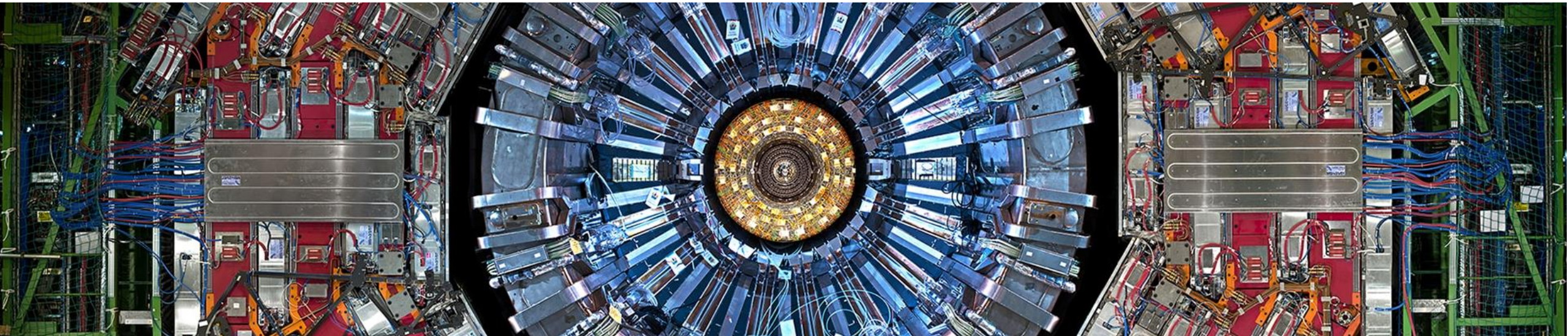
ÖAW AUSTRIAN ACADEMY OF SCIENCES

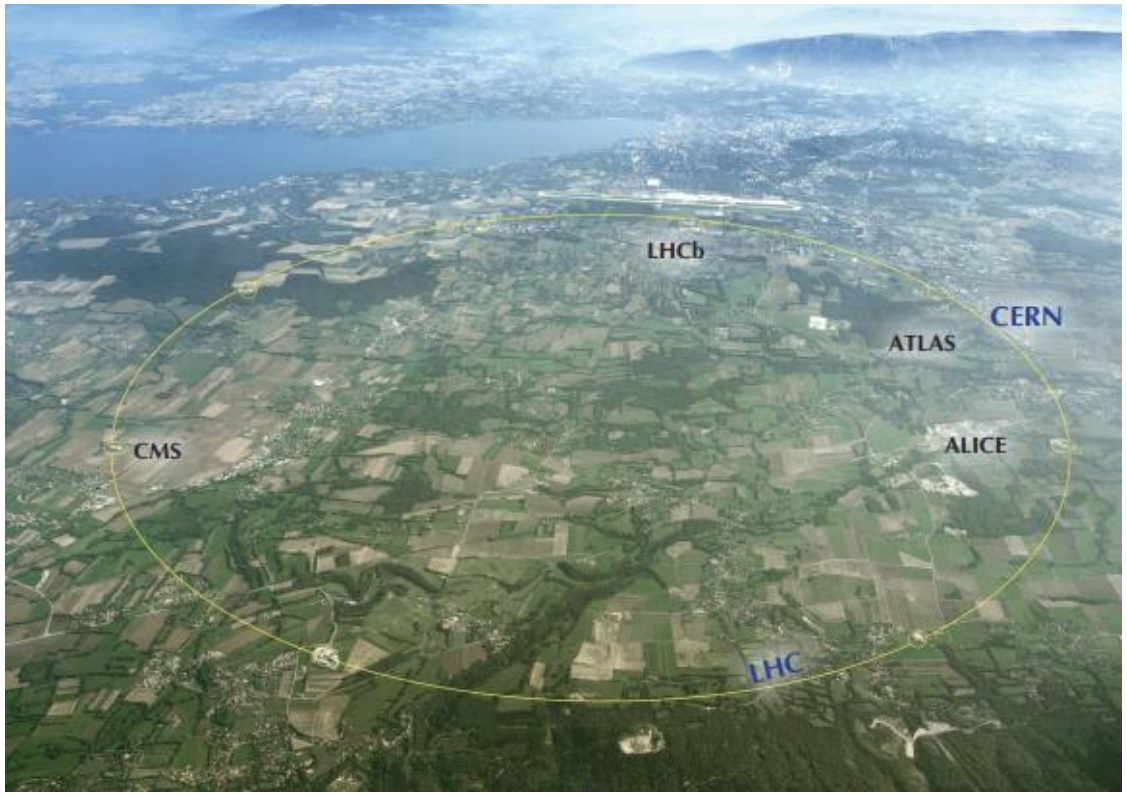
MBI Marietta Blau Institute for Particle Physics

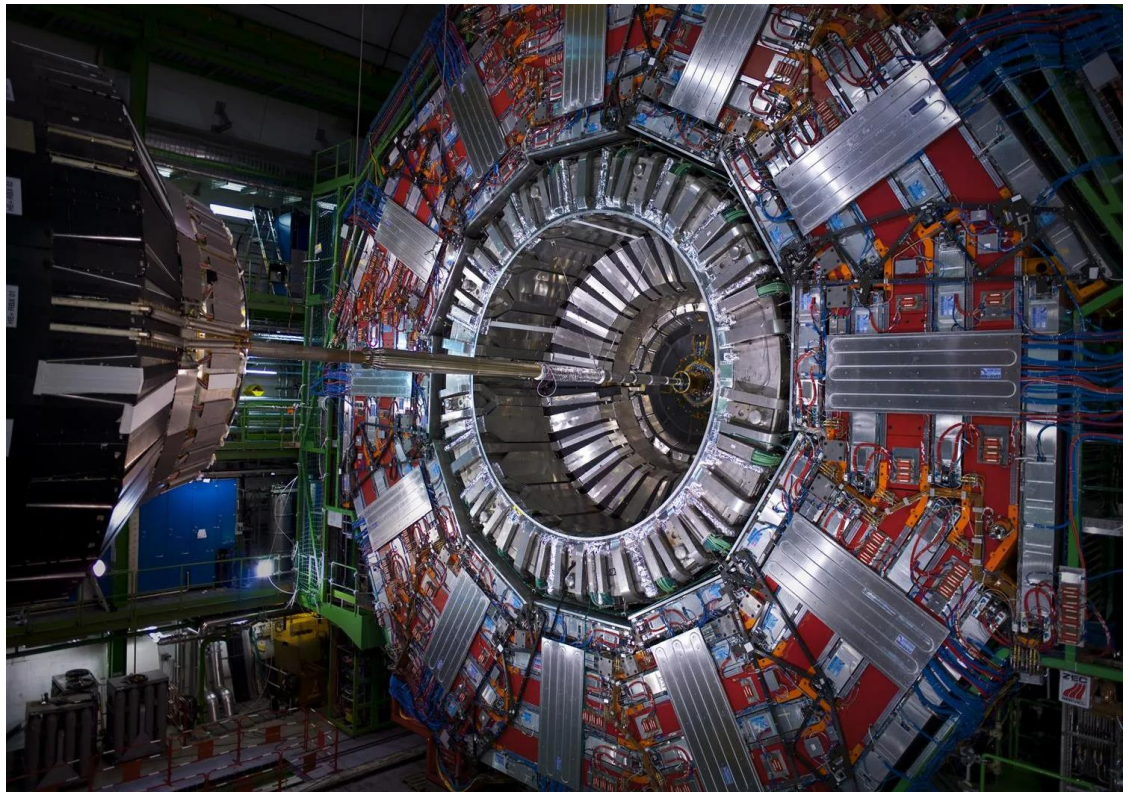
CLIP, the Worldwide LCG, and the Universe

