans for the future
  - Working Groups have all been renamed to Activity Areas

- **Various SG actions** will be detailed in the next slides ...
  - Set up of the Advisory Group
  - First steps on HSF Affiliated Projects and Software
  - HSF input to the European Particle Physics Strategy Update

- RECAP - **mailing list for the SG** is hsf-steering@googlegroups.com

# HSF evolution/organisation: Advisory Group

- **Mandate** and composition described in a <u>public document on the HSF website</u>, following circulation for feedback via the HSF Forum

## Mandate

The purpose of the Advisory Group (AG) is to represent HSF Engaged Communities (ECs) that contribute to HSF activities, in order to provide strategic input to the HSF on their long-term strategies and needs. The AG *is an advisory group* and *not* a decision-making entity; decisions on HSF strategy are taken by the HSF Steering Group (SG). Engaged Communities are envisaged to cover major HEP experiments (including the LHC experiments) and WLCG.

- The **AG is being formed** as we speak
  - The SG received a positive response for membership from the LHC Experiments, WLCG, DUNE, Belle II and MCnet

- **First meeting foreseen for early 2025**

# HSF evolution/organisation: HSF Affiliated Projects and Software

- The proposal documents on "**HSF Affiliated Projects and Software**" and "**Affiliated Projects and Software Guidelines**" have been circulated within the community and finalised. They are now available on the HSF website
  - https://hepsoftwarefoundation.org/projects/affiliated.html
  - https://hepsoftwarefoundation.org/projects/guidelines.html

- RECAP:
  - All affiliated projects have to commit to software sustainability best practises
  - List of affiliated projects (to be) hosted in dedicated web area
  - GitHub badges will express the level of maturity, community support, engagement and implementation of best practises, etc. Example:

- **First reviews !**
  - The prmon package was the first project to be reviewed
    - Public process via a GDoc, advertised via the HSF Forum
  - More in the pipeline :)

# HSF seminar series

- This is a **significant change in the way we organise activities**
  - Let's see what experience tells us - goal is to have these as regular HSF events
- Seminars started officially in October

- Oct. 1st: First seminar on
  "Julia in high-energy physics: a paradigm shift or just another tool?"
- Followed by a second seminar on "Julia as a Statically-Compiled Language" on Oct. 3
  - Note: recordings are available for these seminars
- Both timely with the JuliaHEP workshop that took place that week

- **Future seminars** will appear here

# HSF Training Activity: the Training Center

- Be sure to check the HSF Training Center currently listing over 20 modules!

- Note: Universal praise at the WLCG-HSF workshop in DESY
for the amazing work done
by the Training team, an example
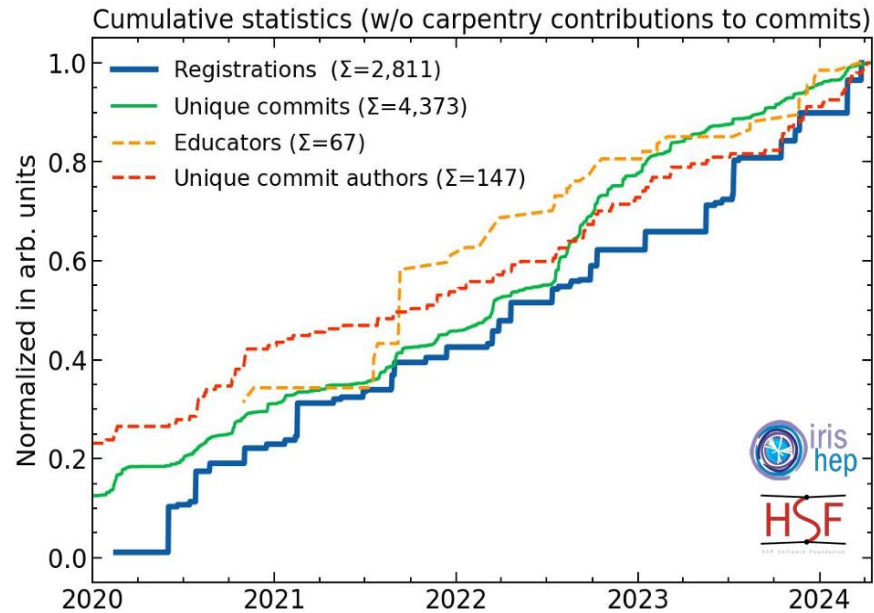
# HSF Training Activity: analysis of event data

