Searching for Anomalies with Graphs

Graziella Russo

SAPIENZA UNIVERSITÀ DI ROMA

PhD seminars
09/11/22

INFN

ATLAS EXPERIMENT

Introduction

1. **Why?** Physics motivation for the search

2. **What** is an **anomaly?**

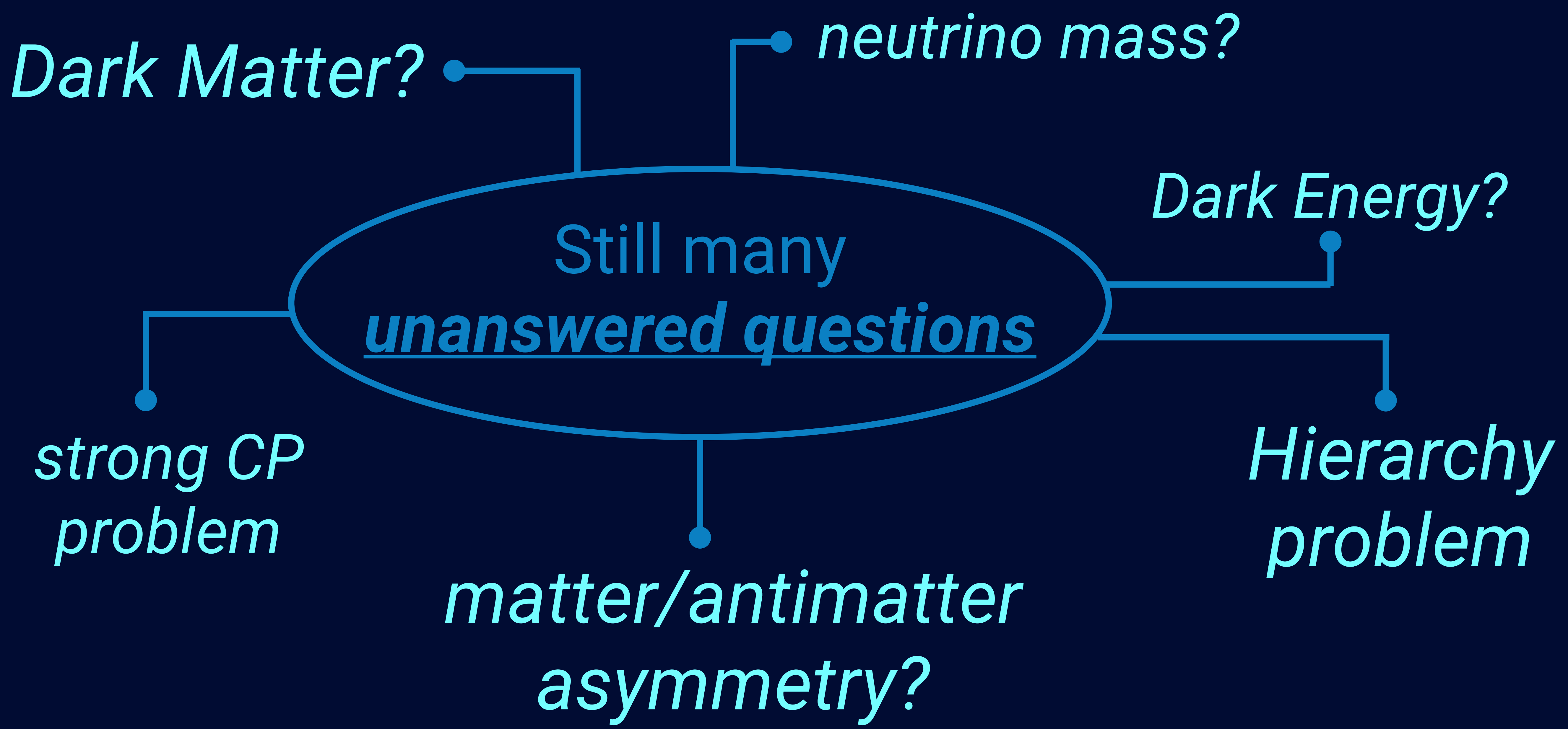
3. **How?** A Machine Learning approach

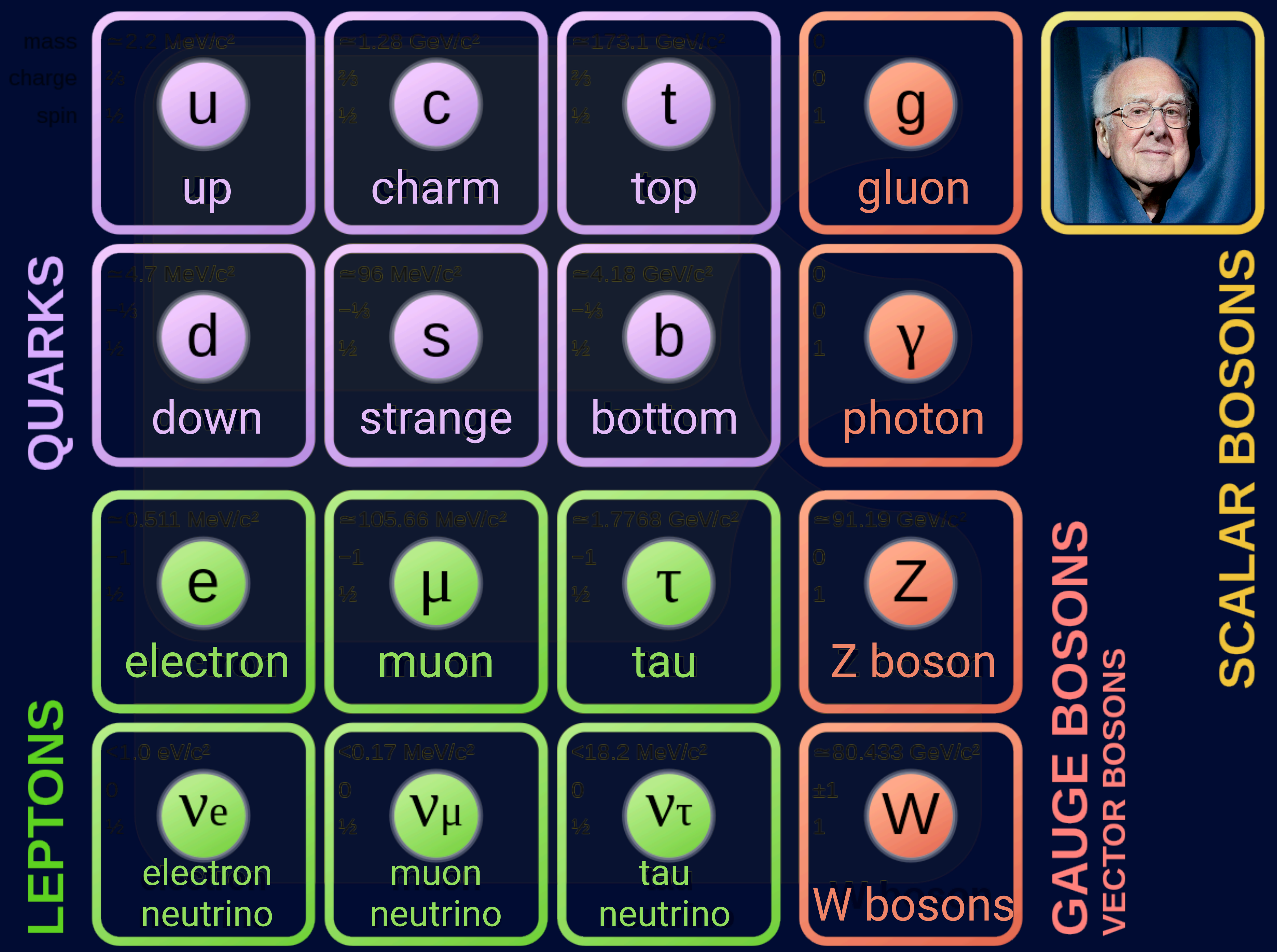
4. Brief introduction to Machine Learning and AutoEncoders

