

AI in fundamental physics: strategy & future directions

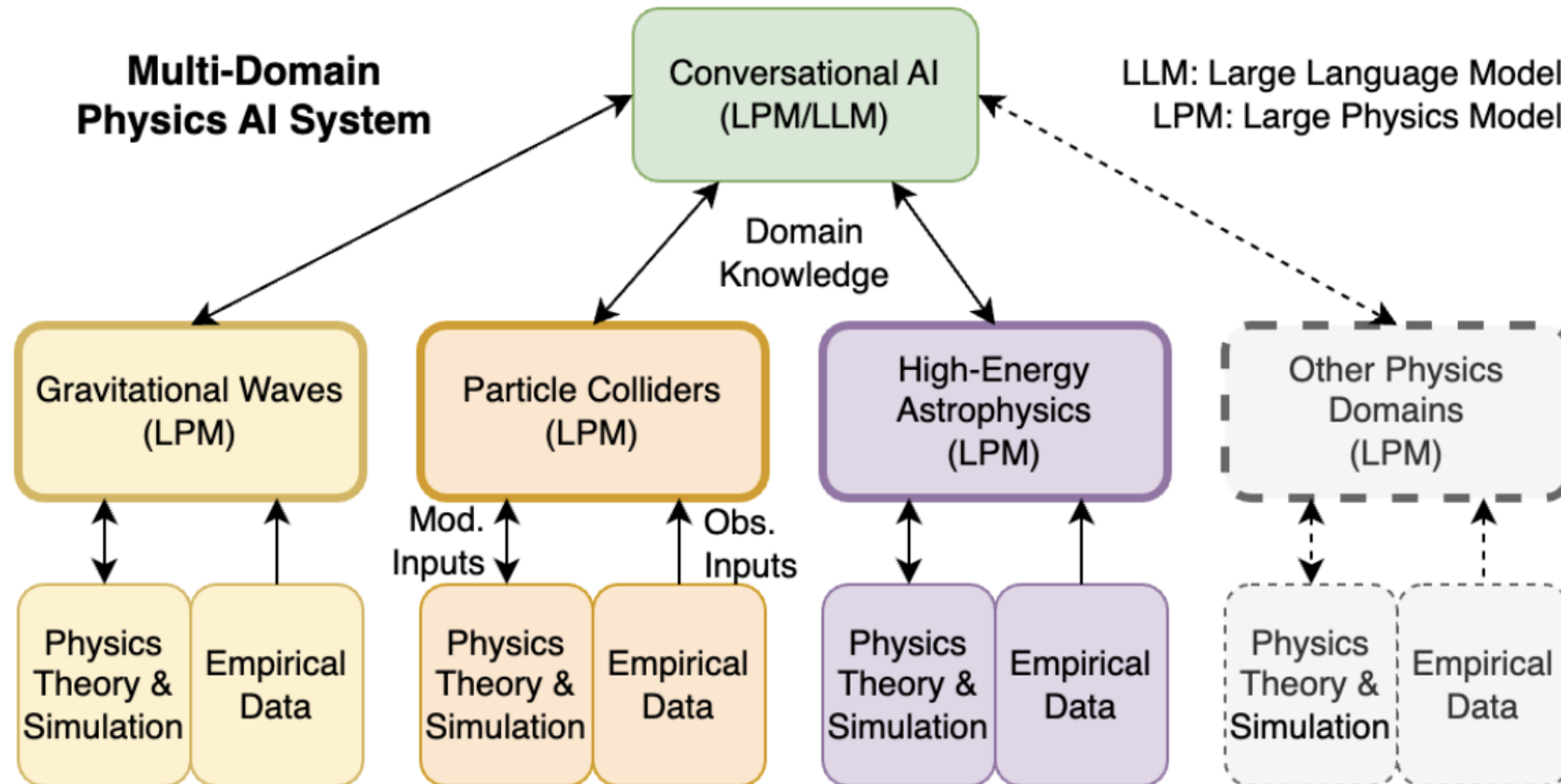
Sascha Caron

(Radboud U and Nikhef)



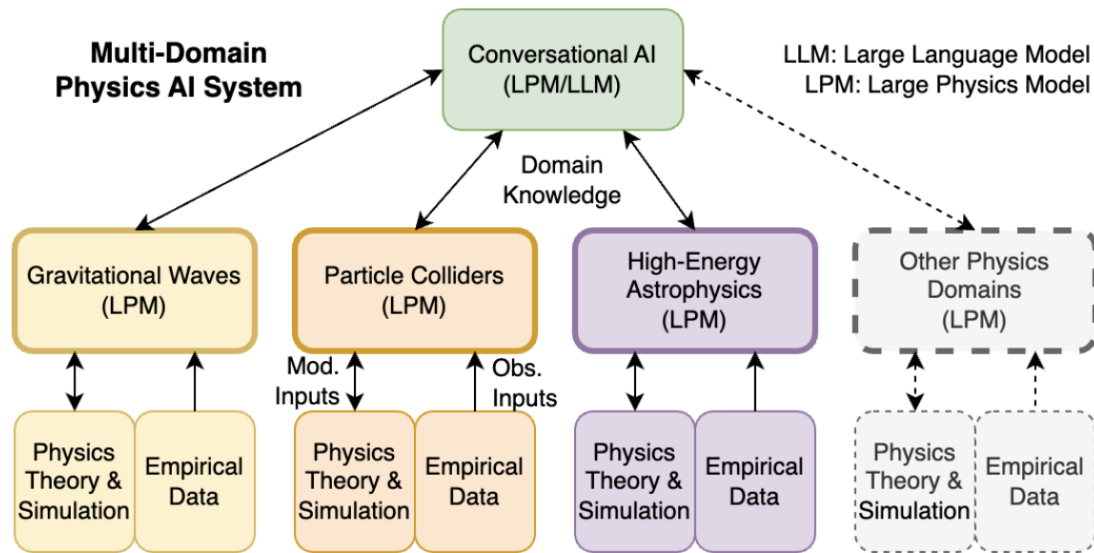
Grand vision 1: Build AI models

Large Physics Models



Eur.Phys.J.C 85 (2025) 9, 1066 (Barman, Caron, et al)

Large Physics Models

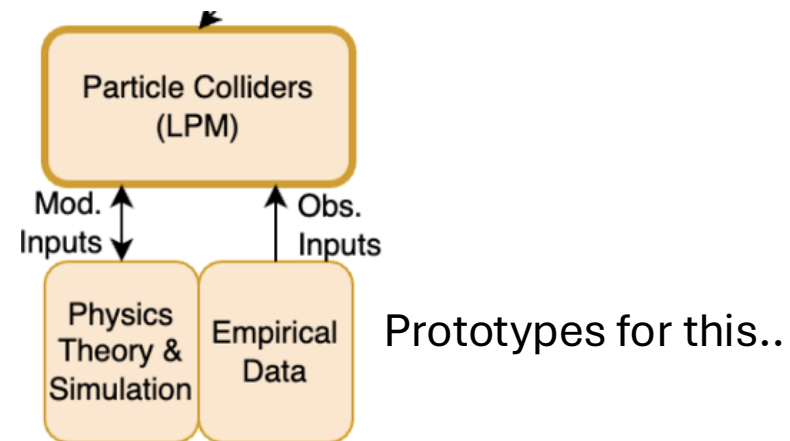


Foundation models adapted for physics
→ integrating symbolic reasoning,
experimental data analysis, and
theory synthesis.

Inspired by the **collaborative structure of particle physics experiments**, we propose the same interdisciplinary approach to **building Large Physics Models**.

