AI/ML for Particle Accelerators: A Look Backward and Forward

Auralee Linscott Edelen
International Particle Accelerator Conference 2026
May 22, 2026

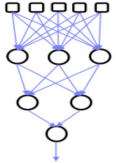
edelen@slac.stanford.edu

Artificial Intelligence (AI)

- *How to enable machines to exhibit aspects of "intelligence"*
- *knowledge, learning, planning, reasoning, perception*

Machine Learning (ML)

- *Use learned representations to complete tasks without being explicitly programmed*
- Tasks: Regression, Classification, Dimensionality Reduction, etc.

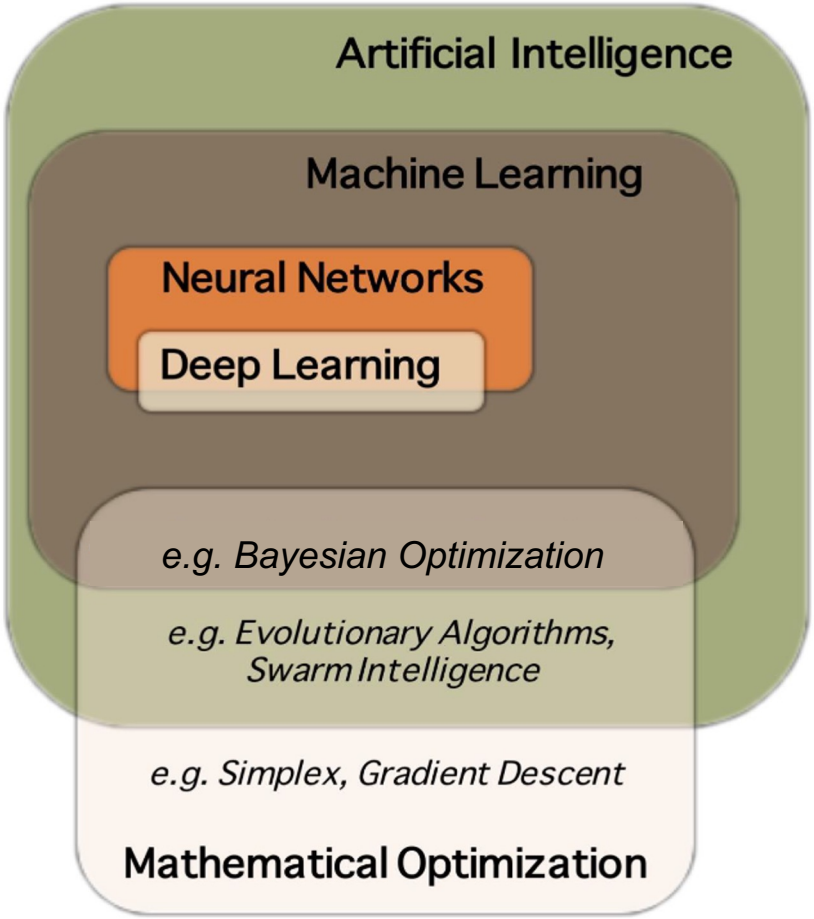


Neural Networks (NNs)

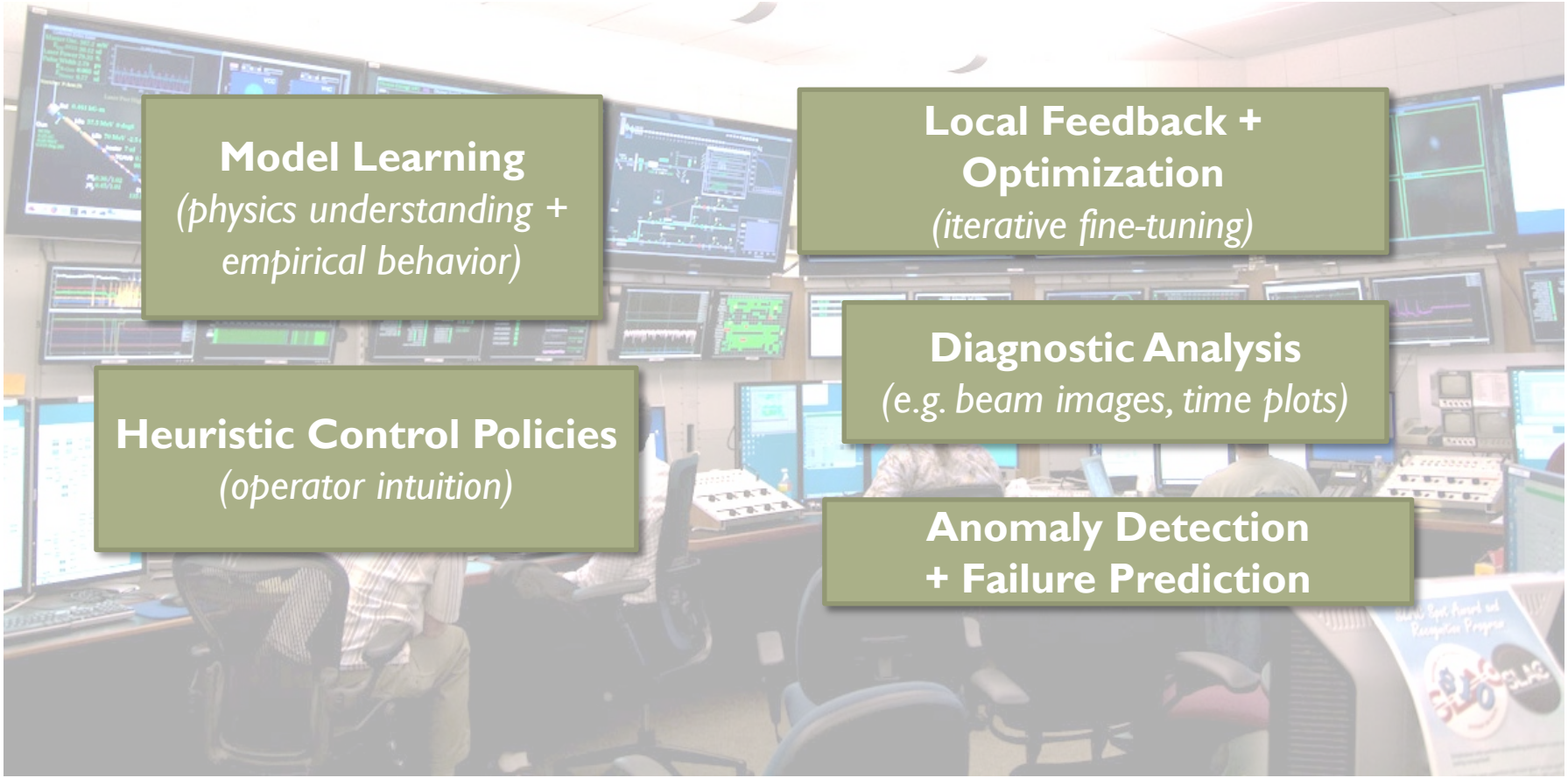
- *Class of ML structures that use many connected processing units to learn input/output maps (used to be called "connectionism")*

