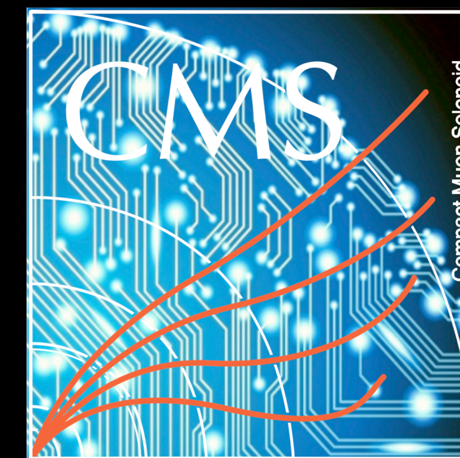
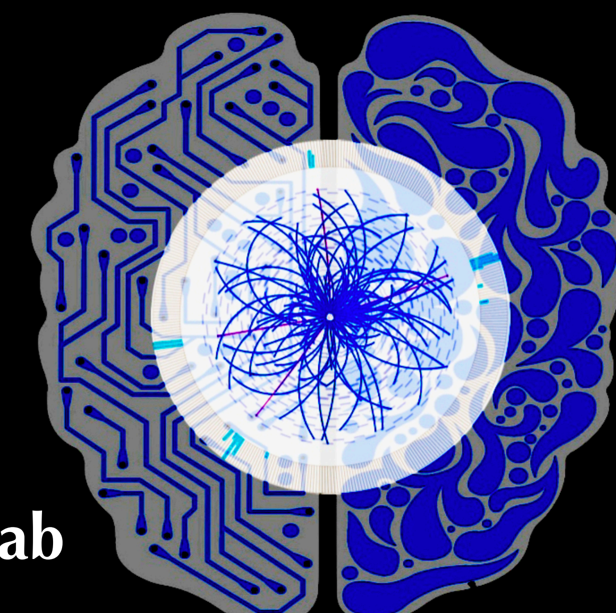
Transforming the LHC Physics Program with AI

Jennifer Ngadiuba (Fermilab)

42nd International Conference on High Energy Physics
July 18-24, 2024
Prague





FastML lab

Fermilab

ICHEP 2024
PRAGUE

ichep2024.org

42nd International Conference on High Energy Physics

18-24 July · 2024 · Prague · Czech Republic

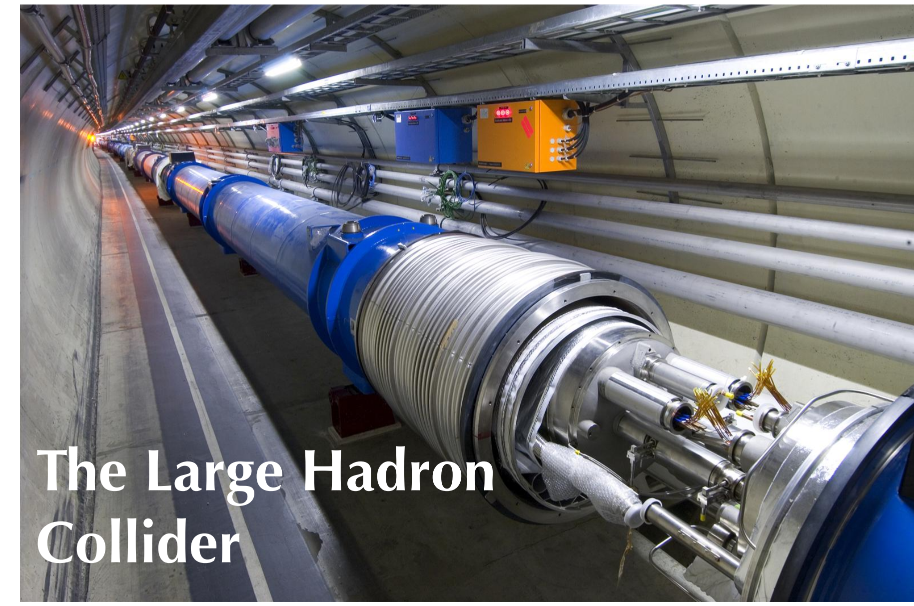
IUPAP International Union of Pure and Applied Physics

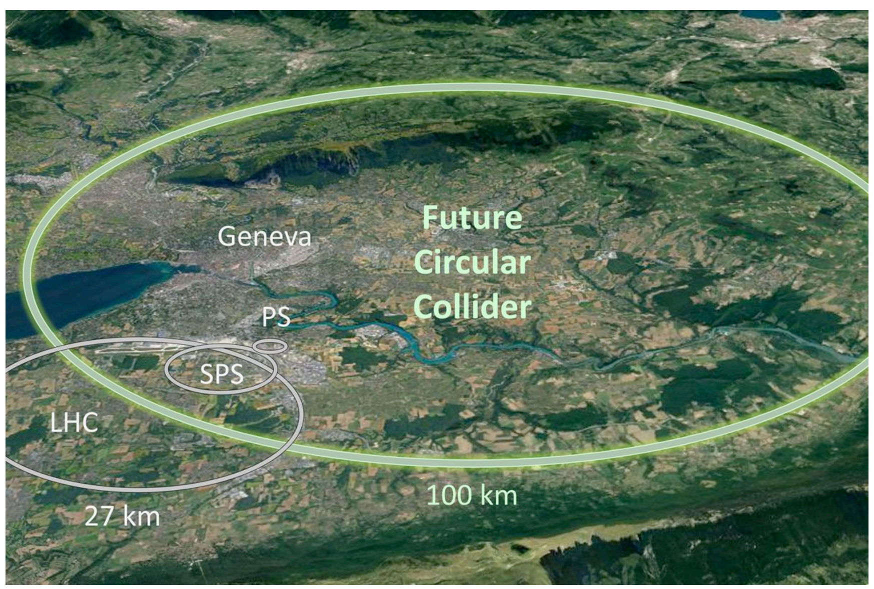
Machine learning for particle physics ⟷ **Particle physics for machine learning**

Thank you!

Big Science in 21st century

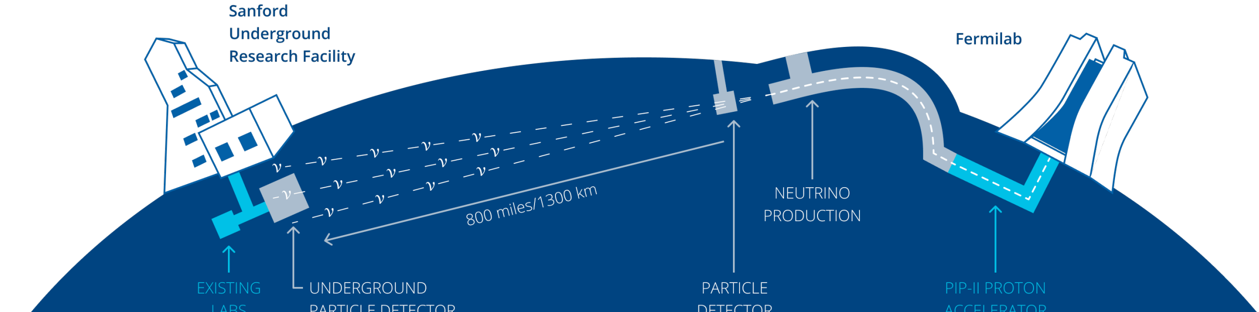
Probing the **fundamental structure of nature** requires complex experimental devices, large infrastructures and big collaborations.





International MUON Collider Collaboration

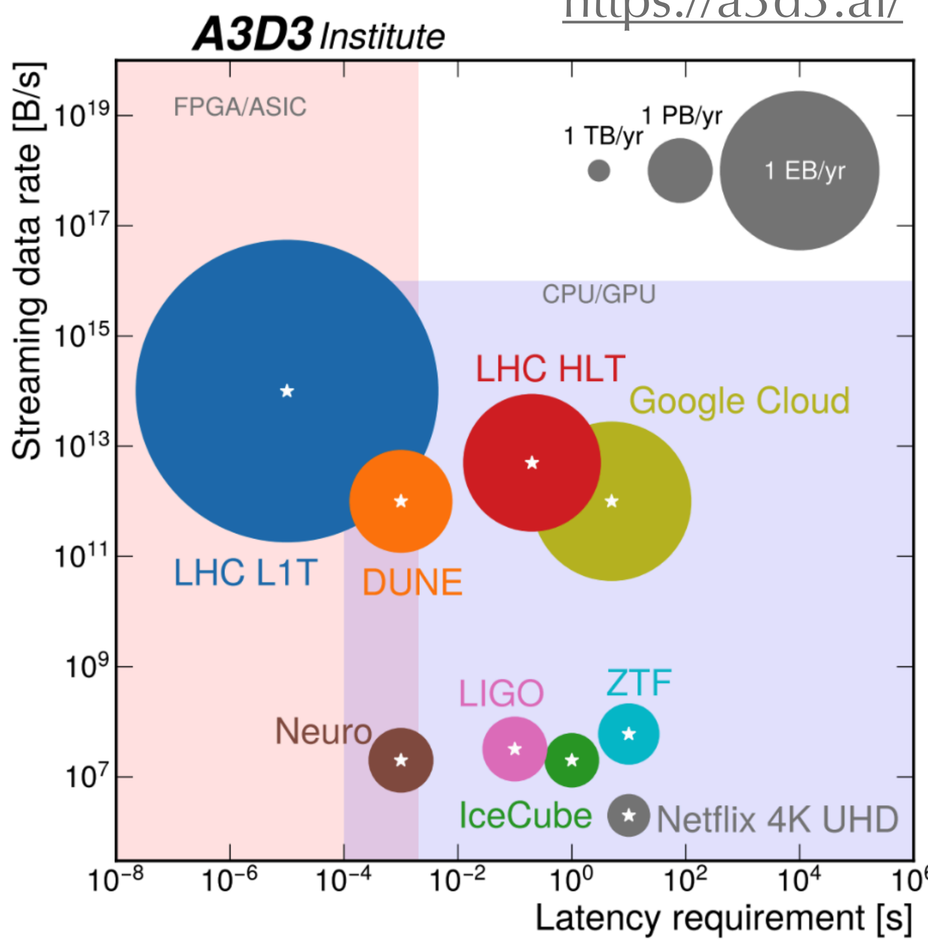
The DUNE neutrino experiment

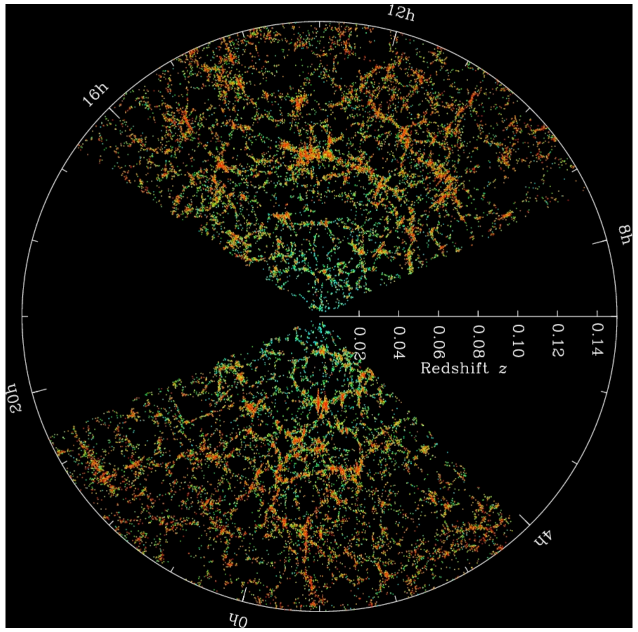


Big Science = Big Data

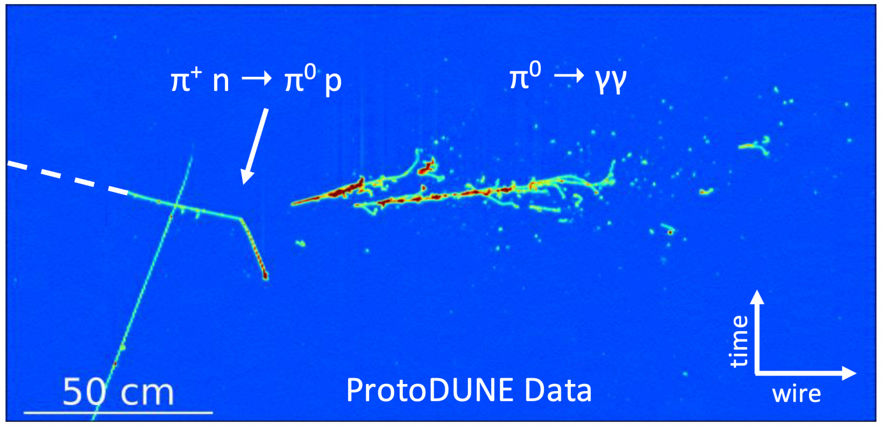
- Increasingly complex data both in **volume and dimensionality**
- Increasing need for **efficient and accurate data processing pipelines**
- Challenge in **simulating expectations** for what experiments may observe
- But also need for innovative **data & discovery driven** physics analyses approaches

https://a3d3.ai/

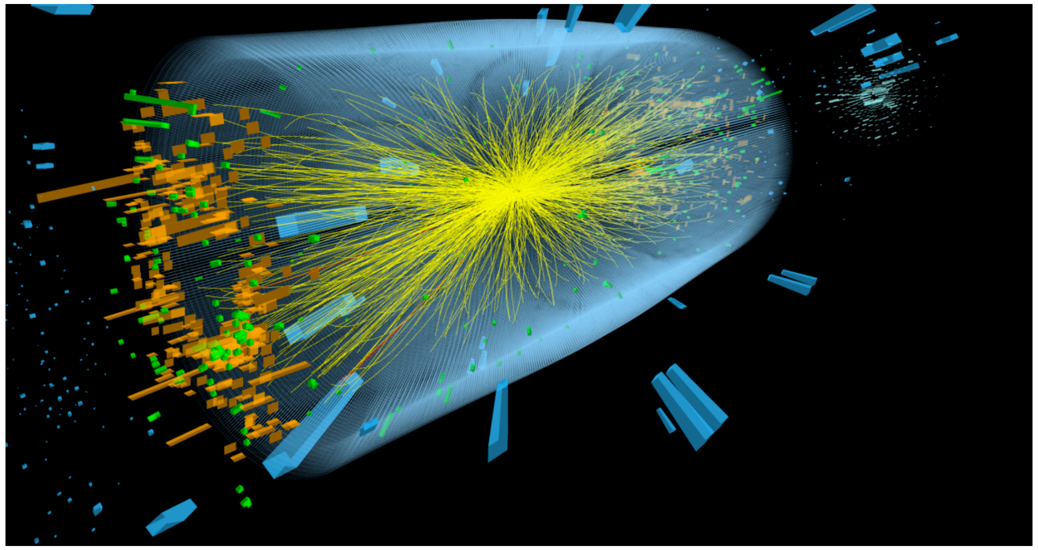




Sloan Digital Sky Survey



Interactions in LArTPC

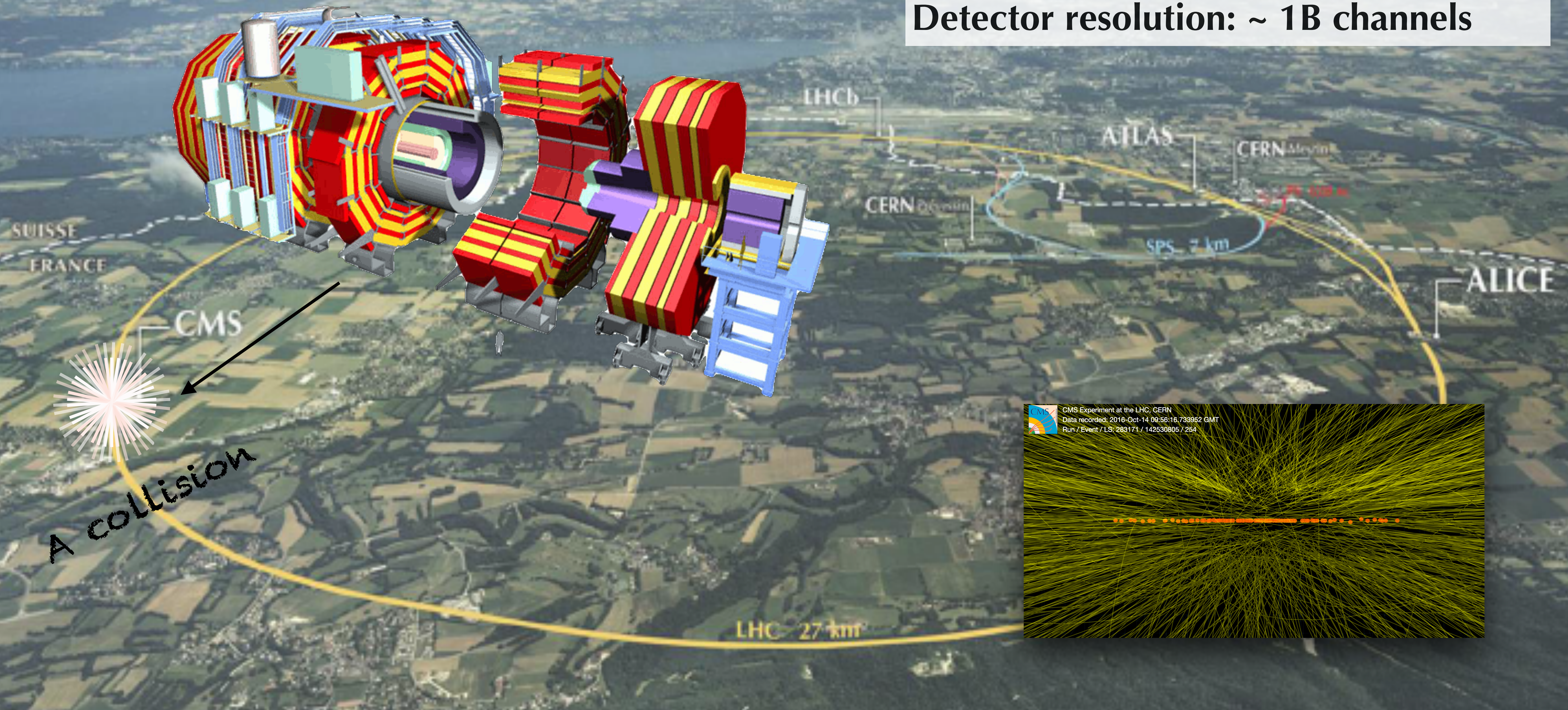


A LHC collision

Big Data @ the Energy Frontier

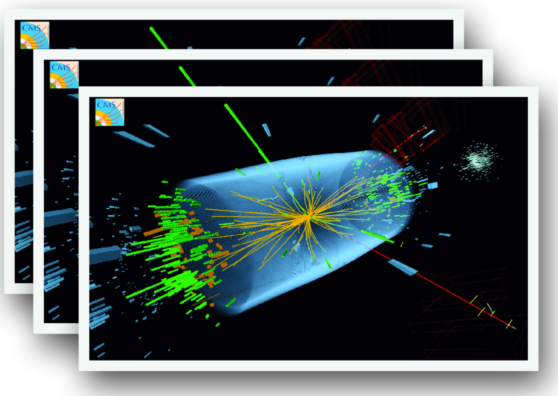
The Large Hadron Collider (LHC)

Collision frequency: 40 MHz
Particles per collision: O(10^3)
Detector resolution: ~ 1B channels



Extreme data rates of ~Pb/s!

