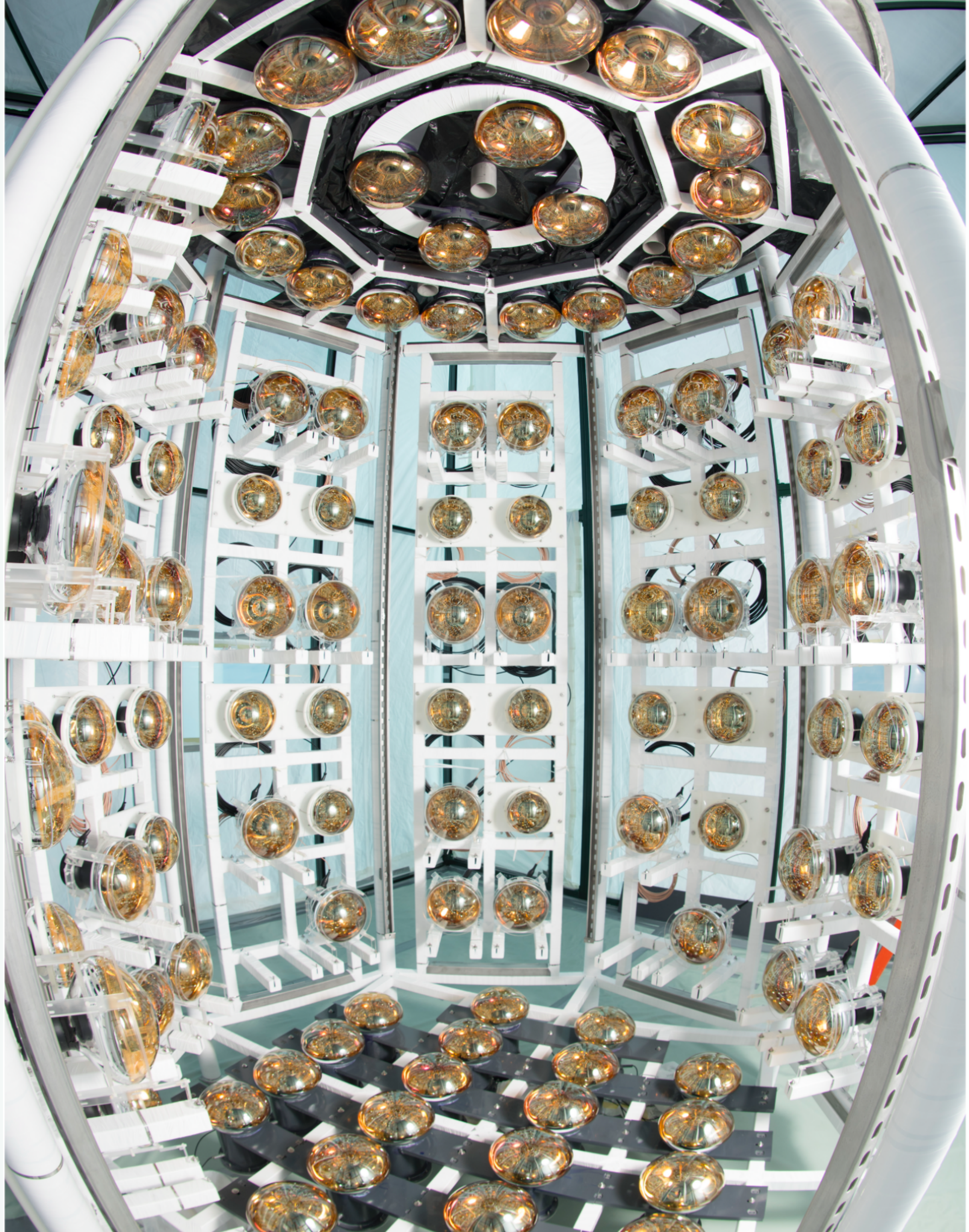
AI for Smaller Experiments

Mayly Sanchez - Florida State University

CLARIPHY - AI Collaboration Meeting - September 14, 2026



What makes an experiment small?

"Small experiments exist on a continuum, but they share the fact that the team doing the analysis is the same that is constructing and running the experiment."

Lindley Winslow, "Thoughts on Small Experiments" at the Blueprint Workshop.

Photo: ANNIE Collaboration

What makes an experiment small?