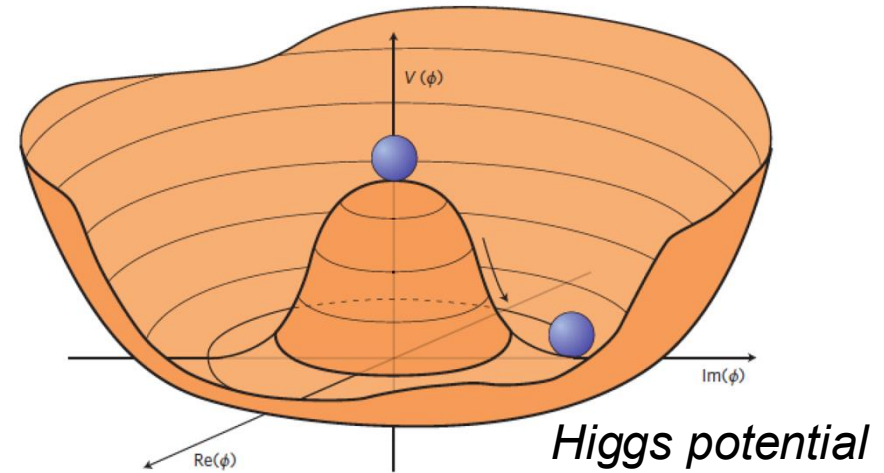
Eur.Phys.J.C 85 (2025) 9, 1066 (Barman, Caron, et al)

- Experimental Physics: Particle Collider Building Block



HL-LHC Physics Goals

- Higgs potential
 - di-Higgs measurement
- Unexpected physics



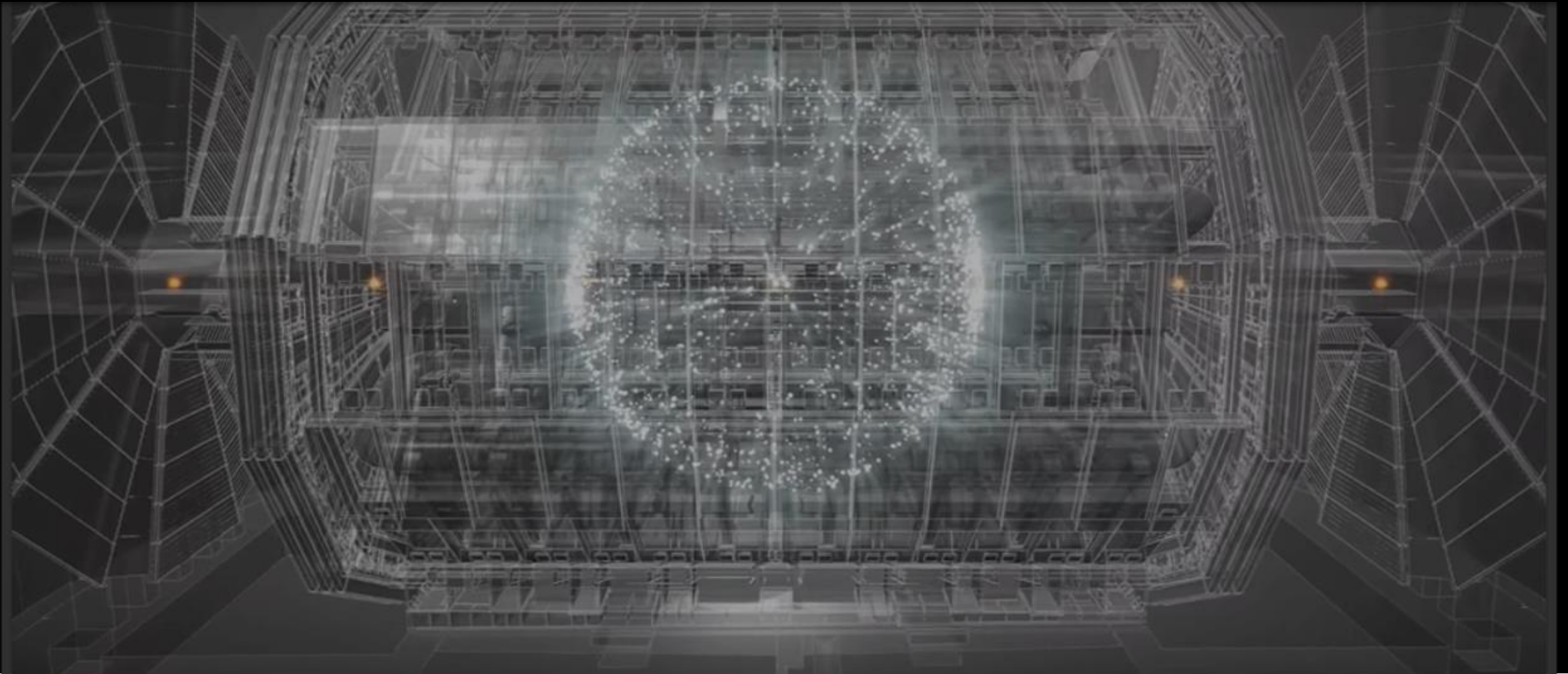
*Still not all information in the experimental data is used:
a challenge across experimental physics*

→ Fundamental question (also beyond LHC):
How do we learn what matters for each specific physics goal?