R. Schoefbeck (MBI)



CERN, CMS, and Austria

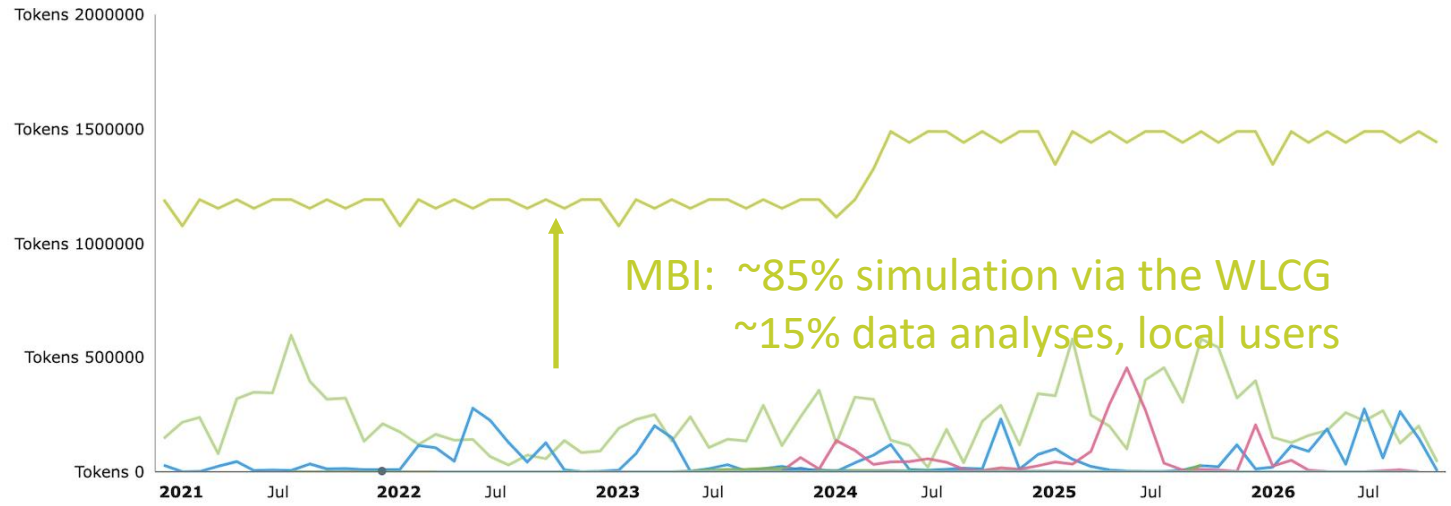




CERN: European particle physics laboratory near Geneva, home to the 27 km LHC.
CMS: Records LHC collisions and co-discover the Higgs boson in 2012.
Austria: CERN member since 1959. MBI (formerly HEPHY) is a founding member CMS.
MBI: Contributed to the silicon tracker and trigger electronics, and leads searches and measurements.

Worldwide LHC computing grid

CLIP hosts "T2": Virtual computing center - part of WLCG



WLCG transfers

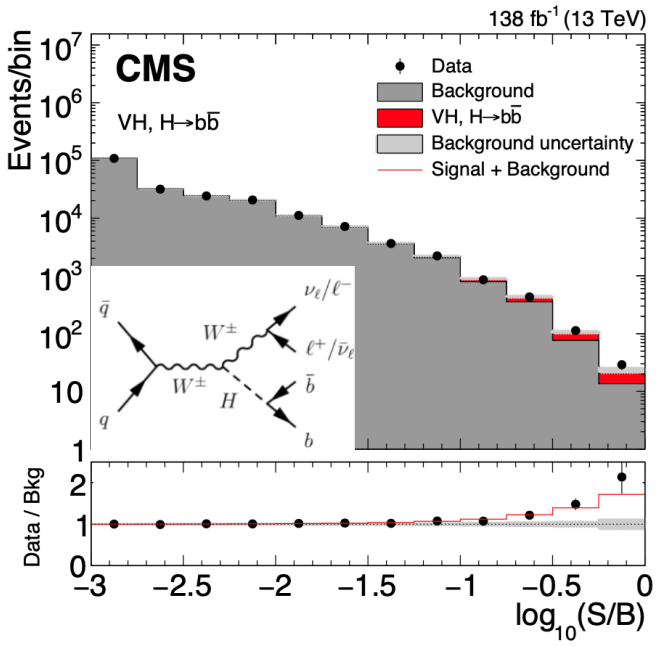


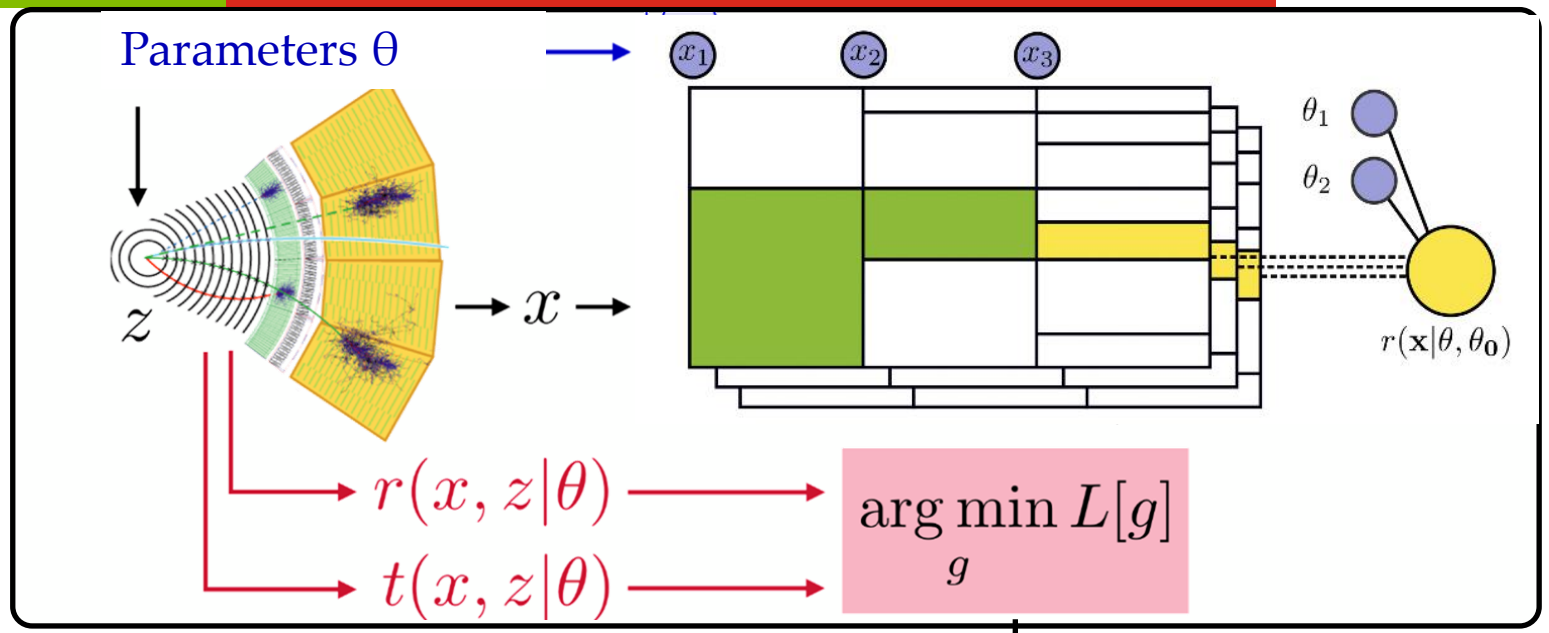
Consumption is constant – enabled by sharing resource - high efficiency