- Training Activity Area performed a **study to better understand who uses the material, which material, etc.**
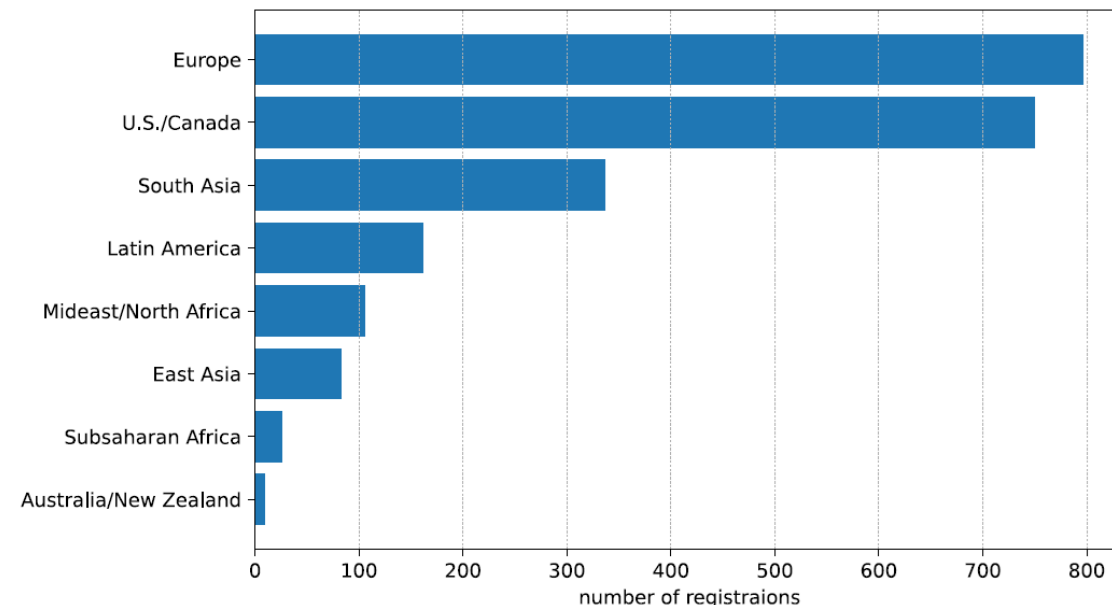  - Also to make better decisions about what events to plan, how, and when.
  - Done as part of the pre-CHEP 2024 Workshop
  - See the full talk
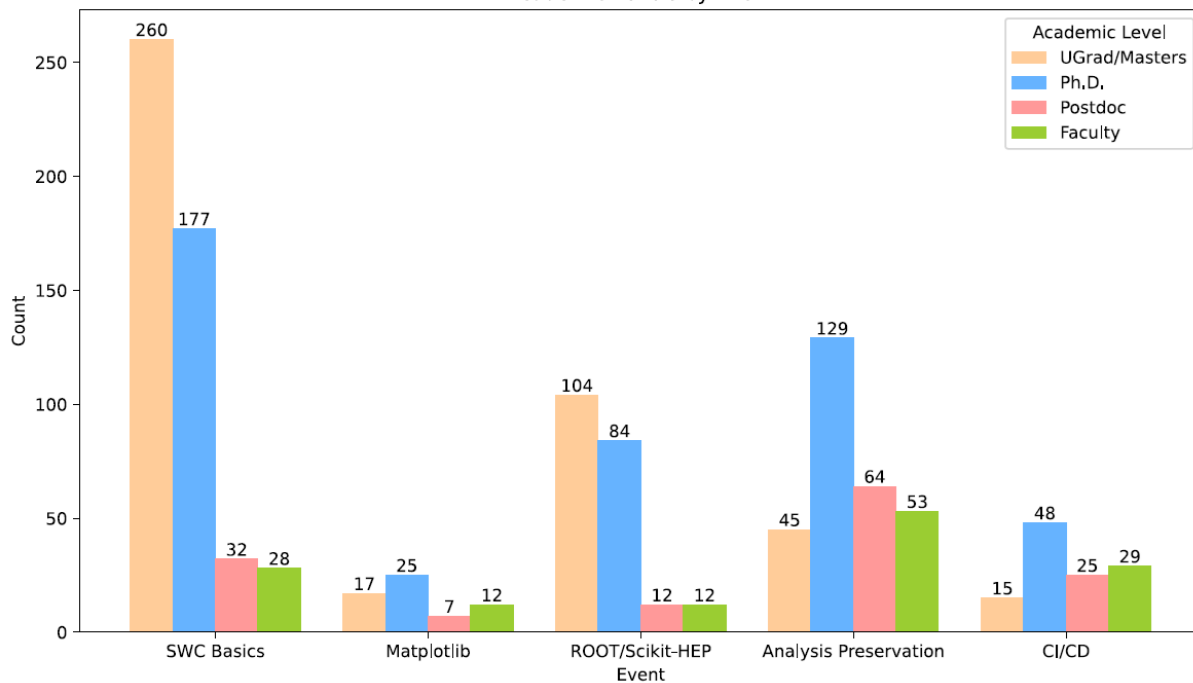
**From Indico registrations:**