"Small experiments exist on a continuum, but they share the fact that the team doing the analysis is the same that is constructing and running the experiment."

Lindley Winslow, "Thoughts on Small Experiments" at the Blueprint Workshop.

Not a number of people threshold! For a small experiment, the impact of AI can appear in the measurement, in the control room, or in the software that makes the experiment possible.

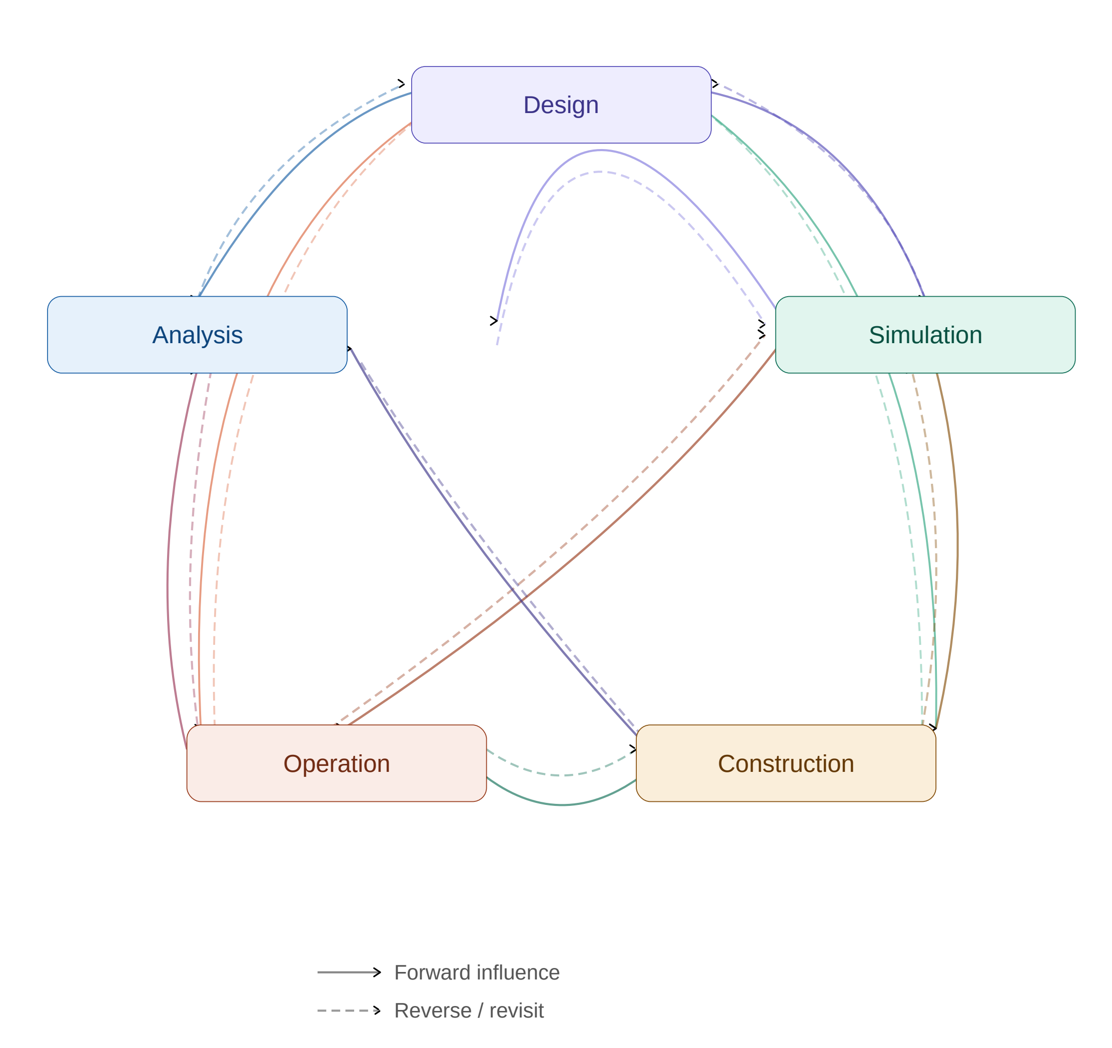
Big impact: improvement in one task frees essential human cycles for the rest

Photo: ANNIE Collaboration

One team, the whole experimental lifecycle

“These steps cannot be treated independently and we need better tools for each task and integration of these tasks.”