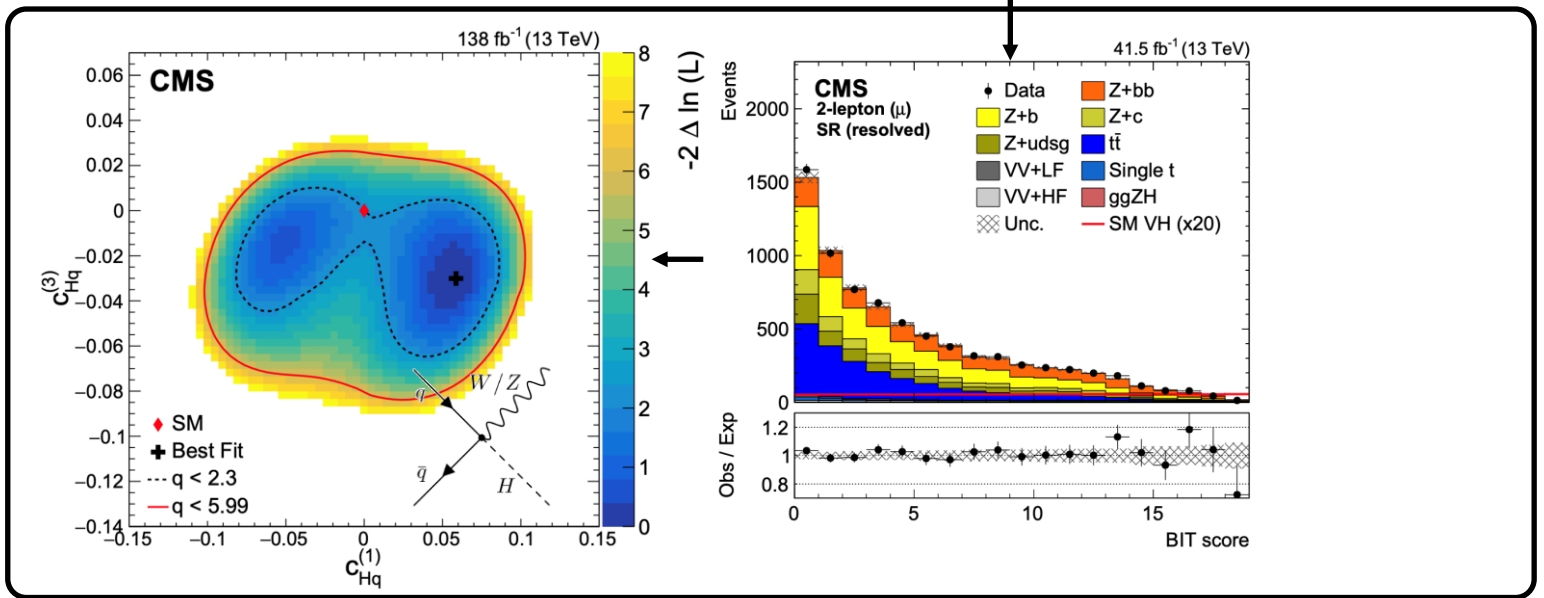
WLCG stores, processes, and shares LHC data.
170+ sites in 40+ countries, over 1 million CPU cores and
2 exabytes of storage [link]

Scientific progress loop:
Theoretical question → *predictions* → *simulation* → *data analysis*

Couplings of the Higgs boson







1. Theoretical prediction
2. Event simulation (T2)
3. ML algorithms
   - Dedicated, in-house, CBE
   - Award winning
4. Up to x8 better Higgs boson coupling measurements

Impossible without CLIP/CBE!

Higgs boson couplings

- Problem: Particle mass break symmetry

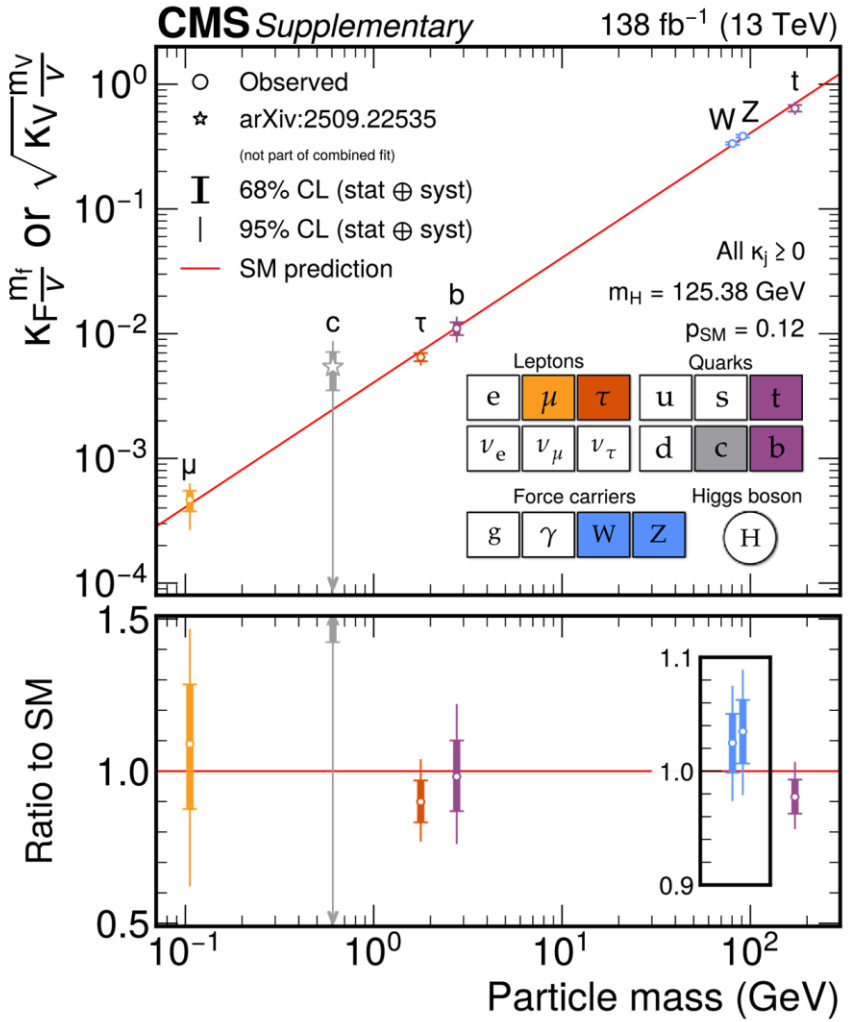
$$\mathcal{L} \supset -\cancel{m}\,\underbrace{\bar{f}f}_{\text{particle}}$$

- Idea (Higgs et al.)
  - Couple to a new field, keep symmetries & shift that field
    → mass term *generated*