**From surveys:**



**Sources of data**

- Indico registrations: 24 events since 2019
- Pre- and post-event surveys: students are required to fill these out (Zoom password is at the end of the pre-event survey), 17 events since 2021

  Surveys have been analyzed before, but not across all events.
- Website traffic: anonymized, but can identify unique visitors *each day*, since May 2023
- Zoom and Slack activity for the February 2024 "Analysis Pipelines" event (at least)

3 / 13

# HSF and connected activities: European PP Strategy Update

- [European Particle Physics Strategy Update](#) due next year:
  - Inputs from the community by March 31st 2025
  - Open Symposium 23-27 June 2025

- The HSF has a strong track record of submitting inputs to these strategic deliberations

- **HSF SG started the preparatory work towards a contribution** on the critical areas of event generation, simulation, reconstruction, analysis and training and careers
  - Will the main focus of the SG in the coming months (aside HSF reviews of projects)

- Contact established with concerned Activity Areas to ensure community involvement in the preparatory work by the end of the year (Event generation, Detector simulation, Reconstruction & software triggers, Data analysis, Training)
  - There will be Activity meetings organized to discuss and scope the goals of the contributions
- **Plan to have a draft public and open for comments end of January**

- **Endorsement by the community**: signature by individuals but also discussion about an endorsement at the level of the experiments (the latter having been suggested to us)

# HSF and connected activities: workshops

- PyHEP workshops - PyHEP 2024 and PyHEP.dev 2024, see next slides …

- JuliaHEP 2024 @ CERN 30 September - 4 October
  - 2 days of plenary, 3 days of hackathons in IdeaSquare (See next slide …)

- Oct. 19-20: HSF Training community event - session as a pre-CHEP 2024 Workshop
  - Talks from HSF-India, LHCb, EVERSE, DUNE, ErUM-Data-Hub, etc.

