UNIVERSITY OF LIVERPOOL

Computing & Software work and R&D

Eduardo Rodrigues, for several colleagues (see names in pages)

Particle Physics Annual Meeting, Liverpool, 22nd – 23rd May 2025
https://indico.ph.liv.ac.uk/event/1995/

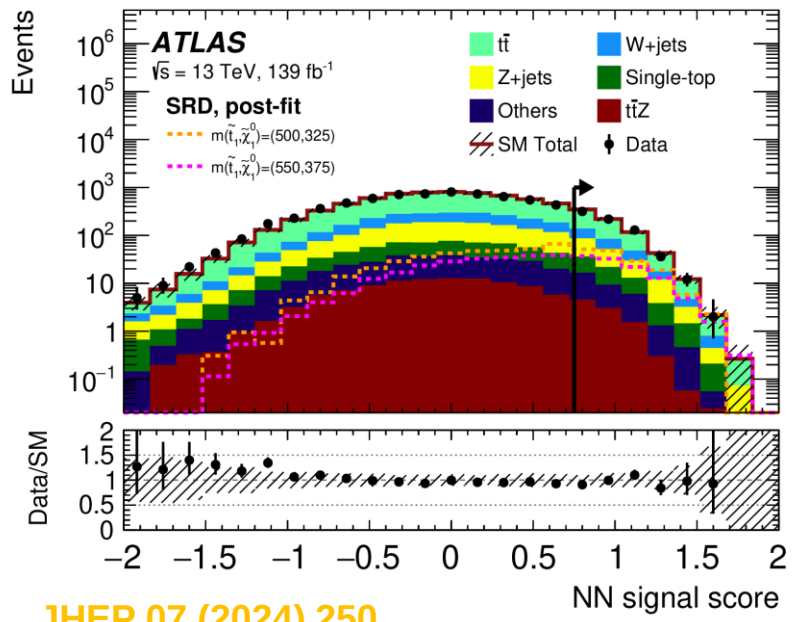
AI in ATLAS

ATLAS (as other PP experiments) has successfully leveraged AI across numerous aspects, including data analysis, theory calculations, detector calibration/monitoring, and real-time data selection.

Different set of problems to be addressed depending on the case:

Offline data analyses → searches for new physics are often a "classification" problem → **event classification** and **process discrimination** in large and diverse datasets for discovery

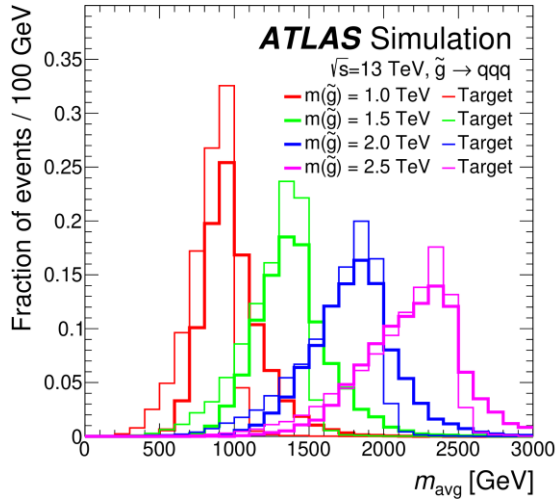
Search for top+charm+MET (arxiv:2402.12137)



JHEP 07 (2024) 250

- Measurement of ttZ with Z→nunu (Internal for now)
  - **Deep NN's trained to isolate signal (ttZ) from background based on event-level variables for potential SM (first) observation**

- GNN with attention network to reconstruct potential BSM Physics particles - JHEP 05 (2024) 003