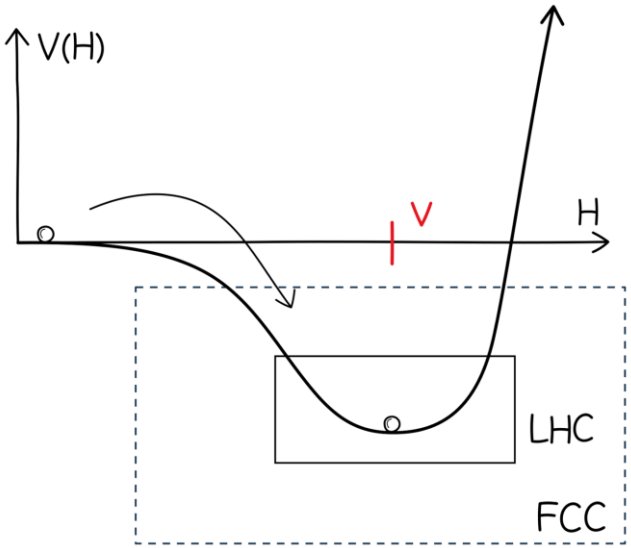
$$-y(\bar{f}H^0 f), \qquad H^0 = \frac{1}{\sqrt{2}}\left(\underbrace{v}_{\text{const}} + h\right)$$

$$\mathcal{L} \supset -\underbrace{yv}_{\text{mass}}\,\bar{f}f - \underbrace{y}_{\text{Higgs coupling}}\,h\bar{f}f$$

- Preserves symmetries
- Prediction: Particle mass ∝ coupling

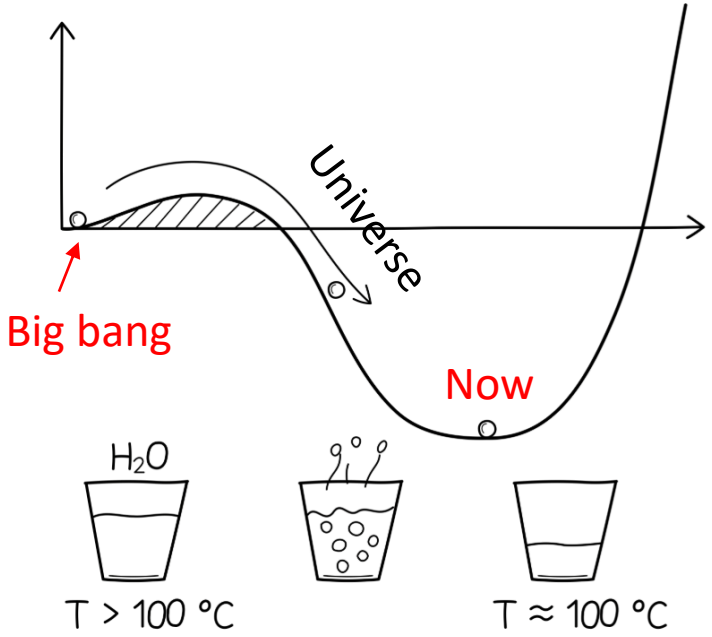


Take this further:
How does the shift come about?
Answer: Add Higgs potential



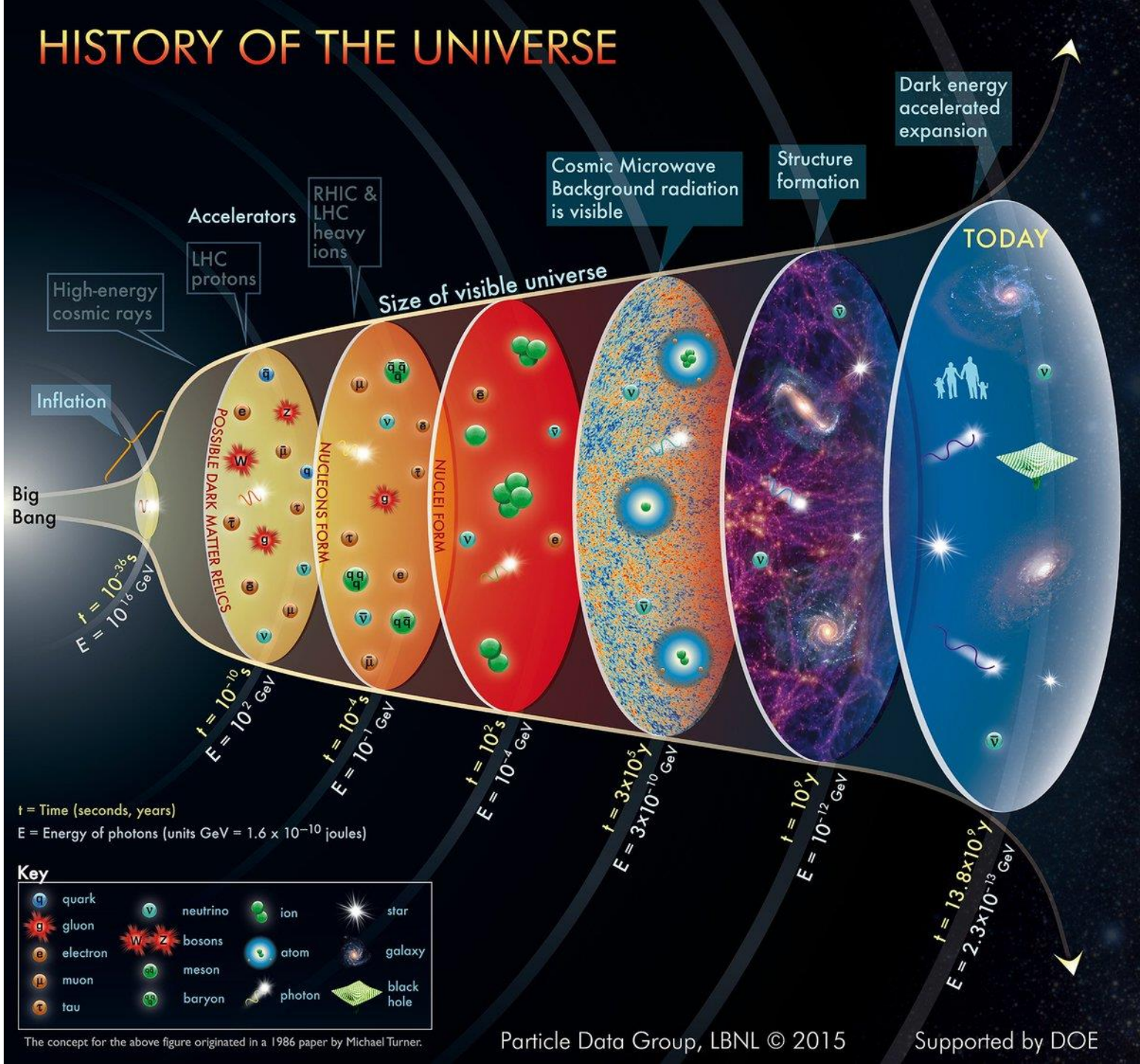
- After the big bang, universe smoothly rolled down
- Presently sits at minimum
- Measure the details of the potential at LHC/FCC
- Is that relevant at all?