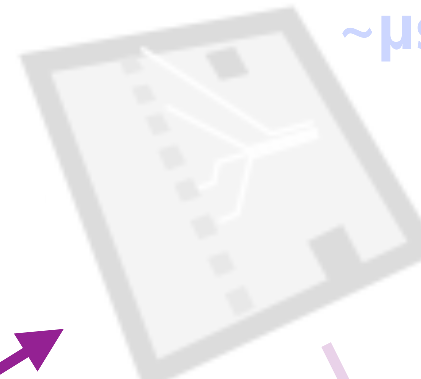
Data reduction workflow @ LHC



CMS Experiment
40 MHz collision rate
~1B detector channels

Worldwide computing grid
Exabyte-scale datasets

FPGA filter stack
~μs latency

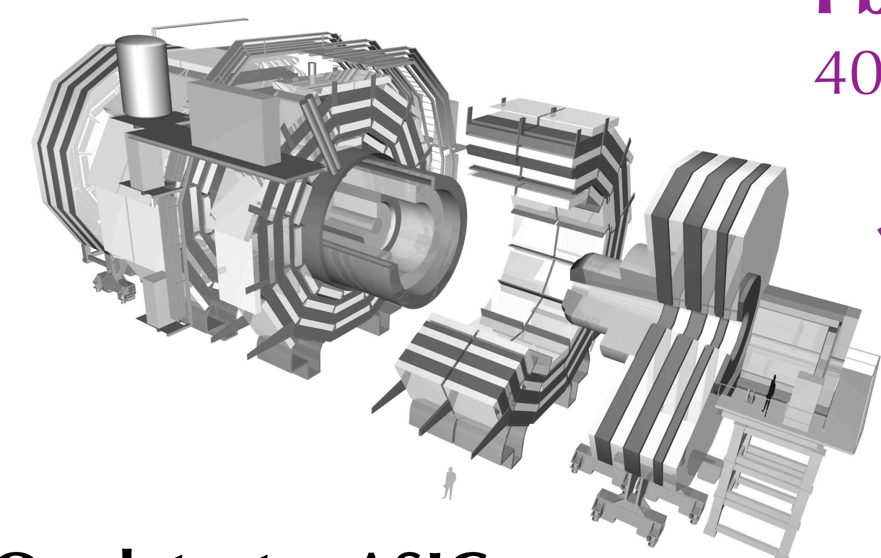


Level-1 Trigger

Pb/s
40 MHz

10s Gb/s
~5 kHz

Offline analysis



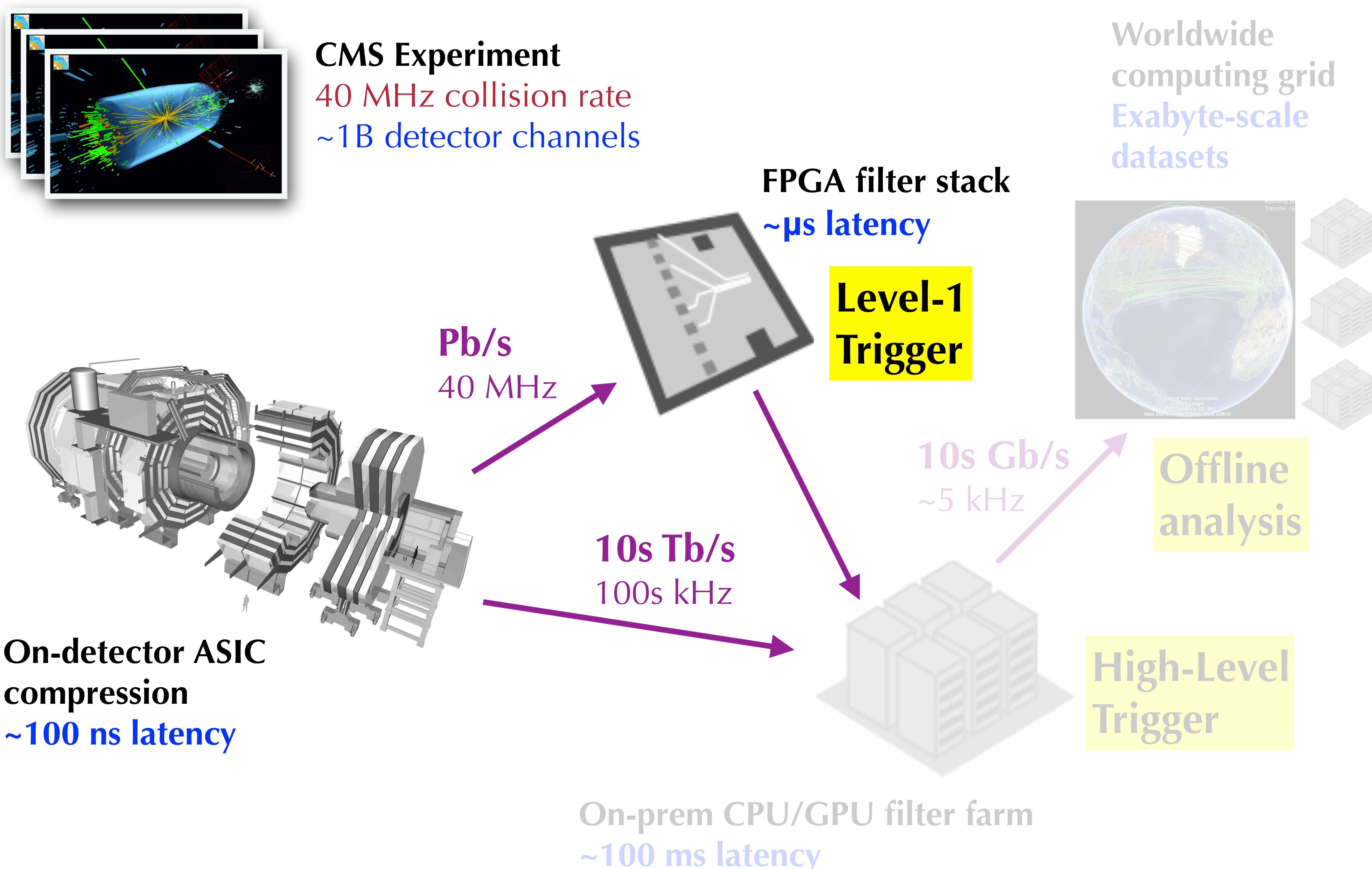
10s Tb/s
100s kHz

High-Level Trigger

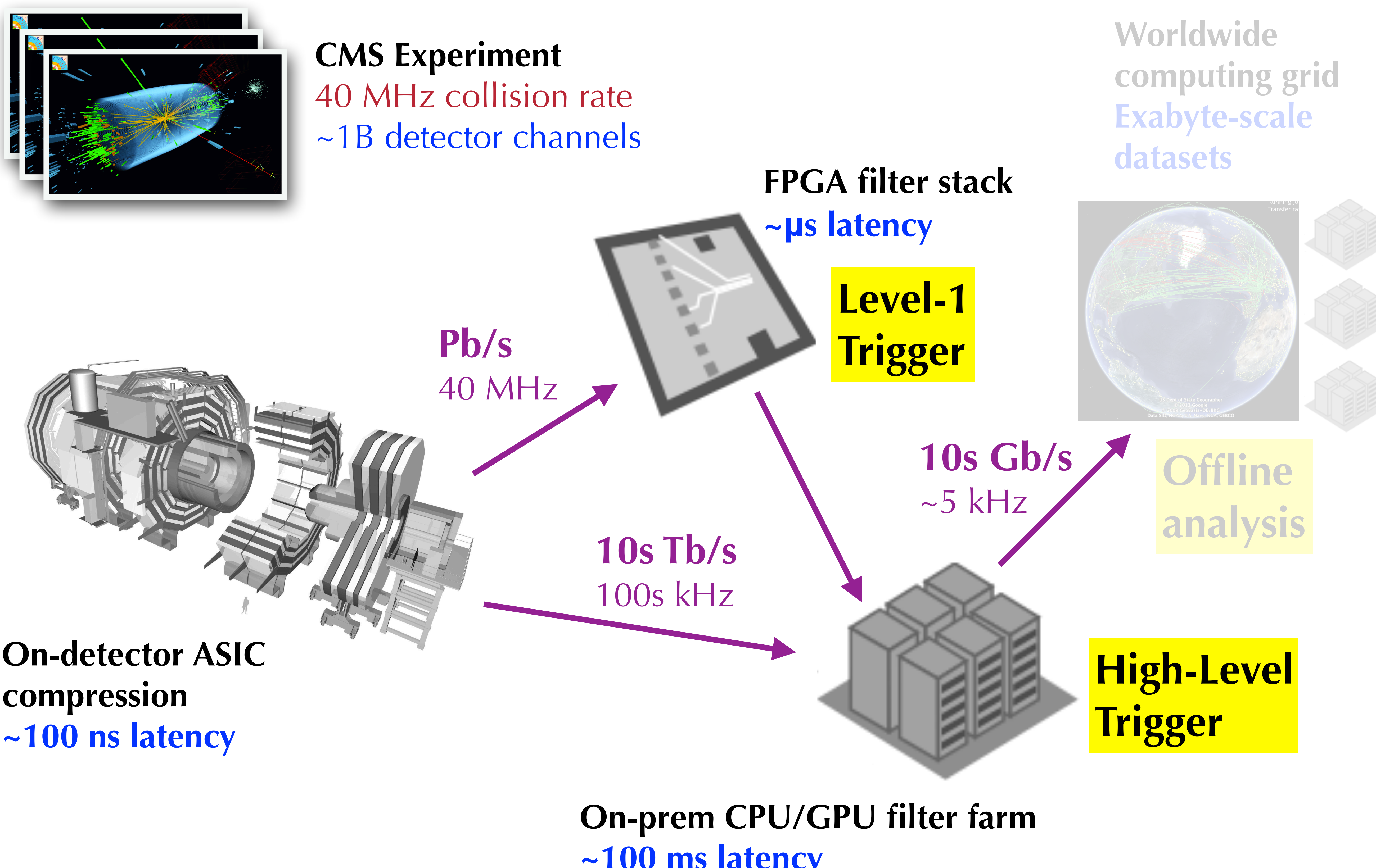
On-detector ASIC compression
~100 ns latency

On-prem CPU/GPU filter farm
~100 ms latency

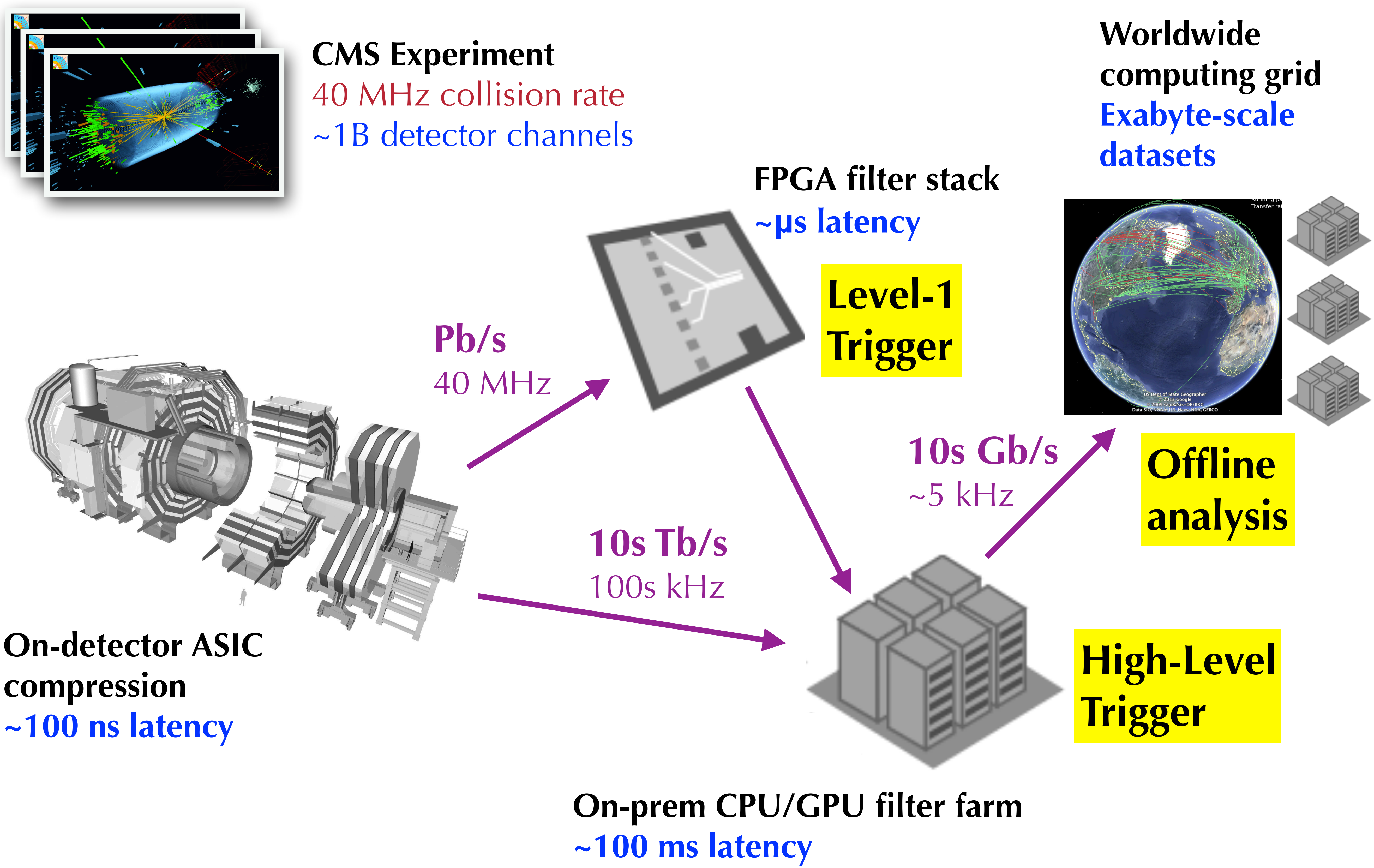
Data reduction workflow @ LHC



Data reduction workflow @ LHC

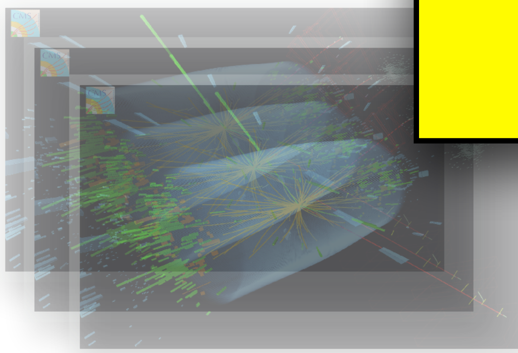


Data reduction workflow @ LHC



Make physics discoveries with 0,0025% of the events!