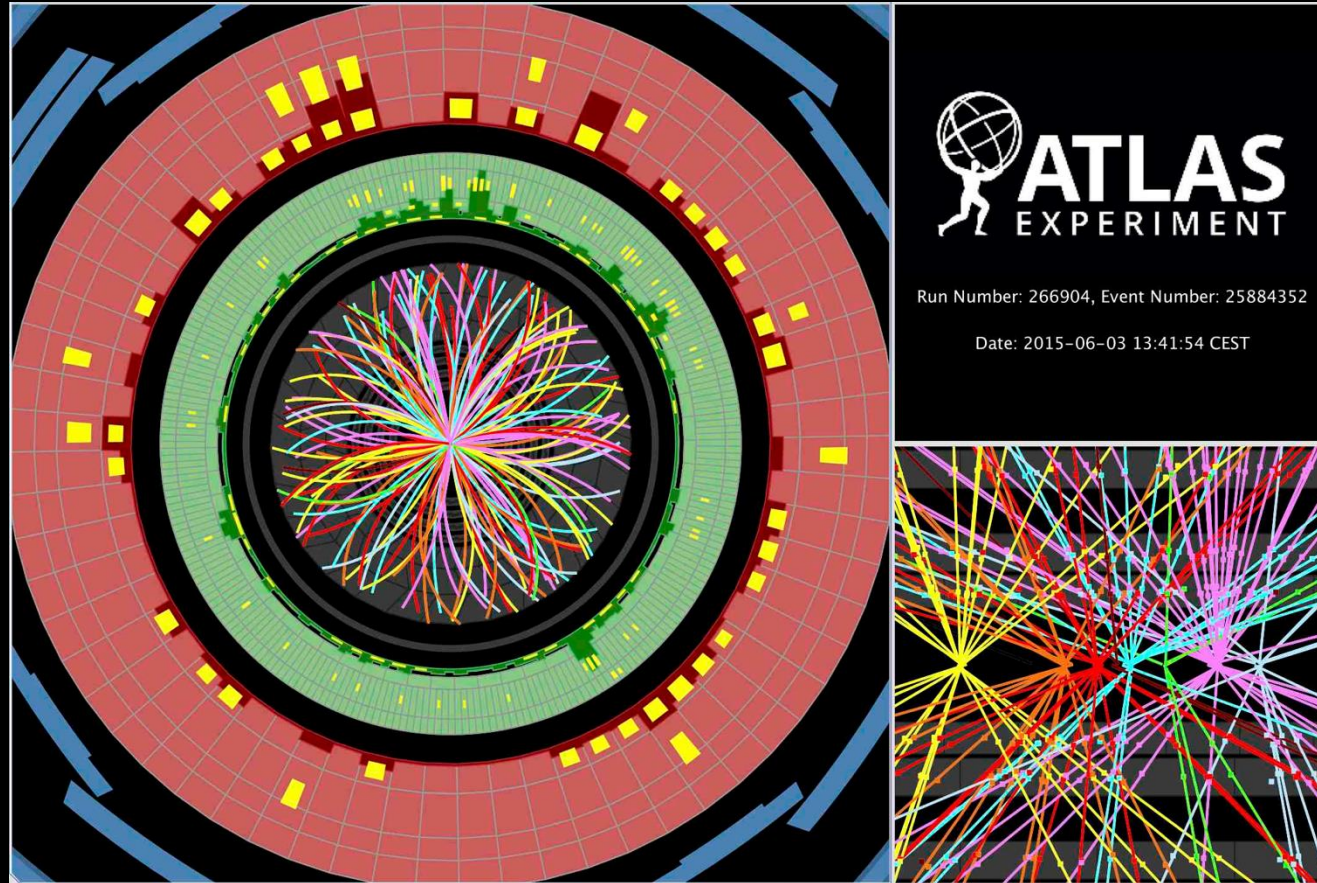
Complex Measurements: LHC Collisions

Collisions every 25 ns

Hundreds of particles $\rightarrow O(10^4)$ measurement points per event

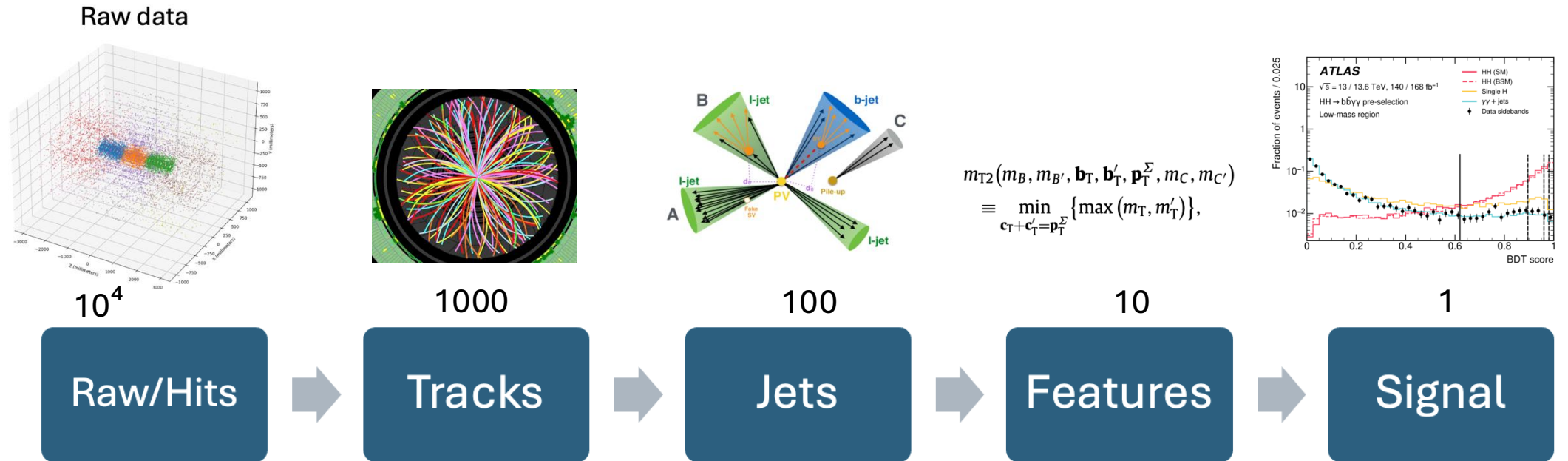


Most events look like this after track reconstruction...
 10^4 hits \rightarrow 10^3 tracks

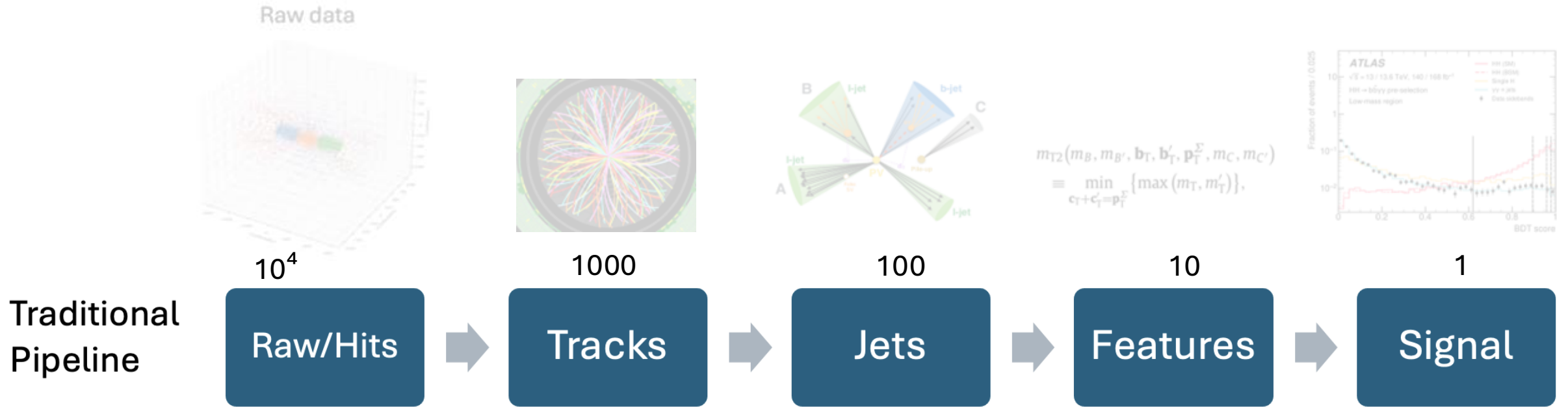


Analysis model in Particle Physics

Traditional Pipeline



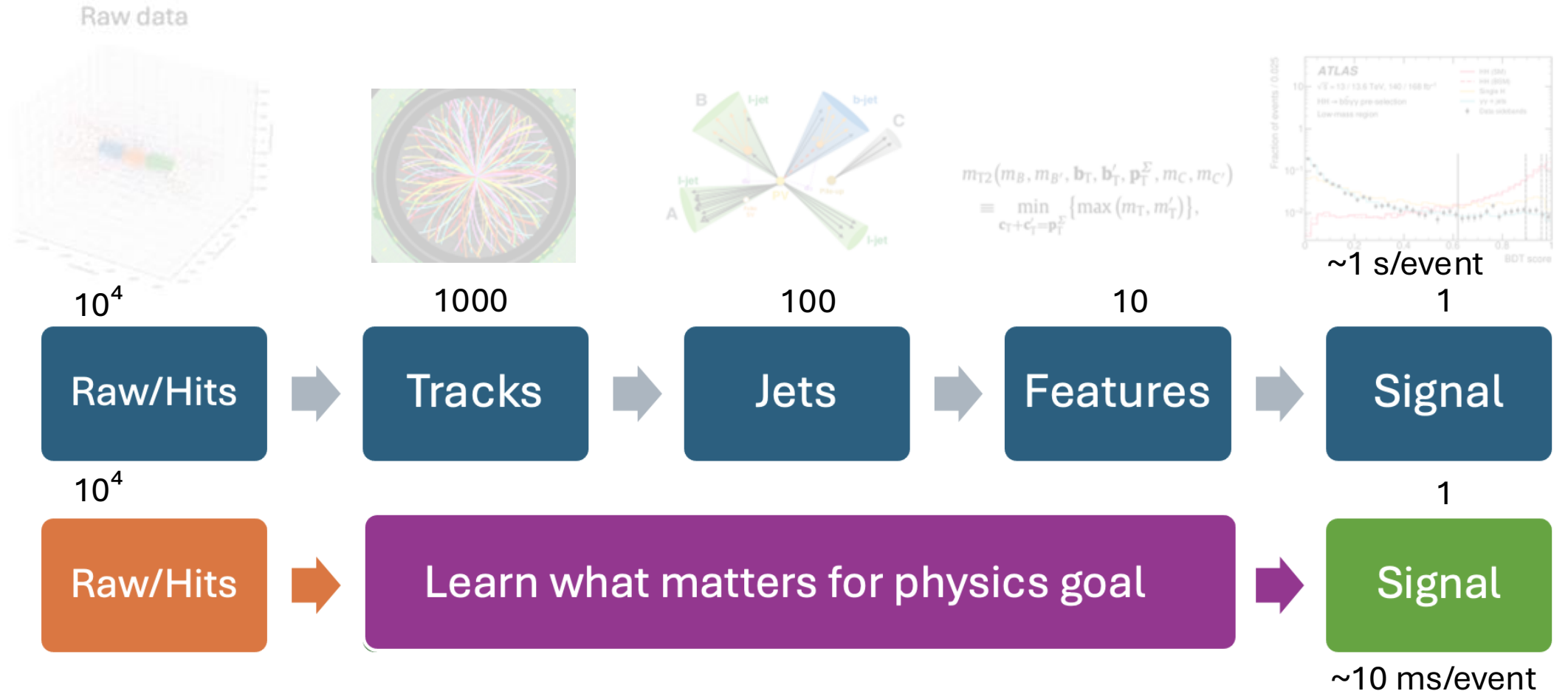
Analysis model in Particle Physics



Why is this not sufficient ?