Ongoing work

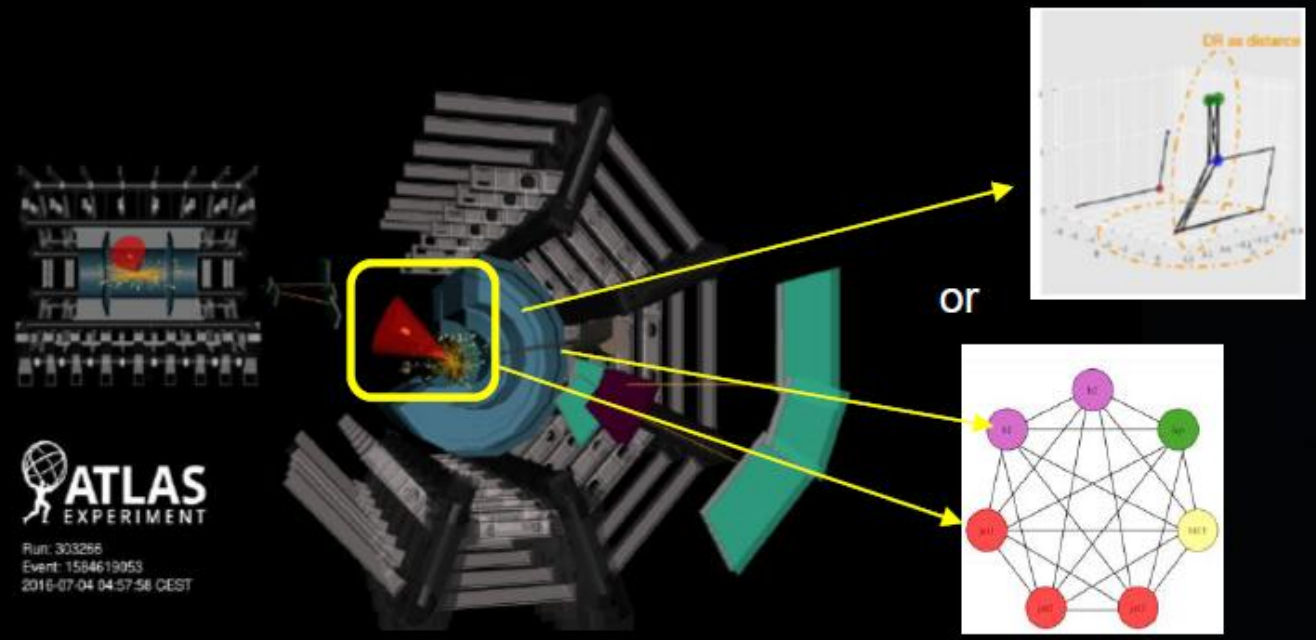
John Anders

AI in ATLAS

ATLAS (as other PP experiments) has successfully leveraged AI across numerous aspects, including data analysis, theory calculations, detector calibration/monitoring, and real-time data selection.

Different set of problems to be addressed depending on the case:

Offline data analyses → searches for new physics are often a "classification" problem → **event classification** and **process discrimination** in large and diverse datasets for discovery



or

How-to: *GNN and beyond:*

- Datasets building with diverse inputs
- Graph pre-processing, graph-based models built and trained
- Model optimization, xAI
- **Simple GNN, GNN attention transformers, autoencoders developed, tested and compared**

Dark photons: Monica D'Onofrio, Cristiano Sebastiani (now at CERN)

Axion-Like Particles: MDO, Nikos Rompotis, Rebecca Irwin (PhD)

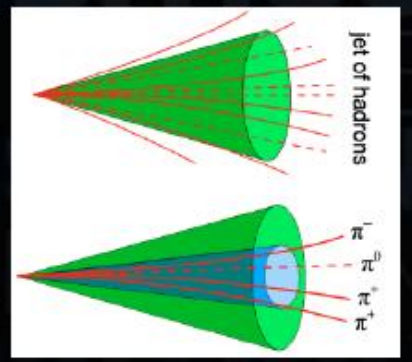
https://arxiv.org/abs/2501.03432

AI in ATLAS

ATLAS (as other PP experiments) has successfully leveraged AI across numerous aspects, including data analysis, theory calculations, detector calibration/monitoring, and real-time data selection.

Different set of problems to be addressed depending on the case:

Offline data analyses → need also to correctly identify "ingredients"
– fundamental physics objects: **Tau leptons** – complex and challenging



- τ_{had} BR > τ_{lep} BR (65% and 35%)
- τ_{had} decays = 1- or 3-prong (1 or 3 $\pi^{\pm}$'s, and maybe some π^0's)

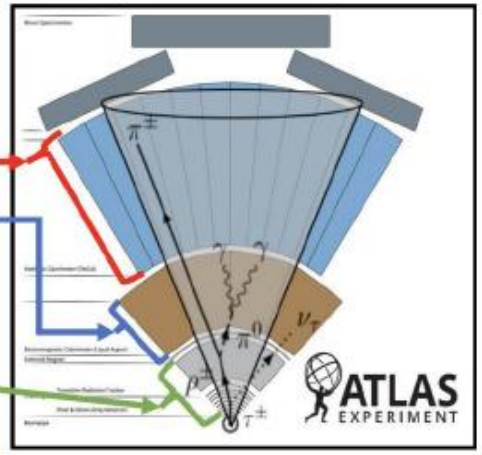
Hadronic Calorimeter
Obtains $\pi^{\pm}$ information