(the rest is lost…)



40 MHz collision rate

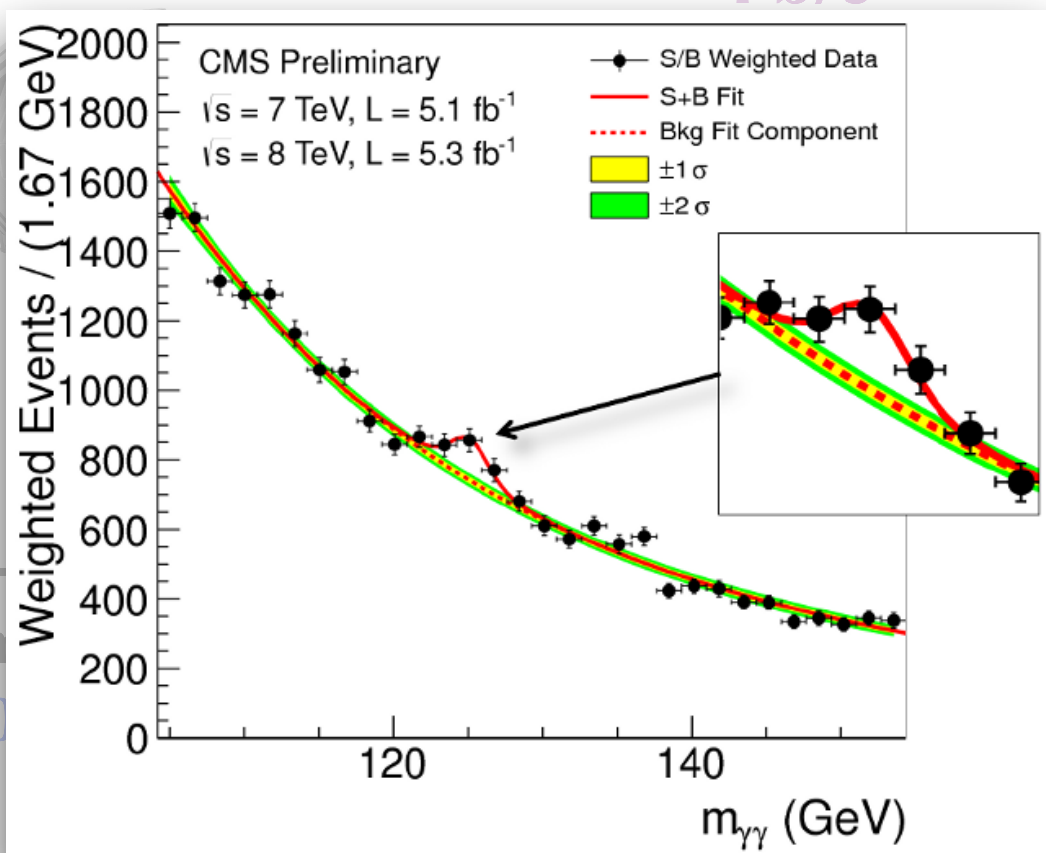
~1B detector channels

FPGA filter stack

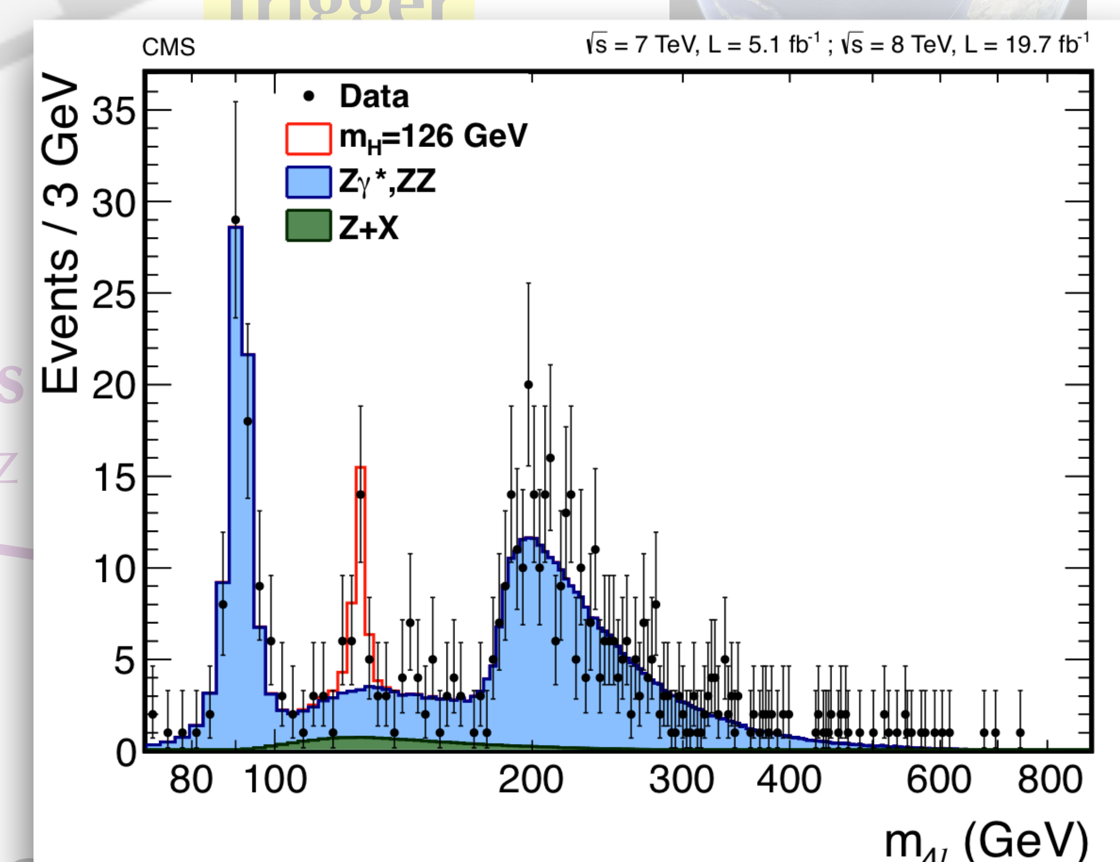
~μs latency

Exabyte-scale datasets

Higgs boson decay to two photons

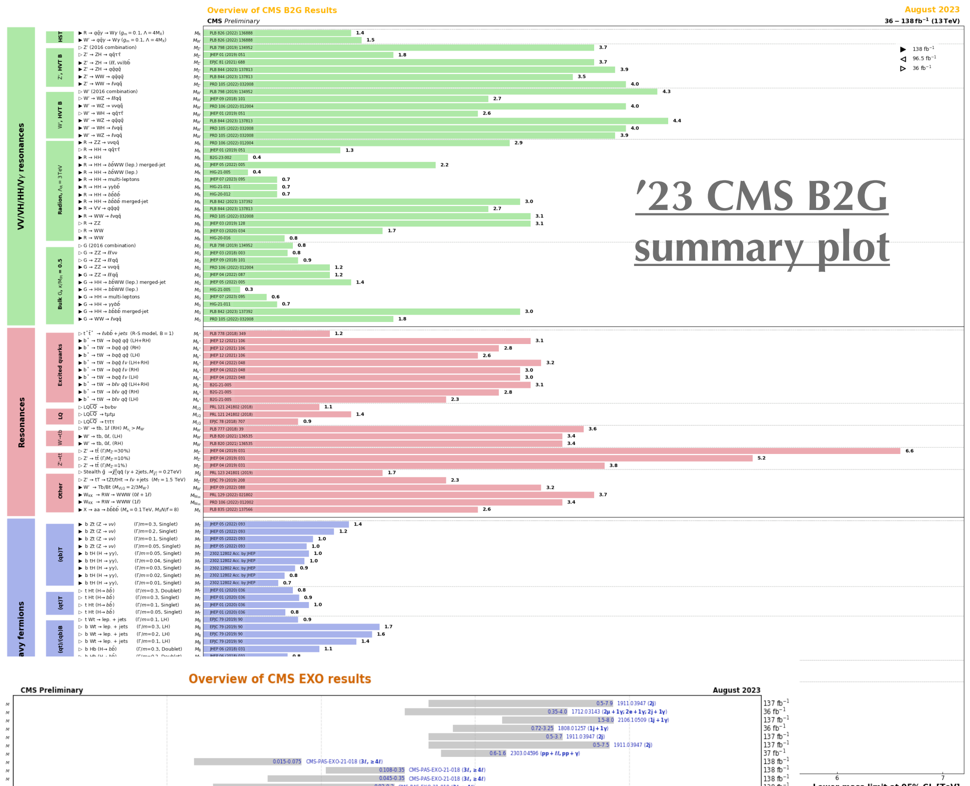


Higgs boson decay to four leptons

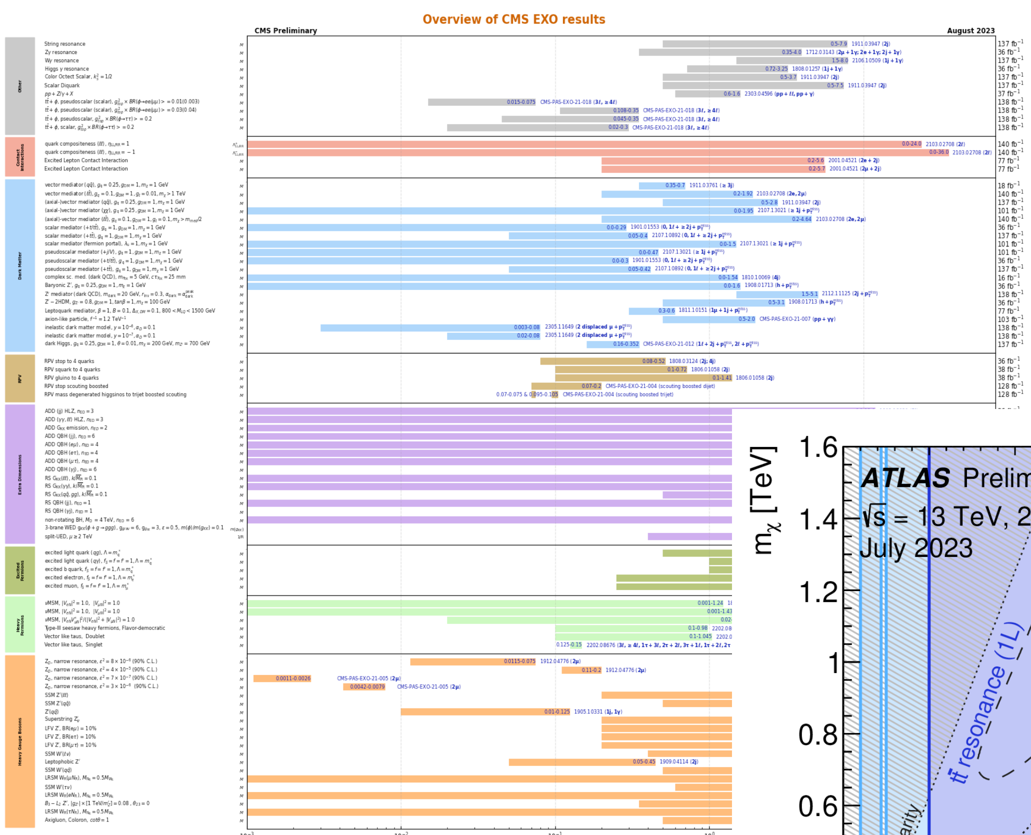


~100 ms latency

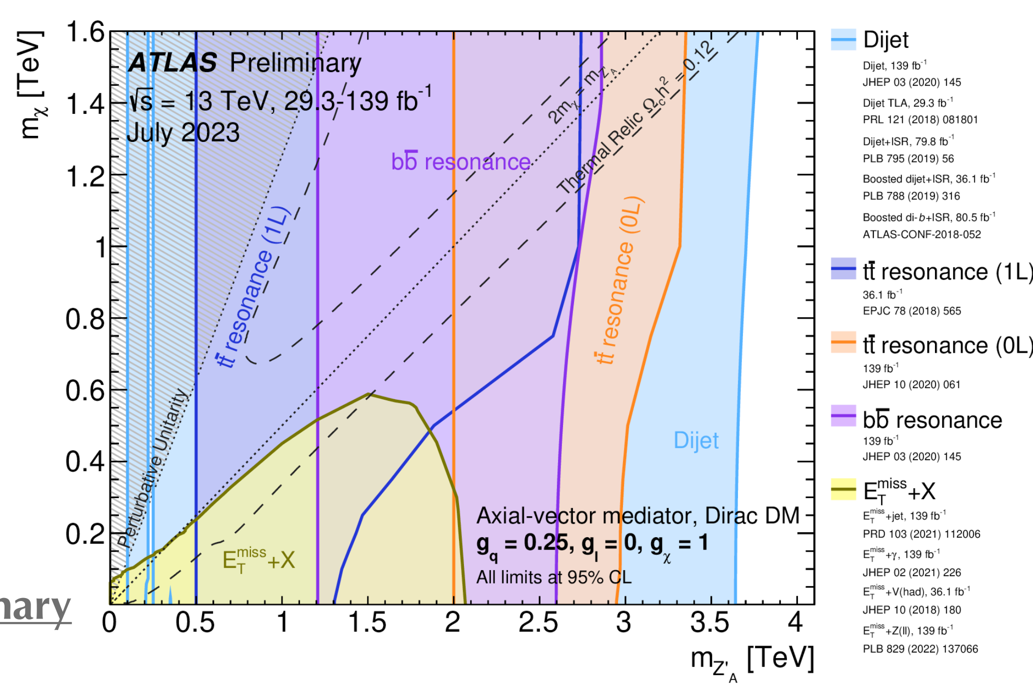
Searching for something



'23 CMS B2G summary plot



'23 CMS EXO summary plot

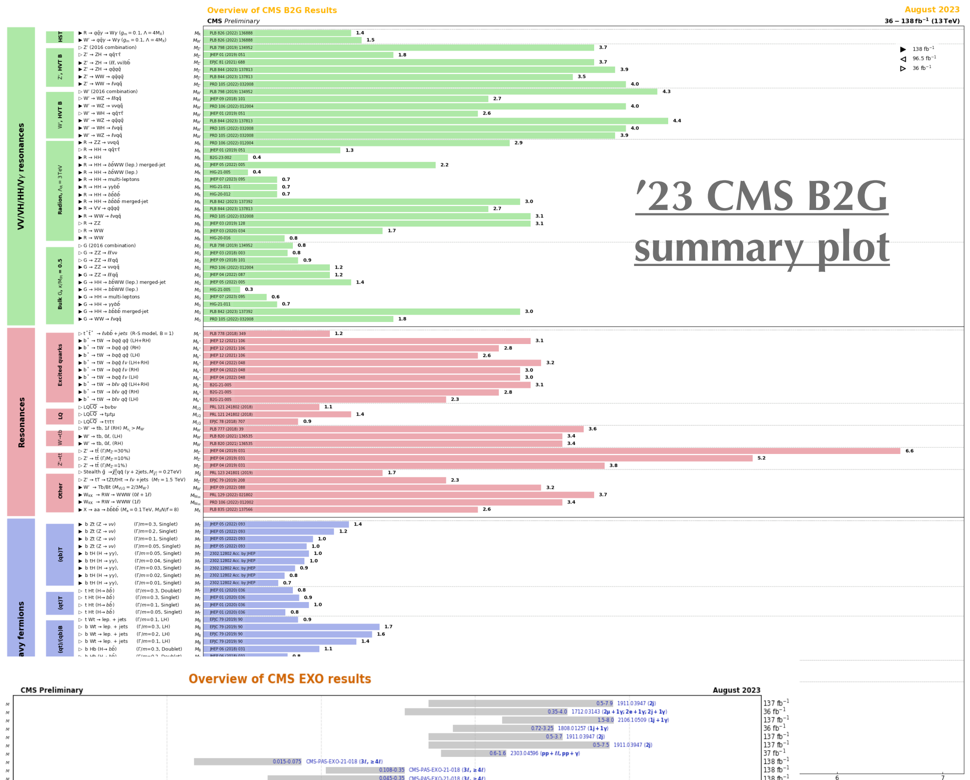


'23 ATLAS Dark Matter summary

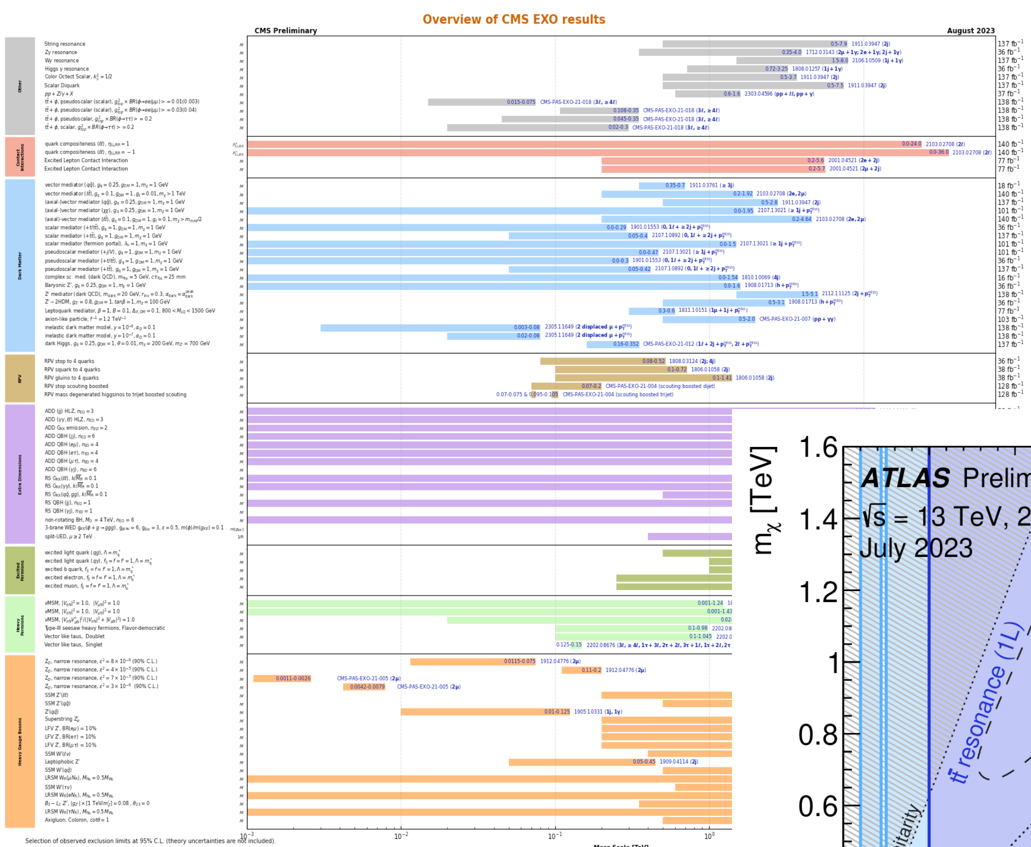
This works well when we know what to search for (eg, the Higgs boson and motivated BSM theories)
➔ strong physics assumptions guide each step of the data analysis workflow:

1. data reduction
2. data preprocessing
3. final high-level features statistical interpretation

Searching for anything

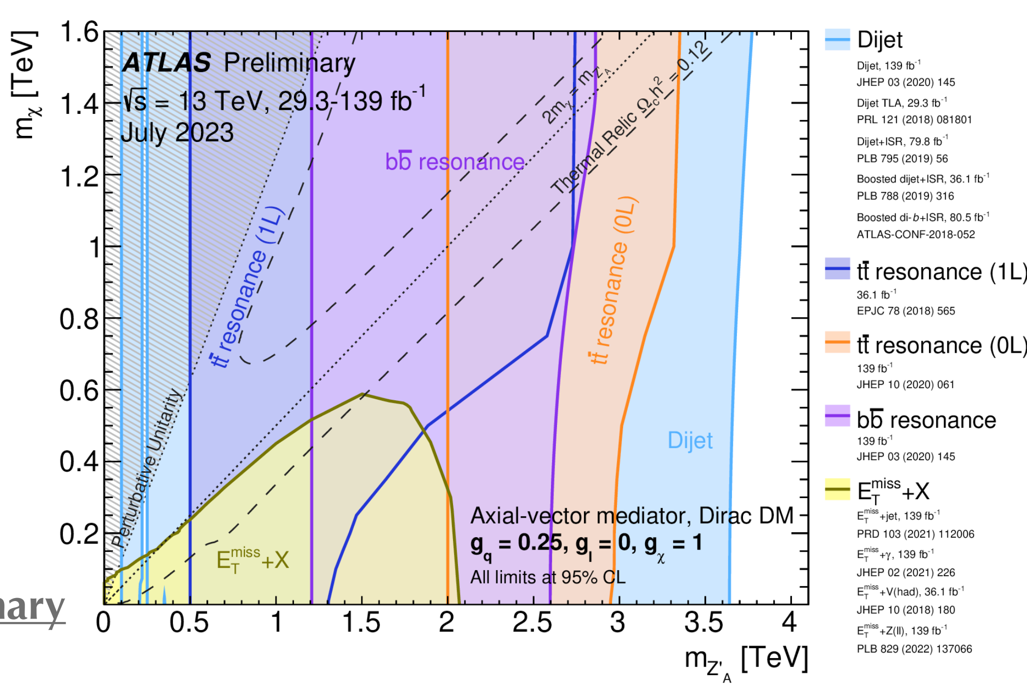


But nature might be different from all of these



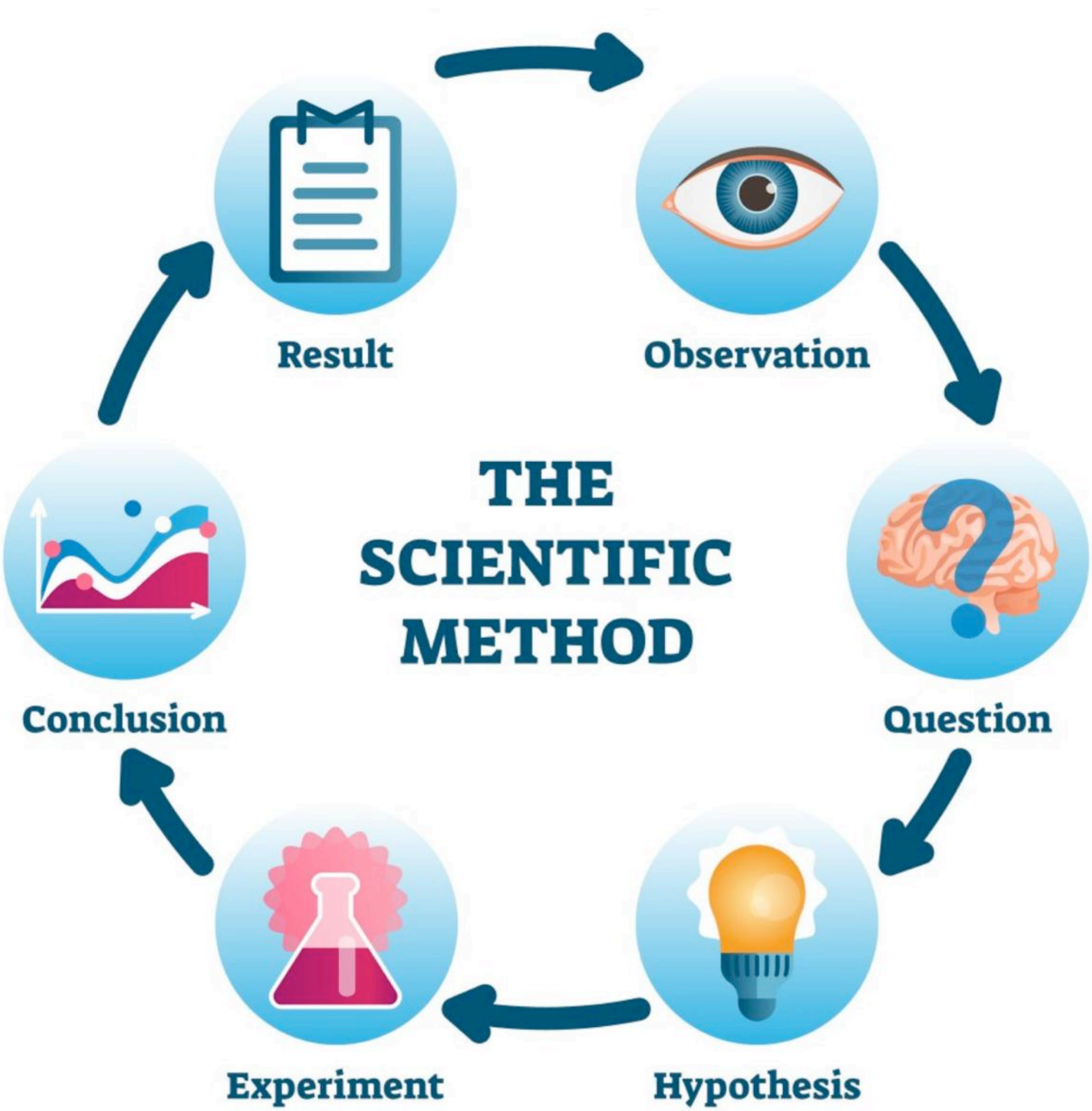
'23 CMS EXO summary plot

It might be that we are not looking in the right corner because we might have not imagined (yet) how new physics look like

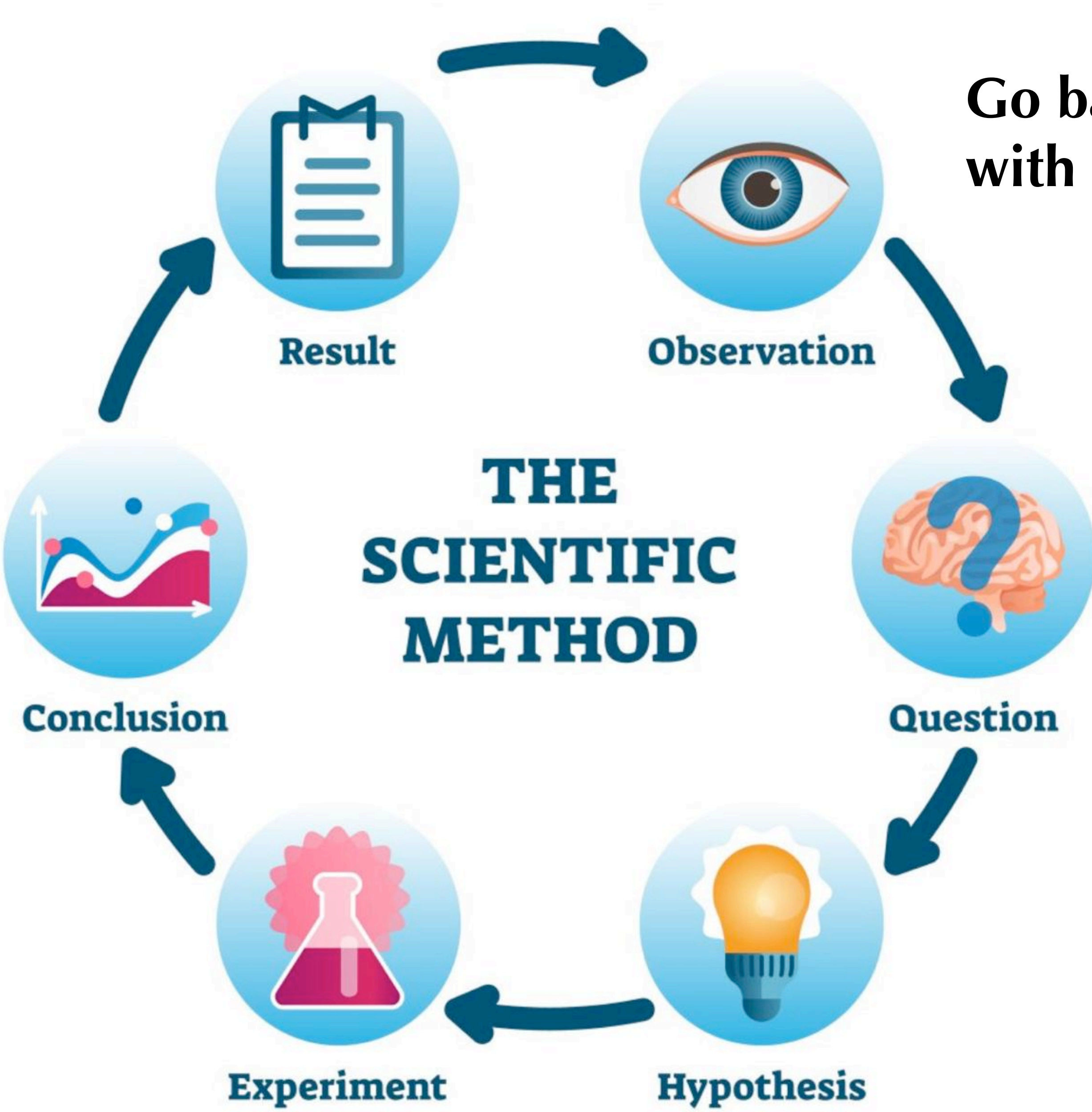


'23 ATLAS Dark Matter summary

Look at all corners



Look at all corners



Go back looking at data with no signal hypothesis

Look at all corners

Go back looking at data with no signal hypothesis

Remain surprise!

Look at all corners

Go back looking at data with no signal hypothesis

Remain surprise!

If we identify an anomalous signal we can formulate an alternative hypothesis

Look at all corners

Go back looking at data with no signal hypothesis

Remain surprise!

If we identify an anomalous signal we can formulate an alternative hypothesis

To be tested with new data and traditional techniques

Look at all corners

Go back looking at data with no signal hypothesis

Remain surprise!

If we identify an anomalous signal we can formulate an alternative hypothesis

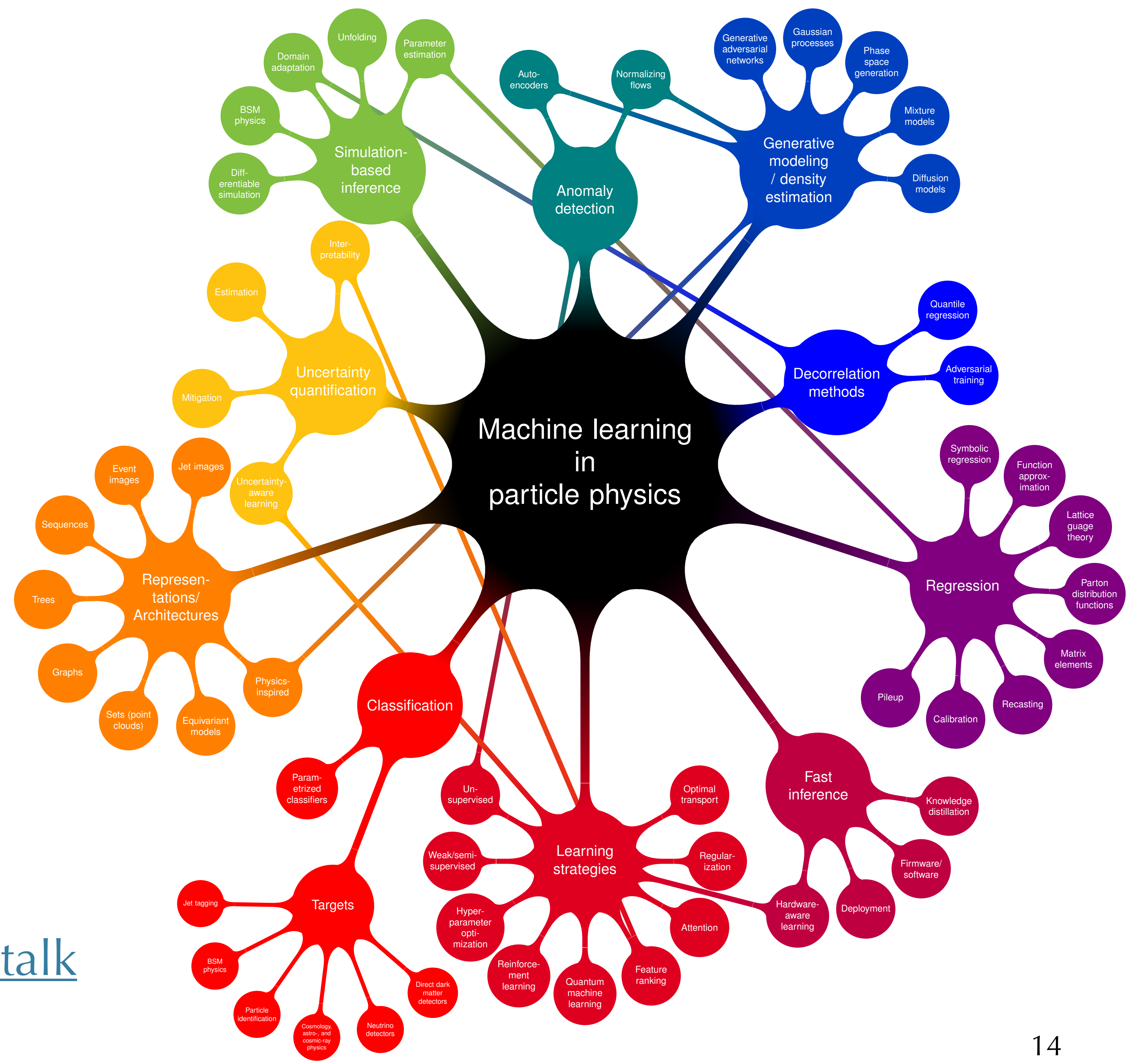
To be tested with new data and traditional techniques

This approach is what we call:
ANOMALY DETECTION

The role of AI

- **Today we can implement anomaly detection efficiently with AI**
- **Machine Learning is used in particle physics since the ‘80s**
  - it was shallow networks back then
- **Over the last decade a rapid progress guided by technological breakthrough led to a revolution in this area**
  - this the era of Deep Learning