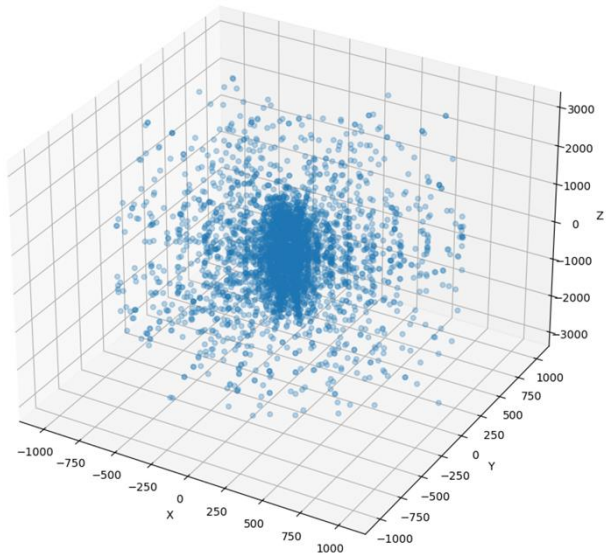
Traditional: Fixed compression designed before knowing the final goal
(and unconventional signatures may fall through the cracks...)

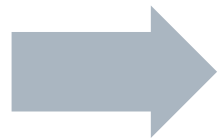
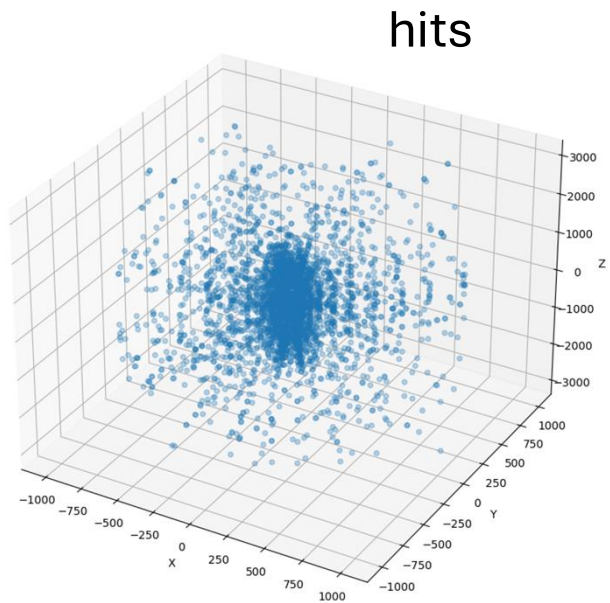
Analysis model in Particle Physics



End-to-end optimization for each physics goal

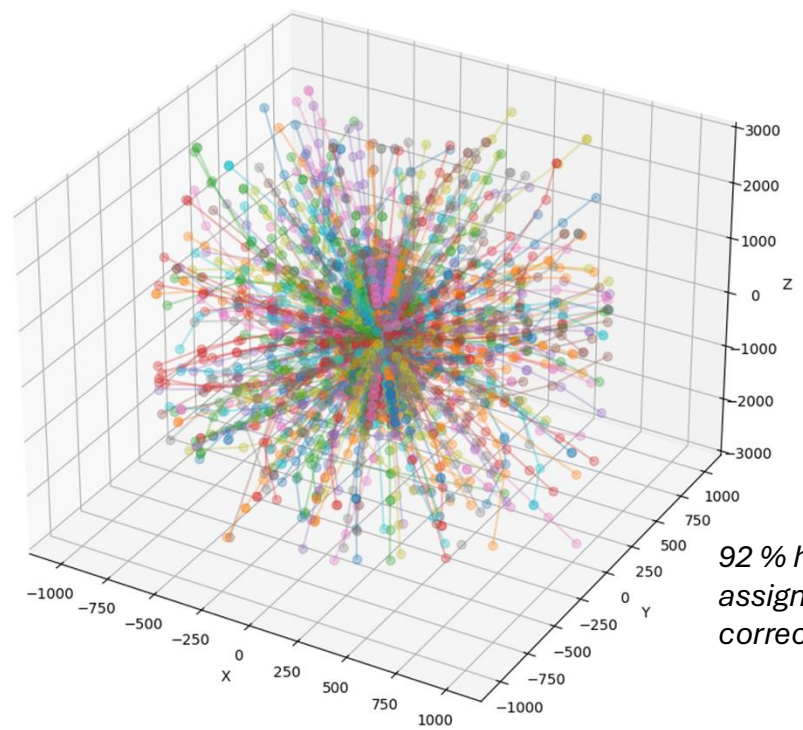
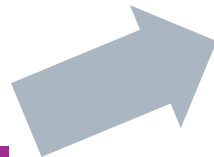
hits



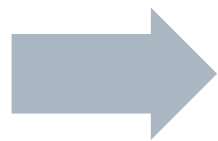
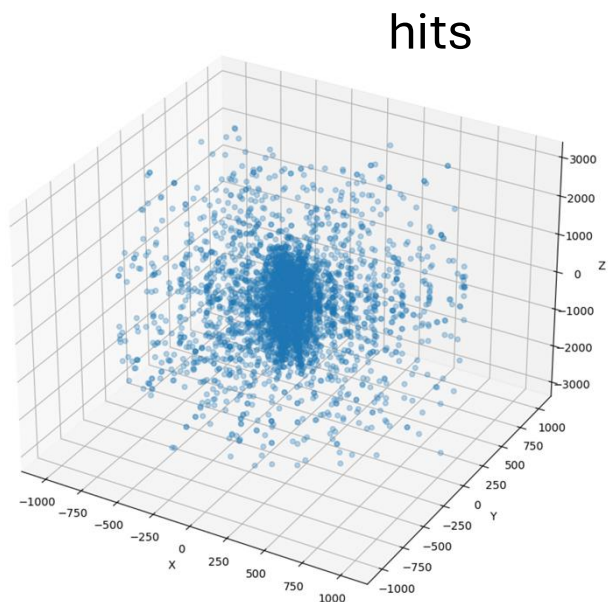


Trackformer/
Higgsformer

tracks

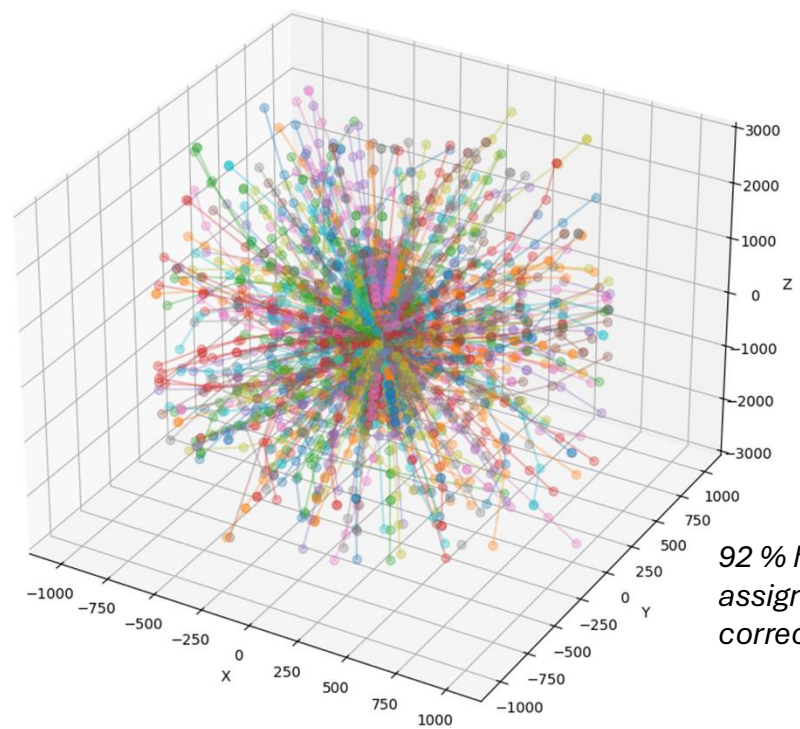
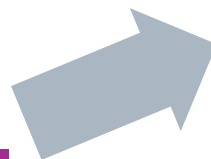