- Next WLCG/HSF joint Workshop will be hosted by IJCLab in Orsay next May 5-9 2025
  - Preparations will start soon
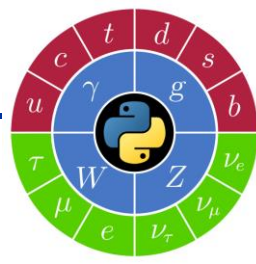
- https://indico.cern.ch/e/PyHEP2024
- **Online event**, 400 registrants but sessions saw 40-80 people at any given time
  - People largely connect for a single session or even presentation:
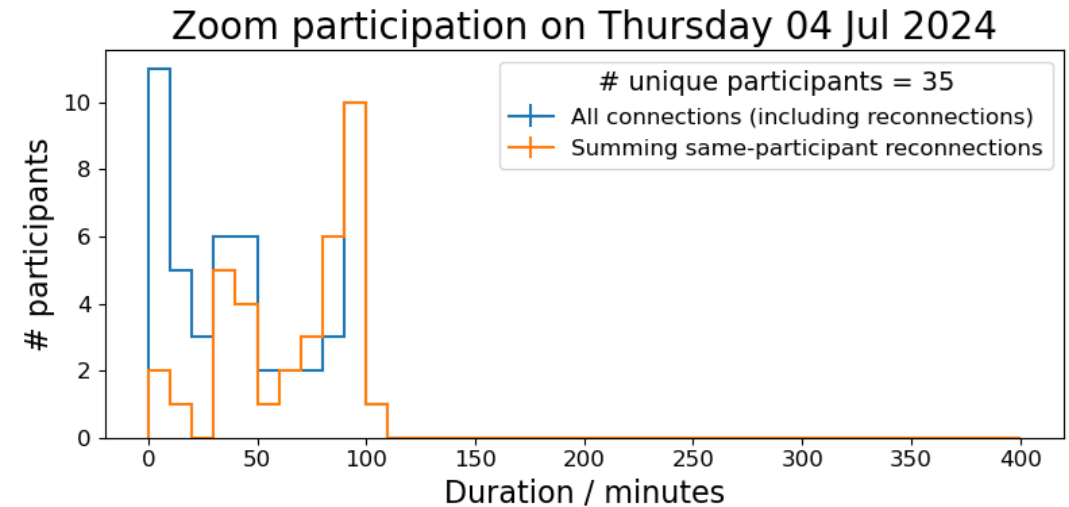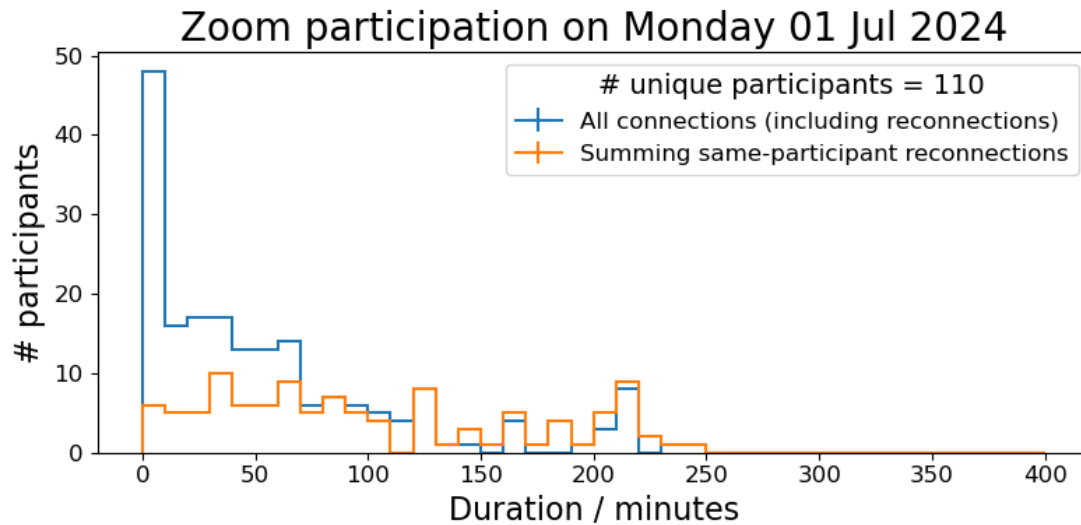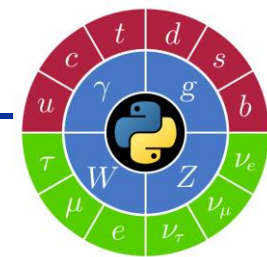
- As always, event hold a rather diverse set of presentations on many topics
  - Good selection of talks from well known and new packages, all on YouTube's dedicated channel
- **Excellent representation from many areas/experiments/projects**