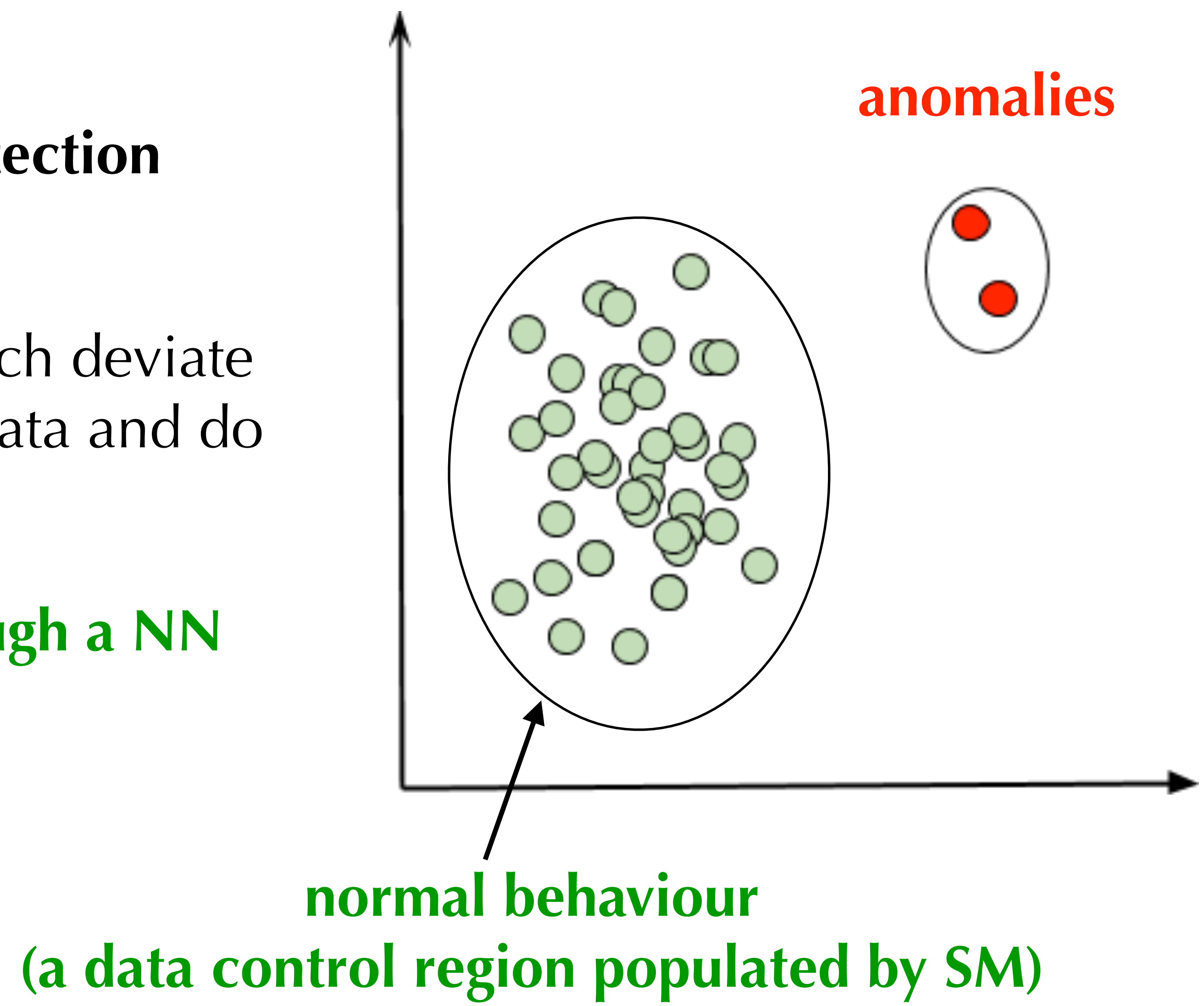
https://iml-wg.github.io/HEPML-LivingReview/



See J. Duarte talk

A data-driven search strategy with AI

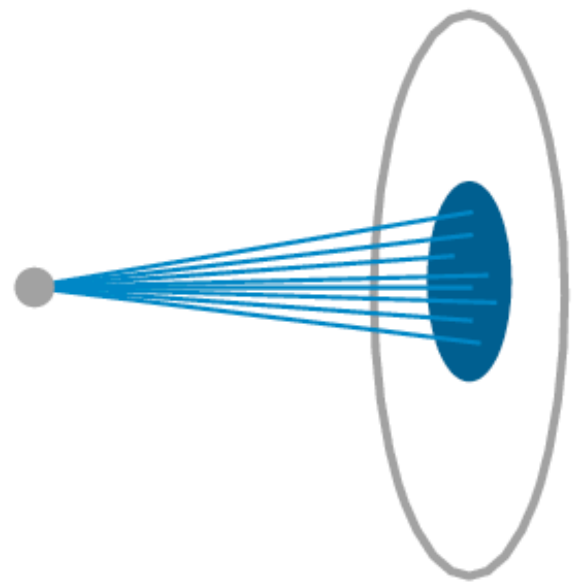
- **Today we can implement anomaly detection efficiently with AI**
- Identifying rare events in data sets which deviate significantly from the majority of the data and do not conform to “normal” behaviour
- **Normal behaviour can be learnt through a NN**



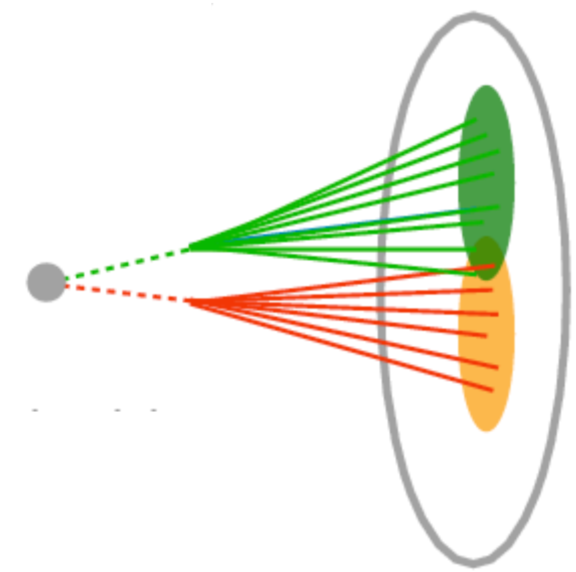
A data-driven search strategy with AI

Take the most studied physics case:
search for a dijet resonance

BACKGROUND JET
(single quark/gluon)

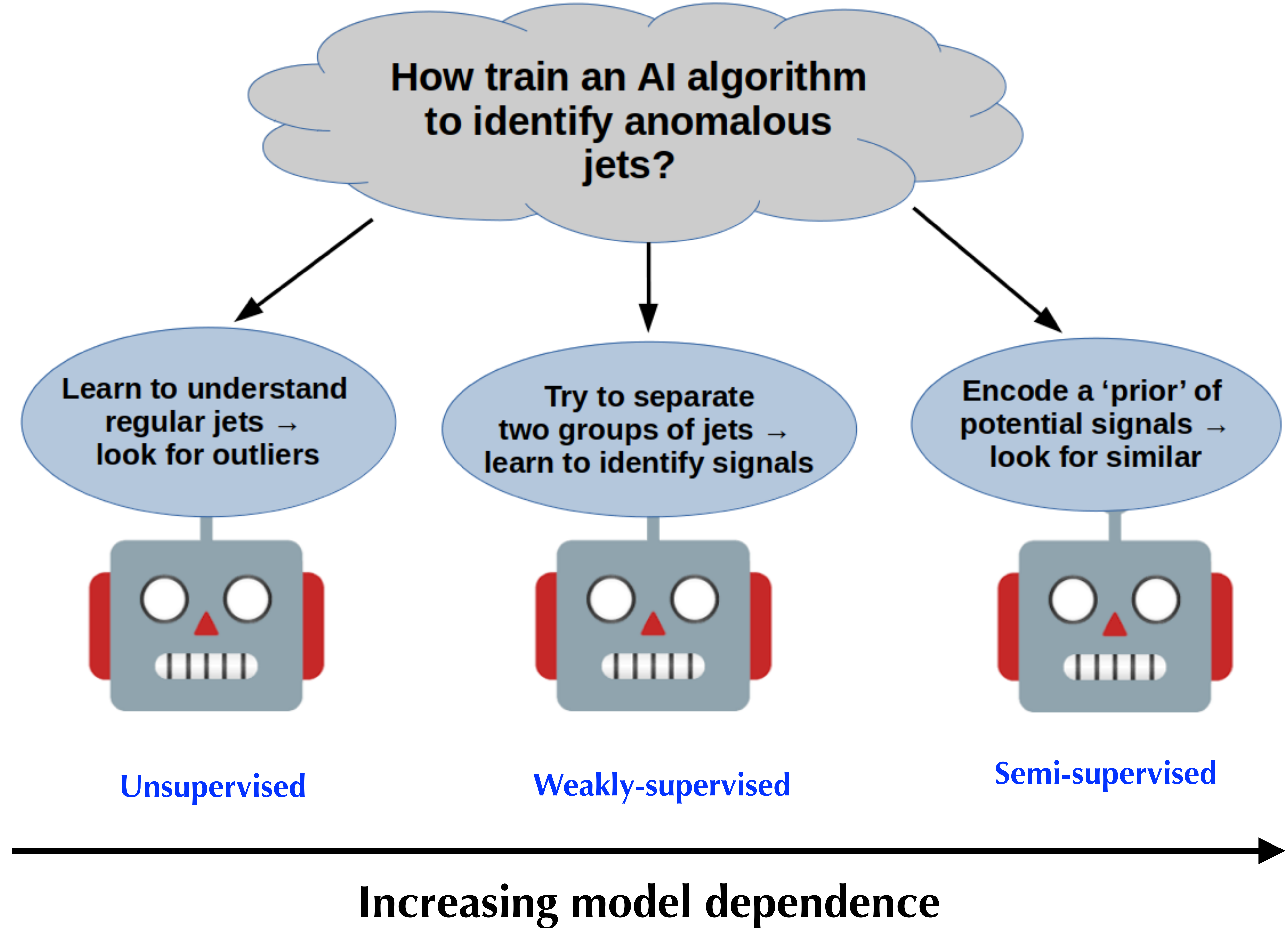


ANOMALOUS JET
(ex, resonance X
decaying to N quarks)



Train a NN to learn the SM QCD background in a control region

Then tag each jet as anomalous in the signal region with no assumption on how it looks like

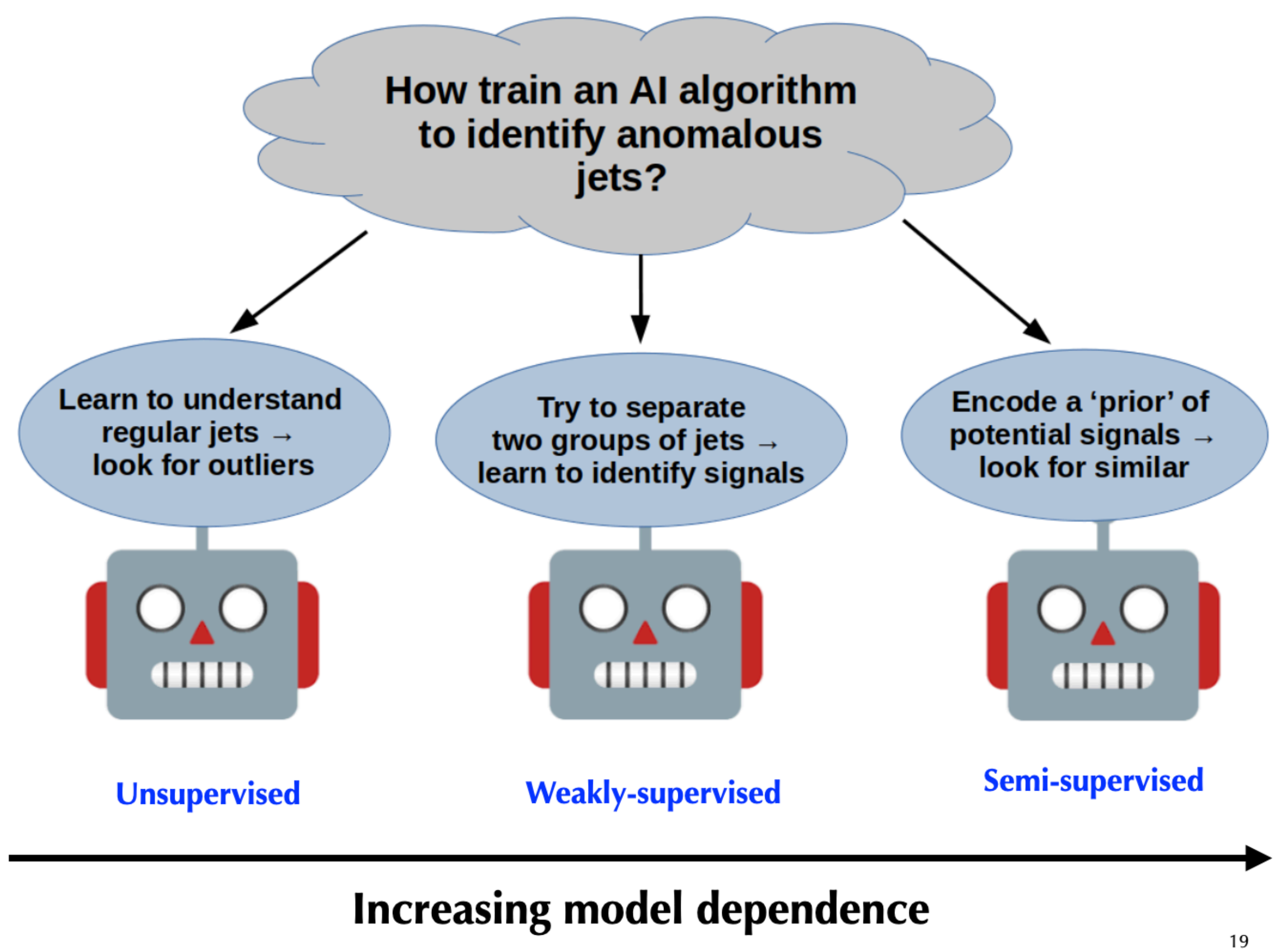

How train an AI algorithm to identify anomalous jets?
Learn to understand regular jets → look for outliers
Try to separate two groups of jets → learn to identify signals
Encode a 'prior' of potential signals → look for similar
Unsupervised
Weakly-supervised
Semi-supervised
Increasing model dependence

First CMS search looking for anomalous dijet events using novel ML techniques released this year!

Model-agnostic search for dijet resonances with anomalous jet substructure in proton-proton collisions at $\sqrt{s} = 13$ TeV

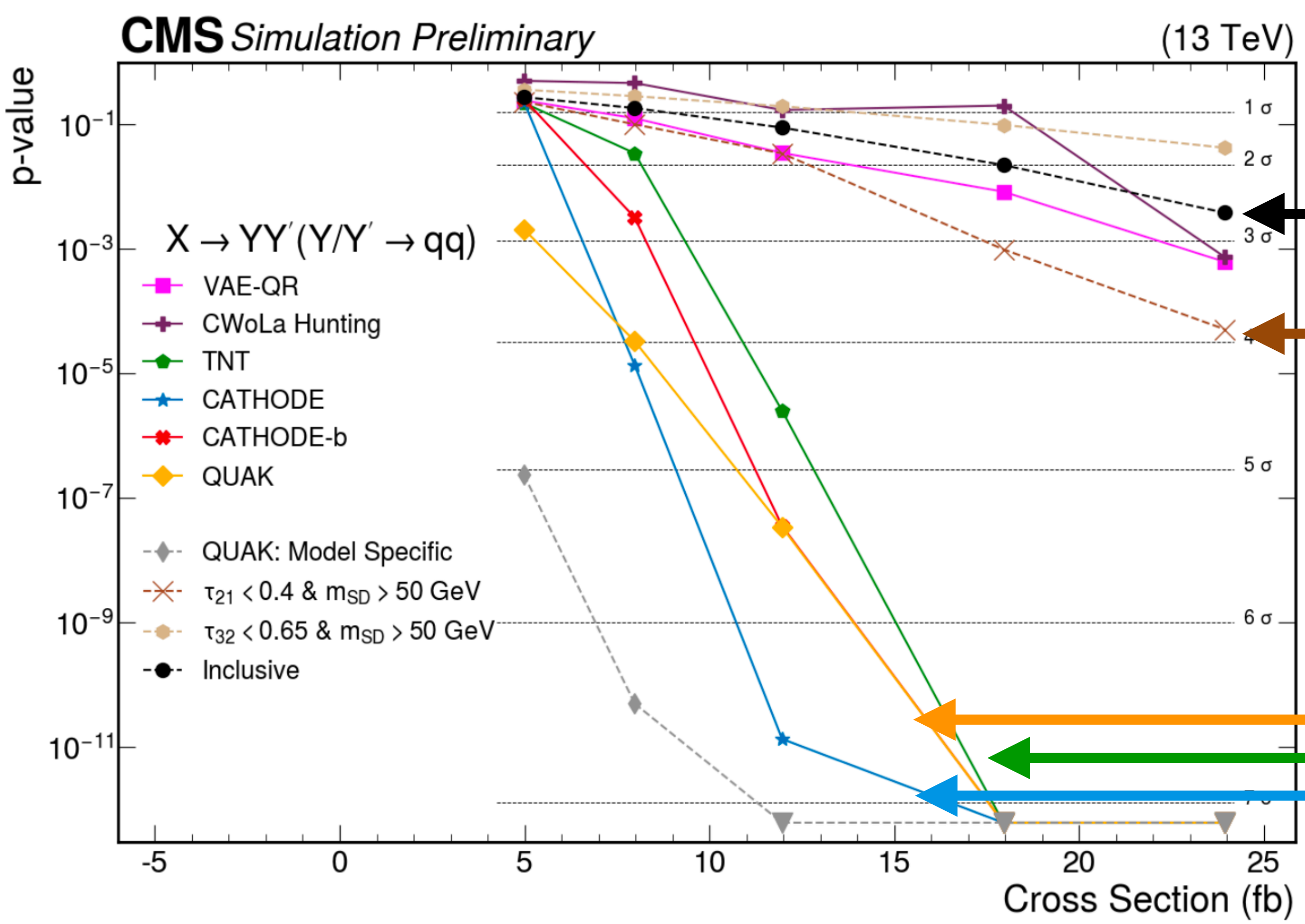
The CMS Collaboration

[CMS-PAS-EXO-22-026]