Electromagnetic Calorimeter
Obtains π^0 information
(via $\pi^0 \rightarrow \gamma\gamma$ and $\gamma \rightarrow e^-e^+$)

Tracking Detector
Collects charged particle track information, e.g., direction and position of $\pi^{\pm}$'s from τ-decay

τ_{had} Decay:

- Highly collimated – narrow cone
- Small cross-section
- Low multiplicity



Mehul Depala and Rob McNaulty (PhD), Jordy Degens, Monica D'Onofrio, Nikos Rompotis

Proposed

Unified Model
Graph Neural Network (GNN)

- ARMA convolutional layers [5]
- Combination of input variables from both ID and DMC
- Inputs can be unordered sets with varying lengths

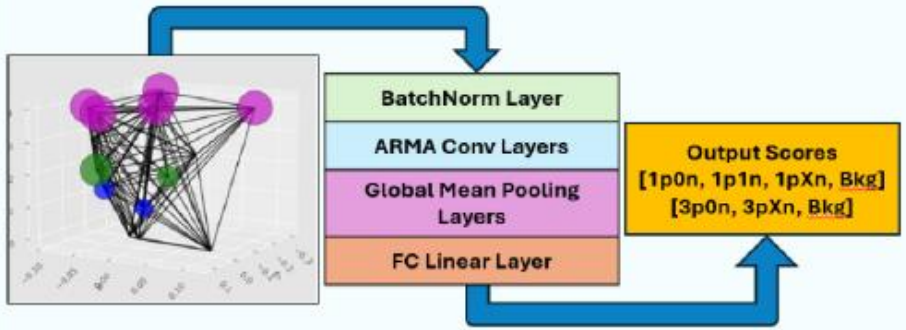


Figure 5.1: Proposed GNN Architecture. From left to right; graph input, passed through the layers of the GNN, which is then given a multiclassification score

AI in ATLAS

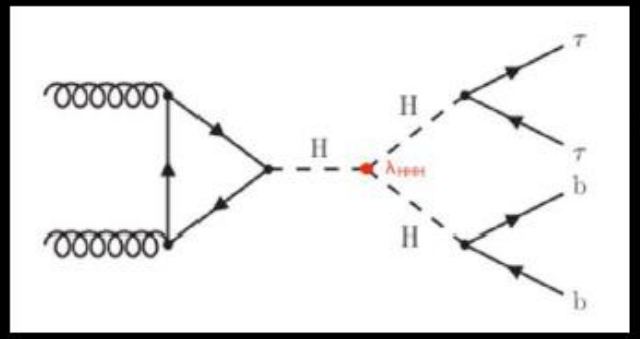
ATLAS (as other PP experiments) has successfully leveraged AI across numerous aspects, including data analysis, theory calculations, detector calibration/monitoring, and real-time data selection.

Different set of problems to be addressed depending on the case:

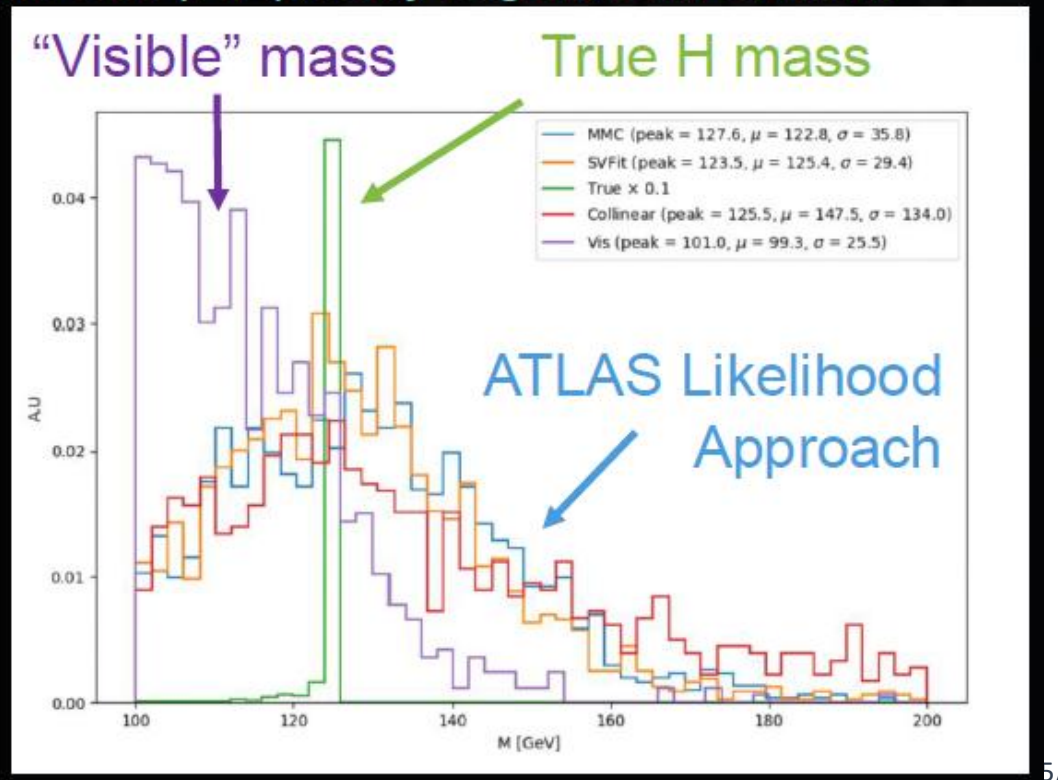
Offline data analyses → or to improve capability to reconstruct recently discovered particles: **the Higgs!!**

Buphesh Dixit (PhD), Jordy Degens, Carl Gwilliam

- Mass of Higgs decaying tau pair cannot be fully reconstructed due to neutrinos escaping the detector without interacting, taking energy with them.



Loss of information



- Current methods solve set of unconstrained equations probabilistically using likelihood minimisation but very slow and room for improvement in mass resolution

AI in ATLAS

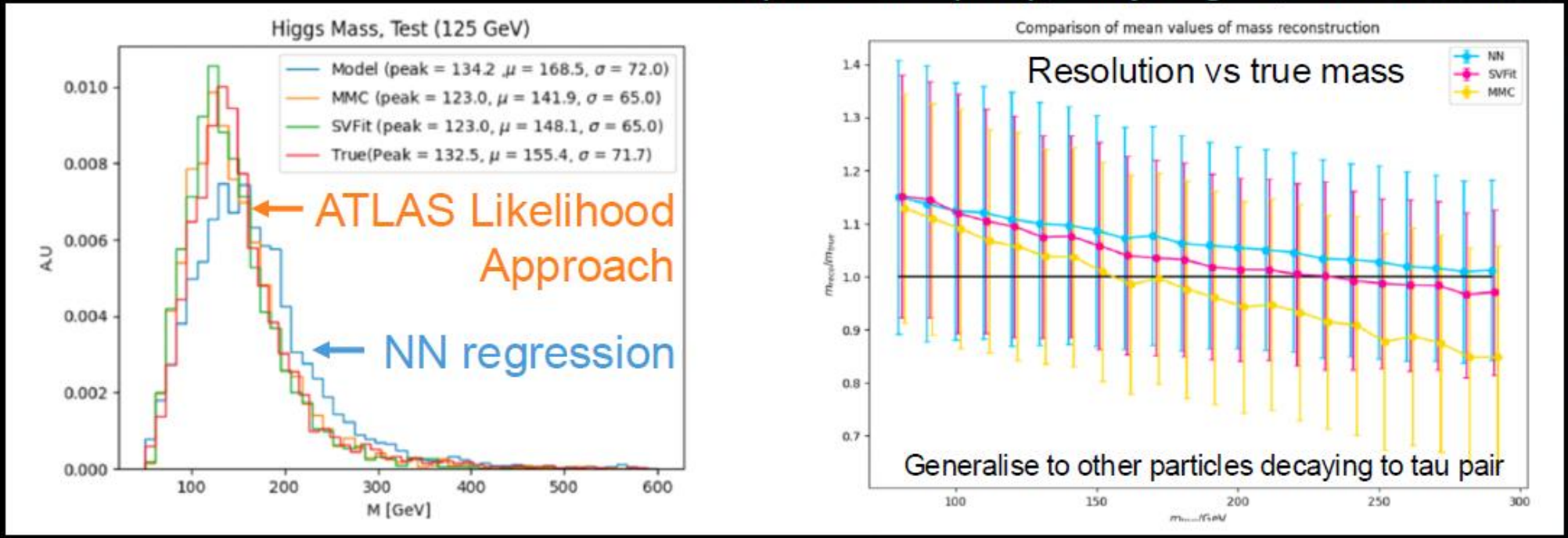
ATLAS (as other PP experiments) has successfully leveraged AI across numerous aspects, including data analysis, theory calculations, detector calibration/monitoring, and real-time data selection.