Deep Learning (DL)

- *Learning hierarchical representations*
- Right now, largely synonymous with deep (many-layered) NNs



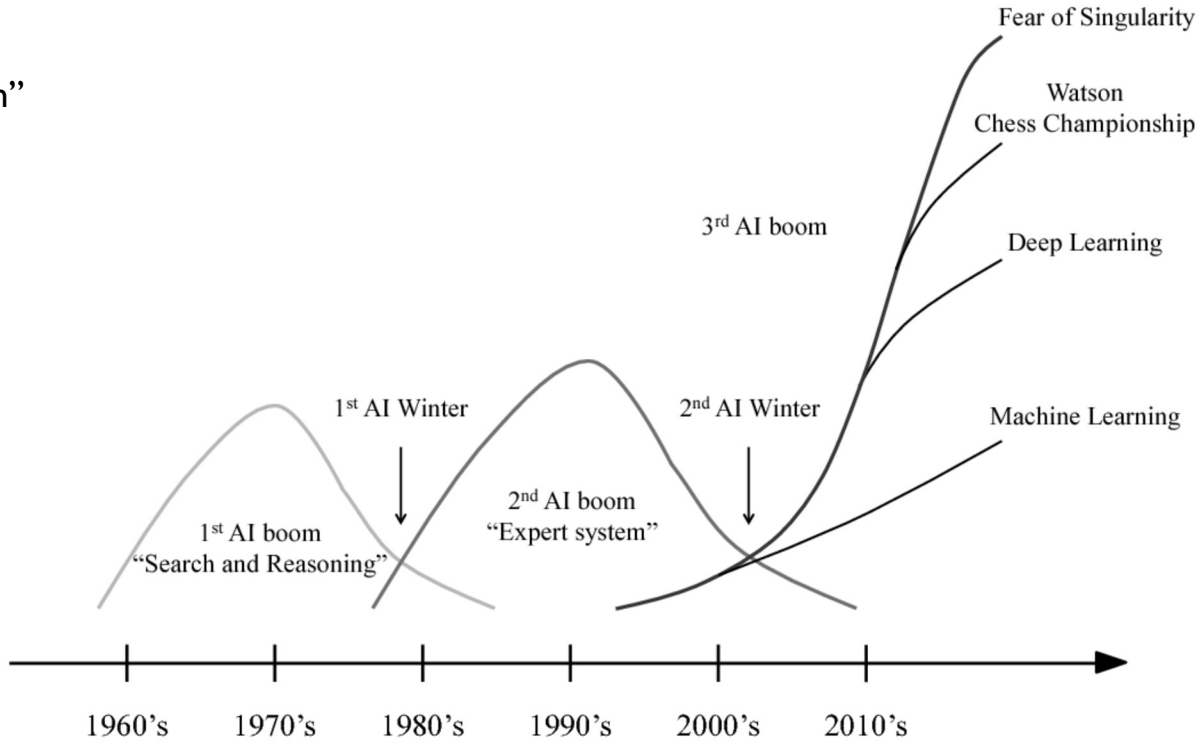

1058
05/20/2010 10:58:32



…many techniques in optimization, machine learning, classical control, etc for these etc.

AI Booms and Busts

- 1950s – 1960s: first neural nets, reasoning / search
- Mid 1980s – early 1990s: "connectionism" neural networks, expert systems
- 1997: Deep Blue beats Gary Kasparov
- 2006: Deep learning breakthroughs at University of Toronto
- 2011: IBM Watson wins Jeopardy
- 2015: Deep learning on GPUs
- 2016: Alpha-Go deep learning software beats best players
- 2022: Tokomak control with reinforcement learning



(Source of picture : 978-4040800202, p.61, Fig. 3)