Anomaly detection in action @ CMS

- Inject signal of varying cross section in background MC and calculate p-value
- Obtain comparison of sensitivity of different methods against standard analysis methods



traditional dijet search

dedicated search for this signal

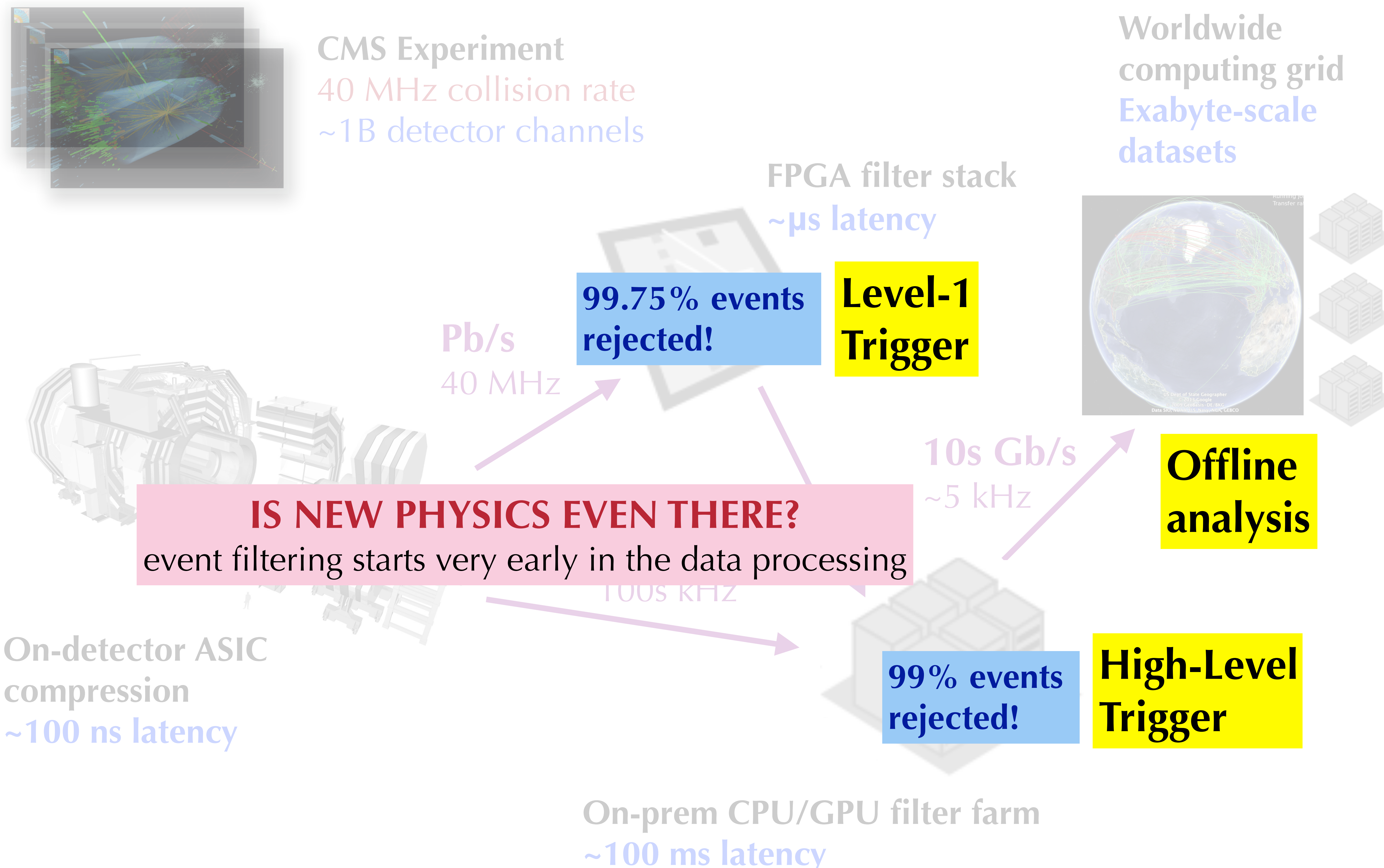
anomaly detection methods

Large discovery potential improvement wrt standard analysis methods for a broad range of signals!

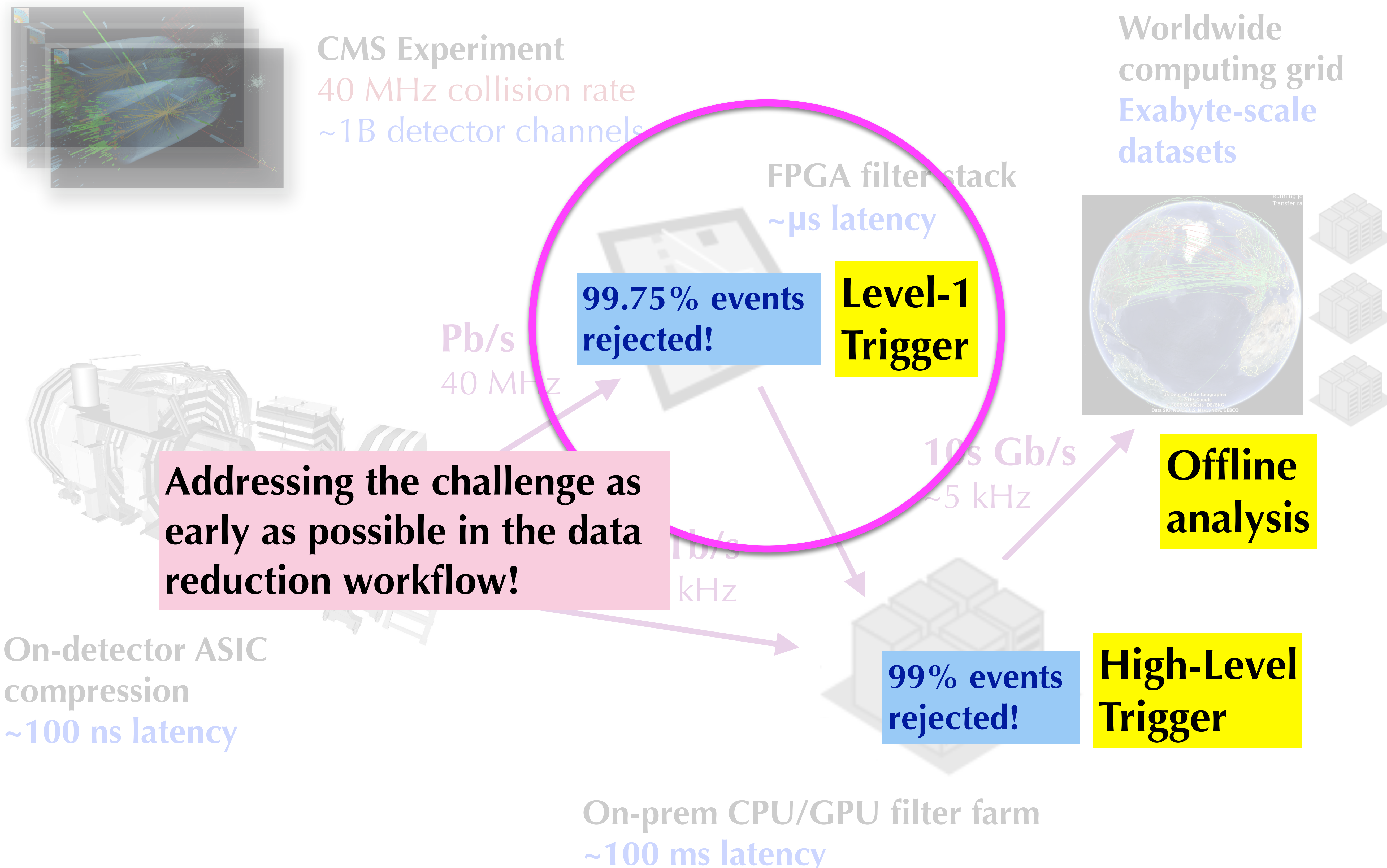
[CMS-PAS-EXO-22-026](#)

See M. Sommerhalder talk

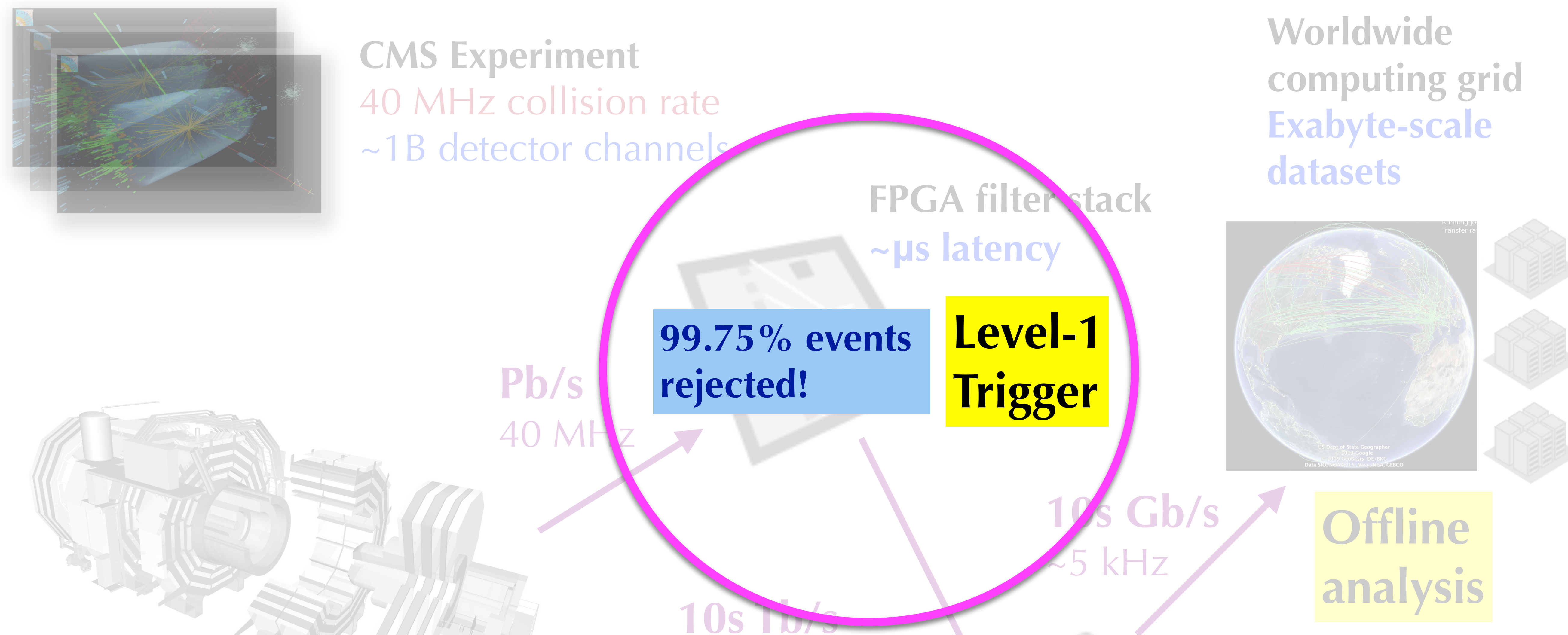
Data reduction workflow @ LHC



Data reduction workflow @ LHC



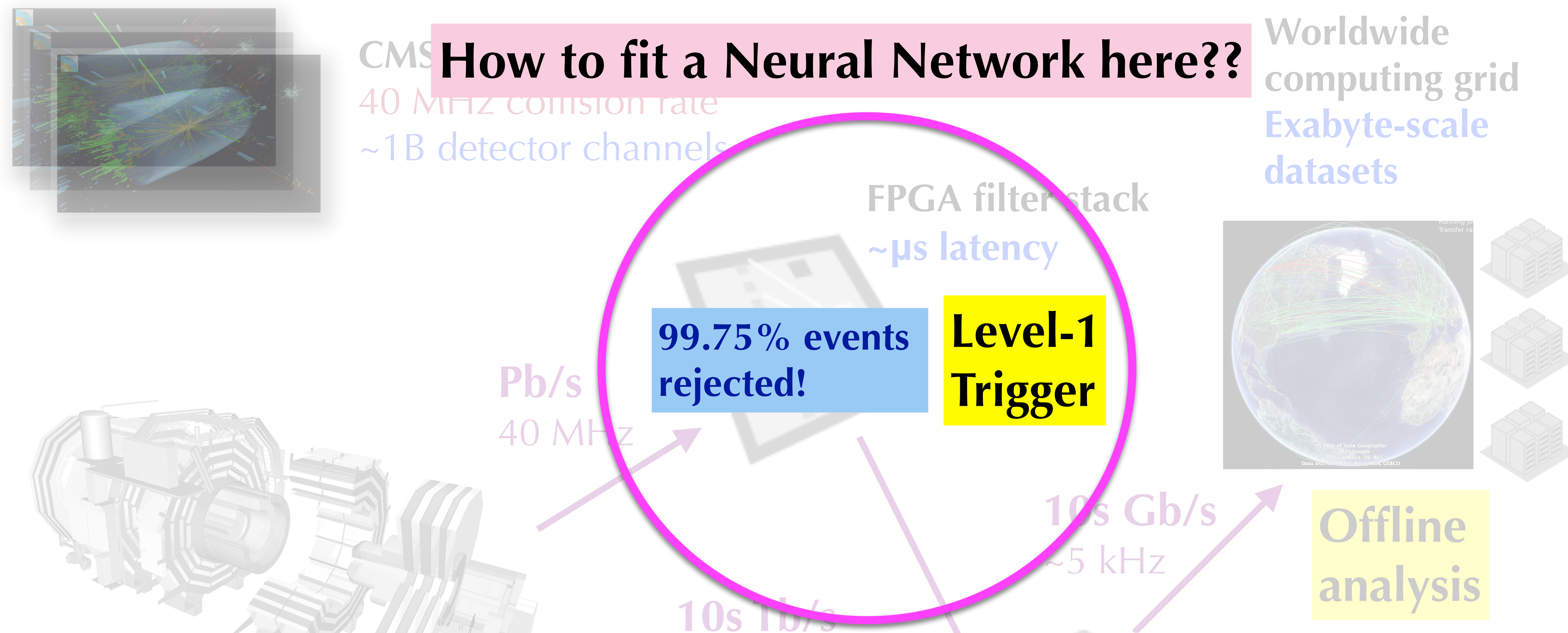
Data reduction workflow @ LHC



- Neural Networks can become relatively large → memory and number of operations required for the inference can easily explode
- **Strict constraints at L1 trigger:**
  - latency of O(μs) → use FPGA hardware
  - scarse resources (mostly occupied to calibrate sensors, build physics objects, etc..)

Data reduction workflow @ LHC

How to fit a Neural Network here??