Different set of problems to be addressed depending on the case:

Offline data analyses → or to improve capability to reconstruct recently discovered particles: **the Higgs!!**

- NN-based **regression** approaches show promise with currently similar performance to LH method but work needed to improve them further

Buphesh Dixit (PhD), Jordy Degens, Carl Gwilliam



From ATLAS to healthcare – a spin out

Pipelines for **AI and xAI** developed for PP can be applicable to diverse fields

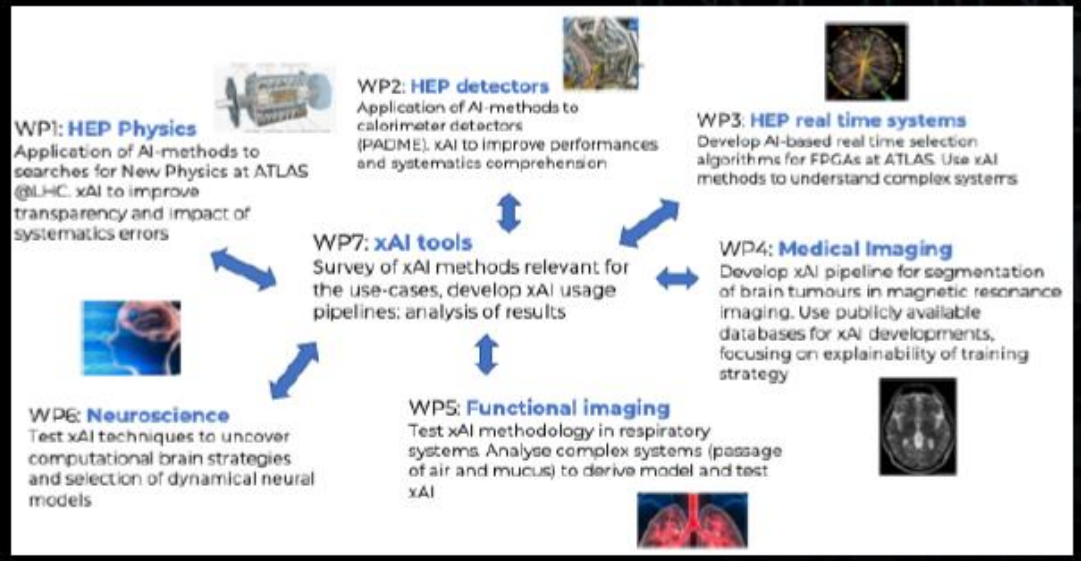
Two examples:

MUCCA project:

Goal to study xAI in heterogeneous cases *quantifying strengths* and solving weaknesses of new and state of the art methods on Deep Learning applications

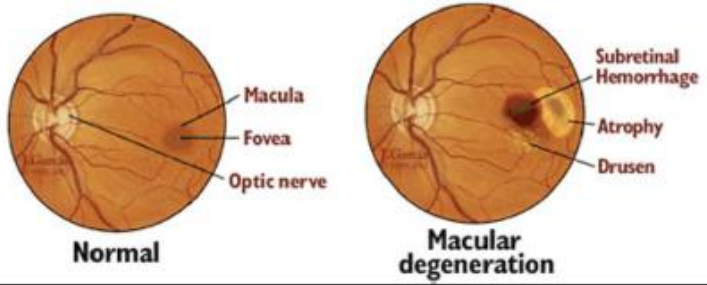
Now wrapping up..