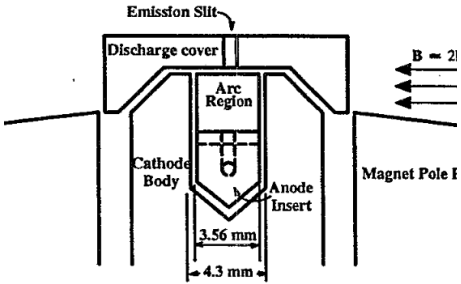
1990 1995 2000 2005 2010 2015 2020 2025

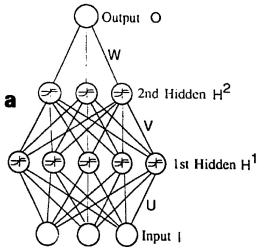
early exploration

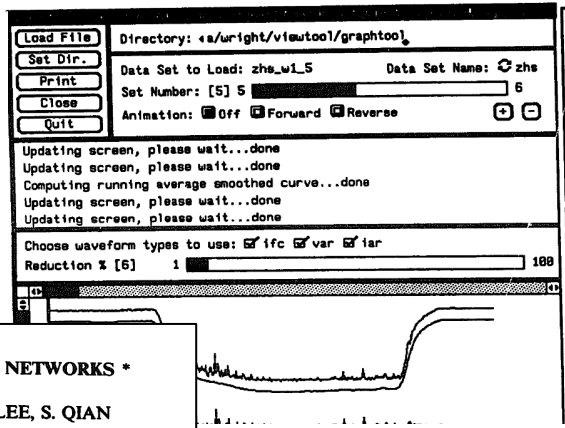
tentative re-exploration

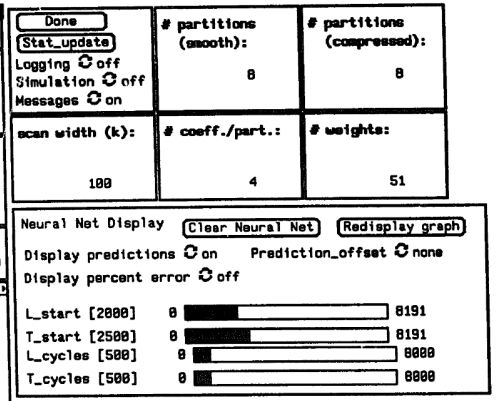
grassroots momentum

serious attention









CONTROL OF A NEGATIVE-ION ACCELERATOR SOURCE USING NEURAL NETWORKS *

J.A. HOWELL, C.W. BARNES, S.K. BROWN, G.W. FLAKE, R.D. JONES, Y.C. LEE, S. QIAN and R.M. WRIGHT

Los Alamos National Laboratory, Los Alamos, NM 87545, USA

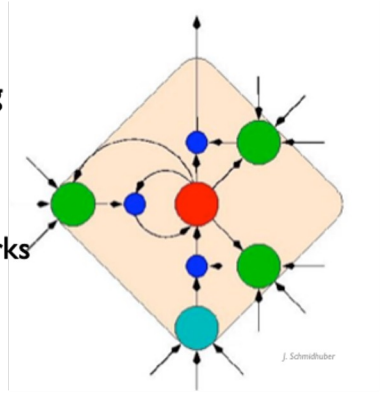
What changed in the 2010s…

Increased computational capability enables more complicated NN architectures and faster training + larger data sets

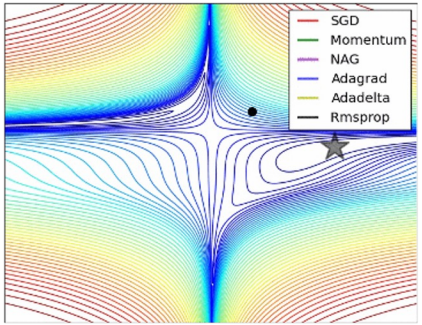
GPUs

Accessibility of HPC clusters

Neural network architectures and training paradigms,
such as long short term memory (LSTM) networks, generative adversarial networks (GANs)

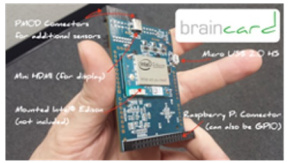


Can **easily share** large data sets, code, and computing setups
(e.g. via cloud computing services)



Better **theoretical understanding** of NNs and improved **optimization methods**

specialized hardware: neuromorphic chips, TPUs



Applications have driven a lot of advancement (both algorithmic and practical/heuristic)

1990 1995 2000 2005 2010 2015 2020 2025

early exploration

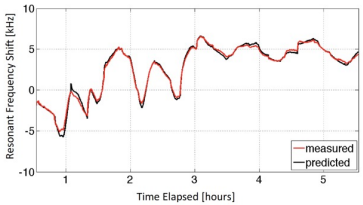
tentative re-exploration

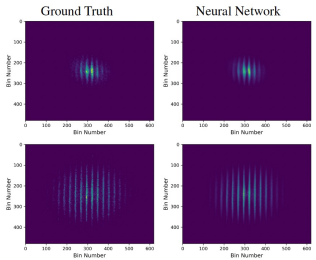
grassroots momentum

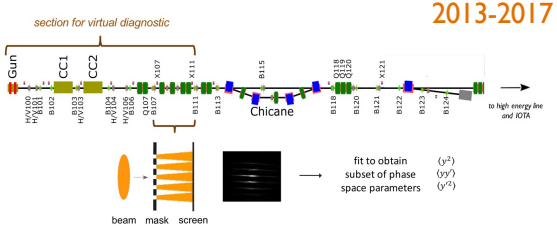
serious attention

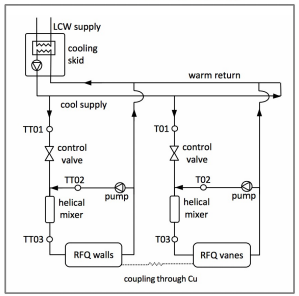
NN surrogates, virtual diagnostics, and sim2real

