

EuCAIFCon 2025

Workgroup discussions

17 June 2025

EuCAIF Workgroup discussions

- WG 1: Foundation models
Tue 12:00
 - WG 2: Co-Design
Wed 16:30
 - WG 3: FAIR and sustainable AI
Wed 15:00
 - WG 4: JENA WP4 (ML and AI infrastructure and resource needs)
Tue 14:15
 - WG 5: Building bridges - Community, connections and funding
Thu 16:15
 - Discussion: Simulation-based inference & Uncertainty quantification
Tue 16:40
-
- Everybody is invited to join the discussion, and encouraged to contribute!
 - Check the conference website (later today) for details about mailing lists, google docs, etc.

WG1: Foundation models (FMs)*

Background: Pioneered in LLMs like ChatGPT or image generators like DALLE

What is a FM? Multimodal FMs centralize information from various data modalities & domains & encode them in a common meaningful latent representation + multi-head fine-tuning

Why a FM? Amortization, automation, acceleration [compute & person-power]:
Narrow task-centric → multi-task, reusable, data+MC-trained backbone, reduce uncertainties

What can we do in WG1? [bottom-up]

- Define potential **transformative impact** of FMs for our community
- Design **strategic roadmap** to foster progress as a community
- **Benchmark data set(s)**, data challenge, success metric, downstream tasks
- Physics-encoding, mitigate domain shifts, explainability, scalability, language & symbolic encoding
- Facilitate **collaboration**, network, training, topical meetings, **funding**

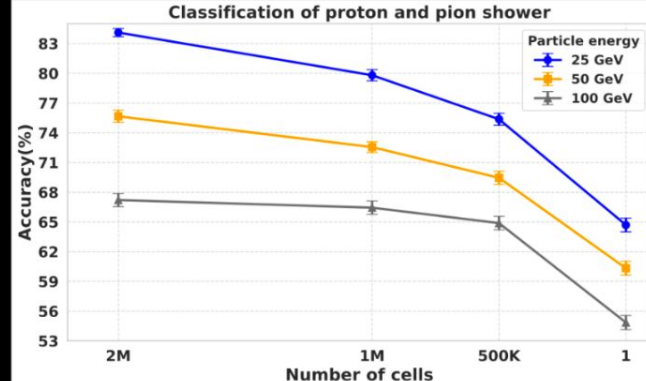
Working Group 2: Co-Design

“End-to-end” is not such if you only optimize the software, or only the hardware. Doing everything together – that’s what we should aim for. It is called **co-design**.

WG2 has 48 members in EUCAIF. Why not you? You could become an author, do joint research, participate in grants, etc. **Let us discuss on Wednesday at 4.30PM** (room T2b)!

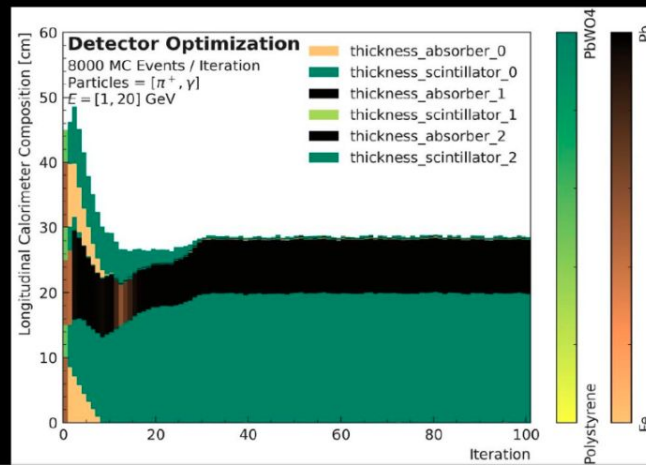
Foreseen in 2025:

- Publish a white paper outlining cases where co-design is essential
- Aggregation in projects (calorimetry, muon tomography, ...)
- Funding applications (ITN, Pathfinder, ...)



Above: p/π discrimination of tailored GNN as a function of number of cells in hadron calorimeter

Below: end-to-end optimization of a sampling calorimeter w/diffusion model



WG3 - FAIR and sustainable AI - status and plans

Organisers: Caterina Doglioni & Gabrijela Zaharias

- WG3 covers two sides of sustainability
 - FAIR (Findable, Accessible, Interoperable, Reproducible)
 - see [FAIR4ML from Research Data Alliance](#)
 - Environmental impact of ML
 - CO2 footprint but also positive impacts in data analysis acceleration
 - Does not cover explainable/trustworthy AI and AI ethics
- What do we want to do as a WG?
 - **Lots of people** outside (and within) our field are **working on this**. We don't want to reinvent the wheel but rather **gather resources that are easy to use** for people from our field
 - Connect to existing initiatives (many!)
 - Eventually, a **whitepaper** with good practices down the line (6mo / one year from now)
 - **Wednesday session: practice with hands-on tools** (volunteers with ML algos welcome!)