- Matter was created (0.1 ns)
  - **We don't understand *how***
  - Matter annihilates anti-matter
  - Far too uniform ⟵ Higgs potential helps
- Assume a *barrier* in the potential



  - After BB: "tunneled" through barrier
    - "1st order phase transition"
    - Violent "boiling" of H-field
    - May have create visible matter

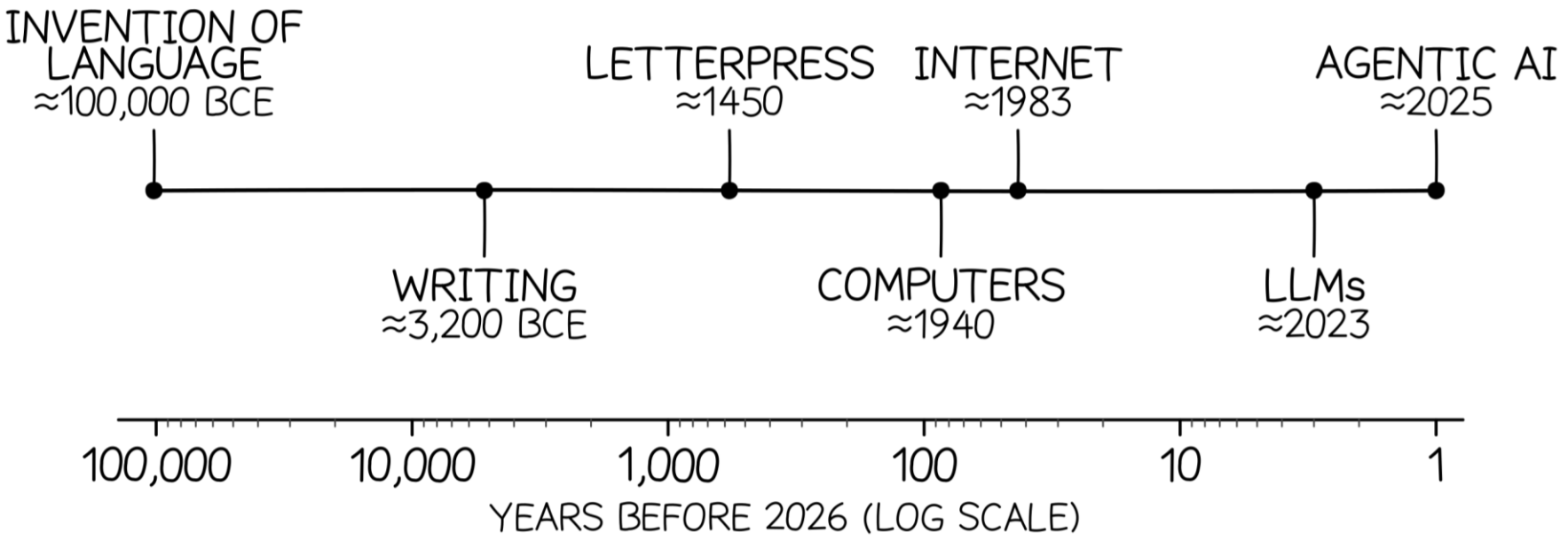
  - Measure potential with Di-Higgs

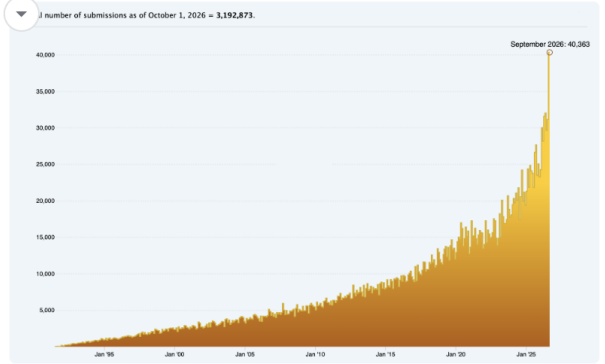


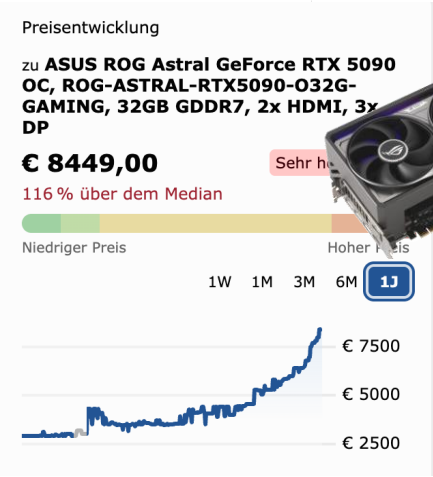




Moore's law, extended









⟶ phase transition in *information processing*

Outlook

Science lives in language

- Papers, explanations, wiki instructions, and the discussions
- Computers now take part
- Phase transitions involve all scales

Small scale

- Data analysis workflows automatized
- Cost of idea generation dropped
- Time scale of implementation ~zero

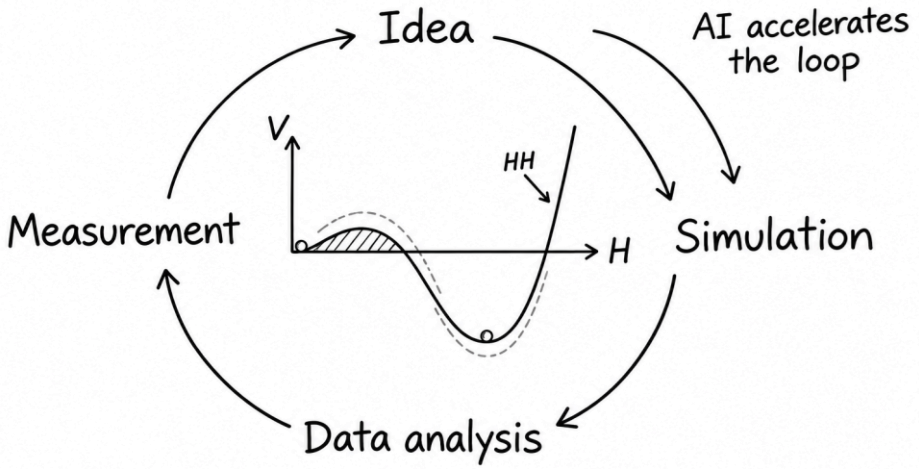
Medium scale

- Papers are "implemented"

Large scale

- Millenium Prize problems solved
- *Companies* establish research groups
- Fields will disappear/emerge

A phase transition of how we study phase transitions



Large research bodies get organized

- CERN AI strategy [link] implemented in experiments
  - 8 Working groups, lighthouse projects (FLARE – Mistral)
- Many activities at the OeAW

This will be one wild ride

- Computing is at the heart of all the changes
- Be hopeful: We have the best people in place

Working groups and lighthouses

Lighthouse	Sector	RCS working groups
01 Common AI Infrastructure & Platform	RCS-IT	WG7 Infrastructure
02 AskCERN	all sectors	WG6 LLM assistants
03 Edge AI for real-time DAQ	RCS	WG4 Online processing · ATS involved in rad-hard ML
04 Universal Fast Simulation Engines	RCS	WG3 Generation & fast sim
05 AI-driven operations	ATS, RCS	WG5 Control & operations
06 Accelerator systems design	ATS	RCS interest: surrogate models, in common with LH-09
07 Physics Foundation Model	RCS	WG2 Global event interp.; WG1 Pattern recognition
08 AI Agent for Particle Physics	RCS	WG6 LLM assistants; WG8 Theory
09 AI-assisted detector design	RCS, ATS	— (a new WG, if the effort starts)
10 Material discovery	ATS	RCS interest: [to confirm]

CERN AI strategy and lighthouses: cio.web.cern.ch/artificial-intelligence

Harry Potter generator

Training time 00:00

```
GPU 0  [||||||||||||||||..]  90%    1.0/20.0 GiB
GPU 1  [||||||||||||||||||] 100%    0.4/20.0 GiB
GPU 2  [||||||||||||||||||] 100%    0.4/20.0 GiB
GPU 3  [||||||||||||||||||] 100%    0.4/20.0 GiB
```

> Eine Eule klopfte gegen das Fenster.

ks wirbelnMundungus merk gezeigt Madame merk Umbrid Unterrichtsstunden Augenliderenz weder heitergebnisseäs» Wen Kommando Wurmschwanz Zaubertränke Gegenflüchegänge …

Curated replay · clock shows original training time

CERN–Mistral: FLARE

- **15 Oct 2026: CERN–Mistral statement of intent**
  CERN and Mistral AI. CERN supplies physics expertise, scientific data and computing. Mistral contributes expertise in large-scale model training.
  Funding/support: Budget and resource split undisclosed (undefined) in the announcement.