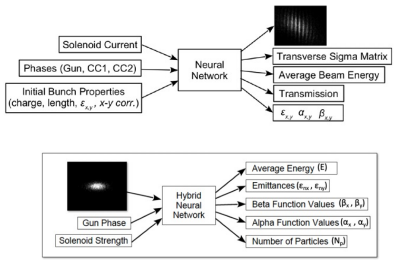
2013-2017

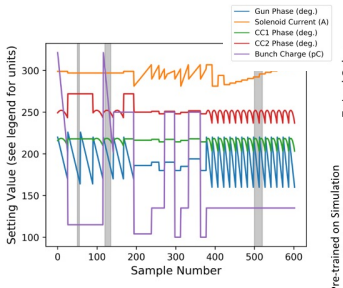


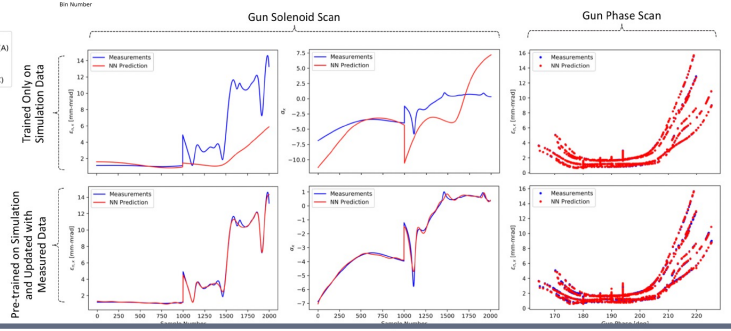












A. Edelen dissertation and IPAC18

1990 1995 2000 2005 2010 2015 2020 2025

early exploration

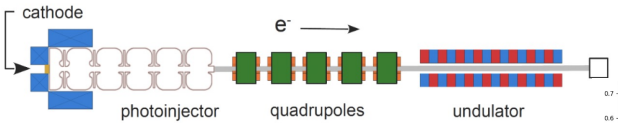
tentative re-exploration

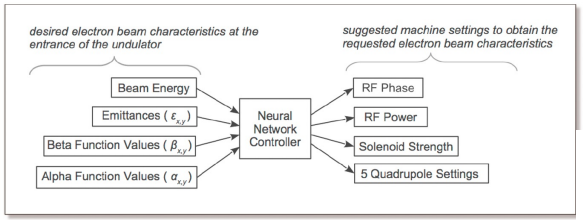
grassroots momentum

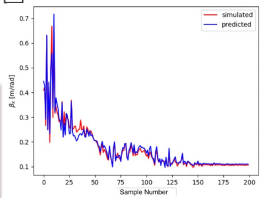
serious attention

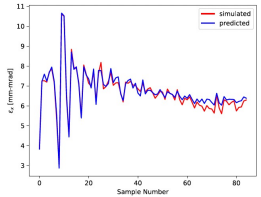
2013-2017

NN control policies / reinforcement learning

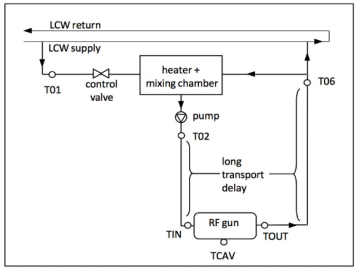


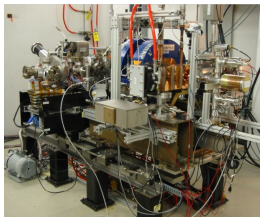




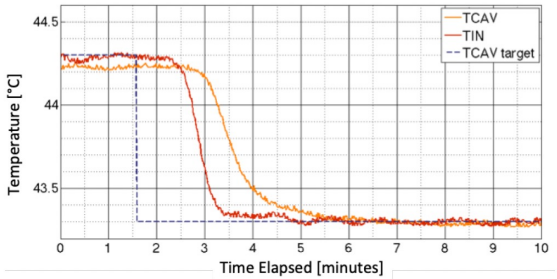


Sim study: map desired beam to settings
Solution in 1 iteration (vs. 150 with simplex)





Model predictive control + time-dependent NN model



5x faster settling + minimize overshoot
→ lower wasted RF power → save $$

A. Edelen dissertation and IPAC18

1990 1995 2000 2005 2010 2015 2020 2025

early exploration

tentative re-exploration

grassroots momentum

serious attention

BERKELEY LAB SLAC NATIONAL ACCELERATOR LABORATORY THE UNIVERSITY OF CHICAGO

USPAS: Optimization and Machine Learning for Particle Accelerators

2019

"Opportunities in Machine Learning for Particle Accelerators" https://arxiv.org/abs/1811.03172

1990 1995 2000 2005 2010 2015 2020 2025

early exploration

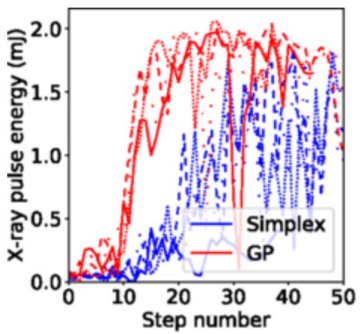
tentative re-exploration

grassroots momentum

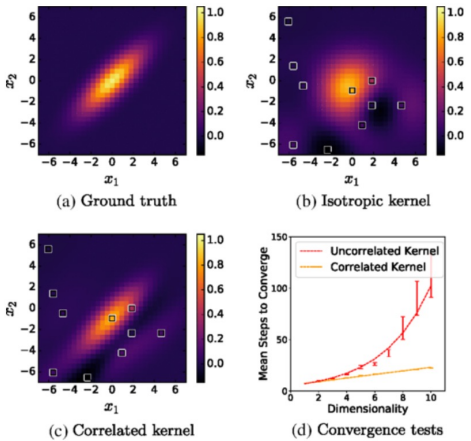
serious attention

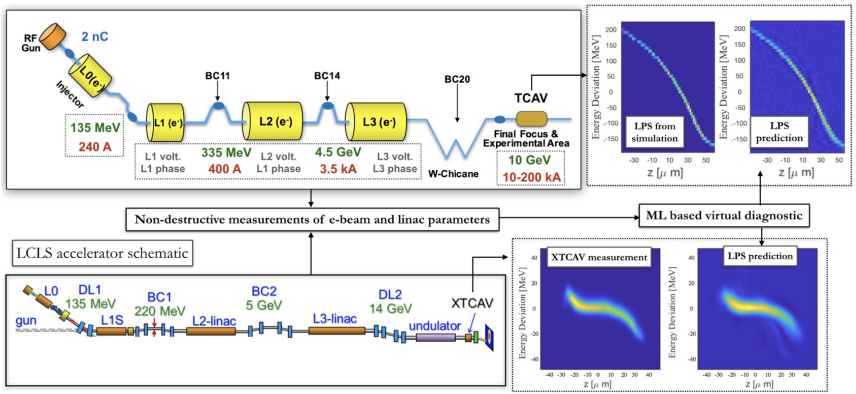
2018-2022

Physics-informed Bayesian optimization @ LCLS 2016-2020



(a) Live 12 quadrupole optimization





ML virtual diagnostic @ LCLS/FACET-II 2018

DOI: https://doi.org/10.1103/PhysRevLett.124.124801, DOI: https://doi.org/10.1103/PhysRevAccelBeams.21.112802