- Where do we stand

- Kick off meeting May 21 at CERN and online, joint with the Interexperiment ML working group [recording [here](#)]

FAIR & Sustainable EuCAIF WG Kick-off (joint with IML)

📅 Wednesday May 21, 2025, 9:00 AM → 10:30 AM Europe/Zurich

📍 222/R-001 (CERN)

9:00 AM → 9:10 AM Introduction and setting the scene

Speakers: Caterina Doglioni (The University of Manchester (GB)), Dr Gabrijela Zaharias (University of Nova

 20252105_Doglioni...

9:15 AM → 9:35 AM DOME: ML Best Practices & the AI Ecosystem from the life sciences perspective

Speaker: Gavin Farrell (University of Padua)

 GavinFarrell_DOME...

9:40 AM → 10:00 AM Environmental sustainability in AI at CERN

Speaker: Dr Sofia Vallecorsa (CERN)

 Sustainable AI EuC...

10:05 AM → 10:20 AM Discussion and next steps

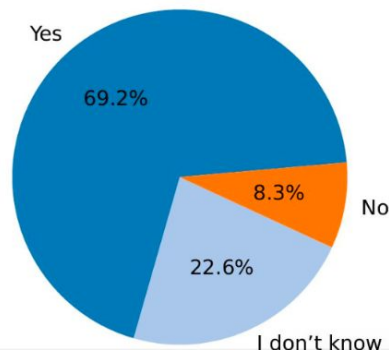
WG4: JENA WP4: Machine Learning and Artificial Intelligence

Andreas Ipp (TU Wien), Sascha Caron (Radboud University, Nikhef)

Since last year: 12 recommendations:

Strategic White Paper on AI Infrastructure for Particle, Nuclear, and Astroparticle Physics: Insights from JENA and EuCAIF

Sascha Caron,^{*a,b} Andreas Ipp,^{*c} Gert Aarts,^d Gábor Bíró,^{e,f} Daniele Bonacorsi,^{g,h} Elena Cuoco,^{g,h} Caterina Doglioni,ⁱ Tommaso Dorigo,^{j,k} Julián García Pardiñas,^l Stefano Giagu,^m Tobias Golling,ⁿ Lukas Heinrich,^o Ik Siong Heng,^p Paula Gina Isar,^q Karolos Potamianos,^r Liliana Teodorescu,^s John Veitch,^p Pietro Vischia,^t Christoph Weniger^u



arXiv:2503.14192

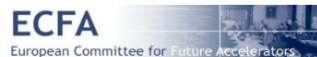
Should we collaborate on large-scale ML models?

How can we enable the recommendations?

- Monitor future development?
- Common applications to EuroHPC, AI factories and AI gigafactories?
- Collaborations and doctoral networks on large-scale ML?
- Should we continue? Other ideas?

Join the working group:

→ <https://indico.scc.kit.edu/event/3813/>



WG 5: Building bridges - Community, connections & funding

- Outcomes so far: eucaif.org website, **EuCAIFCon** conference series



Call for Evidence: „European strategy for AI in science“

Statement by the German Federal Ministry of Research, Technology and Space (BMFTR)

- Furthermore, initiatives focused on networking (e.g. **EuCAIF** at European level, or ErUM-Data-Hub at national level) and on developing explainable AI should be strengthened. Establishing structured and reliable pathways for the transfer of research results into commercial applications is also crucial and should be strategically addressed.
- Identification and exploitation of **consortium/collaboration-level funding schemes**
- How to **strengthen EuCAIF-related research activities in the academic landscape?** (new faculty candidates, new faculty positions, retaining good people)
- How to increase visibility at the European level (--> roadmaps, structural funding)?

Discussion: Simulation-based inference & Uncertainty quantification

- Simulation-based inference plays an increasing role in cosmology, gravitational waves, astroparticle physics and particle/nuclear physics
- Still there are lots of practical hurdles to make SBI a standard workhorse for analysis tasks.
- Goal of the discussion is to establish the most critical needs in the community (common tools, large joined projects, training material, algorithmic gaps, etc).
- Identify clear goals and timelines to form a workgroup