  FLARE / Large Physics Model Lighthouse initiative
  Goal: a single pretrained foundation model trained on billions of simulated proton–proton collisions. Intended tasks include reconstruction, particle identification, anomaly detection and studies of future detector designs.
  Named representatives: Enrica Porcari, CERN CIO, and Johannes Brandstetter, Mistral VP of AI for Engineering. CERN lists FLARE among its AI initiatives.

  - AI discussion within CMS [WGM Oct. 29th]: Similar to our agentic AI meeting

  - CERN AI strategy within CMS: M. Pierini in the WGM on Nov 12th
    - CMS will then have to decide what to do

  - Keep planned OeAW – Mistral agreement separate

  - Token consumption of MBI CMS group in units of API: 32kEUR/month (x36 subsidized)

Scientific news

- Geoffrey Hinton (physics nobel laureate) was in Vienna
  - Organized by the TU Wien [recording]. Shouldn’t have been insiders' knowledge.



“You wouldn't teach your kid to read with serial killer’s diaries”
“Smart things can always find a way of getting around dumb things”
“There is no internal theater”
“I believe that death is good”

AI science results

- **8 Sep: Navier–Stokes singularity (OpenAI)**
  Can a smooth viscous flow develop infinite velocity in finite time? Proposed Millennium solution with smooth forcing and finite energy. Paper and Lean public.
- **6 Oct: Quasi-Riemann hypothesis**
  Zeta/L-function zeros govern fluctuations in the distribution of primes. Claimed zero-free region Re(s) > 7/8 for Dirichlet L-functions, including ζ(s). Full RH remains open.
- **6 Oct: Hodge conjecture for CM abelian varieties**
  Can certain topological cycles be built from subspaces defined by polynomial equations? Claimed proof for CM (complex multiplication) abelian varieties in all dimensions and codimensions, a special case of the Millennium problem.
- **6 Oct: Hilbert's tenth problem over the rationals**
  Can a universal algorithm decide whether an integer-coefficient polynomial equation has a rational solution? Claimed proof that no such algorithm exists, extending Hilbert's original integer problem.

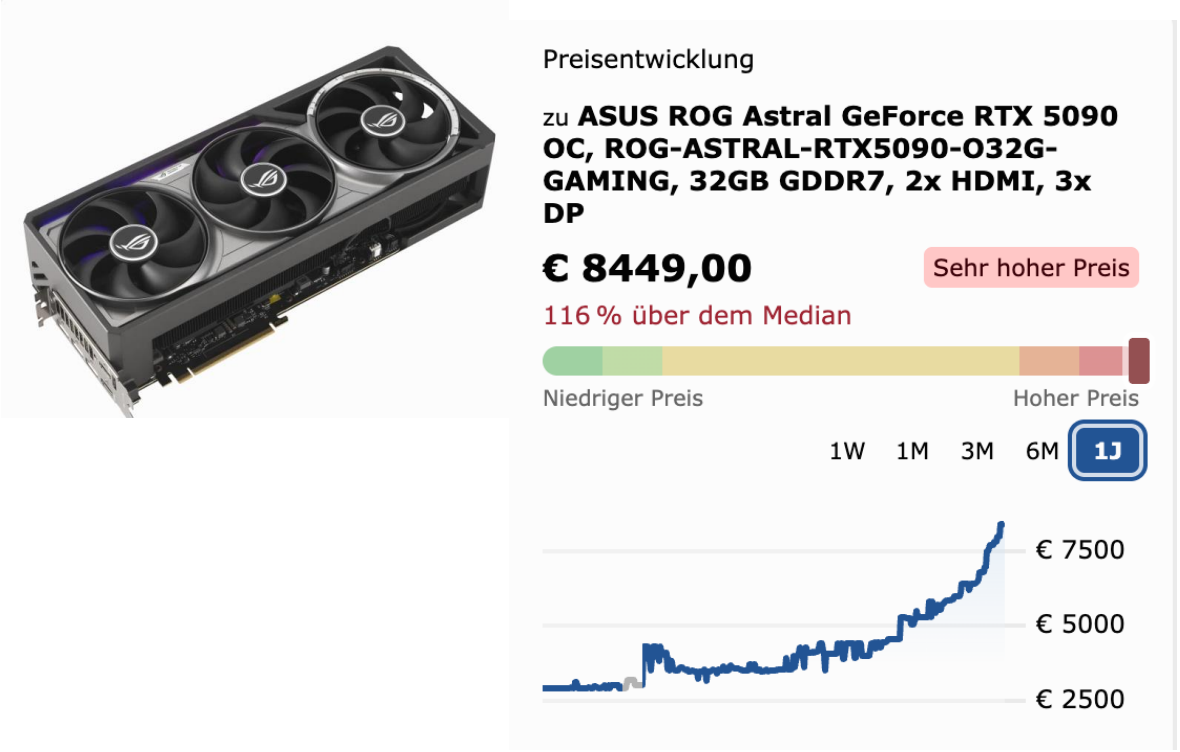
All four: OpenAI internal model. Independent scrutiny continues. Announcement dates: 2026.