92 % hits
assigned to
correct track



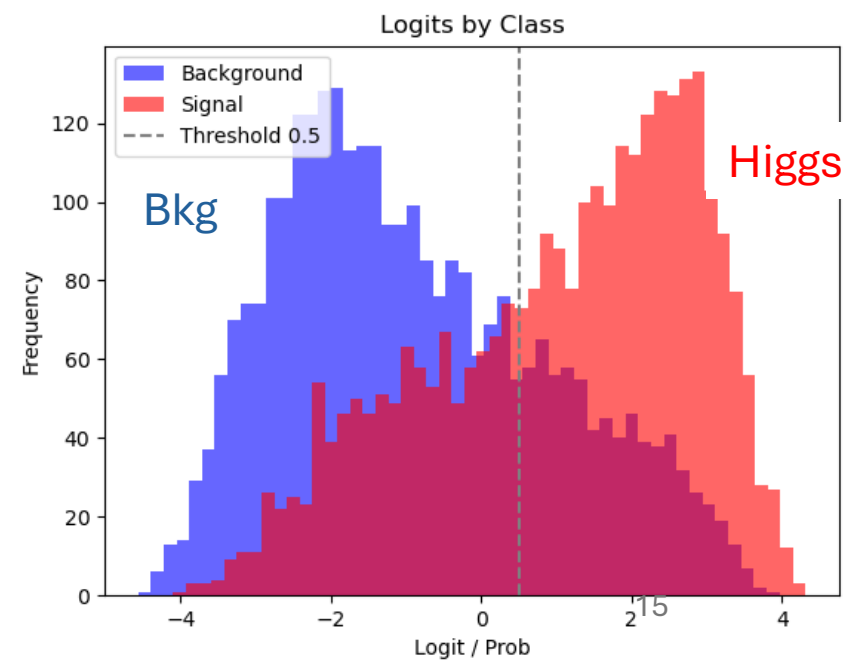
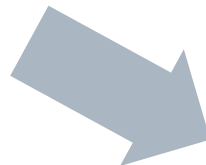
Trackformer/
Higgsformer

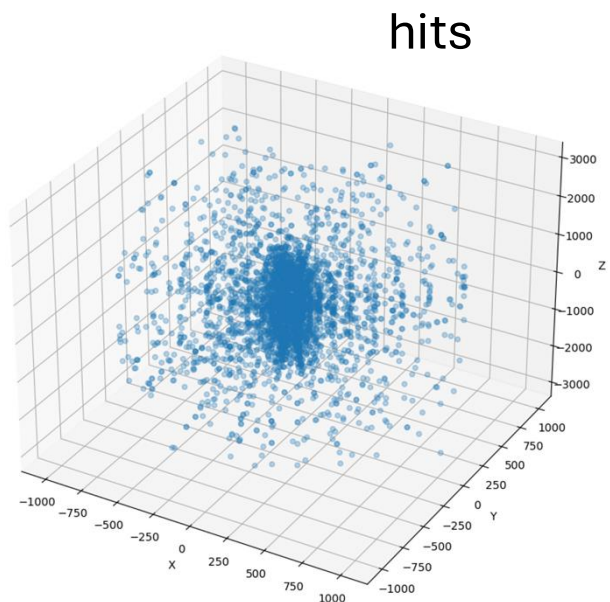
tracks



92 % hits
assigned to
correct track

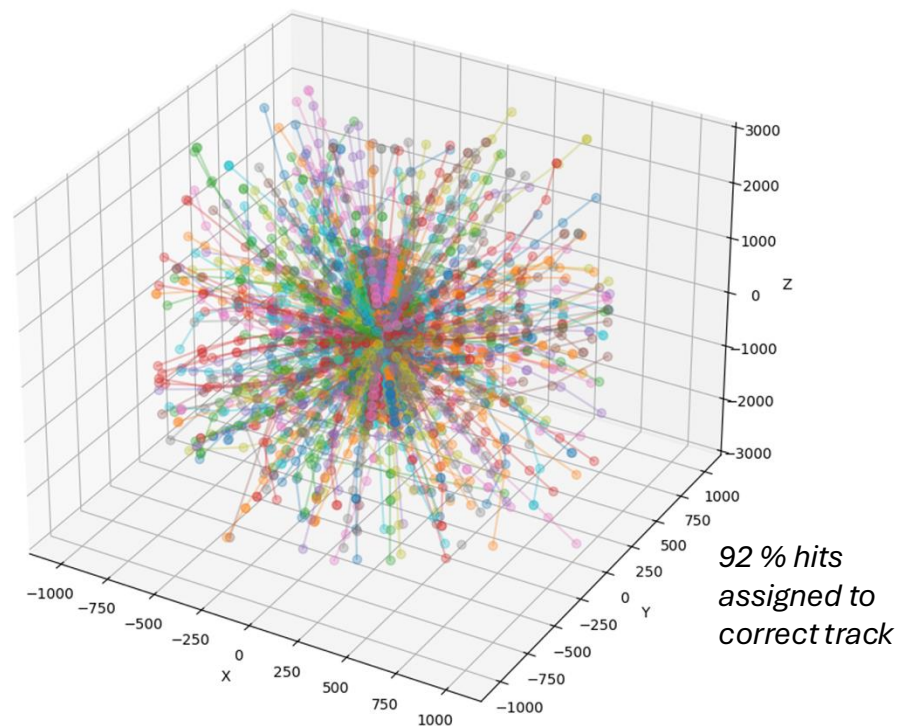
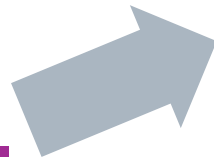
Higgs
classification



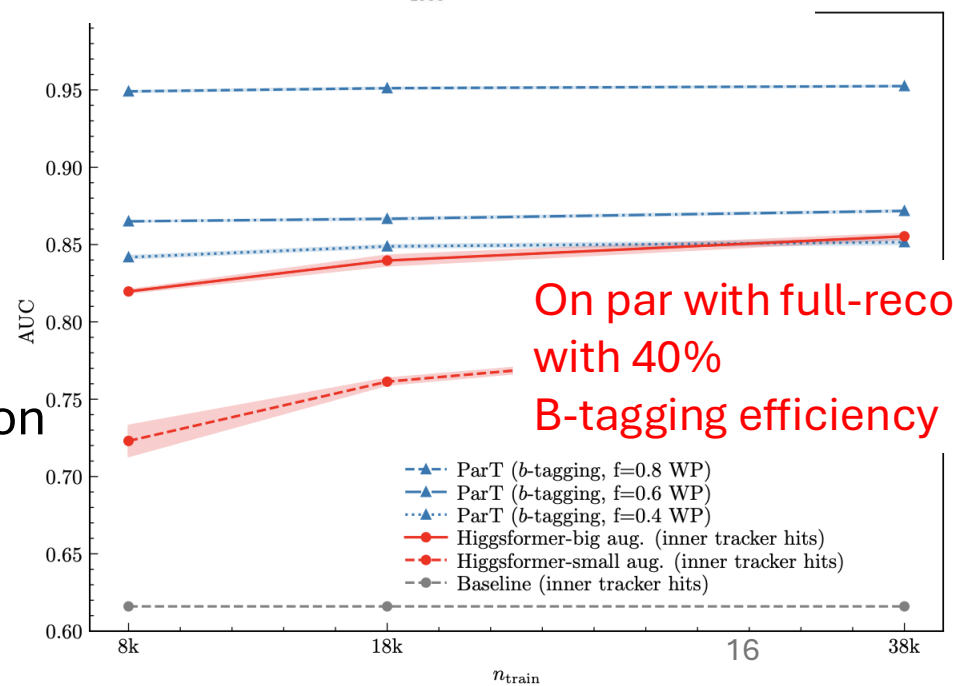
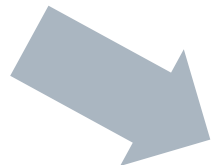