1990 1995 2000 2005 2010 2015 2020 2025

early exploration

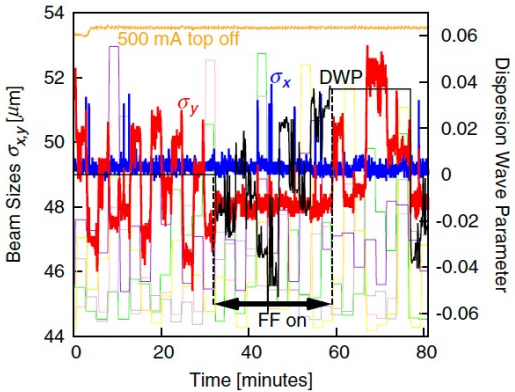
tentative re-exploration

grassroots momentum

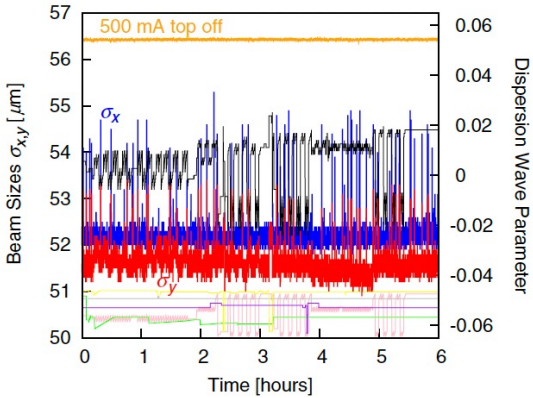
serious attention

2018-2022

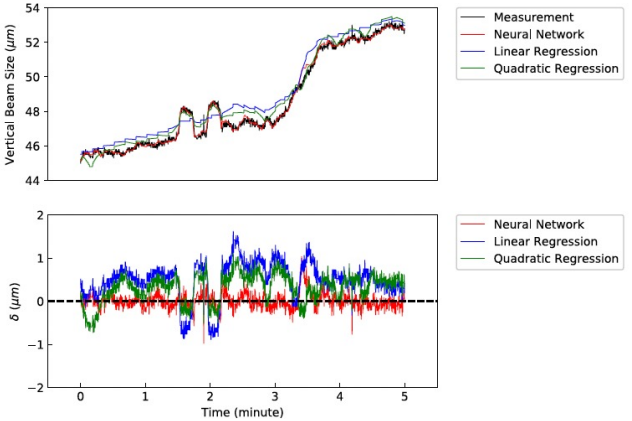
Insertion device compensation @ALS 2018



Uncorrected



NN feedforward



DOI: 10.1103/PhysRevLett.123.194801

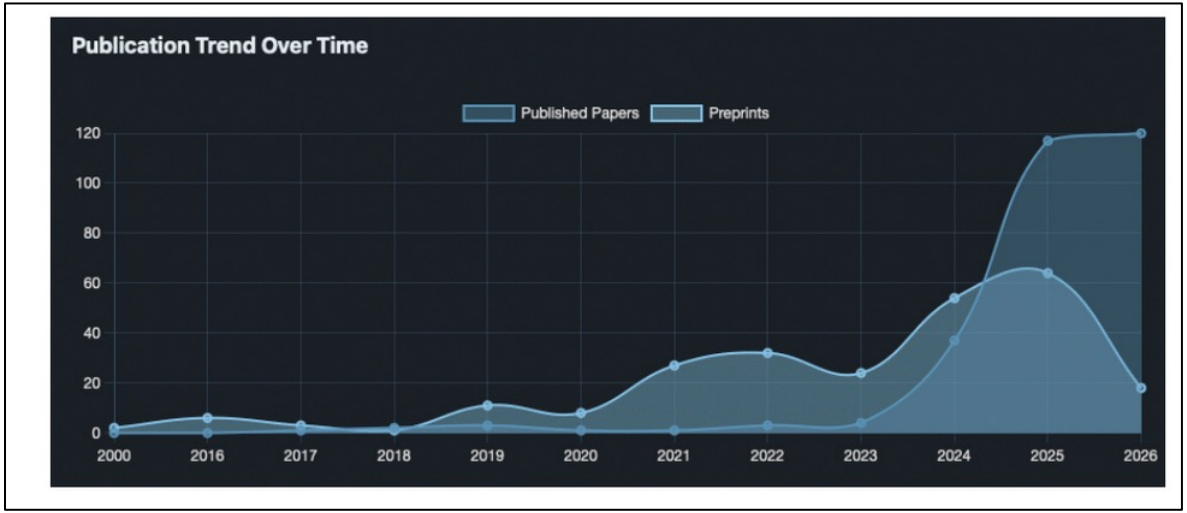
1990 1995 2000 2005 2010 2015 2020 2025

early exploration