5. **Graph Anomaly Detection** in High Energy Physics analysis

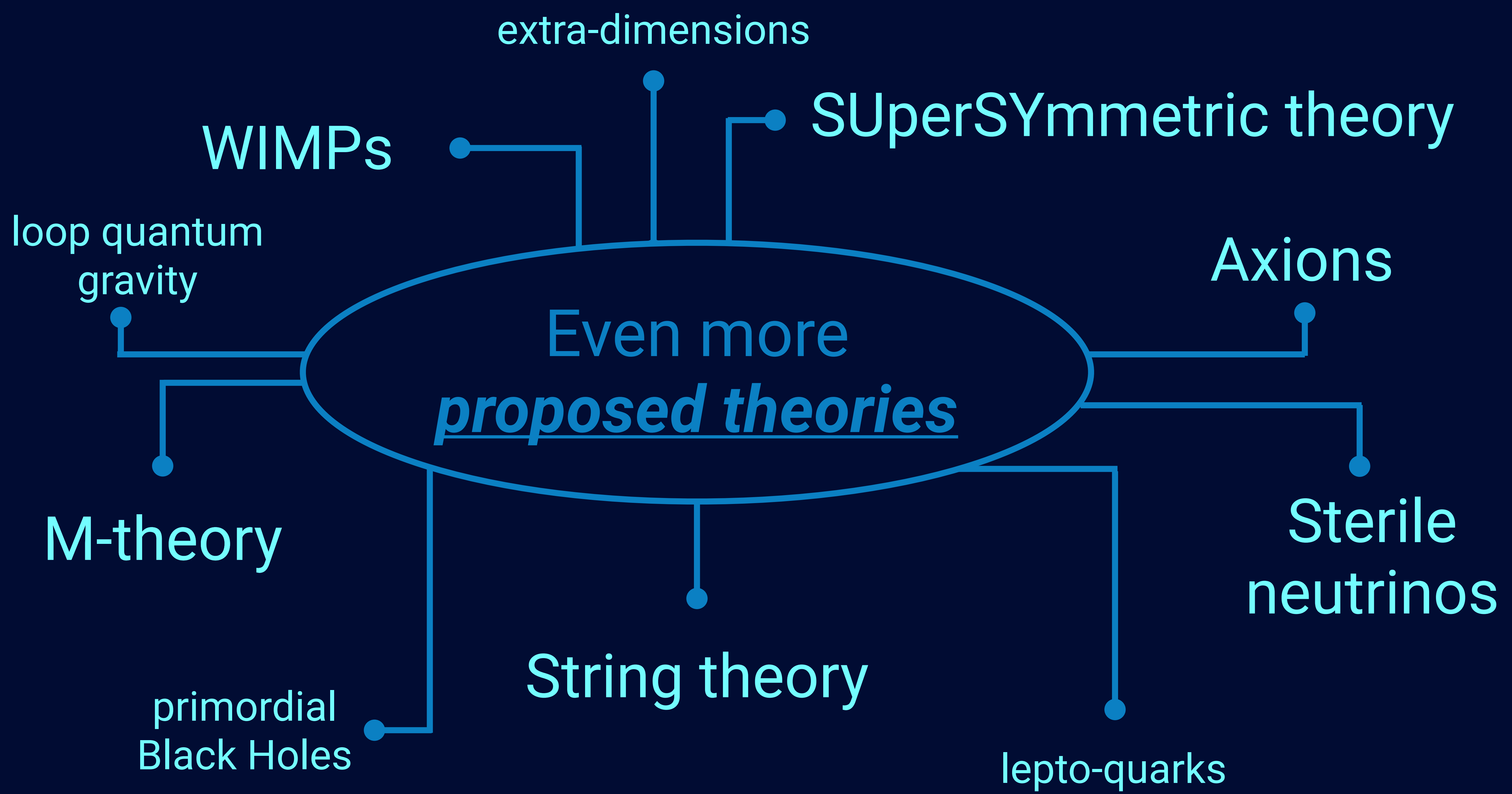
Why? Starting from the beginning

- Up to 2012 Standard Model was not complete: some particles still missing
- 2012: discovery of the Higgs boson 