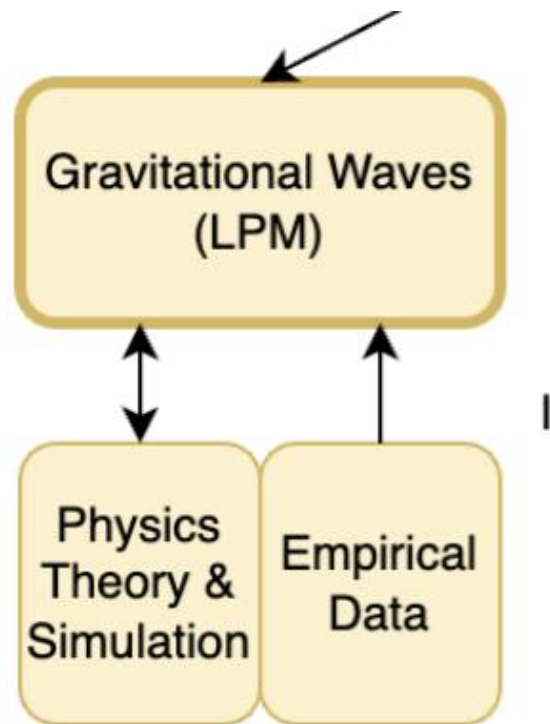
Trackformer/
Higgsformer

tracks

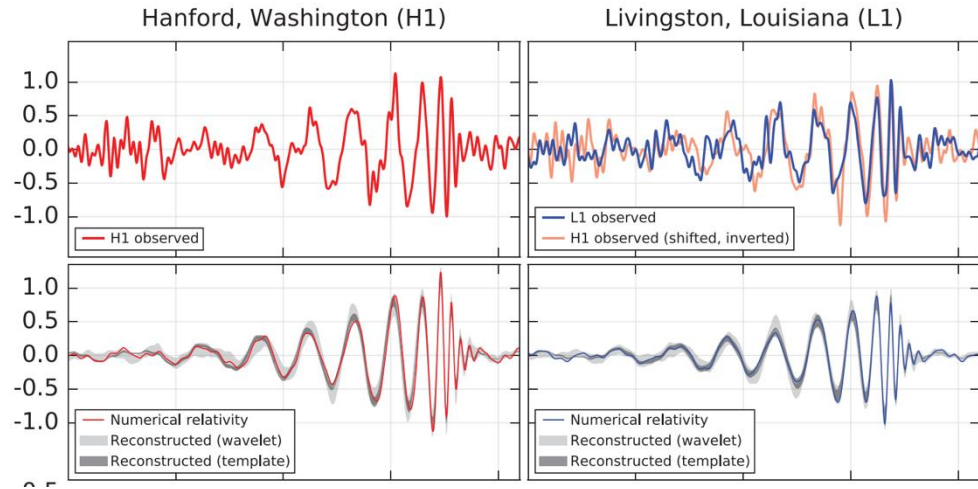


Higgs classification





How close are You ?

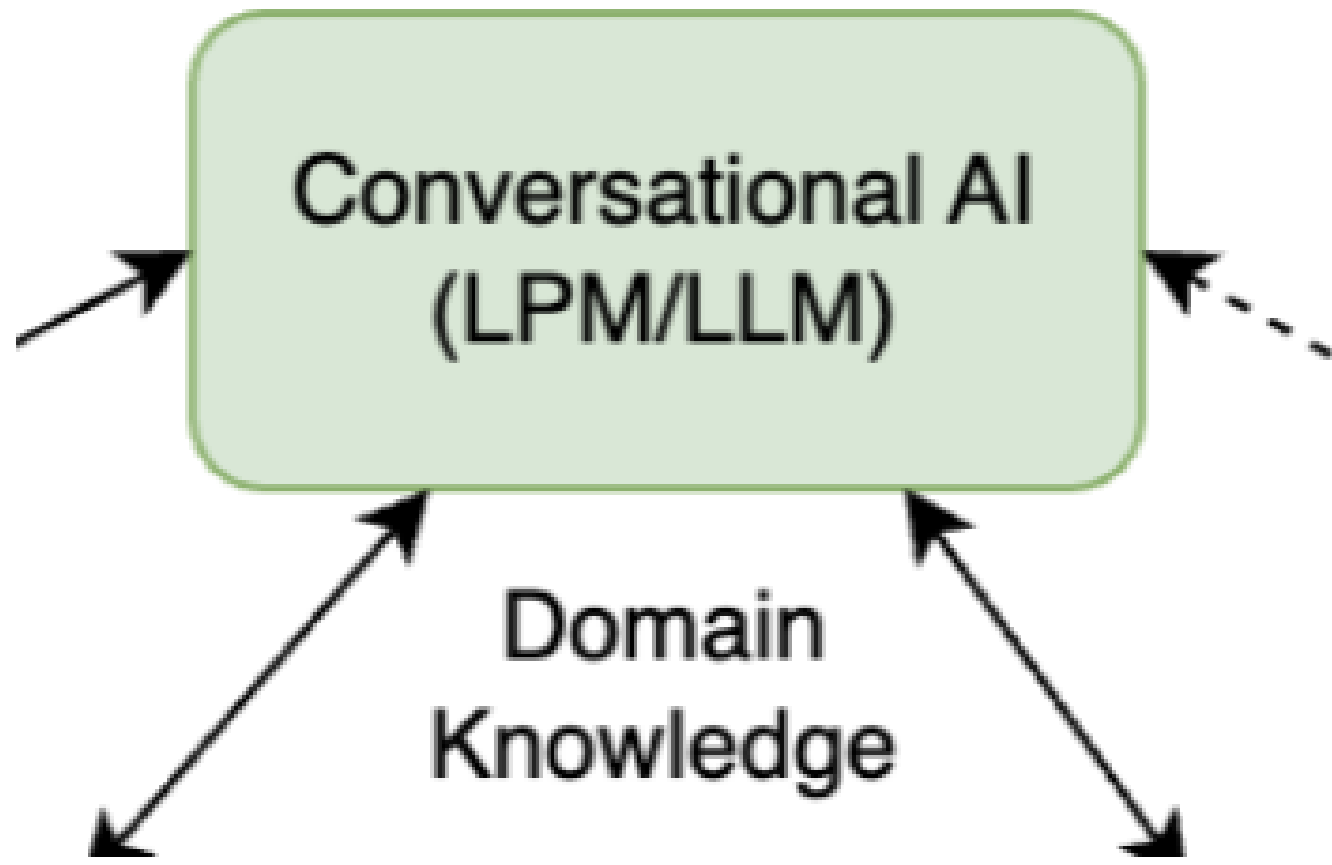


Mostly Noise-dominated time series



Masses
Spins
Signal model
GR consistency?
etc.

$$h(t) \longrightarrow \text{Representation} \longrightarrow \text{Physics}$$



Conversational AI
(LPM/LLM)

Domain
Knowledge

Towards a Benchmark for Scientific Understanding in Humans and Machines

Kristian Gonzalez Barman^a, Sascha Caron^{b c}, Tom Claassen^d, Henk de Regt^a

^a *Institute for Science in Society, Faculty of Science, Radboud University, the Netherlands.*

^b *High Energy Physics, Faculty of Science, Radboud University, the Netherlands.*

^c *Nikhef, Science Park 105, 1098 XG Amsterdam, the Netherlands.*

^d *Institute for Computing and Information Sciences, Faculty of Science, Radboud University, the Netherlands.*

E-mail: kristian@gonzalezbarman@ru.nl , scaron@nikhef.nl , tomc@cs.ru.nl , henk.deregt@ru.nl