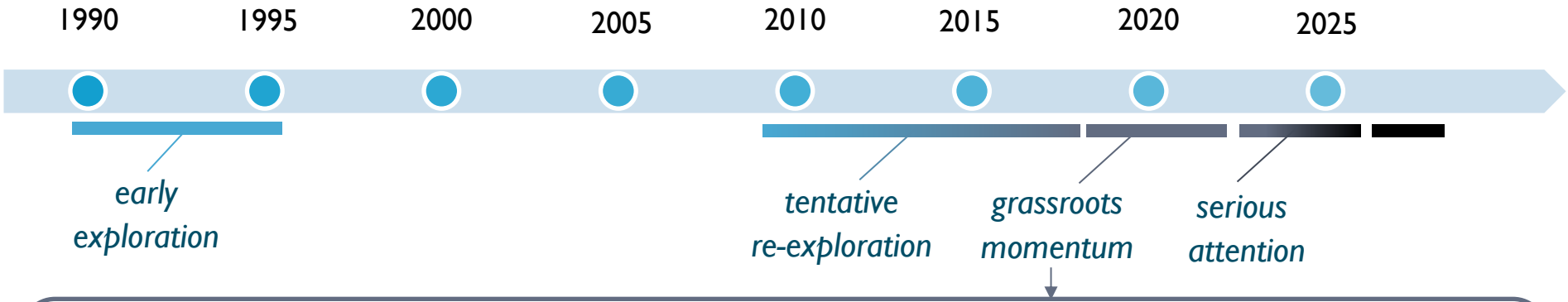
tentative re-exploration

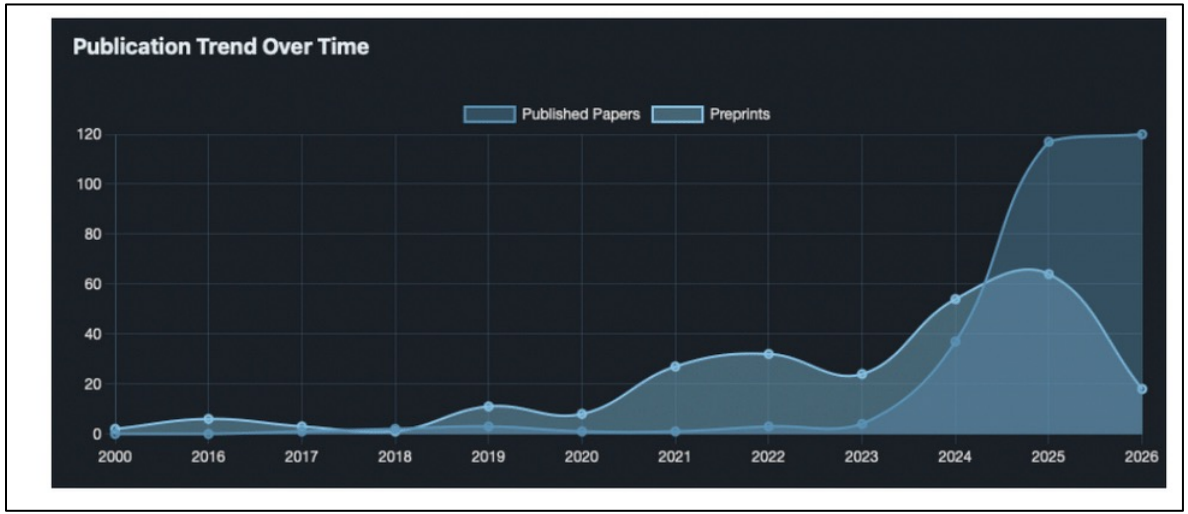
grassroots momentum

serious attention



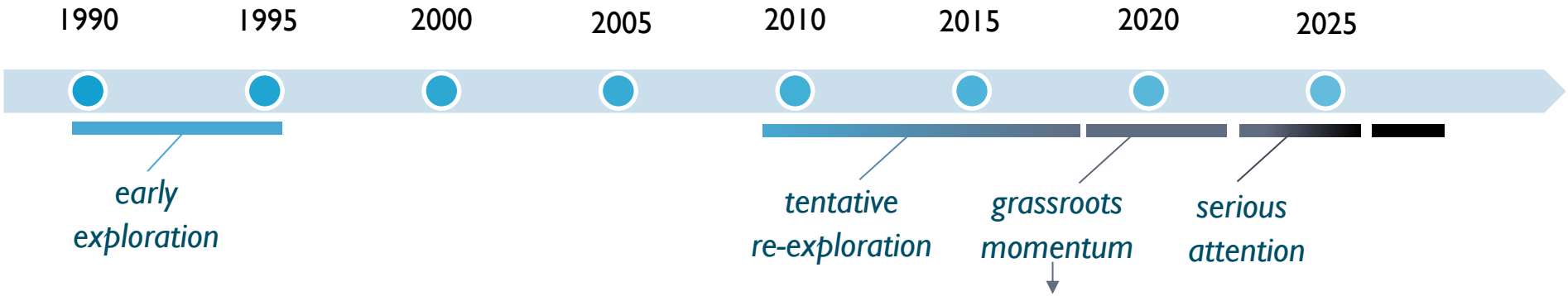
A. Ghribi https://aghribi.github.io/acc-ml-living-review/



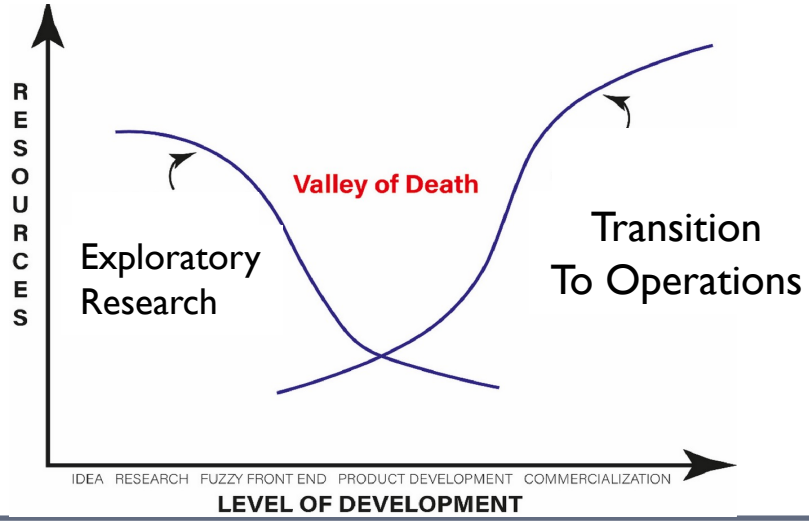


Paper to production can be very hard

A. Ghribi https://aghribi.github.io/acc-ml-living-review/



Biggest blockers not AI/ML maturity but rather infrastructure and continuity of effort…