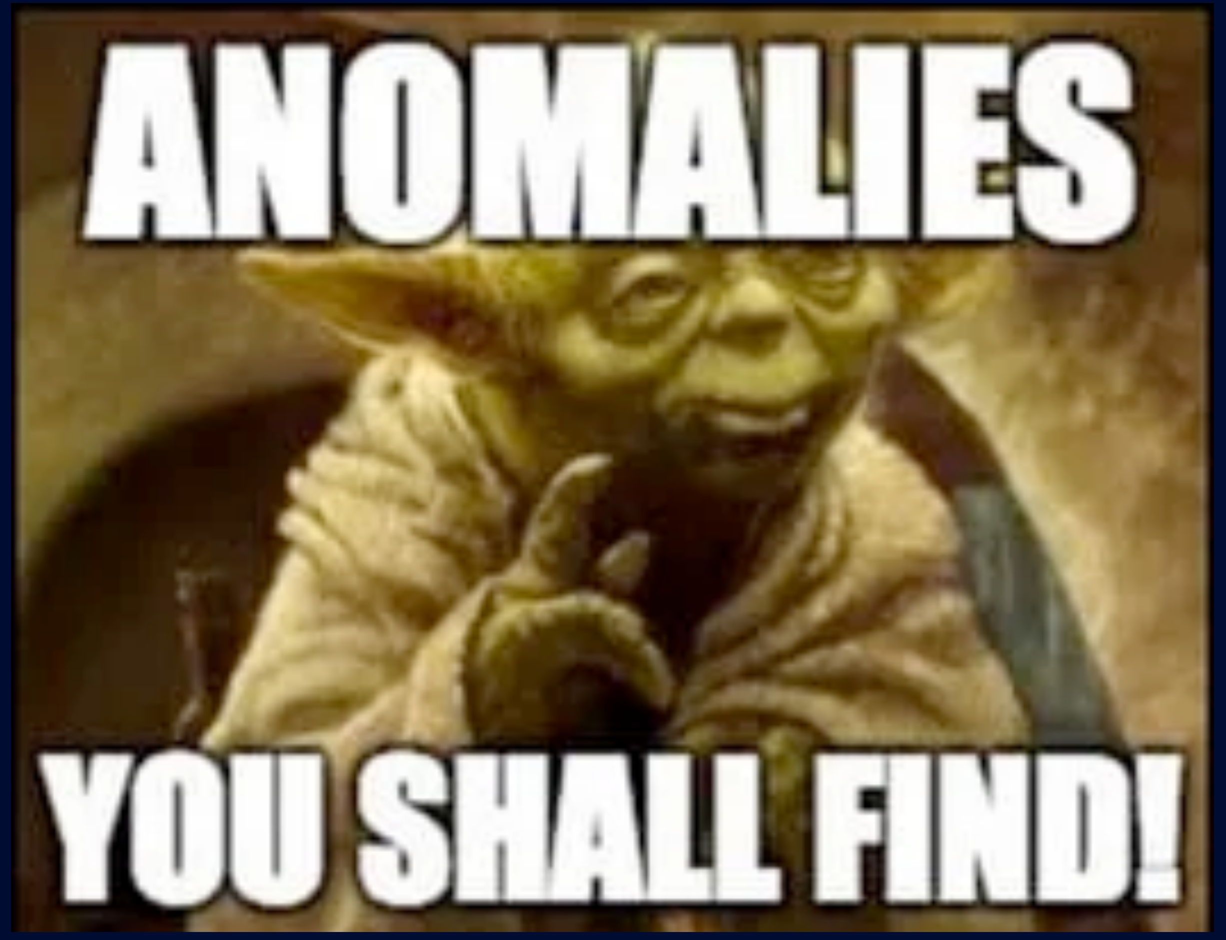




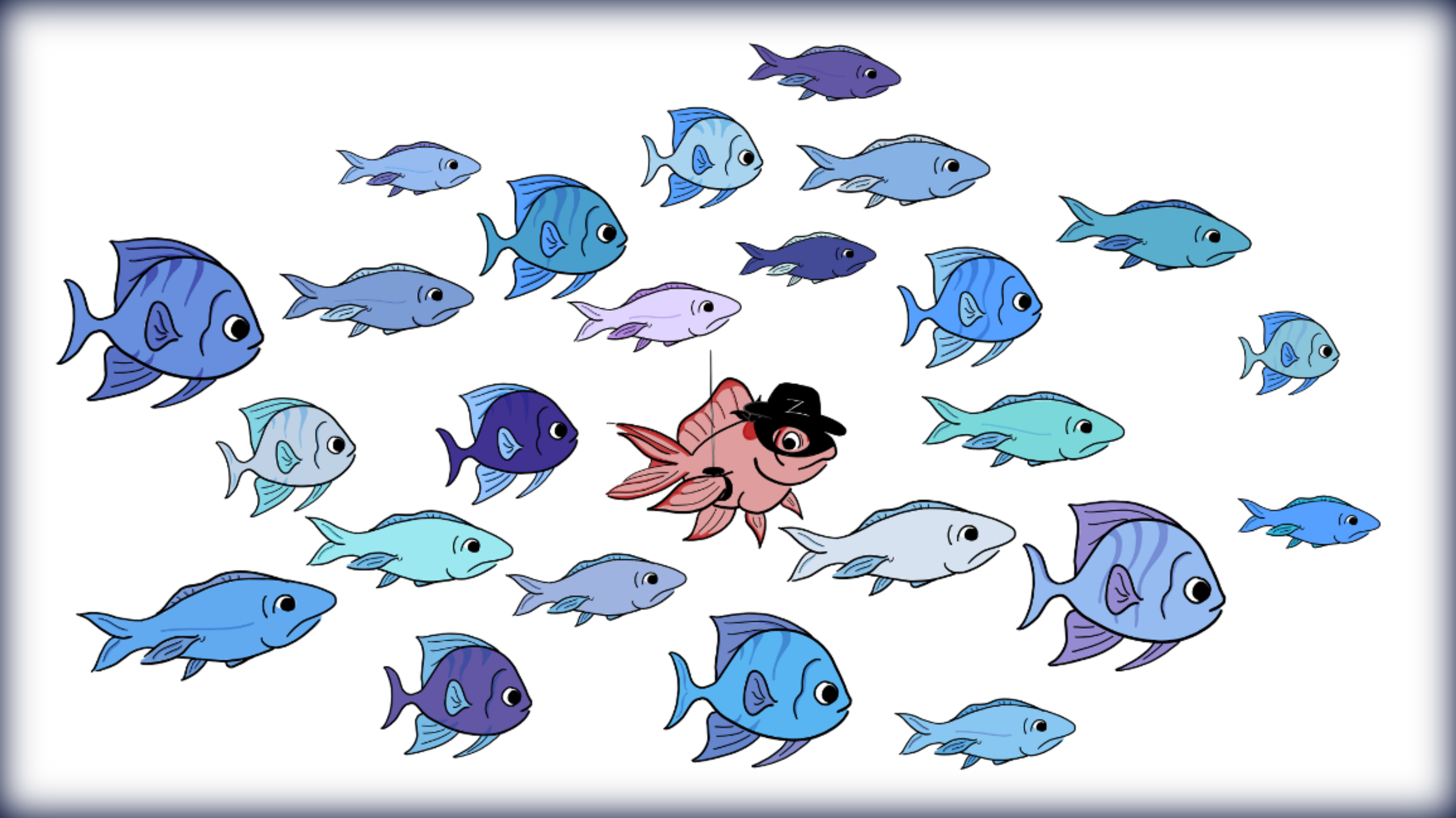
Why? Looking for New Physics



At LHC experiments: **hunt for New Physics**



What is an Anomaly?



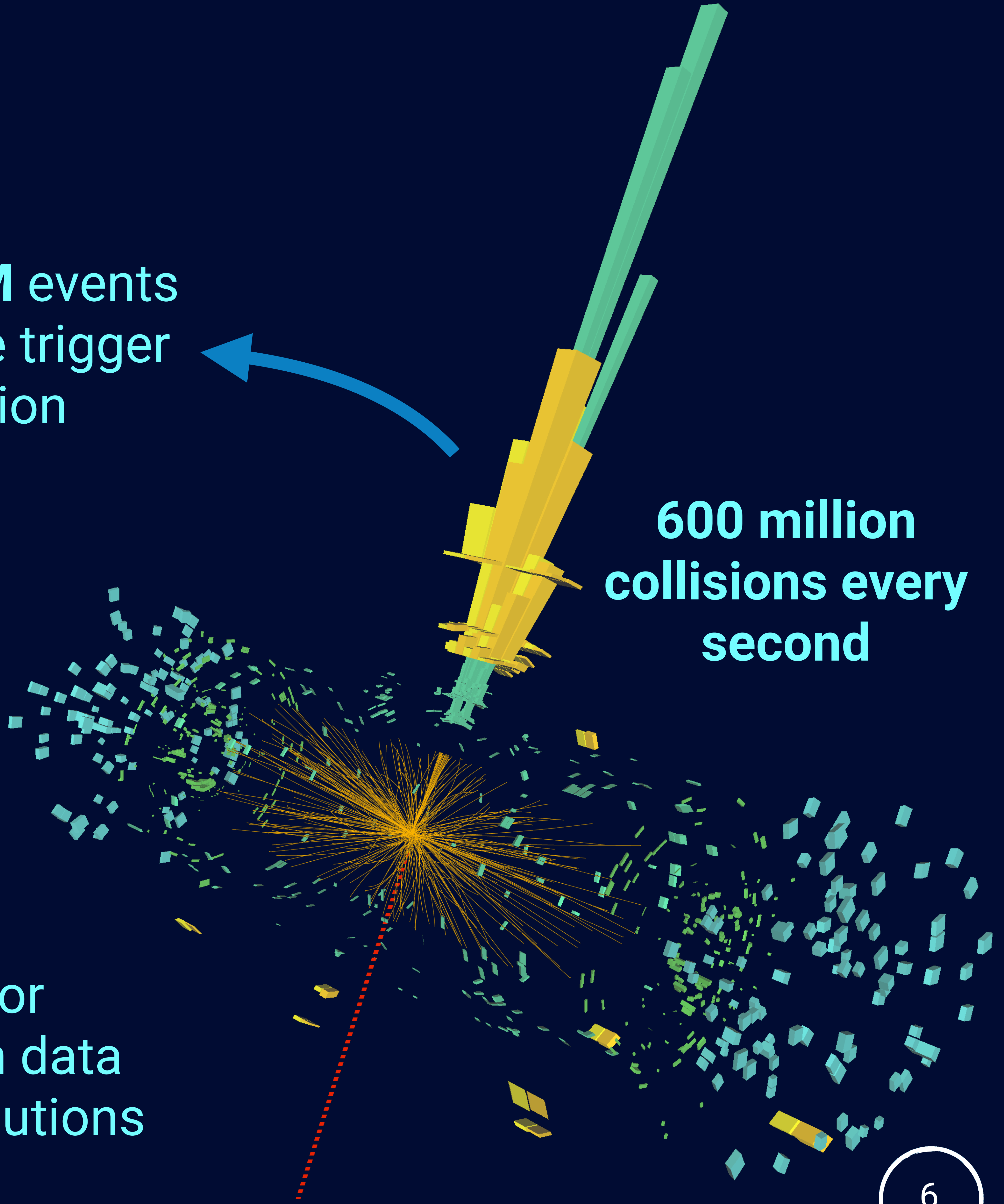
Anomalies are items, events or observations that are anomalous if compared to an established pattern