Lifecycle of the experiment from L. Winslow



AI opportunities follow the work of the experiment.

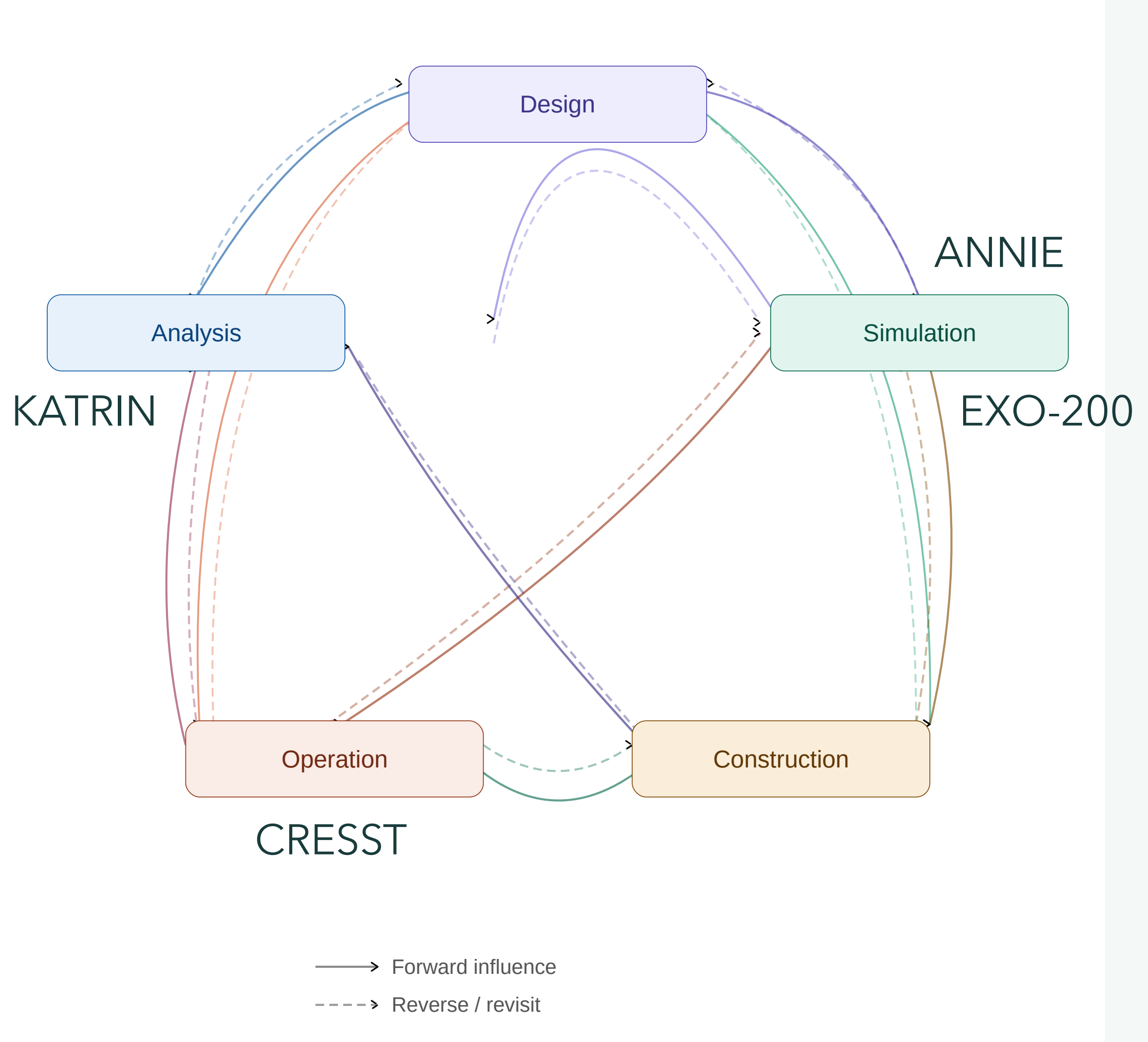
One team, the whole experimental lifecycle

"These steps cannot be treated independently and we need better tools for each task and integration of these tasks."

I will use a few concrete examples on how AI is being used in small experiments, and what it takes to make the gains scientifically useful.

AI opportunities follow the cycle of the experiment.