Abstract

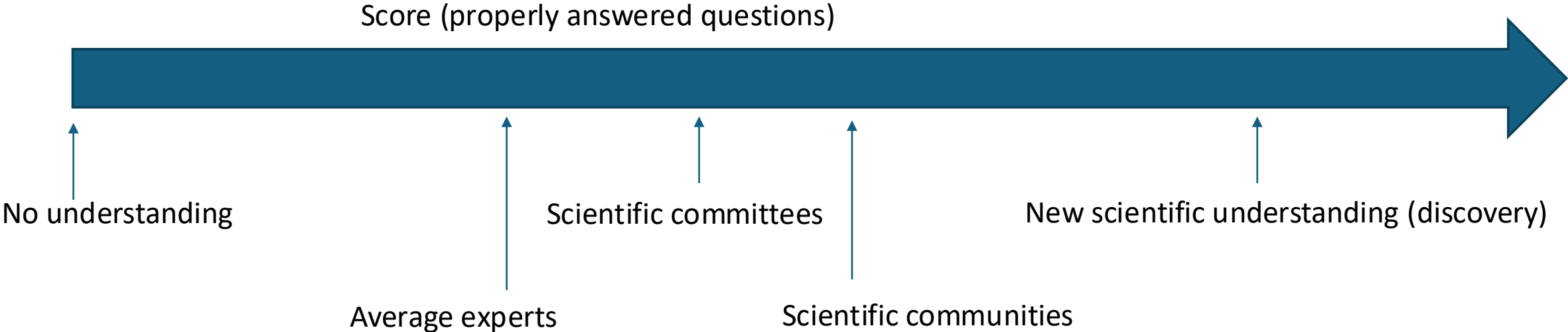
Scientific understanding is a fundamental goal of science, allowing us to explain the world. There is currently no good way to measure the scientific understanding of agents, whether these be humans or Artificial Intelligence systems. Without a clear benchmark, it is challenging to evaluate and compare different levels of and approaches to scientific understanding. In this Roadmap, we propose a framework to create a benchmark for scientific understanding, utilizing tools from philosophy of science. We adopt a behavioral notion according to which genuine understanding should be recognized as an ability to perform certain tasks. We extend this notion by considering a set of questions that can gauge different levels of scientific understanding, covering information retrieval, the capability to arrange information to produce an explanation, and the ability to infer how things would be different under different circumstances. The Scientific Understanding Benchmark (SUB), which is formed by a set of these tests, allows for the evaluation and comparison of different approaches. Benchmarking plays a crucial role in establishing trust, ensuring quality control, and providing a basis for performance evaluation. By aligning machine and human scientific understanding we can improve their utility, ultimately advancing scientific understanding and helping to discover new insights within machines.

Instead of presupposing that internal mental states and representations are required for understanding,

we suggest to identify understanding with an agent's ability to reason about and manipulate objects of investigation.

Also: Understanding is not binary ! → Score !

Benchmark for “Scientific Understanding” of agents (humans and AI)



Towards a Large Physics Benchmark

Kristian G. Barman*¹, Sascha Caron*^{2,3}, Faegheh Hasibi², Eugene Shalugin²,
Yoris Marcet², Johannes Otte², Henk W. de Regt², and Merijn Moody^{4,5}

¹Ghent University, ²IMAPP and ICIS, Radboud University, ³Nikhef, NL, ⁴Dutch Institute of Emergent Phenomena, University of Amsterdam, ⁵Institute of Physics, University of Amsterdam

July 30, 2025

<https://arxiv.org/pdf/2507.21695>

We propose a framework for a **benchmark for fundamental physics**:

Collecting 3 types of (difficult/deep) questions:

1. A, B, C, D
2. Open end (what is the result of X,Y,Z)
3. Score → Higher is better (Code)

*If you like to propose/evaluate a question and become co-author
→ Send email to scaron@nikhef.nl
and eugene.shalugin@ru.nl
for the access token*

Grand vision 2: Build AI communities

Motivation

- AI is becoming core infrastructure for physics (HEP, APP, GW, theory).
- Not just method development important → but **organization, coordination, and strategy**.

Questions

- *How can we fund AI in physics via national + EU ?*
- *How do we structure AI in physics at European levels ?*
- *How do we connect to and develop AI initiatives ?*



Home

Activities

Members

Material

How to join

Contact



European Coalition for AI in Fundamental Physics

 **Upcoming**

EuCAIFCon 2026

August 24 - 28, Heidelberg

European: EuCAIF

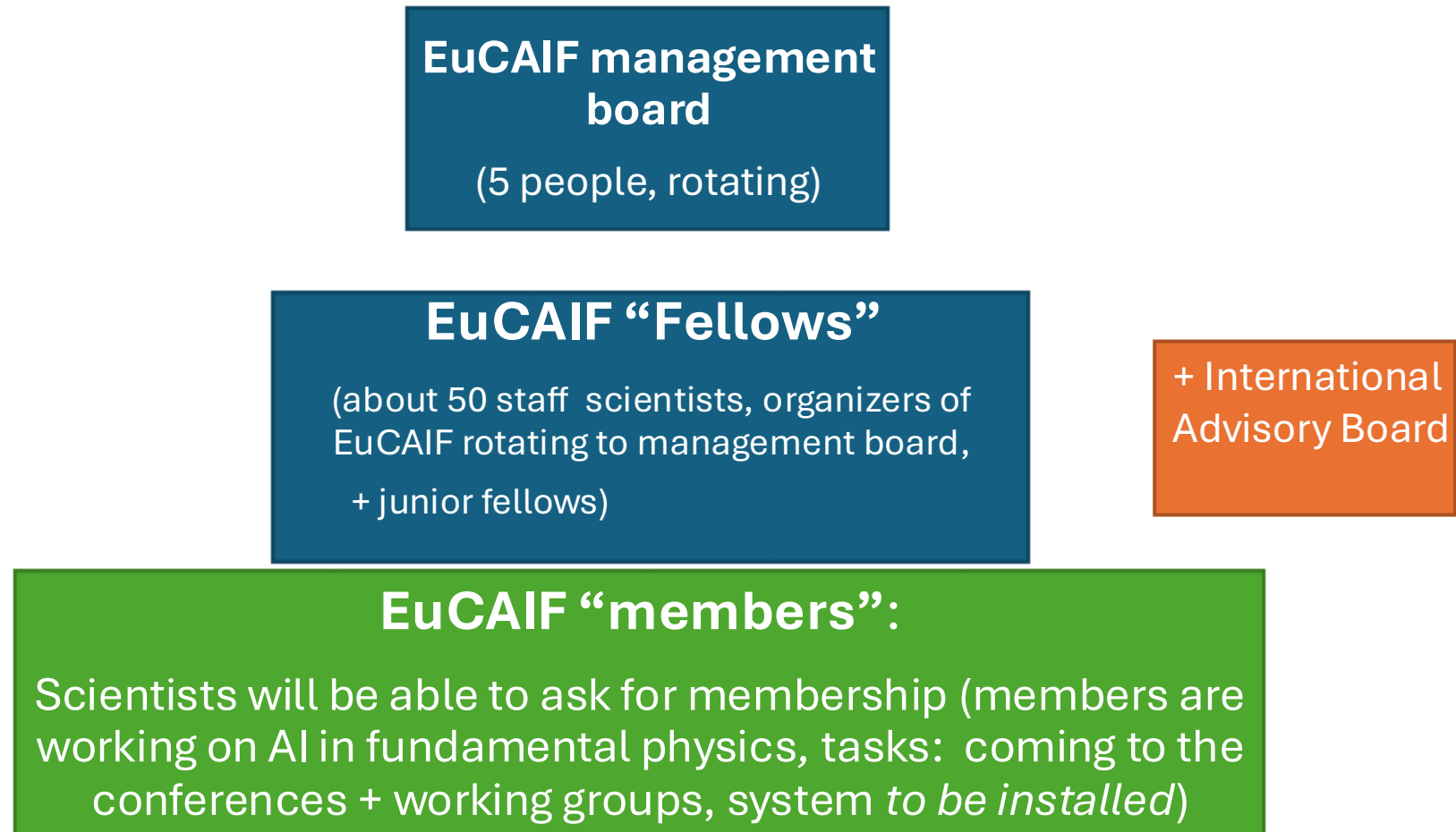
European Coalition for AI in Fundamental Physics (*ECFA, APPEC, NuPECC endorsed*)

- ~60 Fellows + Junior Fellows
- De facto coordinating body for AI in fundamental physics in Europe

Objectives

- EuCAIFCon (European AI-in-Physics conference, in [Heidelberg August 2026](https://indico.physi.uni-heidelberg.de/event/1277/), <https://indico.physi.uni-heidelberg.de/event/1277/>)
- Strategic engagement (EU strategy & funding)
- Large-scale joint proposals
- Scientific working groups
- ...