cyber-security intrusion detection, medical diagnosis, system health monitoring

Anomaly Detection (AD) is an Machine Learning (ML) technique to identify such abnormal behaviours

How? LHC, a big data factory

1 out of 1M events passes the trigger selection



600 million collisions every second

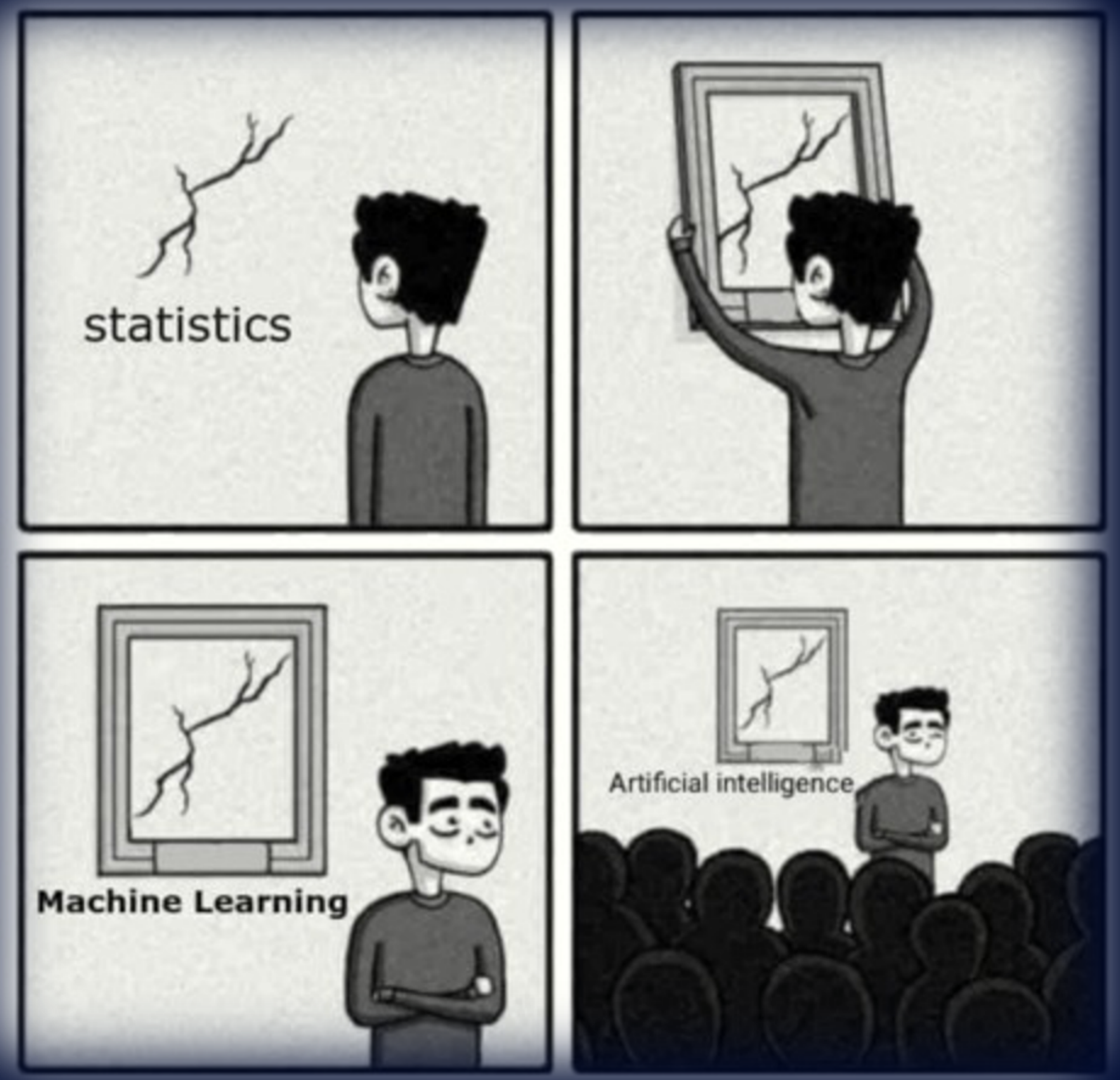
1.7 PB data per experiment per year

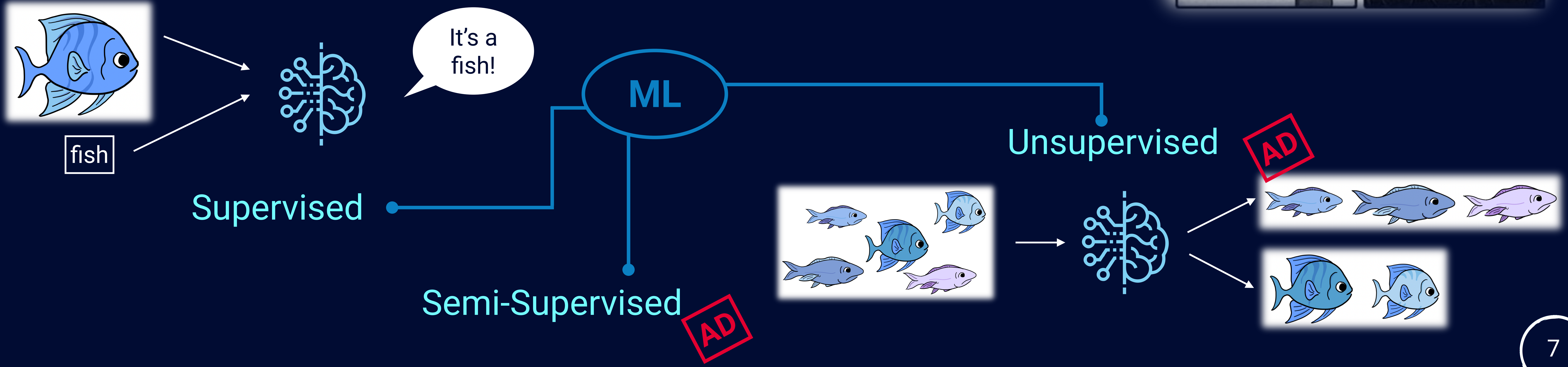
MC simulations for comparison between data and theoretical distributions