Lifecycle of the experiment from L. Winslow

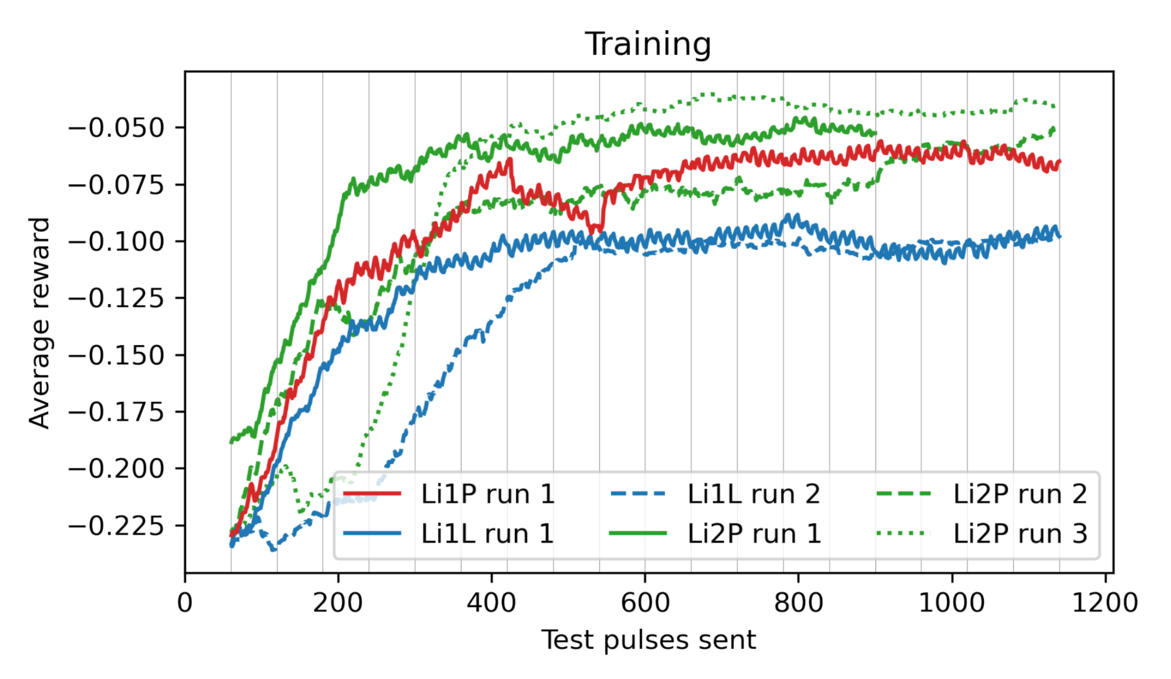


CRESST: Tuning the detectors

- Cryogenic calorimeters depend on heating and bias settings, with a response that depends on the path taken through parameter space.
- A deep reinforcement-learning controller selects heating and bias settings from the measured test-pulse response.
- Repeatable tuning with less manual intervention, while retaining a clear measure of detector performance.

3 detectors
6 live optimization runs

Convergence within ~1.5 h



G. Angloher et al., arXiv:2311.15147, Fig. 7

Live convergence demonstrated; operating points below the human benchmark. (in 2023!)