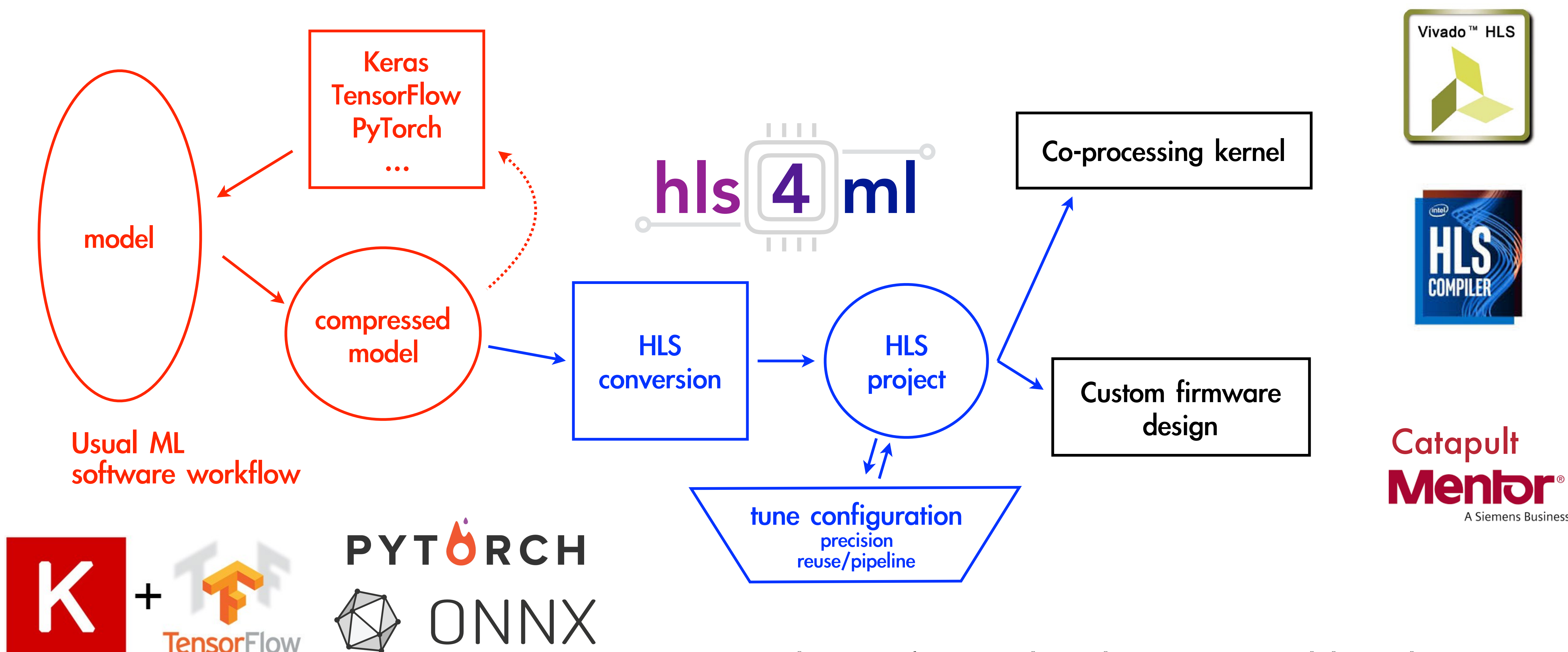


- Neural Networks can become relatively large → memory and number of operations required for the inference can easily explode
- **Strict constraints at L1 trigger:**
  - latency of O(μs) → use FPGA hardware
  - scarse resources (mostly occupied to calibrate sensors, build physics objects, etc..)

Bring ML models to hardware for real-time AI

high level synthesis for machine learning

A tool to efficiently program the FPGA hardware for Neural Networks with experimental constraints in mind!



https://fastmachinelearning.org/hls4ml/

Bring ML models to hardware for real-time AI

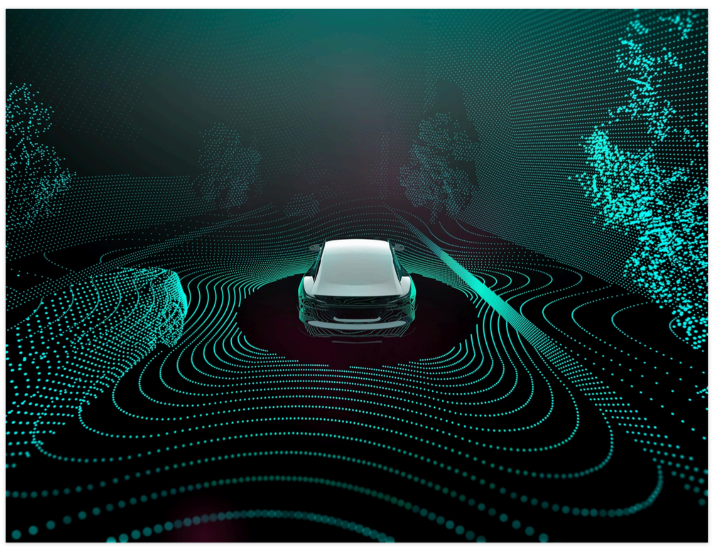
high level synthesis for machine learning

Sparking the interest of industry
(e.g., Google, Volvo, Siemens, AMD, …)

Colliding particles not cars: CERN's machine learning could help self-driving cars

CERN and software company Zenseact wrap up a joint research project that could allow autonomous-driving cars to make faster decisions, thus helping avoid accidents

25 JANUARY, 2023 | By Priyanka Dasgupta



CERN's expertise in machine learning could help the field of autonomous driving (Image: Zenseact)

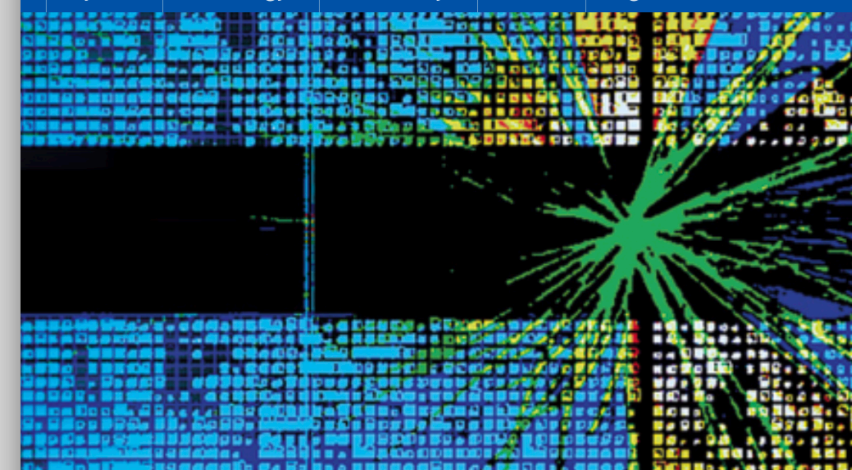
zenseact

VOLVO

CERN

CERNCOURIER | Reporting on international high-energy physics

Physics ▾ Technology ▾ Community ▾ In focus Magazine



COMPUTING | FEATURE

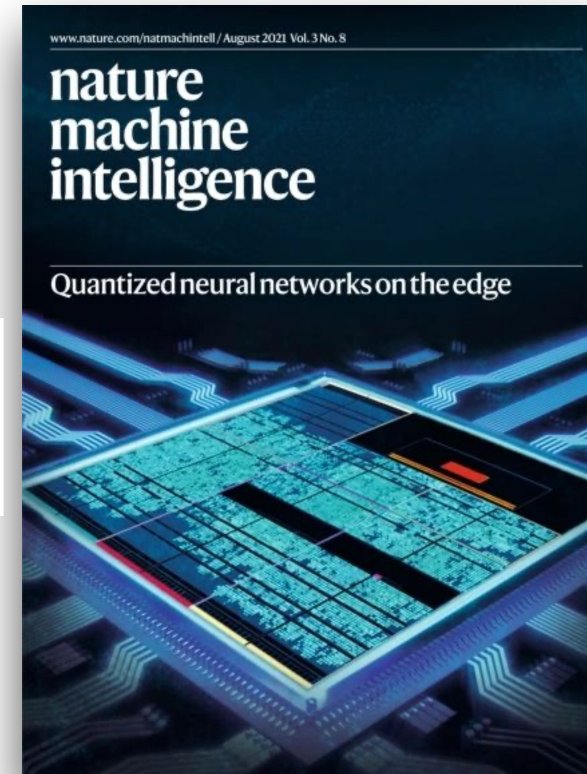
Hunting anomalies with an AI trigger

31 August 2021

Jennifer Ngadiuba and Maurizio Pierini describe how 'unsupervised' machine learning could keep watch for signs of new physics at the LHC that have not yet been dreamt up by physicists.

CALIFORNIA INSTITUTE OF TECHNOLOGY 1891

CERN



Siemens Digital Industries Software Newsroom Overview All news Blogs

PRESS RELEASE

Siemens simplifies development of AI accelerators for advanced system-on-chip designs with Catapult AI NN

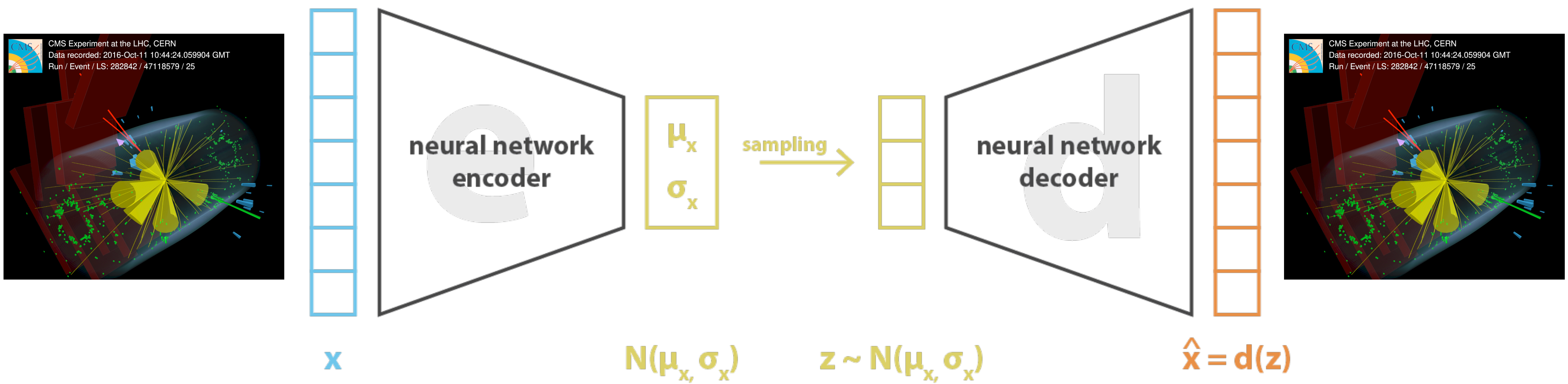
May 21, 2024
Plano, Texas

SIEMENS

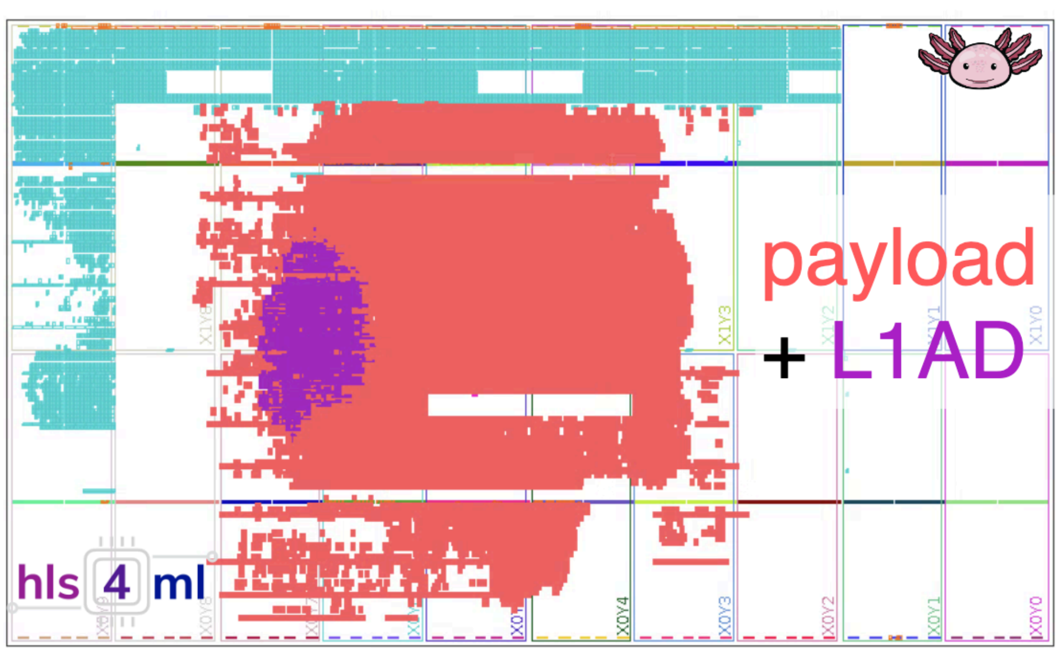
Fermilab

Ultra-fast anomaly detection @ CMS

CMS establishing a new trigger paradigm with sub-μs autoencoders for anomaly detection!



See A. Gandrakota talk



	Latency	LUTs	FFs	DSPs	BRAMs
AXOL1TL	2 ticks 50 ns	2.1%	~0	0	0

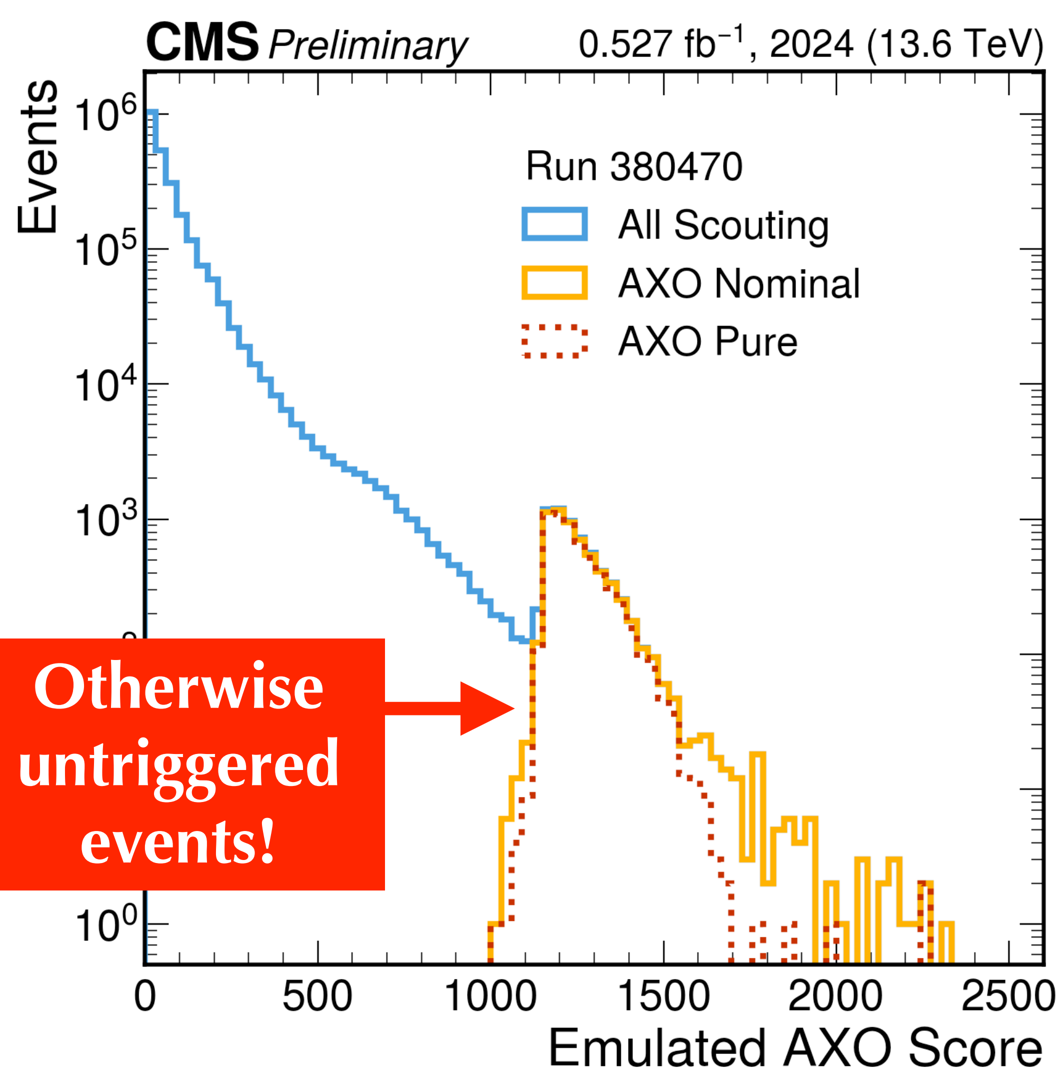
Ultra-fast anomaly detection @ CMS

Now taking data!