Monica D'Onofrio, Cristiano Sebastiani (now at CERN)



Age-related Macular Degeneration studies

- **Age-related Macular Degeneration (AMD)** = Common eye disease that blurs central vision
  - Occurs when aging causes damage to the macula
  - Leading cause of vision loss for older adults (typically 50+ years) – but doesn't cause complete blindness



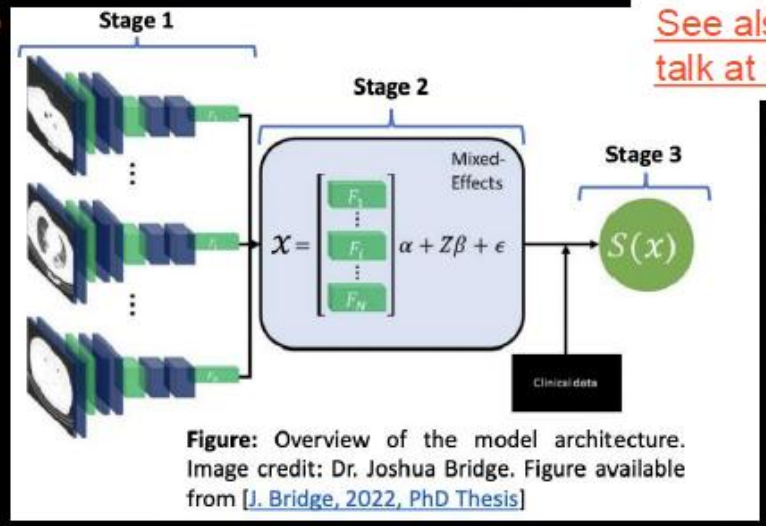


Figure: Overview of the model architecture. Image credit: Dr. Joshua Bridge. Figure available from [J. Bridge, 2022, PhD Thesis]

See also Rob McNulty talk at IoP

Rob McNaulty (PhD), Monica D'Onofrio, Nikos Rompotis

CNN used so far – moving to GNN soon

LHCb – Analysis Facilities

Eduardo

(As mentioned briefly in the LHCb report.)

- **Discussions started back in Nov. 2024 with several IT people to understand the best strategy, minimizing efforts towards a prototype AF to eventually leverage the LHCb CPU and GPU resources**
  (they are special, since data-taking/online resources)
  - Part of Liverpool's work plans for LHCb UK Upgrade 2

- Converged in the **proposal of a pilot project in collaboration with CERN-IT**: set-up of a **prototype AF leveraging as much as possible what has been designed and put in production by the SWAN team for the benefit of the community** – basically, compute is offloaded to (LHCb) external resources
  - **Duration of 1 year**, possibility of an extension subject to relevance and mutual interest
  - Delivery of a production-ready AF for LHCb is planned in 3 Phases (pilot project is only the first 2)

- Status: formal IT Steering Committee approval of project & allocation of IT person-power (2 x 0.1 FTE) should happen today!

- *But work has informally started …!*

- NOTE FOR DISCUSSION: could envisage something similar for Liverpool's Tier-2! Interest …?

CERN

RCS-ICT PSO

DRAFT

Utilizing LHCb Farms as Analysis Facilities through SWAN

Authors:

Pilot-Proposal
Eduardo Rodrigues (University of Liverpool) and *Francesco Sborzacchi* (CERN)

PSO
Andreas-Joachim Peters & *Xavier Espinal* - CERN RCS-ICT Engagement Manager,
Simone Campana CERN IT Business Engagement Lead

Responsibles:
See section *Personpower*

Submission Date: 27.2.2025

This proposal aims to leverage the SWAN infrastructure to transform LHCb trigger farms into Analysis Facilities (AFs), benefiting both LHCb analysts and CERN-IT. Discussions with CERN-IT's SWAN team have confirmed mutual interest in this initiative.

A one-year pilot project will develop a prototype AF using SWAN, laying the foundation for a future production-ready AF. The pilot focuses on prototyping and technical feasibility, with long-term production requiring additional planning for maintenance and support.

The ultimate objective is to **develop a production-ready Analysis Facility** (AF) **for LHCb** through a three-phase approach. The pilot project focuses on the first two phases, prioritizing prototyping and technical validation before moving toward full production, which will necessitate additional planning for maintenance and support.

This initiative **serves as a potential model for other communities** seeking to repurpose existing computing resources for advanced analysis and machine learning applications.