KATRIN: Accelerating the calculation

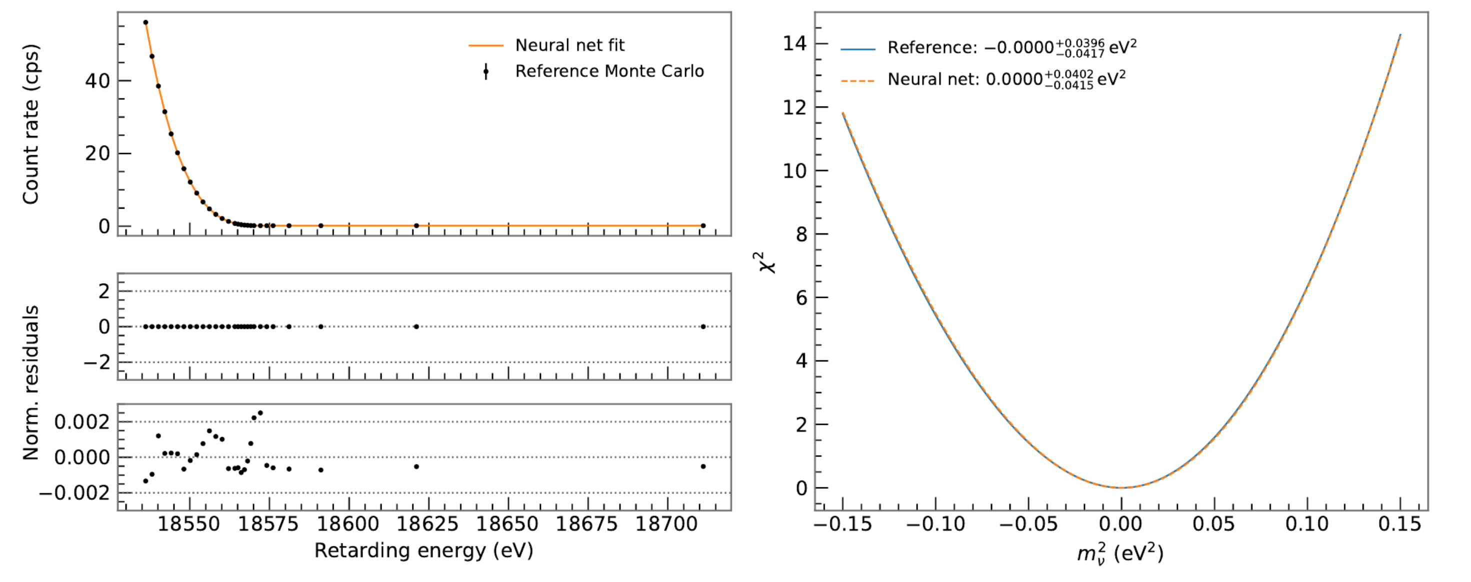
- KATRIN trained a neural network to predict the beta-decay spectrum from the relevant physics and nuisance parameters.

- They report 1000x faster model calculation while meeting relative accuracy requirements below ~10^{-4}.
  Used in the analysis yielding
  mv < 0.45 eV (90% CL)

Independent direct calculation remains a physics cross-check.

~1,000× **faster neural model evaluation**

Spectrum and likelihood cross-check on simulated data



C. Karl et al., arXiv:2201.04523, Fig. 4 | KATRIN, Science 388 (2025), arXiv:2406.13516

ANNIE: Porting the simulation and analysis

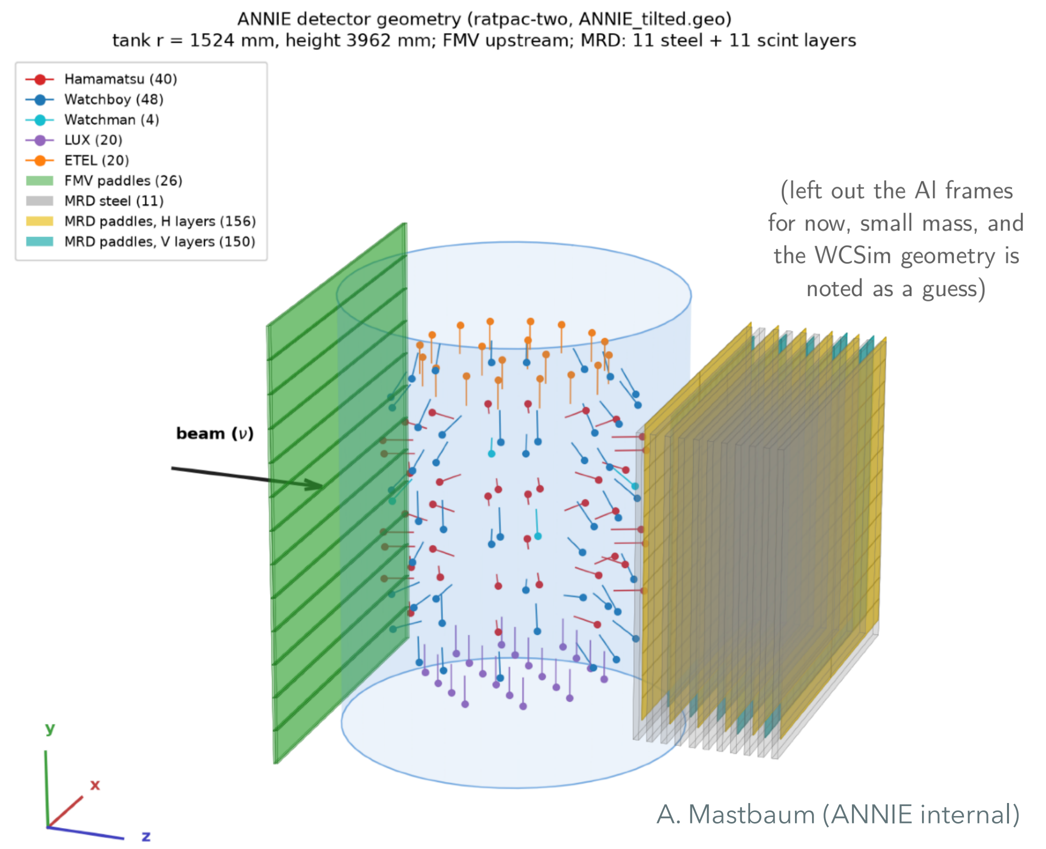
- Recently used agentic coding to port simulation/analysis workflow from WCSim to RATPAC.