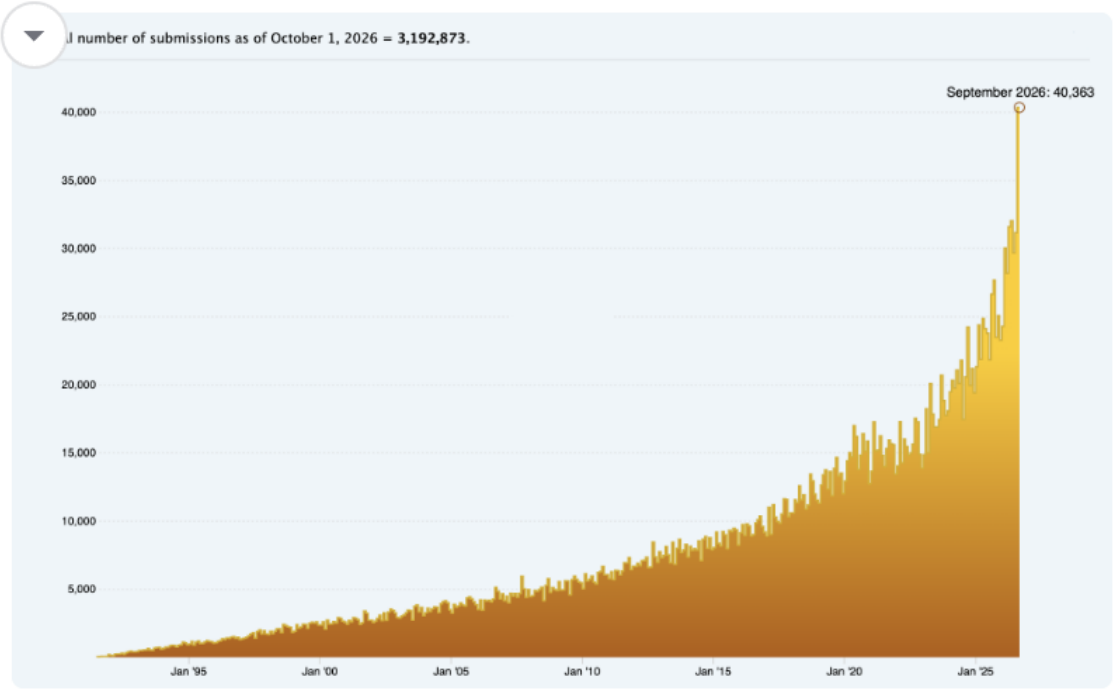
AI science results

- **14 Oct: AI-designed proteases (RFdiffusion)**
  Can we create enzymes that cut proteins at chosen sites? Chen, Wu, Baker et al. used RFdiffusion: 36% of designed zinc proteases active in lab tests. Peer-reviewed (Nature).
- **26 Oct: Local smoothing, Kakeya and restriction**
  Must extreme wave focusing be short-lived? OpenAI claims local smoothing in all dimensions, implying Kakeya and Stein restriction. Paper and Lean public.
- **26 Oct: Exponential Time Hypothesis (ETH)**
  Must solving logical constraints on true/false variables take exponential time? OpenAI claims a subexponential 3-SAT algorithm. Review: Impagliazzo, Paturi, Williams. Paper and Lean public.
- **26 Oct: Yang–Mills / QCD existence**
  In pure Yang–Mills, massless gluon fields should produce a spectrum with a massive lightest excitation above the vacuum, closely linked to confinement. OpenAI claims rigorous 4D existence only. Review led by David Gross. Artifacts await release. Mass gap remains open.

OpenAI's 26 Oct results remain claims under independent examination. Announcement dates: 2026.

Correlation = Causation





Arxiv submissions
(before constraints)

AI science landscape

- **8 Oct: Anthropic / Genesis Mission**
  Anthropic, DOE/national labs, NASA, NIH and NSF.
  Field: sciences including fusion and quantum computing.
  Funding/support: $150m over 3 years for Claude access, API credits and technical support.
  FNAL/LPC intro by Daniel Diaz [link]
  -> Very concrete integration in CMS workflows

- **15 Oct: Simons / Anthropic / OpenAI**
  Simons Foundation, Anthropic and OpenAI.
  Field: mathematics and basic science.
  Workshops, assessment of AI results and access to tools.
  Funding/support: $15m combined funding/resources, including $5m from OpenAI. Simons manages awards.

- **15 Oct: OpenAI / Rosalind Academic Plan**
  OpenAI and academic labs, including UCSF and Stanford/Arc Institute.
  Field: life sciences. Sponsored GPT-Rosalind access.
  Funding/support: Up to 30 free seats per lab for up to 12 months. Total budget undisclosed.

AI science landscape

- **21 Oct:** OpenAI / PerturbAI / Scripps Research

  OpenAI, PerturbAI and Scripps Research. Thousands of agents analyse mouse-brain gene-perturbation data and propose experiments.
  Field: neurobiology.
  Funding/support: Amount undisclosed. Joint research and a public reasoning atlas.

- **22 Oct:** Anthropic / guest investigators (!!)

  Anthropic researchers with academic, nonprofit or government-lab groups. Direct 8–12-week collaborations.
  Field: life sciences and drug discovery.
  Funding/support: Free model/harness use. Total budget undisclosed. Joint publication and open code/data planned.

- **26 Oct:** OpenAI / academic research groups

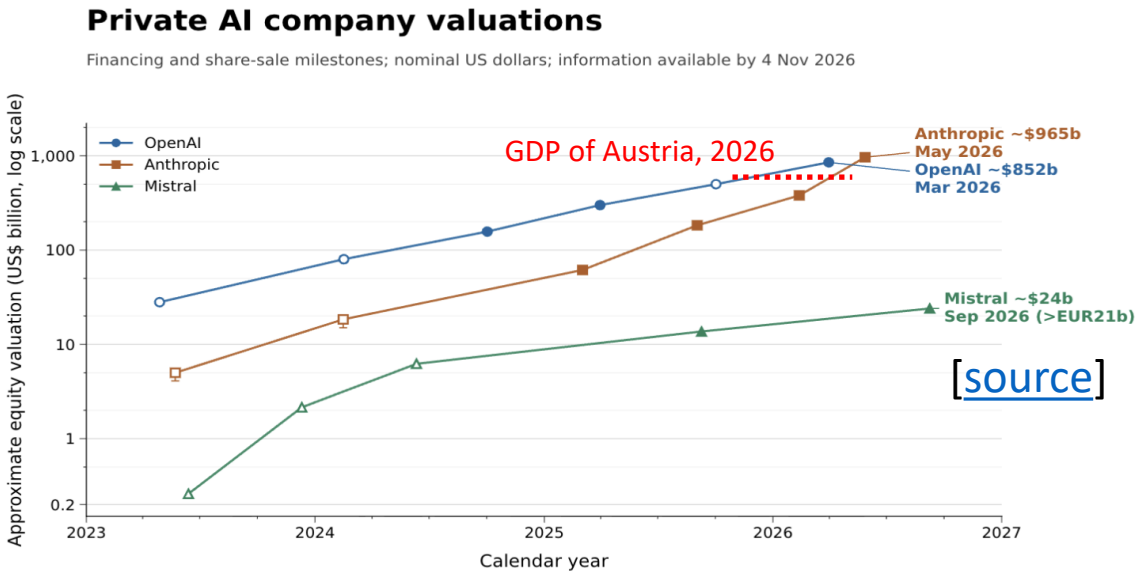
  OpenAI with independent groups. Gross, Douglas, Jaffe, Klebanov and Nekrasov examine claimed QCD/Yang–Mills existence proofs.
  Field: mathematics, complexity theory and QFT.
  Funding/support: Unpaid participation and conference support. Budget undisclosed.

AI science landscape

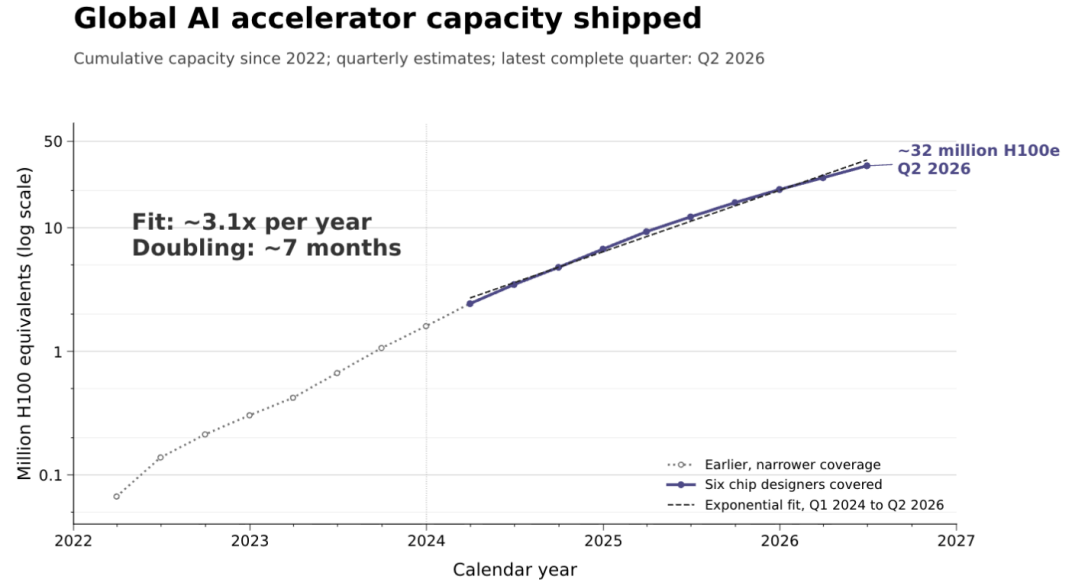
- **27 Oct:** [Anthropic / six drug-development partners]
  BigHat Biosciences, Bristol Myers Squibb, Divergence Labs, Enveda, Generate Biomedicines and The Column Group, with Anthropic. Molecular design and development workflows, including laboratory-tested CAR-T component designs.
  Field: drug discovery, mostly preclinical.
  Funding/support: Amounts undisclosed. Direct research partnerships with Anthropic scientists and engineers.
- **28 Oct:** [Anthropic / AI for Science expansion]
  20 funders, institutions and groups, including EMBL, Sanger, Crick, STScI, Sloan, Simons and Schmidt Sciences. Related work with HHMI Janelia develops AI control of lab instruments. JWST, NASA, …
  Field: biology, astronomy, climate, mathematics and other sciences.
  Funding/support: Free Claude credits through institutions and selected projects. Total value undisclosed.