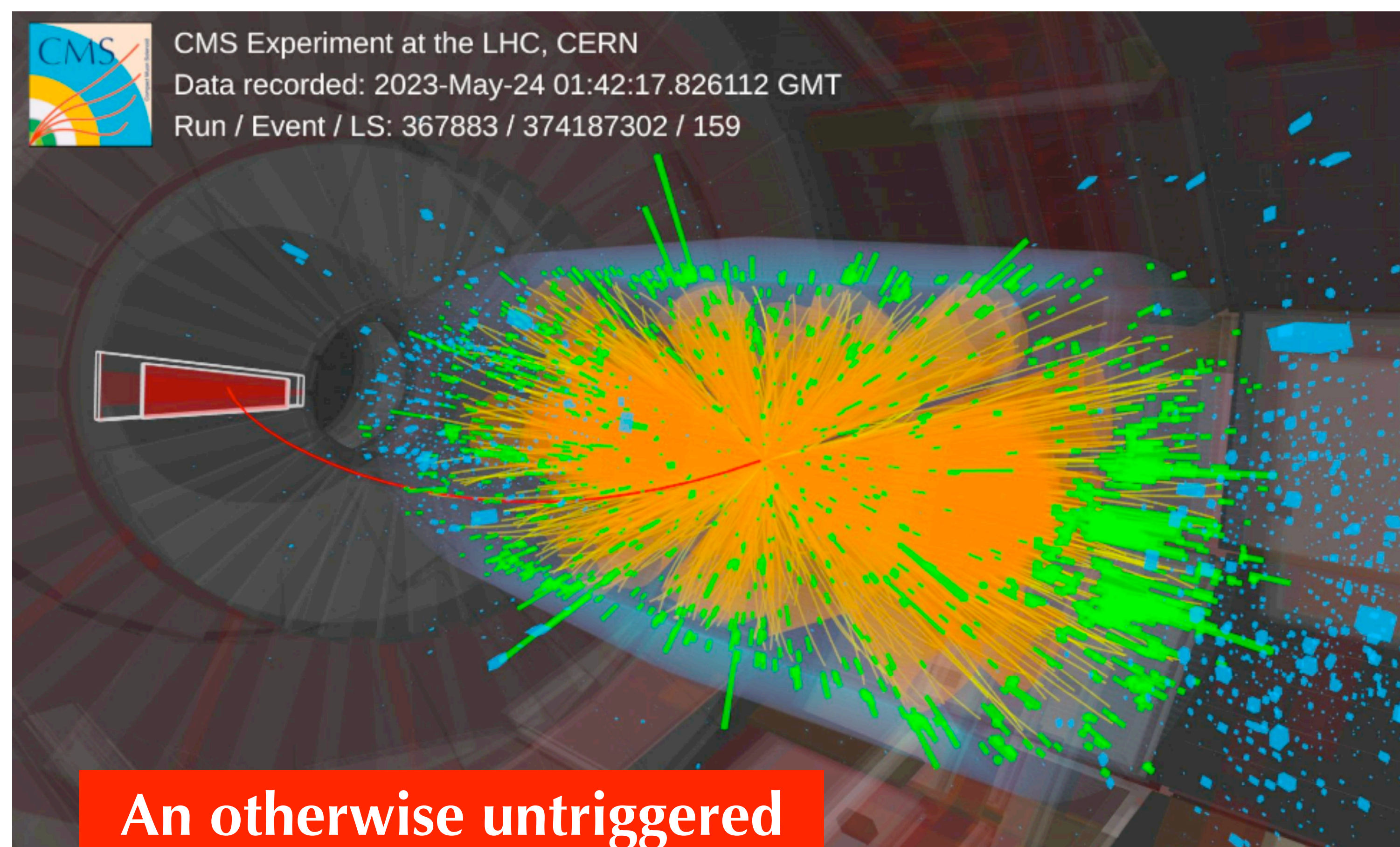
Anomaly eXtraction Online Level-1 Trigger aLgorithm

CMS-DP-2023-079
CMS-DP-2024-059

See A. Gandrakota talk



Otherwise untriggered events!



An otherwise untriggered high-multiplicity event!

A look into the future

Smartpixels: Towards on-sensor inference of charged particle track parameters and uncertainties

Real-time Inference with 2D Convolutional Neural Networks on Field Programmable Gate Arrays for High-rate Particle Imaging Detectors

Machine learning evaluation in the Global Event Processor FPGA for the ATLAS trigger upgrade

Smart sensors using artificial intelligence for on-detector electronics and ASICs

Neural network accelerator for quantum control

Neural-network-based level-1 trigger upgrade for the SuperCDMS experiment at SNOLAB

Link to articles

Low latency optical-based mode tracking with machine learning deployed on FPGAs on a tokamak

- The CMS anomaly detection trigger is a **first step towards a revolutionary new paradigm for data acquisition at colliders** (and beyond)

A look into the future

Smartpixels: Towards on-sensor inference of charged particle track parameters and uncertainties

Real-time Inference with 2D Convolutional Neural Networks on Field Programmable Gate Arrays for High-rate Particle Imaging Detectors

Machine learning evaluation in the Global Event Processor FPGA for the ATLAS trigger upgrade

Smart sensors using artificial intelligence for on-detector electronics and ASICs

Neural network accelerator for quantum control

Neural-network-based level-1 trigger upgrade for the SuperCDMS experiment at SNOLAB

Link to articles

Low latency optical-based mode tracking with machine learning deployed on FPGAs on a tokamak

- The CMS anomaly detection trigger is a **first step towards a revolutionary new paradigm for data acquisition at colliders** (and beyond)
- **AI at the edge of experiments is a game changer** — HEP at the forefront of R&D

A look into the future

Smartpixels: Towards on-sensor inference of charged particle track parameters and uncertainties

Real-time Inference with 2D Convolutional Neural Networks on Field Programmable Gate Arrays for High-rate Particle Imaging Detectors

Machine learning evaluation in the Global Event Processor FPGA for the ATLAS trigger upgrade

Smart sensors using artificial intelligence for on-detector electronics and ASICs

Neural network accelerator for quantum control

Neural-network-based level-1 trigger upgrade for the SuperCDMS experiment at SNOLAB

Link to articles

Low latency optical-based mode tracking with machine learning deployed on FPGAs on a tokamak

- The CMS anomaly detection trigger is a **first step towards a revolutionary new paradigm for data acquisition at colliders** (and beyond)
- **AI at the edge of experiments is a game changer** — HEP at the forefront of R&D
- **Requirements for AI in particle physics** go far beyond industrial and commercial applications because of extreme environments:
    - speed, throughput, fidelity, interpretability, and reliability

A look into the future

Smartpixels: Towards on-sensor inference of charged particle track parameters and uncertainties

Real-time Inference with 2D Convolutional Neural Networks on Field Programmable Gate Arrays for High-rate Particle Imaging Detectors

Machine learning evaluation in the Global Event Processor FPGA for the ATLAS trigger upgrade

Smart sensors using artificial intelligence for on-detector electronics and ASICs

Neural network accelerator for quantum control

Neural-network-based level-1 trigger upgrade for the SuperCDMS experiment at SNOLAB

Link to articles

Low latency optical-based mode tracking with machine learning deployed on FPGAs on a tokamak

- The CMS anomaly detection trigger is a **first step towards a revolutionary new paradigm for data acquisition at colliders** (and beyond)
- **AI at the edge of experiments is a game changer** — HEP at the forefront of R&D
- **Requirements for AI in particle physics** go far beyond industrial and commercial applications because of extreme environments:
  - speed, throughput, fidelity, interpretability, and reliability
- **Demonstrators TODAY crucial to learn NOW lessons for the future**

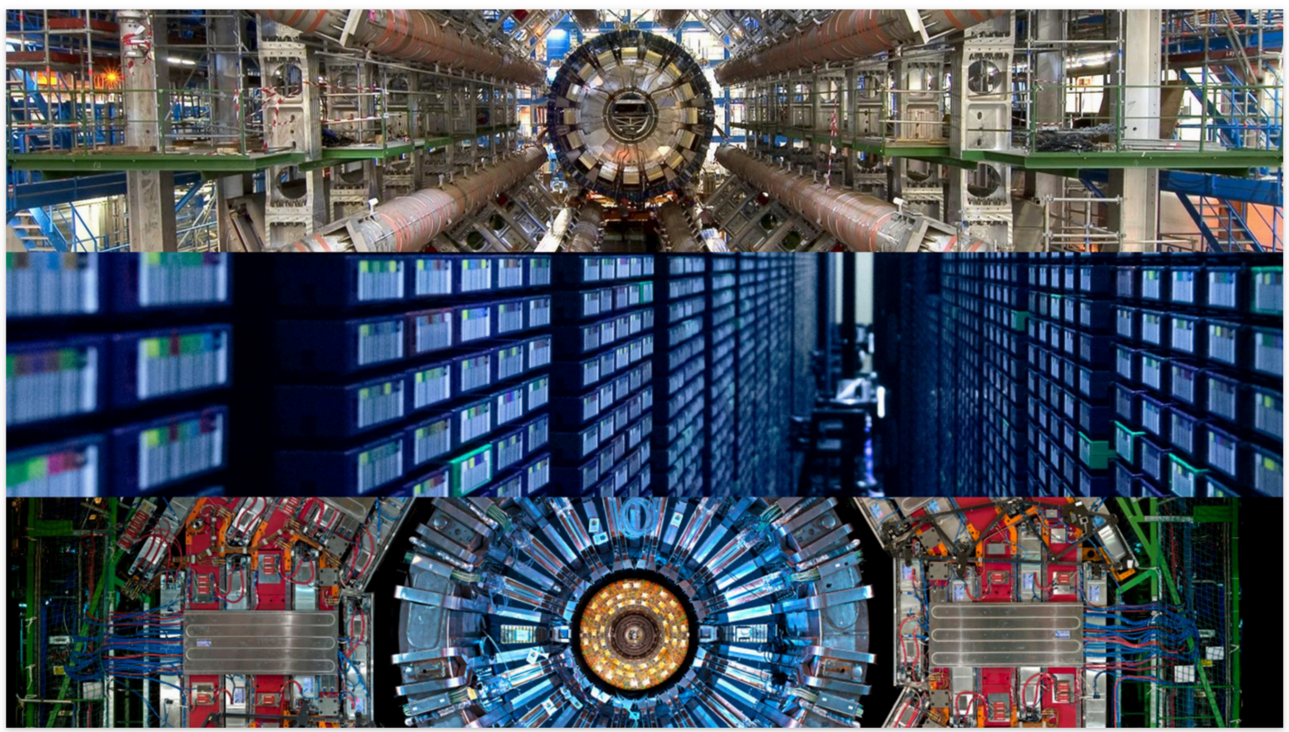
A look into the future

CERN news

The next-generation triggers for CERN detectors

The recently launched Next-Generation Triggers project is set to remarkably increase the efficiency, sensitivity and modelling of CERN experiments

11 APRIL, 2024 | By Antonella Del Rosso



From top to bottom: ATLAS, CERN Data Centre and CMS (Image: CERN)

- Eric & Wendy Schmidt foundation fund a CERN project that will *enhance the physics reach of the ATLAS and CMS experiments at HL-LHC and beyond* using novel technologies:
  - neural network optimisation
  - quantum-inspired algorithms
  - high-performance computing and FPGAs
  - theoretical modelling

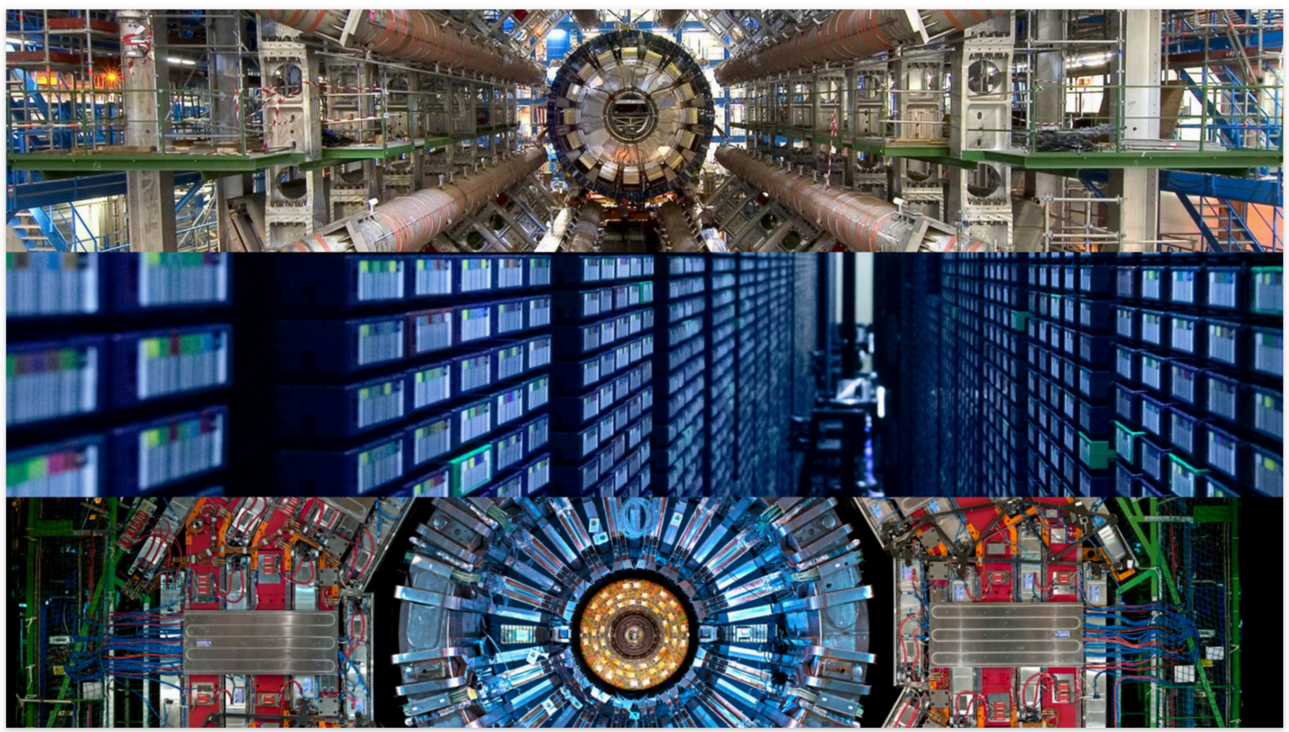
A look into the future

[CERN news](#)

The next-generation triggers for CERN detectors

The recently launched Next-Generation Triggers project is set to remarkably increase the efficiency, sensitivity and modelling of CERN experiments

11 APRIL, 2024 | By Antonella Del Rosso



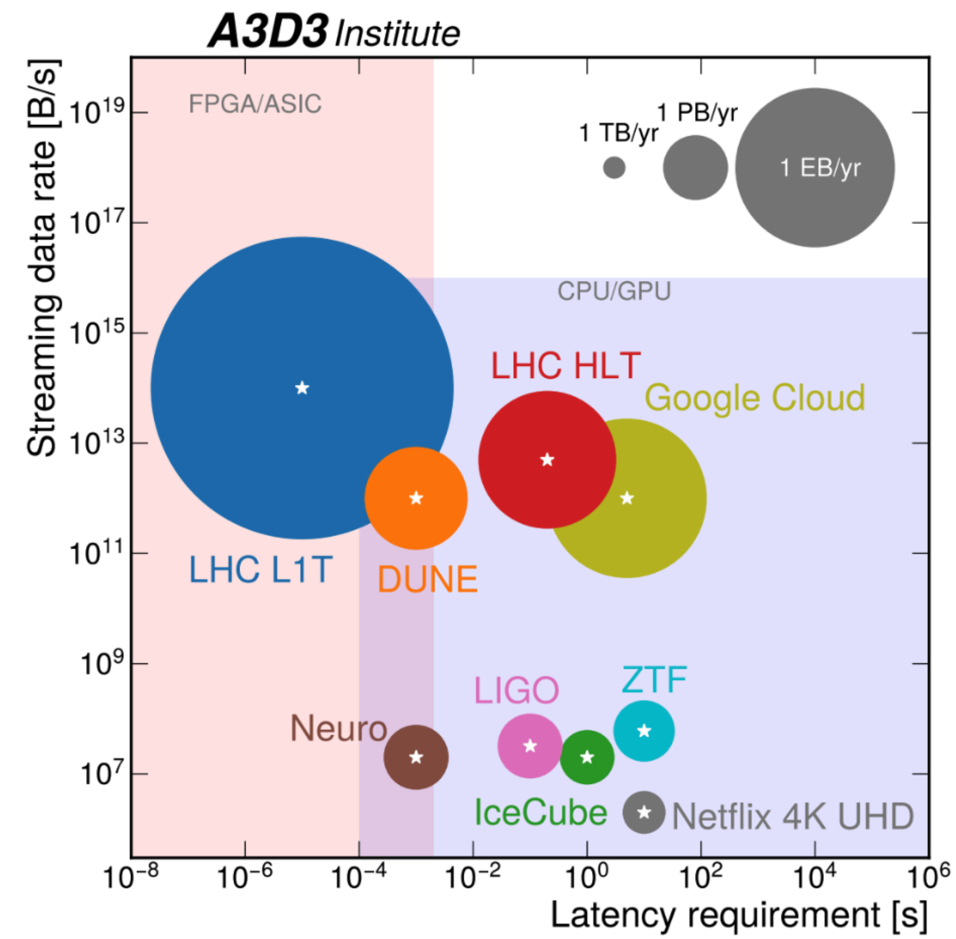
From top to bottom: ATLAS, CERN Data Centre and CMS (Image: CERN)

- Eric & Wendy Schmidt foundation fund a CERN project that will *enhance the physics reach of the ATLAS and CMS experiments at HL-LHC and beyond* using novel technologies:
  - neural network optimisation
  - quantum-inspired algorithms
  - high-performance computing and FPGAs
  - theoretical modelling

The NextGen Triggers project will mark a new chapter in high-energy physics, leveraging upgraded event-selection systems and data-processing techniques to unlock a realm of discoveries.

Summary & Outlook

- **Our scientific mission is beautiful and engaging:** answer to fundamental questions about the universe with very advanced particle detectors
  - enormous challenges due to large volume of data and experimental system constraints
  - AI offers a solution because of its scalable capabilities
  - but not without an intense effort to change our approaches at every data processing stage



- **Promote interdisciplinary collaborations:** *physicists, data scientists, computer scientists, electrical and computer engineers, software engineers*

- **To invent solutions that would not be available in the world otherwise**

- **Then inject back into society the technology innovation fuelled by our unique physics requirements**

ICHEP 2024

PRAGUE

ichep2024.org

42nd International Conference on High Energy Physics

18-24 July · 2024 · Prague · Czech Republic

IUPAP

International Union of Pure and Applied Physics

Machine learning for particle physics ⟺ **Particle physics for machine learning**

Thank you!