PyHEP 2024 - physics areas of registrants

# collaboration affiliations = 417
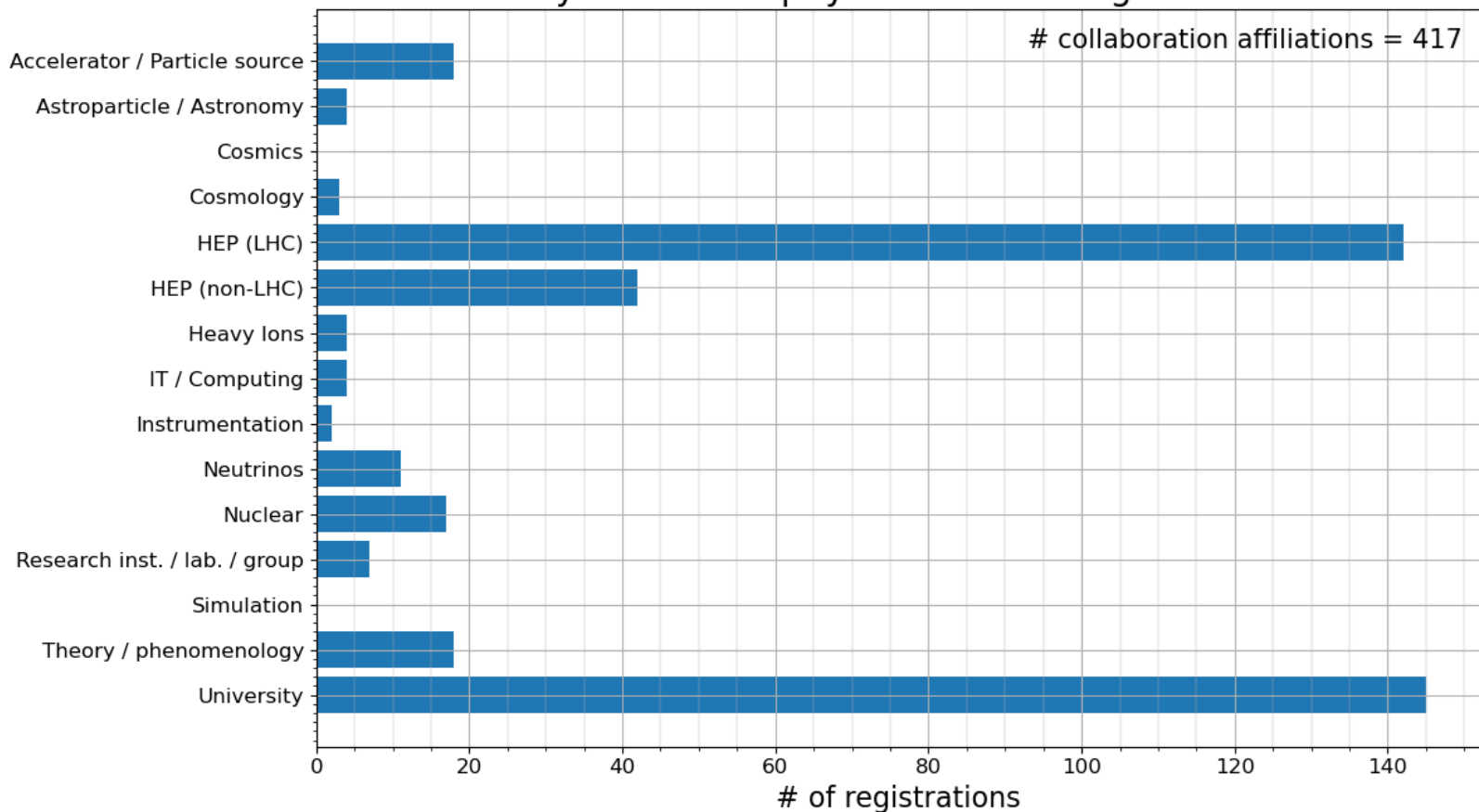


PyHEP 2024 - collaborations represented
HEP (LHC) [142 registrants]