How? With Machine Learning

Machine Learning is building algorithms that learn a task by using experience

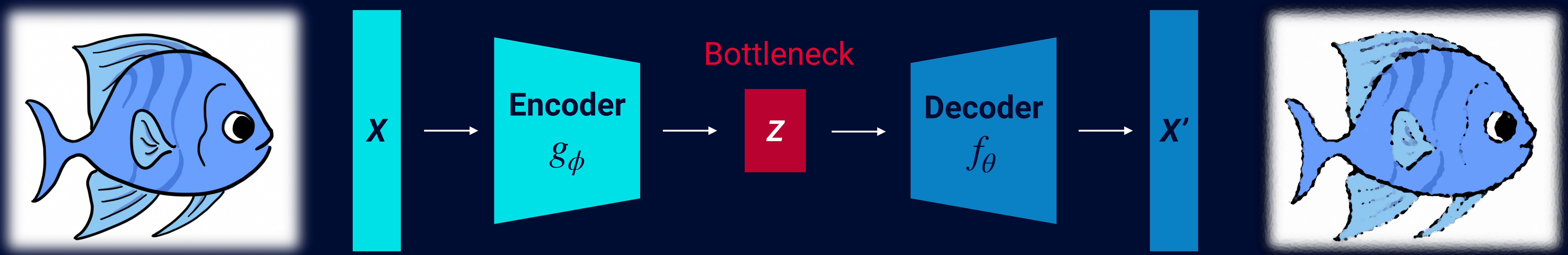
- 2 main phases: training and testing
- given a metric (loss), optimising the parameters





AutoEncoders

AutoEncoders are models that learn how to recreate the input (generative models), ignoring insignificant data (noise)



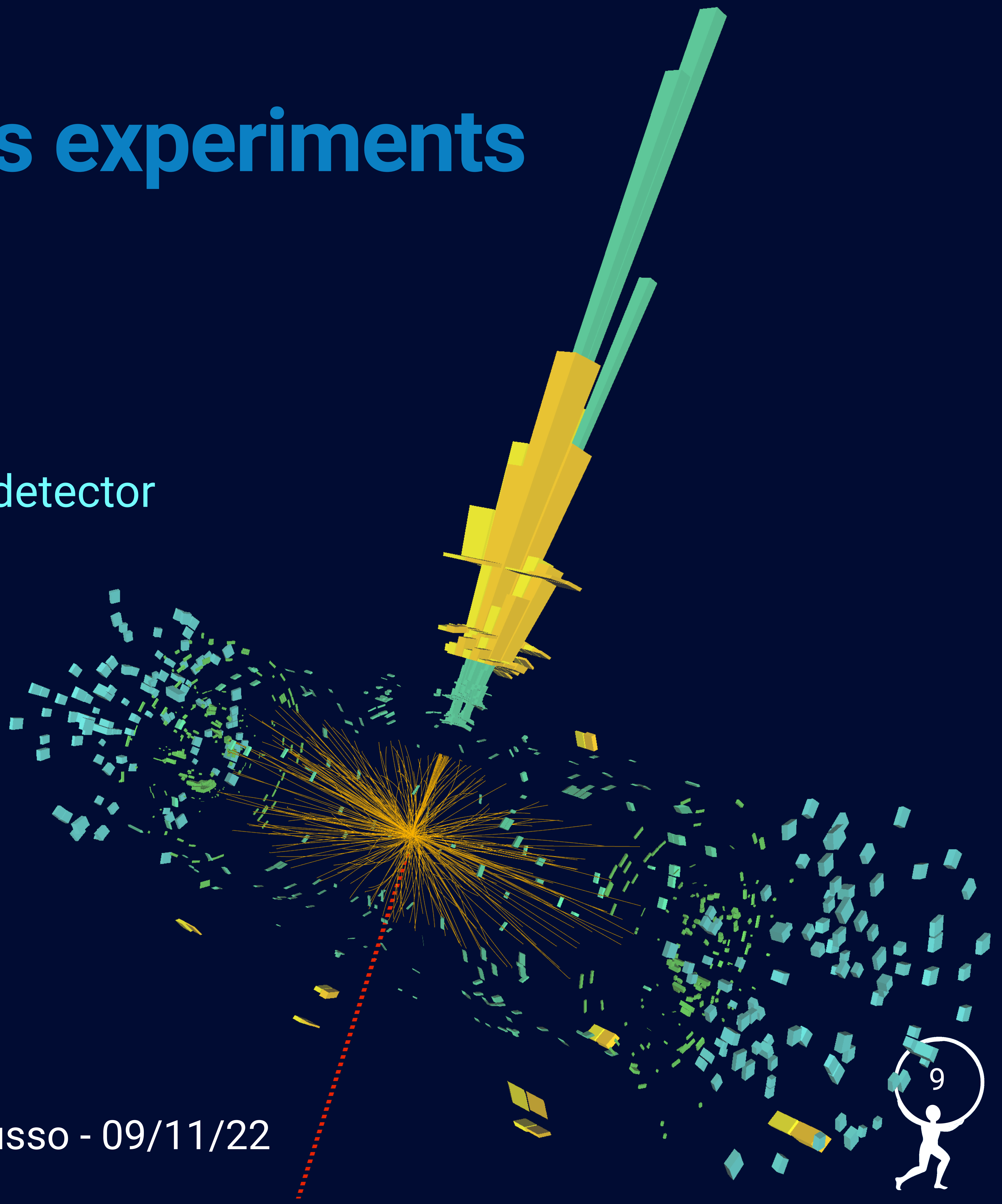
Task: minimise distance(X, X')

How can we use it in High Energy Physics experiments?

