



Utrecht University

Nikhef

# AI in ALICE

Marta Verweij

# Current impact AI

1.1) What are the **physics areas of application** where AI is having a major impact internationally in production use in your research area?

- Briefly sketch the **purpose of application** and what **ML/AI architectures/techniques** are currently used (LLM, GNNs etc...) along any with major **computing resource requirements** (for training / for application)

- Cut optimization: simple models
- Heavy-flavour tagging: deep neural networks
- Track reconstruction in Inner Tracking System and MFT: cellular automaton
- Jet energy calibration and background subtraction: deep neural networks
- Rare signal hunting and anomaly detection:
  - Unsupervised because physics and type of signal is unknown
  - reinforcement learning (for example LSTM), transformers

# Involvement Nikhef scientist

2.1) Are you **involved in** or **internationally leading in** R&D\* on the application of AI/ML methods in your research area? Or are you mainly 'end-users'?

- We did development of an algorithm from scratch for one project, but this is better left to the computing scientist.
- There are interesting opportunities for us combining different tools. For example a transformer with One Class SVDD.

2.2) Who are the **in-house experts**? What is their **expertise**?

Please distinguish staff vs PhD/PD expertise.

- Alessandro Grelli: DNN on FPGA, domain adversarial neural net (DANN)
- Marta Verweij: LSTM, One Class methods, transformers. All supervised
- Plenty of PhDs, and BSc+ MSc students

# Involvement Nikhef scientist

2.3) To what extent use [MSc/BSc projects](#) in your group focus on AI/ML use?

Do you have experience with Comp.Science (AI) and/or Math students (Bsc/Msc/PhD) in your group working AI/ML in physics?

- Many of our students (~70%) at all levels do projects involving AI
- Collaboration with mathematicians at Utrecht University. Especially fruitful with our physics+math TWIN bachelor students

2.4) Are you (at Nikhef) using [AI-based coding assistance tools](#)?

- Some of our researchers use it to speed up their progress. Expect to get better and more important for day-to-day activities.

# AI next years – midterm (2-4 years)

What is the expected ML/AI use in your research area in the next years?

- 3.1) Extrapolating from current experience, where will **future/improved deployment of ML/AI** bring **strong gains in physics performance** in the next few years?

What AI/ML **techniques** do you expect to perform strongly in the near future

(LLMs, foundation models, simulation-based inference, generative models etc etc)

**Inference and explainability.** Especially when looking for unknown signal it is important to understand why the machine is making certain decisions.

**Broader and more-controlled use in physics data analysis on all topics, especially considering the much larger data rates in ALICE in run 3.**

- 3.2) Are there major **novel ML/AI application** areas in the pipeline? (i.e. areas where ML is now not used)

**Areas to explore:**

- **detector simulation.** GEANT simulation are currently a bottleneck so if AI can help in this area that would be very beneficial. Generative AI.
- **tracking:** here ML is already used in the subsystems. New development needed for future faster detectors. Transformers, foundational methods.
- **Hardware data reduction using AI:** event trigger on FPGA with neural net (pp only, Pb-Pb no online trigger)

# AI next years – midterm (2-4 years)

- 3.3) What are the expected future [computing resource requirements](#) (for training / inference)?

No clear picture at the moment

Challenge identified for tracking: parallization on of the GPU for full event tracking. Especially in Pb-Pb it is a very high density track environment.

# AI next years – midterm (2-4 years)

What is the involvement of Nikhef scientists in ML/AI deployment R&D for next years?

4.1) What is the ambition of your Nikhef research program for AI/ML use and deployment R&D?

Be part of the new developments on the data analysis and tracking side.

4.1) Are there [clear leading institutes/consortia](#) in the international field in these R&D efforts?

It is scattered over many institutes worldwide

4.3) How do [Nikhef efforts compare](#) to these in a) expertise, b) person-power, c) infrastructure?

a) Expertise: we lack computer scientists and could profit from more collaboration with mathematicians and AI experts.

b) Person-power: see a) + PhDs

c) Up to now Nikhef cluster was enough. Expect higher need of resources in the future.

4.4) Who are your [partners nationally and internationally](#) in your efforts & ambitions?

Changes per project, especially on the data analysis side.

4.5) What [expertise, person-power, infrastructure are you missing](#) to compete effectively and/or realize your ambitions?

PhDs, postdocs, computing scientist.

4.6) What are your expectations and/or plans related to AI-driven [coding assistance](#) and/or [ChatGPT-style AI](#) (physics) knowledge services?

Importance will grow a lot in time.

# AI next years – longterm (5-10 years)

## C – Long-term future (5-10 years) – future directions

Given the pace of developments in AI/ML, predicting the evolution of its future possibilities on a 5–10y scale is rather speculative.

Instead, here we focus on future application areas where disruptive improvements in AI/ML can make a difference

- Thinking ‘Disruptively Big’: what research applications are a good target for future AI methods
  - Pb-Pb tracking (~10 thousand tracks in one event)
  - Detector simulation and design
- 5.1) What are **problems ‘of interest’ that are currently unsolvable** (from practical computational point of view) but could be solvable with disruptive AI methods.

In other words, are there **paradigm-changing** way of thinking about **solving physics problems if ‘unlimited computational’ abilities** were to be available?

Clearly there is physics in tails of distributions (rare processes and anomalies). A combination of unlimited amount of data and computing will find them. Difficult to speculate.



# Other

6.1) Is there any other information regarding the use and potential of ML/AI in your research area that is relevant to be discussed in the task force?

ALICE has a statistics and machine learning topical group where we try to steer the various efforts.