Community work – HSF, PyHEP

Eduardo

HEP Software Foundation (HSF)

- Continue as **member of the Steering Group**
  - Great way to get to know about (community, but not only) activities across experiments and non-HEP communities, etc.
- Related – act as **community software liaison on the WLCG Management Board**
  - In particular, report to the LHCC on HSF/community Software (4 times a year)
- **WLCG-HSF joint workshops**: co-organiser (May 2025 edition, IJCLab, Orsay)
- Also, presentations at the UK SWIFT-HEP project workshops (Nov. 2024, April 2025)

https://hepsoftwarefoundation.org/

HSF PyHEP group

- **Co-convened** (from day-1) until end of 2024
- Likewise, **co-organised the PyHEP and PyHEP.dev workshops** from 2018 till end of 2024 (9 events in total!)
- Now was « time to let it go»
- See here for the 2025 workshops …

Related publications

- **The Critical Importance of Software for HEP**
  Christina Agapopoulou et al., HSF prepared inputs for the European Particle Physics Strategy Update
  arXiv:2504.01050 [hep-ex], Zenodo doi:10.5281/zenodo.15097159

- **PyHEP.dev 2024 Workshop Summary Report**
  A. Alshehri et al., Report of PyHEP.dev 2024 developer's workshop, Aachen (Germany), 26-30 August 2024
  arXiv:2410.02112 [hep-ex]

- **Analysis Facilities White Paper**
  arXiv:2404.02100 [hep-ex], HSF-TN-2024-01

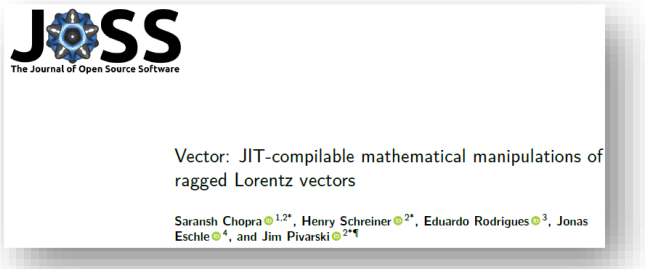
Community work – Scikit-HEP

"Scikit-HEP is an open-source community-driven and community-oriented project with the goal of providing an ecosystem for particle physics data analysis in Python, fully integrated with the wider scientific Python ecosystem.

It expands the typical Python data analysis tools for particle physicists with packages spanning the spectrum from general scientific libraries for data manipulation to domain-specific libraries."

Admin aspects

- April 2025: **Scikit-HEP joins the Scientific Python Ecosystem Coordination Core Projects** as 1 of the first 2 domain specific stacks accepted
- March 2025: we got **accepted as an organisation on PyPI**
- December 2024: **Coffea package (key in IRIS-HEP) moved to our org**



Talks & publications

- **JOSS paper on the Vector package is about to be out!**
- **Project paper** started to get prepared
  - Will replace my CHEP proceedings (close to 50 citations so far!)
- **Talk "The Scikit-HEP project - news and future directions"**, WLCG/HSF Workshop 2025, IJCLab, Orsay (France), May 2025

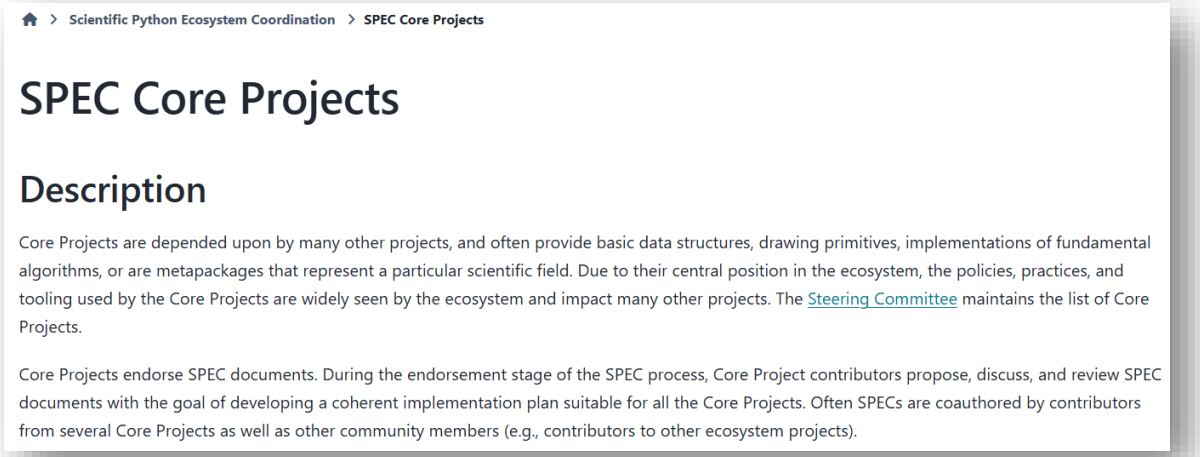
Community work – Scikit-HEP - engagement with the wider community