Data in High Energy Physics experiments

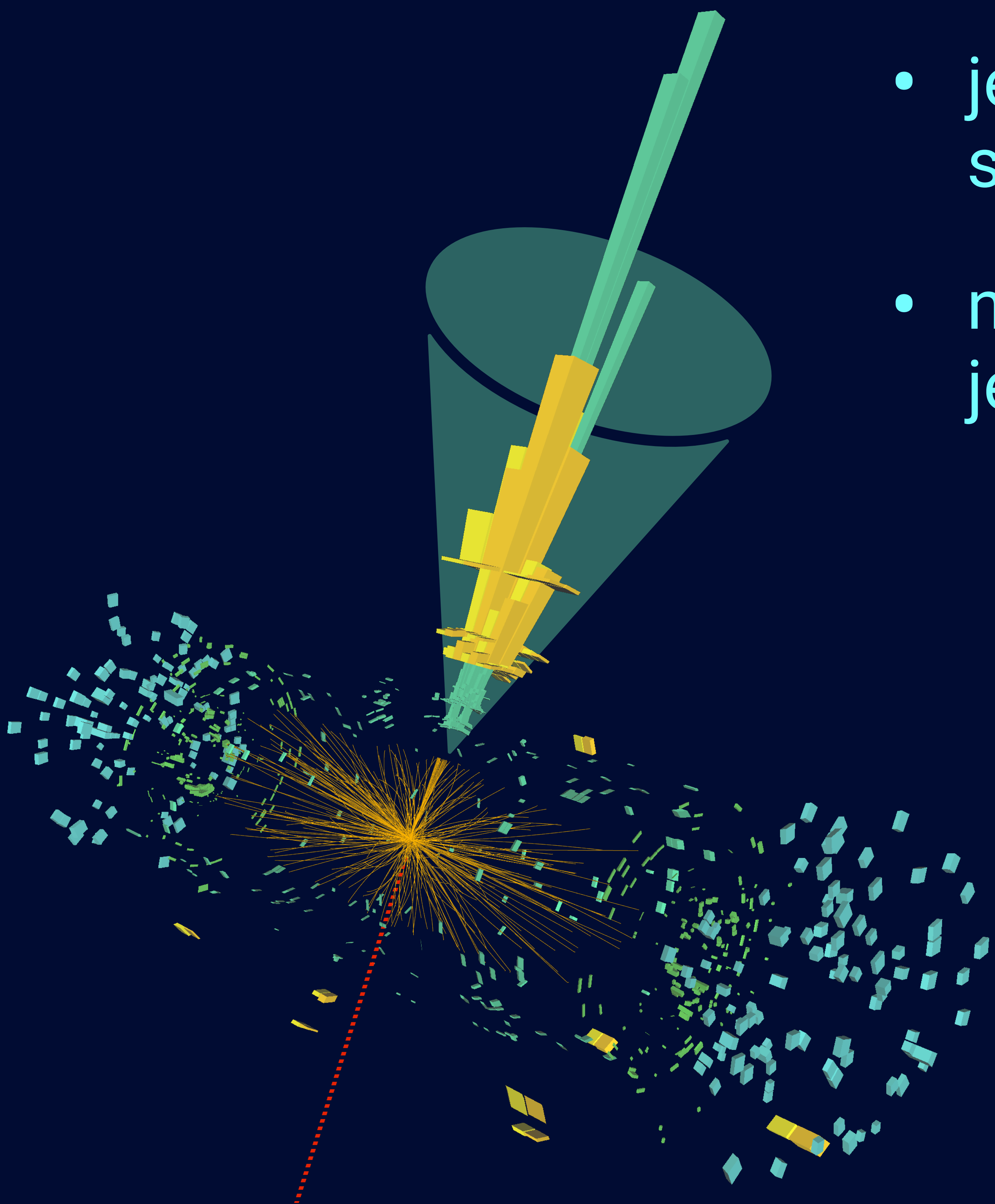
- Many sub-detectors to identify different particles
- Particles leave **tracks** and **energy deposits** in the detector

We can rearrange data in order to feed a ML algorithm



Jets and Graphs

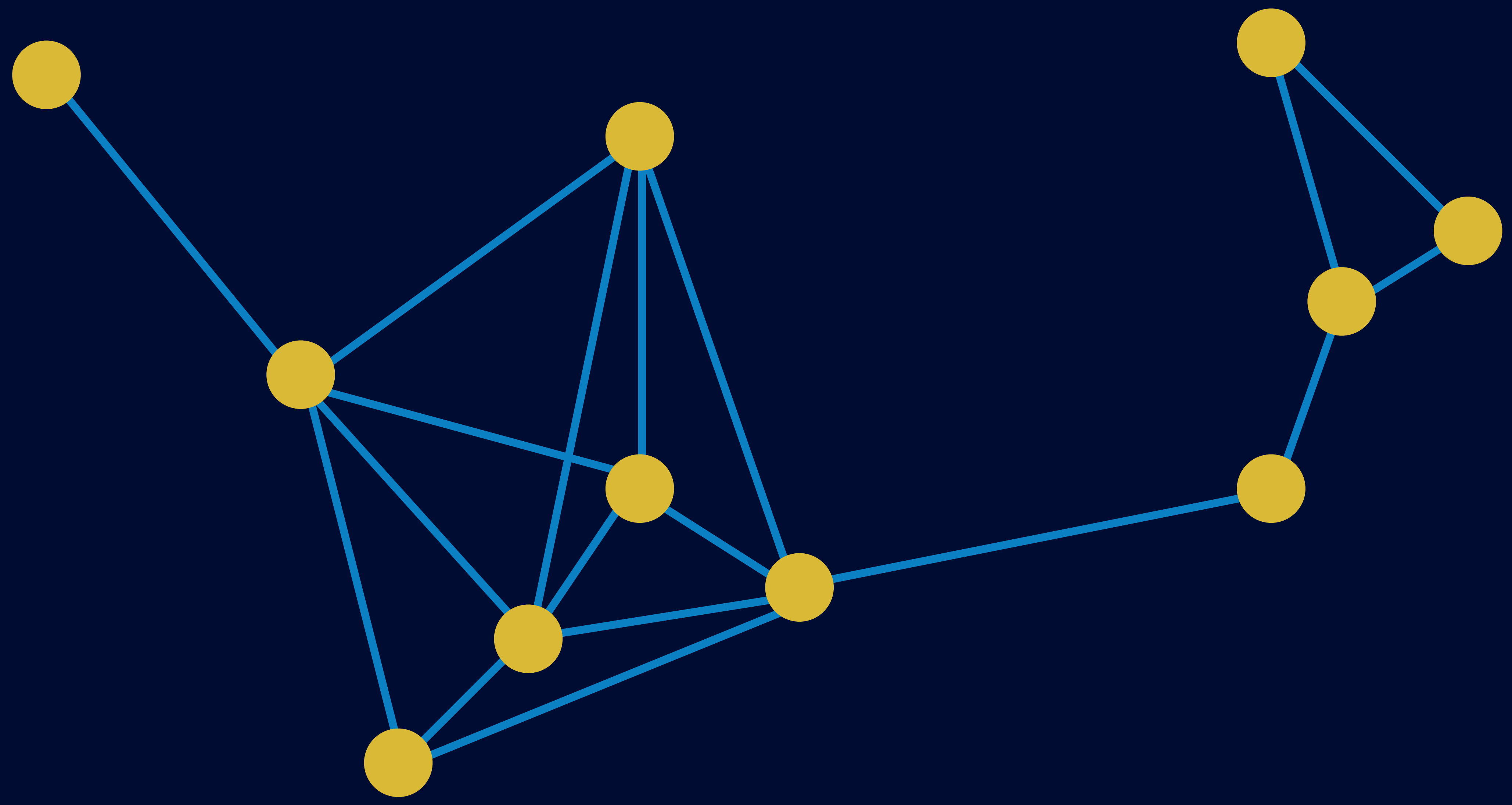
Jet: narrow cone of particles produced by the hadronization of quarks and gluons