# PyHEP.dev 2024 "Developers" workshop, Aug 26-30
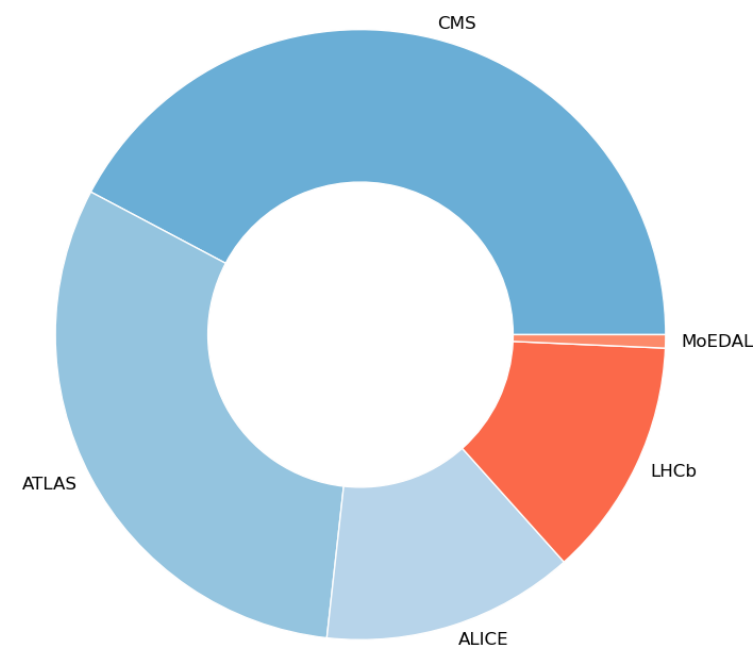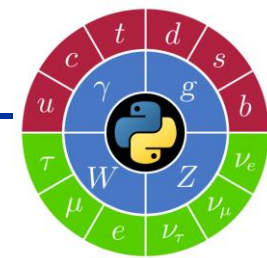
- https://indico.cern.ch/e/PyHEP2024.dev
- **In-person event** in Aachen, Germany
- 26 attendees

- Grand ideas:
1. Ensure developers work and collaborate towards software products that tie well, enhance each other and interoperate together as a coherent ecosystem
2. Plan a coherent roadmap and make priorities for the upcoming year

# PyHEP.dev 2024 "Developers" workshop, Aug 26-30

- **Main topics of focus**, in presentations and/or discussion sessions:
  - What is a HEP analysis? What does PyHEP cover?
  - Challenges - IRIS-HEP's Analysis Grand Challenges & beyond, their recent (May) 200Gbps challenge
  - Building and evaluating likelihoods
  - Statistical models serialisation
  - Analysis workflows
  - Histogramming
  - RDataFrame/Coffea analyses (at scale)
  - Future of PyHEP.dev

- A live Google Doc on "Questions and Discussions" is available from (the top of) Indico

- A **workshop report** (executive summary) started being written by all on the last day of the workshop
- **Published on the ArXiv**