Eduardo

- We started in 2024 to engage with Scientific Python organisation
- Conversations have continued and evolved …
- **April 2025:**
  Scikit-HEP joins the Scientific Python Ecosystem Coordination Core Projects as 1 of the first 2 domain specific stacks accepted

Scientific Python

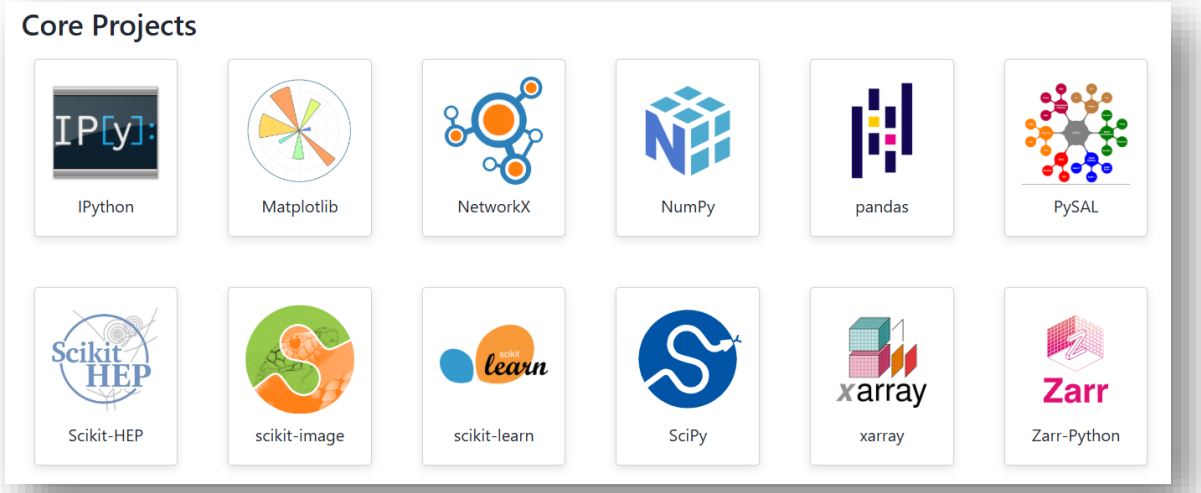
https://scientific-python.org/



What are the characteristics of a Core Project?

- Widely used in scientific research
- Widely used in the scientific Python ecosystem
- Developed using shared community practices
- Developed in the open by their communities
- Well documented and well tested
- Open source

- **Present list of core projects**,
  see Scientific Python - SPEC Core Projects:



Cross-cluster efforts

- **Created last year by the department heads, cf. the several posts by Monica**:
  - AI and Data Science for Physics
  - Sustainable Technologies
  - Quantum Science and Technologies
  - Medical Physics

- **Many contributions by Monica and other PP colleagues**

Talks from me (hence a biased selection)

- Sustainability in Computing - news from Particle Physics, Jan. 2025, Liverpool meeting on Sustainable technology in Physics
- Personal Interests, Nov. 2024, Liverpool kick-off meeting on Data Science and AI
- Sustainability in Computing - "Views" from Particle Physics, Sep. 2024, Liverpool kick-off meeting on Sustainable technology in Physics
- QC/QML @ LHCb, Sep. 2024, Liverpool kick-off meeting on Quantum Science and Technologies Development (QSTD)
- Also, mini-course "Big Data Python ecosystem for HEP analysis“ @ STFC School on Data Intensive Science 2024, Liverpool, July 2024

Cross-cluster efforts – AI / Data Science in Physics

Beyond PP:

- Accelerator Science: design, optimization and control systems of particle accelerators (see Alex Hill's talk)
- Cancer research: infrared spectral images analysis using ML. *Patent on machine learning algorithm*.

- Attention on AI policy, strategy (UK, EU, worldwide)
  → See for example this document ("Enabling AI for High Energy Physics Experiment and Theory") highlighting key challenges, opportunities, and a preliminary **action plan for AI in PP**
  → **Submissions to EPPSU by several colleagues**

- Focus on challenges, barriers and opportunities in:
  - Getting **research funding** (consortia, EPSRC, EU initiatives etc)
  - **Software, hardware** needed → facilitate access to resources, collaborate with experts at national laboratories (Hartree, SciML..)
  - **Knowledge exchange** → industry partnerships (NVIDIA, IBM)
  - **Skills/training** and capacity building → enhance training pipelines

Thank you for listening !

And do get in touch if any of this catches your interest ☺ !