Source: https://www.sciencedirect.com/science/article/pii/S2666188822000119

siloed funding streams and projects

1990 1995 2000 2005 2010 2015 2020 2025

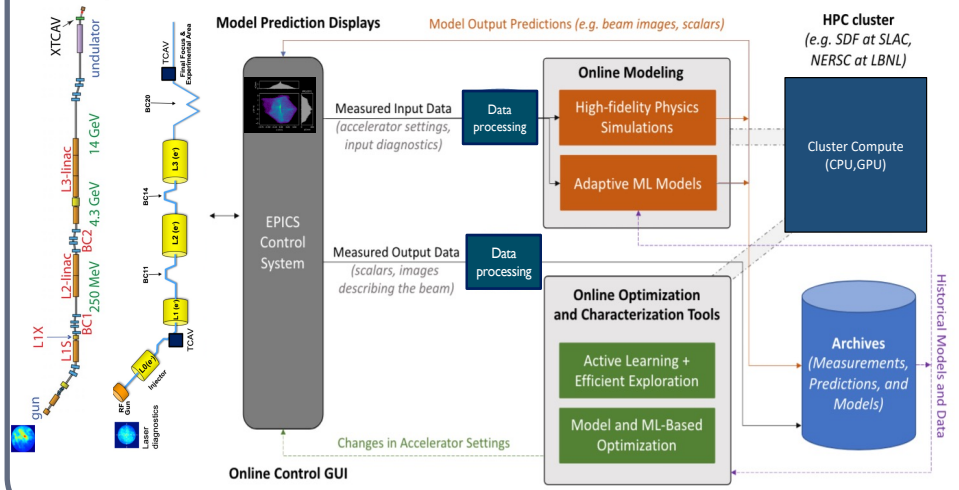
early exploration

tentative re-exploration

grassroots momentum

serious attention

local facility infrastructure



modular/transferrable tools

1990 1995 2000 2005 2010 2015 2020 2025

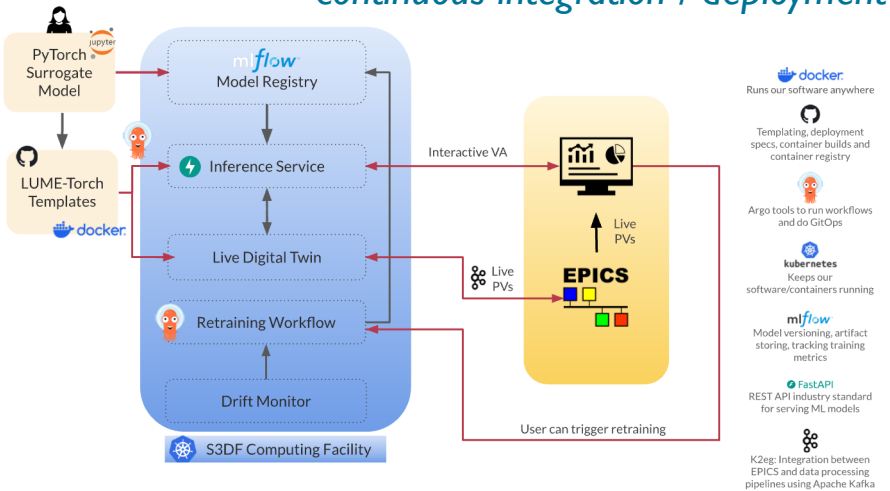
early exploration

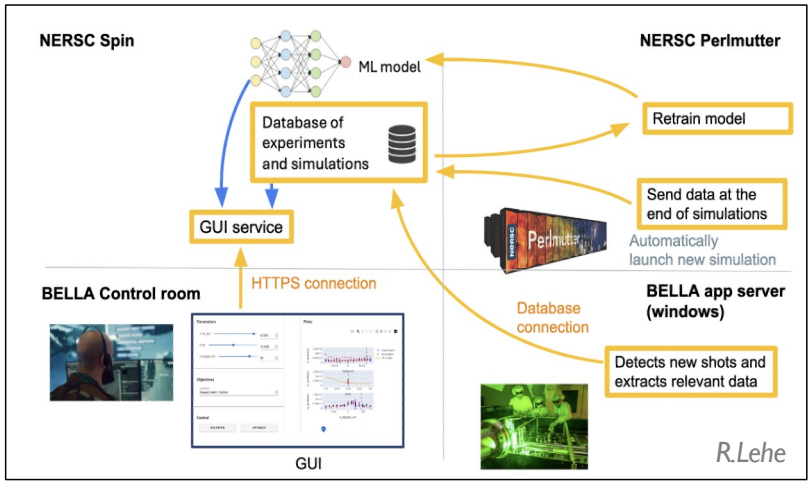
tentative re-exploration

grassroots momentum

serious attention

continuous integration / deployment and live model updating and retraining





github.com/slaclab/inference-service, github.com/slaclab/lume-model-deployment-template

1990 1995 2000 2005 2010 2015 2020 2025

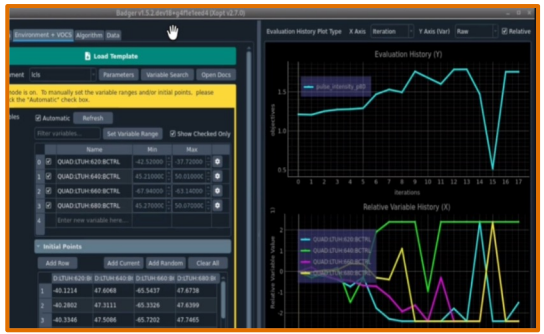
early exploration

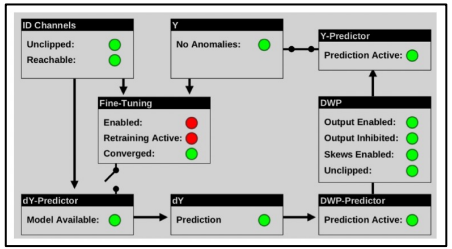
tentative re-exploration

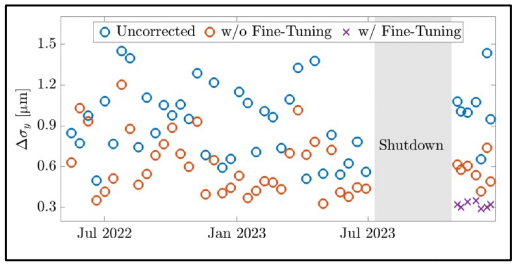
grassroots momentum

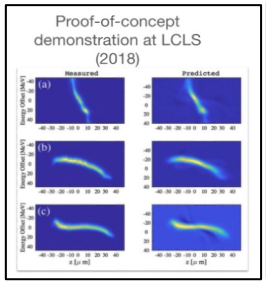
serious attention

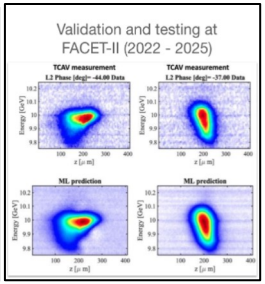
Now a fair number of routine operational deployments worldwide

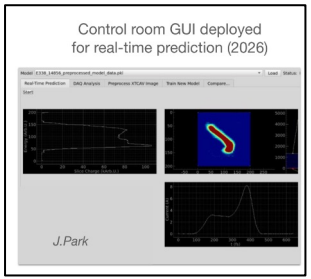












Virtual XTCAV @ FACET-II

Xopt *BO and advanced variants*

FF NN correction + continual fine-tuning @ALS

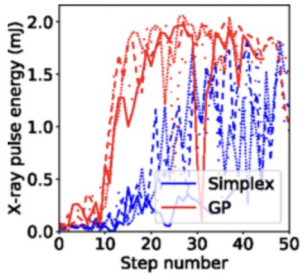
DOI: https://doi.org/10.1103/PhysRevAccelBeams.27.084801