# JuliaHEP 2024
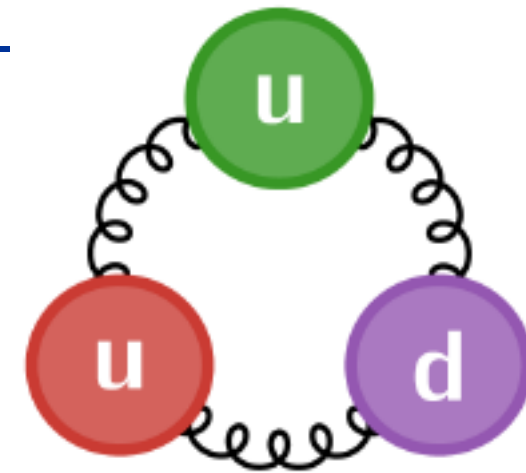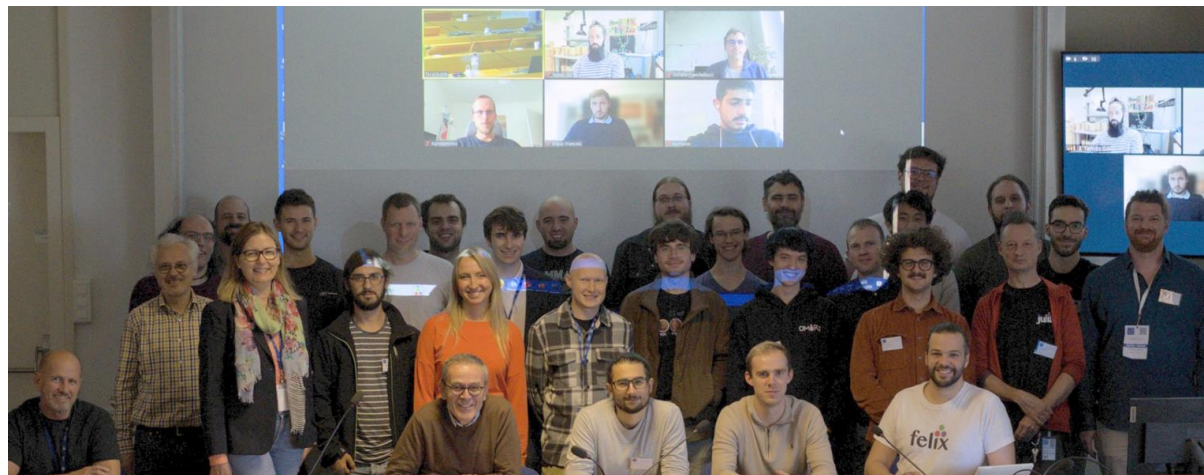


- **The second JuliaHEP workshop was organised at CERN 30 Sep - 3 Oct**
  - 40 in person participants and 40 remote
- **Reviewed the progress of Julia as a promising language for HEP**
  - Including the first HSF Seminar: <u>Julia in high-energy physics: a paradigm shift or just another tool?</u> (Uwe Hernandez Acosta)
  - Considerable progress in access to HEP data: reading ROOT files (including RNTuple), access to Geant4 and Pythia8, jet reconstruction in Julia
- **Two days plenary followed by 3 days hackathon in the IdeaSquare**
  - Focus on interfaces to basic types (Lorentz vectors, particles) as well as ROOT wrappers

# HEP-CCE

- **DOE Project started its Phase 2** earlier this year
  - *Mission: exploit HPC systems to increase HEP physics reach*
  - The 7th biannual all-hands meeting took place in July 2024
  - Results from Phase 1 presented at CHEP 2023
    - Notably completed study of portable parallelization methods for HEP

- **Phase 2 new directions**
  - Portable workflows, with an eye to DOE's Integrated Research Infrastructure initiative
  - Exploit HPC to address projected storage requirements of next generation experiments
    - Intelligent Compression
    - Columnar Data
      - Organized ROOT RNTuple API review and feedback with ATLAS/CMS/DUNE
  - Scalable Machine Learning
    - Exploit modern GPU-based HPC for more challenging training and optimization workflow
      - Simulation-based Inference
      - Quantization-aware Training
      - Large graph networks, etc
  - MC simulation on GPU (Celeritas, Photon Optical Transport, MadGraph, Pepper, etc.)