- Rewrote entire chain in a week or so: Implemented tools, repeated the calibration ladder, reproduced previous analysis work.

Data import, detector geometry, reconstruction, calibration analysis tools now exist.

Work in progress: human validation and energy-scale closure issue remain.

WCSim + ToolAnalysis → RATPAC



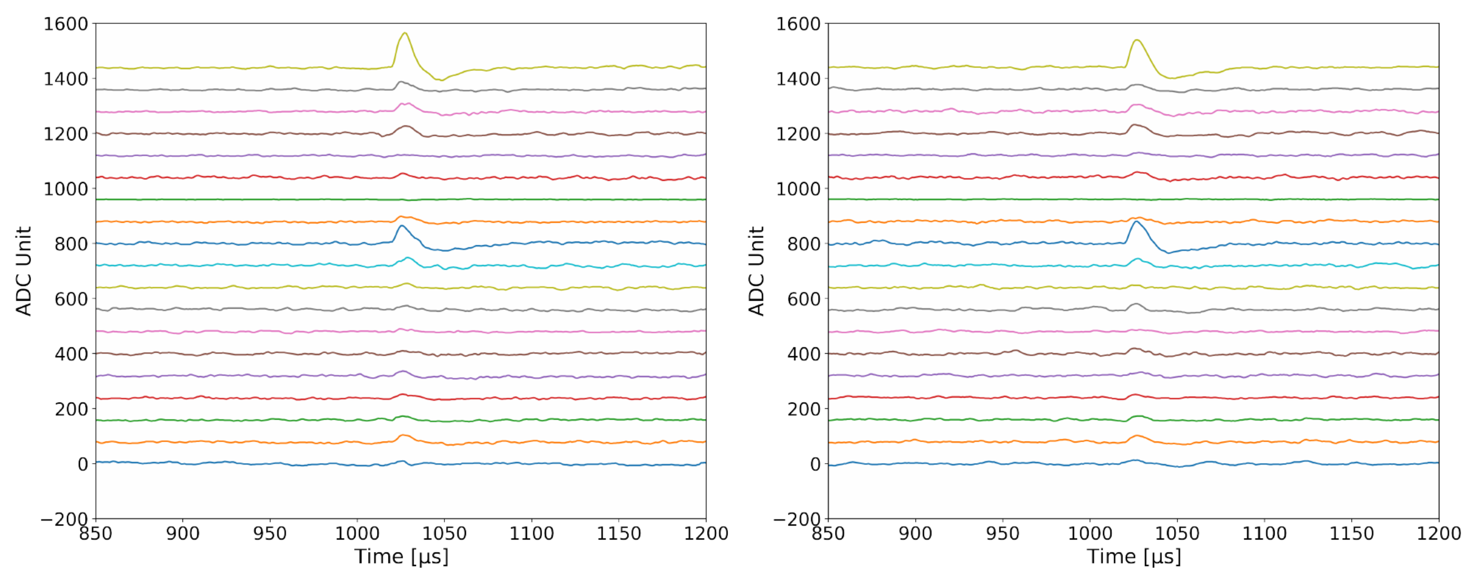
(left out the AI frames for now, small mass, and the WCSim geometry is noted as a guess)

A. Mastbaum (ANNIE internal)

EXO-200: Learning detector response from data

- EXO-200 trained a conditional generative model on calibration data to produce scintillation waveforms given energy and position.
- The model reproduces useful spatial-response features, including the pattern associated with inactive channels, although small residual signals remain.