- jets have a complex substructure
- not simple to identify a jet

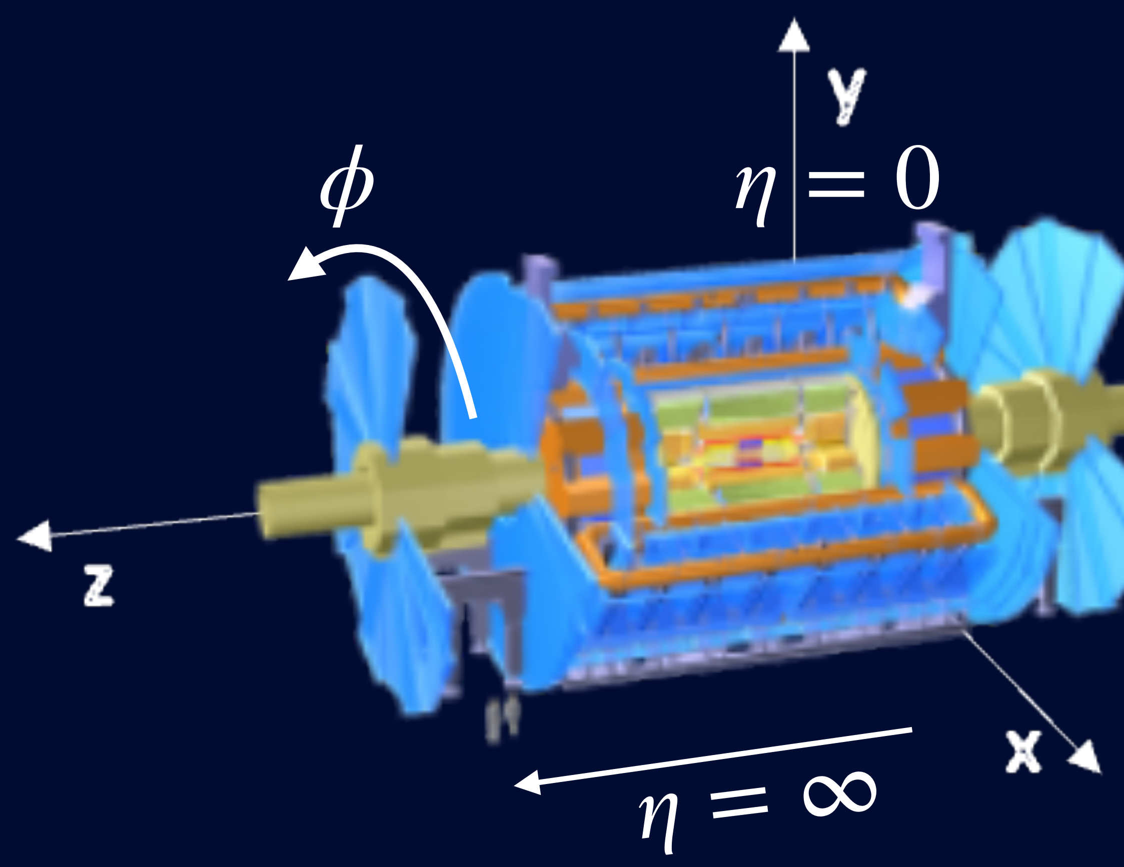
New Physics hidden inside the jet?

Graph: set of points (**nodes**) that can be connected (**edges**)