Where are we headed?

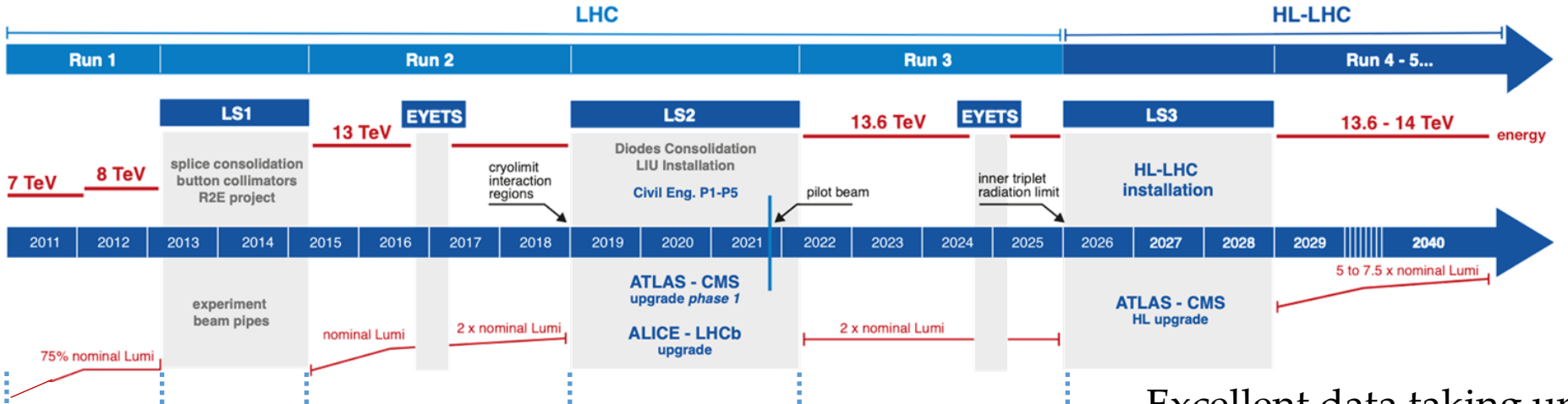
- A factor 37 of subsidy is not sustainable
- Prepare for paying the true cost
  - MBI CMS consumption amounts to approximat
  - The rational response will be to keep 1-2 FTE w
  - *Token inequality* is a consistent pattern in the AI
  - *Diversify* our where we source from
  - *Bundle* share resources



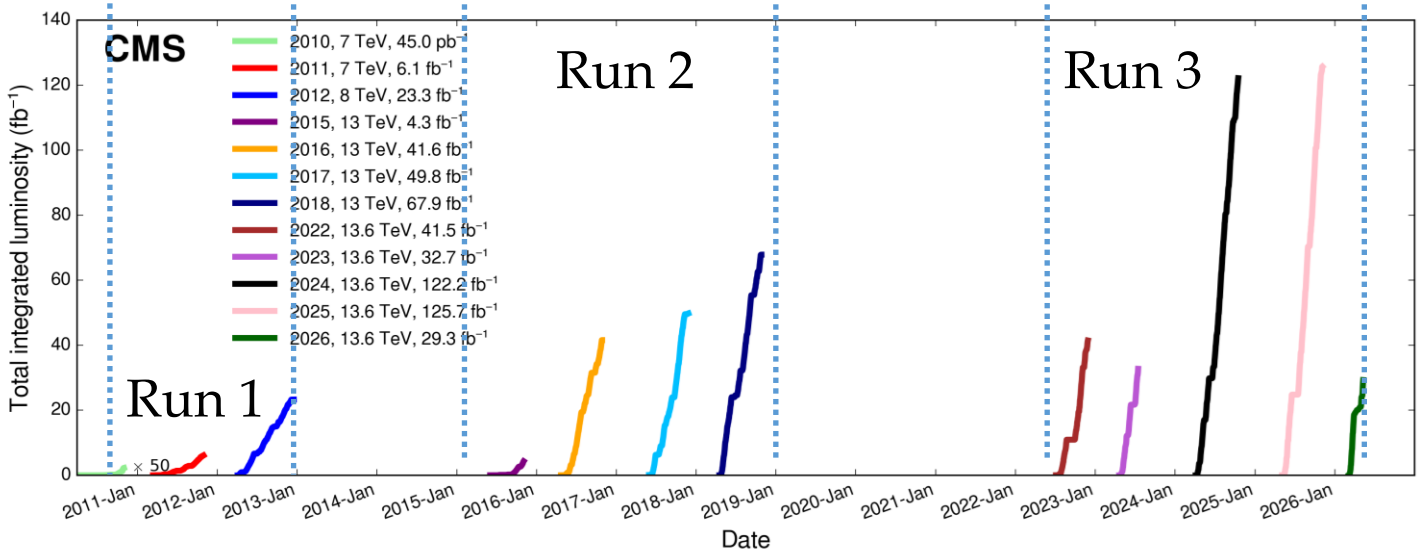
[source]



HL-LHC long term schedule

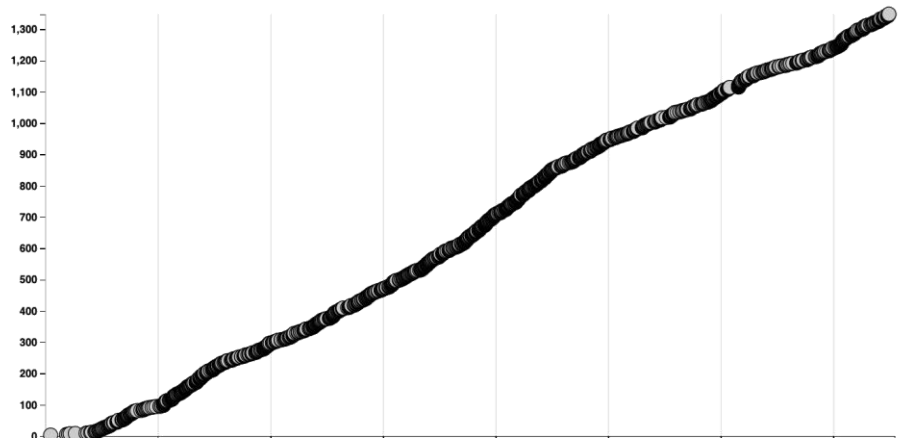


√s=7/8 TeV √s= 13 TeV √s= 13.6 TeV



Excellent data taking until Run 3 >500/fb

- HL-LHC until 2041 – x6 more data!
- ATLAS & CMS: ~1500 collider-data papers



2012 2024