Generated **Measured**



EXO-200 Collaboration, arXiv:2303.06311, Fig. 7

Evaluated using the standard reconstruction.

Response model that learns features of the real detector; generated events retain poorer reconstructed-energy resolution than data.

Sharing across experiments

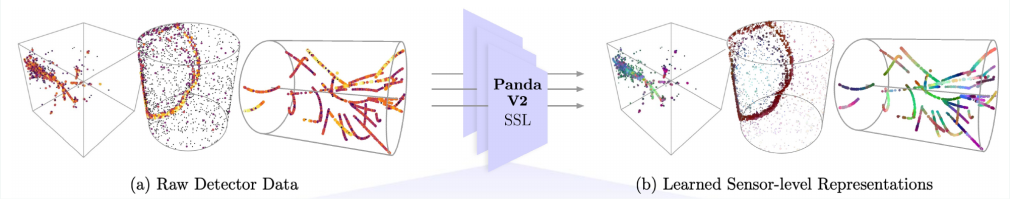
OmniLearned: pretrained weights

Collider pretraining → **MINERvA adaptation**

G. Krzmanc et al., arXiv:2604.12364

OmniLearned transfers collider-pretrained weights to reconstructed objects from simulated MINERvA events. At the same downstream compute, pretrained models outperform comparable models trained from scratch.

Panda V2: shared learning method



S. Young et al., arXiv:2609.00611, Fig. 1

Panda V2 instead applies a common sensor-level architecture and self-supervised objective, with independent pretraining for each detector modality.
Adaptation with 1,000 labeled events.

Water-Cherenkov performance remains below specialist reconstruction.

A practical question for a small experiment is what tools can we reuse, and what adaptation and validation must each experiment still do?

How to address the bottlenecks?

Bottleneck	Promising application	Success metric
Expert-intensive operations	Bounded control and monitoring	Stable data with less expert intervention
Expensive simulation or fits	Validated surrogates and generative response	Lower total cost with controlled error
Reconstruction and limited data sets	Pretrained representations	Better physics performance and affordable adaptation
Software and knowledge upkeep	Coding and source-grounded assistants	Less maintenance effort and reproducible workflows

These choices depend on the experiment and what people, data, and tools already available. Improvement to the key bottleneck tasks provide biggest gains.

More science that a small team can sustain

Design	Simulation	Construction	Operation	Analysis
	ANNIE · EXO-200		CRESST	KATRIN

- The examples in this talk span the experimental lifecycle we started with. They show concrete progress (albeit somewhat fragmented) in control, calculation, software, and detector-response modeling, with different levels of maturity.

- Experiments should choose the tasks that limit the science, reuse what can be shared, and plan for validation and integration across the experimental lifecycle.

- At same time, small experiments can help shape the community effort through problems that connect the experiment, its operation, and the analysis while connecting to broader support.

The metric of success is the additional science that a small collaboration can deliver and sustain with limited human and computing resources.

The collaboration effort is key for small experiments

Community Vision, arXiv:2602.17582

Building an AI-native Research Ecosystem for Experimental Particle Physics: A Community Vision

February 20, 2026 - Version 1.00

Four grand challenges apply to small experiments too!

Experimental design
Sensing and instrumentation
Autonomous experiments
From data to discovery

- Access to sustained AI and software expertise, including hands-on help adapting tools to an experiment.
- Shared/reusable tools that remain practical with limited training data, modest computing resources, and small teams.
- Support for validation and integration across the experimental lifecycle, including maintenance after the first demonstration.

Small experiments contribute concrete physics problems and realistic demonstrators, and should be involved in defining the common tools.

At this workshop… raise your hand if you do some of your science in a small experiment

Community Vision, arXiv:2602.17582

Building an AI-native Research Ecosystem for Experimental Particle Physics: A Community Vision

February 20, 2026 - Version 1.00

Four grand challenges apply to small experiments too!

Experimental design
Sensing and instrumentation
Autonomous experiments
From data to discovery

- Record the needs of small experiments in each relevant breakout report.
- Document the missing support, the science it would enable, and the personnel and computing resources needed to make progress.
- Identify a few candidate projects, with experiment and AI/software contacts, a clear validation target, and an estimate of the effort required.
- Agree on who will coordinate this work and a first follow-up date, with the aim of potentially developing joint funding proposals.