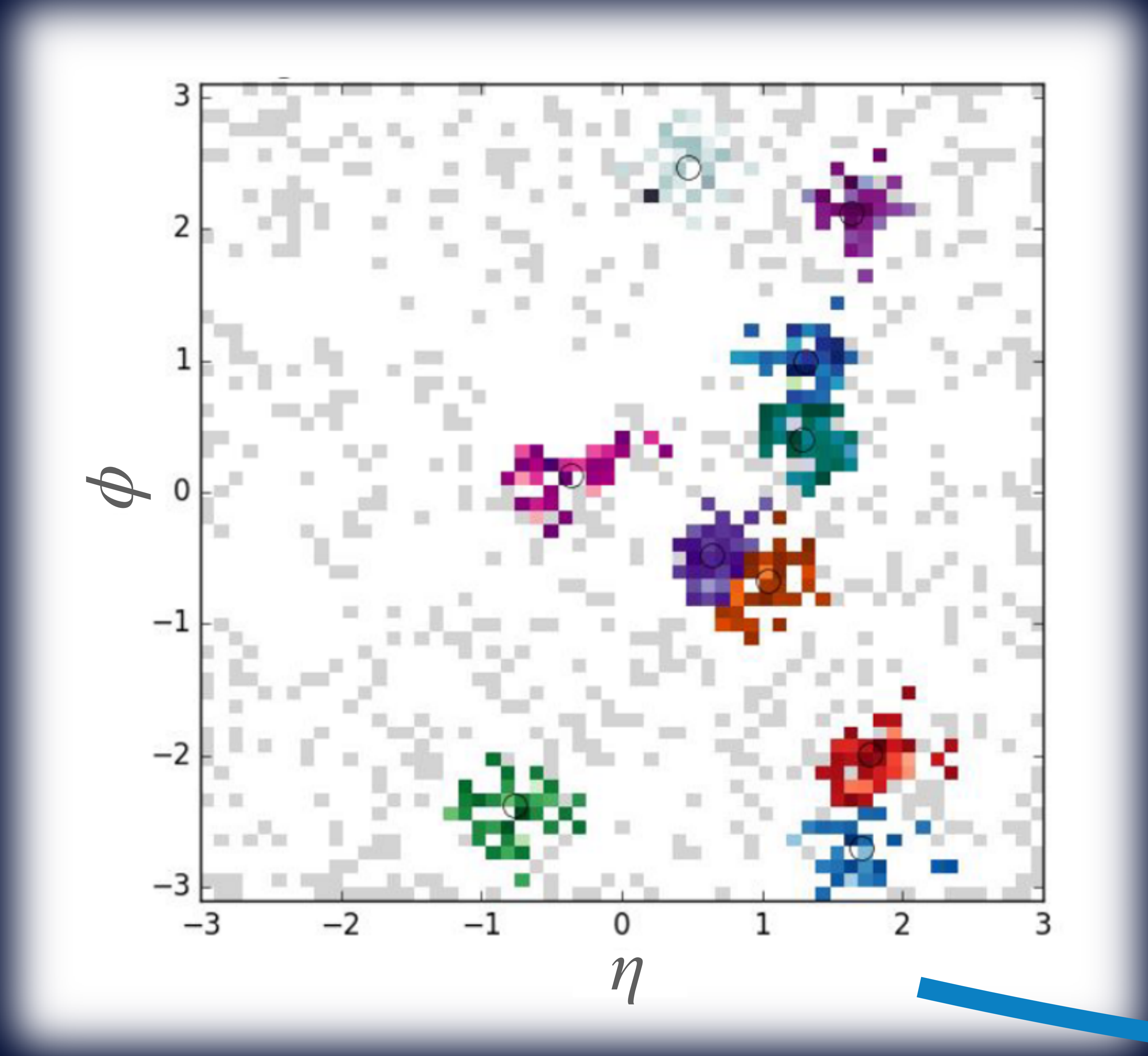


Useful for sparse data and whenever links are important

Graphs of Jets

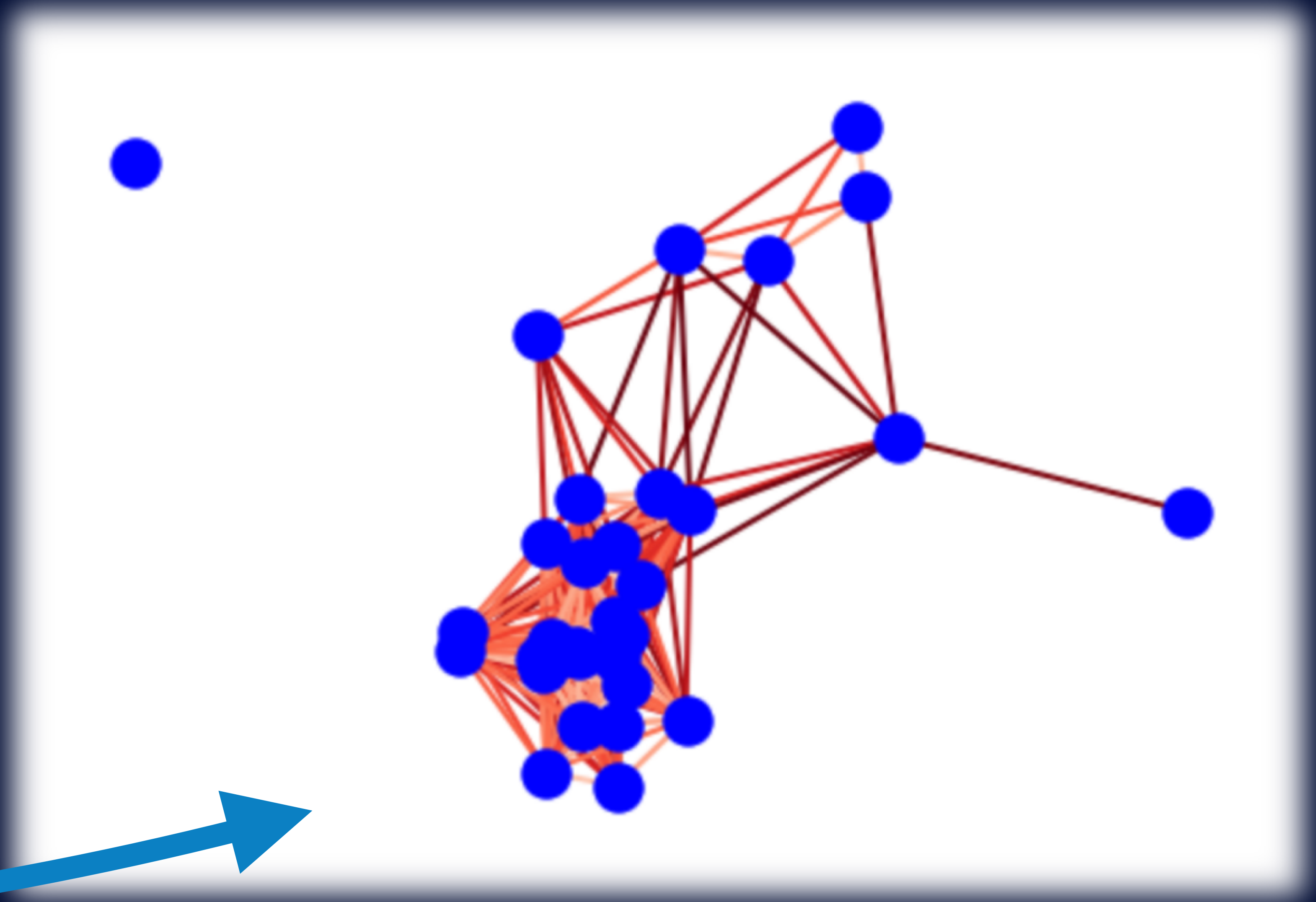


In the (η, ϕ) plane jets can be considered as graphs

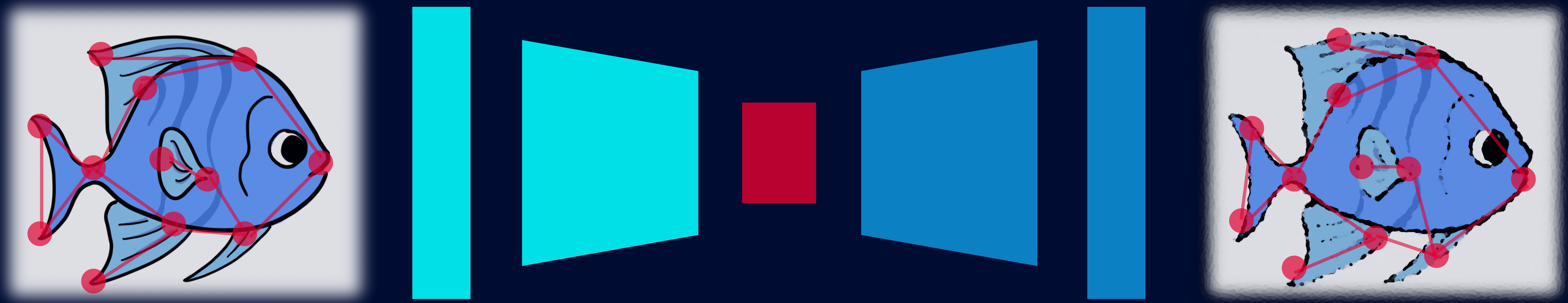


- nodes: (η, ϕ)
- node features: transverse momentum
- edge: $\Delta R_{ij} < \Delta R_{cut}$

$$\Delta R_{ij} = \sqrt{(\phi_i - \phi_j)^2 + (\eta_i - \eta_j)^2}$$



Semi-Supervised approach



dist(,) ≈ 0

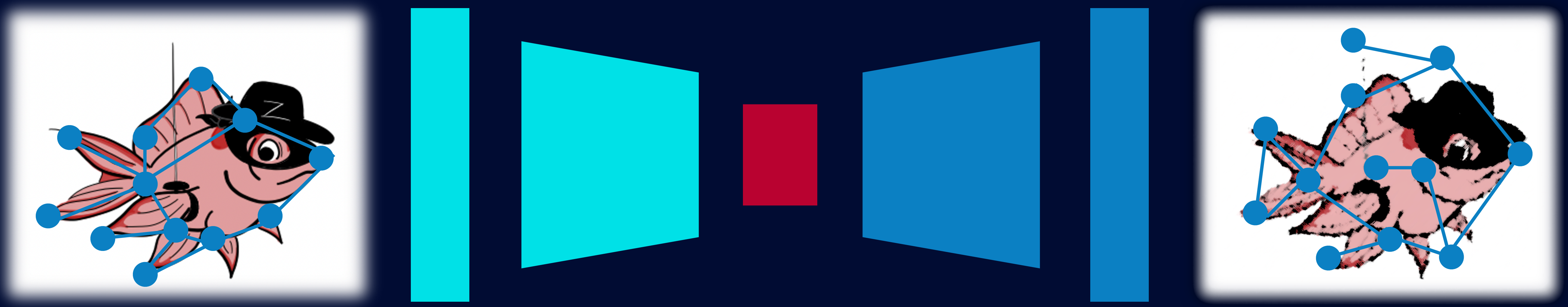
Training only on Standard Model jets

The algorithm learns to minimise the distance between input and reconstructed standard graphs

Different distance distributions between Anomalous and Standard!

Test with also anomalous jets

With anomalous jets distance between input and reconstructed graphs won't be minimal



dist(,) $\neq 0$

Take-home messages

- We experimentalist are not giving up searching for New Physics

- **So many events and data that ML algorithms definitely helps**

- **Jets** may be the **key to New Physics** and can be reinterpreted as graphs for **Graph Anomaly Detection**