# IRIS-HEP

- **IRIS-HEP was renewed** for a second 5 year term, and will thus be active until at least 2028

- The **200 Gbps Analysis Challenge** first presented at the WLCG-HSF joint workshop at DESY in May continues to be key in putting together software components and infrastructure and testing them at scale
  - This is of particular interest to SWIFT-HEP

- The annual IRIS-HEP Institute Retreat took place on 4-6 September at the Univ. of Washington (Seattle).
- The goals of the retreat were to:
  - Checkpoint the status of the IRIS-HEP efforts to date (through Year 6 of the project)
  - Clarify the gaps between where we are now and what will be needed for the HL-LHC startup
  - Update our plans for delivery of IRIS-HEP products to our partners (experiments, ops programs and the LHC and HEP community)
  - Elaborate a vision for what IRIS-HEP can achieve in Year 7 and beyond

# IRIS-HEP

- *Recent (many!) results shown at CHEP are a great summary of recent work.*

- Recent work connected with **Analysis Facilities Design and Optimization**
  - Plenary talk on the 200 Gbps Demonstrator by Alex Held
  - The CMS Instance of Coffee-Casa for the 200 Gbps Demonstrator by Oksana Shadura
  - Operating the 200 Gbps Demonstrator for ATLAS by Rob Gardner
- Also working on building a coherent **analysis ecosystem** that spans languages, summarized by the talk *Navigating the Multilingual Landscape of Scientific Computing: Python, Julia, and Awkward Array*, by Ianna Osborne.
- IRIS-HEP also had several contributions around **analysis**: *GIL-free scaling of Uproot with Python 3.13 subinterpreters* by Jim Pivarski, and *Building a Columnar Analysis Demonstrator for ATLAS PHYSLITE Open Data using the Python Ecosystem* by Matthew Feickert
- **DOMA** efforts also figured strongly in our recent work, as can be seen in *Enhancing Network Analytics through Machine Learning* by Petya Vasiliv and *CMS Token Transitions* by Brian Bockelman.
- Our recent work on **tracking** was also summarize in the talk *traccc: GPU track reconstruction library for HEP experiments* by Beomki Yeo and *Integration of the ACTS track reconstruction toolkit in the ATLAS software for HL-LHC operations* by Carlo Varni.
- Two talks given on the use of LLM's in HEP, *HEP-Help: a first-stop helpline for particle physics software* by Jim Pivarski and *Leveraging Language Models to Navigate Conference Abstracts: An Open-Source Approach* by Gordon Watts, as we continue to explore this area.
- We helped the HSF run a pre-CHEP Training Event, establishing goals and accountability for training efforts

# MCnet / Generators community

- The LPCC MC WG got set up earlier in the year

- It organised its kick-off meeting on Nov. 14
  - Presentations from the LHC experiments on their "perspectives"

# Scikit-HEP project

- **A fair amount of org packages being tested within IRIS-HEP's Analysis Grand Challenges**.
  - These materialise as enhancements and/or bug fixes, all released on PyPI and Conda

- Interest from "externals"/colleagues to join with packages/projects remains, and is excellent to see

- Also work towards welcoming a couple more developers and maintainers to improve on sustainability

# Scikit-HEP project

- [Coffea package](#) being migrated to the organisation together with [fsspec-xrootd](#)
    - Coffea = Columnar Object Framework For Effective Analysis, a library that has acted as an incubator for different design patterns for CMS and ATLAS style analyses
    - It already uses and integrates many tools in the Scikit-HEP toolset
    - Many developments in Coffea have in the past been migrated out into Scikit-HEP packages as they matured
    - Coffea has been a core part of many Pythonic analyses in CMS and is a center part of the IRIS-HEP Analysis Systems pipeline, as well as an entrypoint to the Dask ecosystem for many
    - fsspec-xrootd is an XRootD implementation for fsspec to allow fsspec to use XRootD accessible storage systems

# WLCG

❑ **The Management Board has recently created a new Technical Coordination Board (TCB)**

❑ **The mandate has been approved and the TCB composition is been prepared**

❑ **Further info to appear at https://wlcg.web.cern.ch/organisation-boards/technical-coordination-board**

# Thanks for listening !

❑ **Several activities of general interest, many of particular interest to SWIFT-HEP**

❑ **Let's talk and see what we can plan and areas where cross-engagement would be most relevant** ☺