EuCAIF organizational structure



EuCAIF Working groups

WG 1: Foundation models & discovery

WG 2: AI-assisted co-design of future ground- and space-based detectors

WG 3: FAIR-ness & Sustainability

WG 4: Machine Learning and Artificial Intelligence Infrastructure (JENA WP4)

WG 5: Building bridges - Community, connections and funding

Which Foundation Models Working Group are you interested in joining? *

- Gravitational Waves
- Multi-Messenger
- Cosmology Cross-Correlation
- Neutrinos + Gamma-Rays
- Galaxy Evolution
- Cross-cutting methods: Detector Simulation & Fast Inference
- Cross-cutting methods: Quark-Gluon Plasma Physics

https://docs.google.com/forms/d/e/1FAIpQLScnMPFN5jfkf84979xPVt_NQ5ikayg3zmNtWsQSEB8G4p5zTrA/viewform?pli=1

Roberto di Austri & Ik Siong Heng

EuCAIF info

Is “AI in GW” your main research topic ? Visit www.eucaif.org to see how you can become a Fellow or Junior Fellow.

If you like to follow the activities of EuCAIF please join the following e-group: eucaif-info@cern.ch

.

.

European RAISE

“**RAISE – the Resource for Artificial Intelligence Science in Europe**”

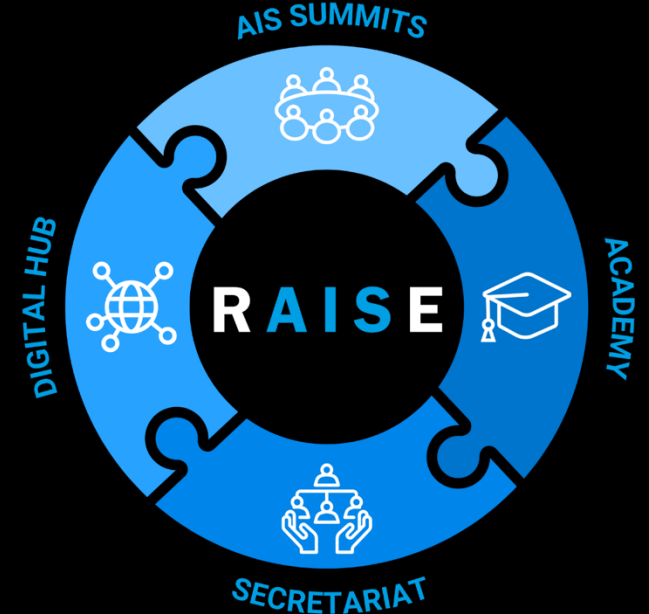
- Flagship European initiative
- Pillars:
 - **AI in Science**
 - **Science for AI → This is us !** ~€100M per year (various funding calls)
 - Computing infrastructure
- Funding instruments:
 - **Small-scale projects (3 M)**
 - **~€15M network initiatives**
- **Next major opportunity: 2027 network call, end of 2026 for synergy grant scale**

SCIENCE Overview

SCIENTIFICALLY
CREDIBLE

AI in Science & AI for Science

STRATEGICALLY
ALIGNED



SCIANCE and RAISE

SCIANCE is coordinating an support action for RAISE
Nikhef responsible for fundamental physics

Activities:

- Set up AI in physics expert panel for europe
- Set up Strategic adviosory board
- Define / summarize scientific priorities
- Organize AI in science summits
- Others (see backup slides)

Summary

- **Vision 1: Toward Large Physics Models**
- **Vision 2: Building a European AI Infrastructure for Physics**

--- Extra slides ---

Why SCIENCE?



Bottom-up approach for AI-enabled scientific research across Europe, involving the scientific community directly



Facilitate cross-domain collaborations between scientific disciplines, AI research, industry and EU infrastructures



Improve data access, infrastructure capabilities, and skills, through feasible upgrades scenarios



Provide structured guidance, cooperation mechanisms and strategic priorities, via a SRIA and implementation roadmap



Increase visibility and integration of AI research and innovations for scientific research

SCIENCE Overview

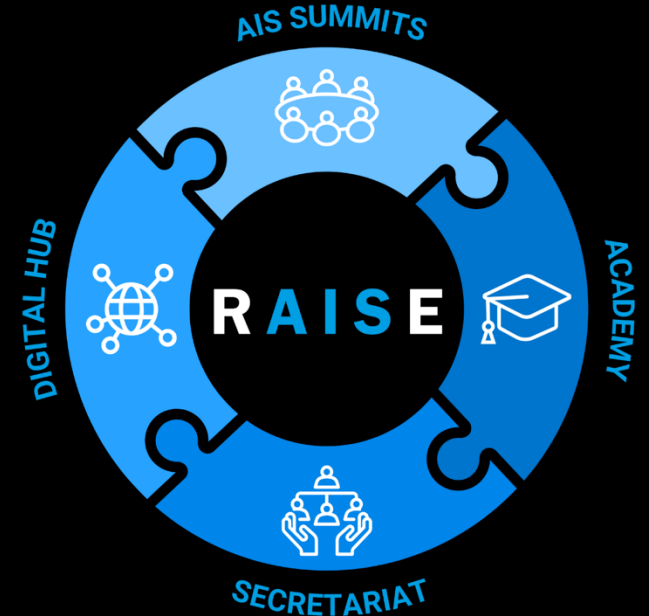
SCIENTIFICALLY
CREDIBLE



AI in Science & AI for Science

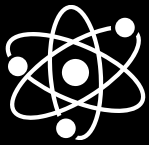


STRATEGICALLY
ALIGNED



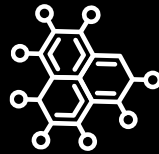
AI in Science Priorities

Scientific Pilot Areas



Physics &
Astronomy

Nikhef &
University of
Manchester



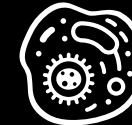
Materials
Science

Constructor
University



Earth
Sciences

Twente
University ITC



Life
Sciences

Euro-
BioImaging



Social Sciences
& Humanities

CNR-ISTI
(SoBigData)



Evidence
Base



SRIA

Co-creation and Alignment with EU priorities



AI in Science Working Groups

Co-creation process involving AI in Science Working Groups to identify SRIA priorities, infrastructure upgrade requirements and cooperation mechanisms, supported by the [RAISE Secretariat](#)



Community Consultations

Validation through broader consultations via [RAISE Digital Hub](#), partner networks and dedicated events at AIS summits



Strategic Advisory Board

Continuous feedback with key European actors including research infrastructure, AI networks and infrastructures, EC & member states (ERA action) and industry.



Policy Alignment

Coordination with European policymakers, including European Commission, Member States, and relevant European initiatives

AI in Science Community Building and Sustainability



RAISE Secretariat pilot for AI in Science

- Support RAISE AI in science coordination activities and SCIANCE AI in Science Working Groups
- Facilitate cooperation and community building within the AI in Science Ecosystem
- Explore operational scenarios and potential funding models to support AI-enabled science beyond the project
- Ensure continuity and sustainability of SCIANCE key exploitable results and project assets for RAISE



RAISE Digital Hub

- Online platform for project resources, best practices, consultation tools, outputs from domain-specific pilots, and training resources
- Long-term access portal for RAISE AI in Science, interface with broader scientific communities and infrastructures



RAISE Academy pilot and Bootcamp

Targeted capacity-building to equip decision-makers with knowledge and tools needed to implement the SRIA



AIS Summits Events

Co-organised sessions, community events, roundtables and consultations at AIS summits.