DOI: 10.1103/PhysRevAccelBeams.27.074602

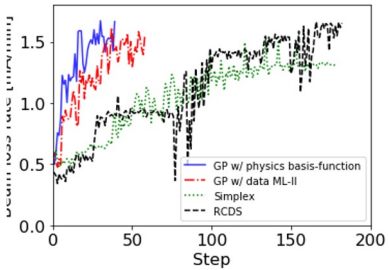
C. Emma, J. Park

Many examples with Bayesian Optimization and algorithmic improvements

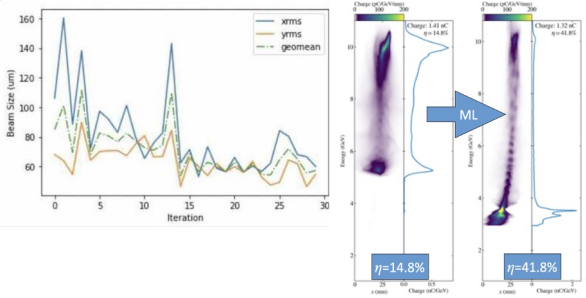
FEL pulse energy at LCLS



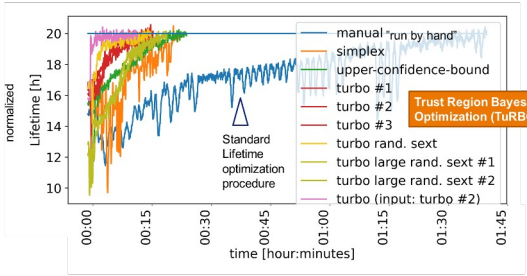
Loss rate at SPEAR3



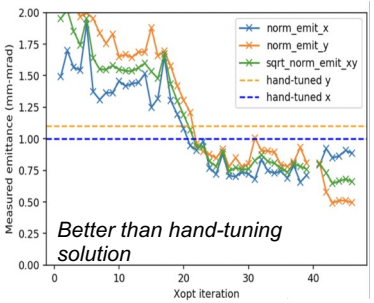
Sextupole tuning at FACET-II
to improve efficiency of acceleration in plasma



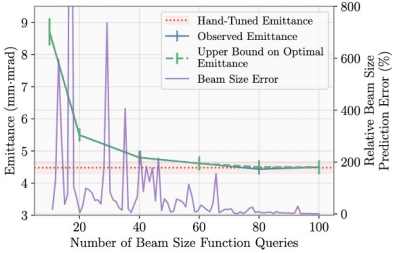
50x faster tuning and best observed lifetime at ESRF



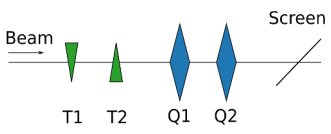
Emittance @ LCLS-II injector

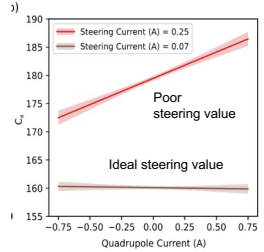


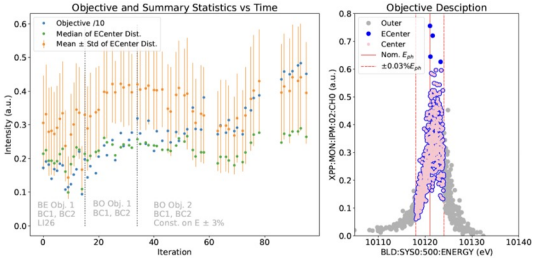
20x faster emittance tuning (BAX)



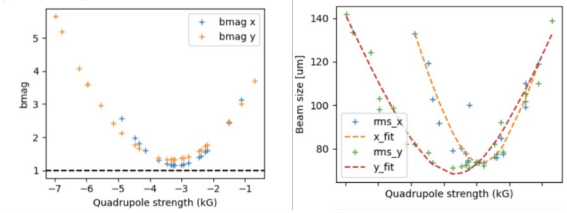
Automated beam alignment at AWA
- 20-30 minutes by hand
- 5 minutes with BAX





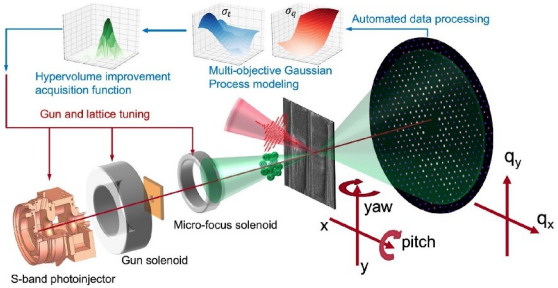


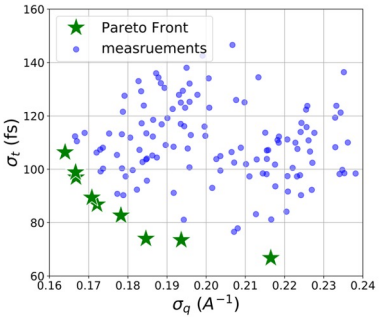
Automatic emittance measurement



Tuning on monochrometer signal

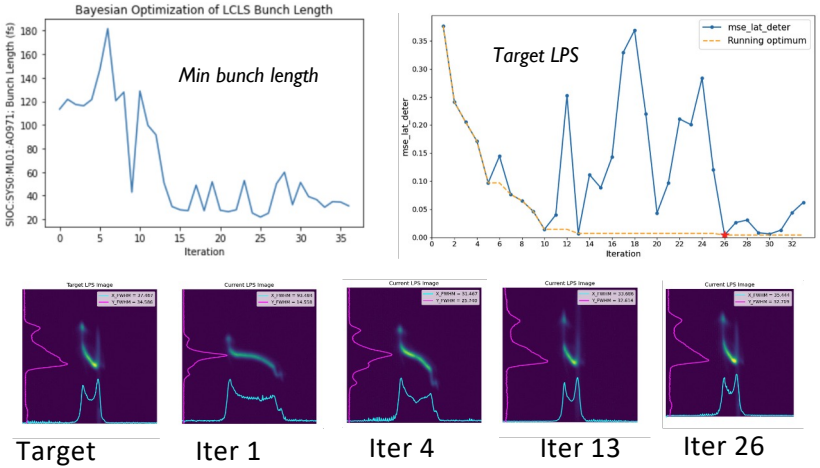
Multi-objective BO → experimental Pareto front



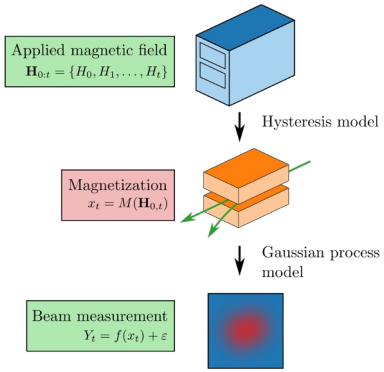


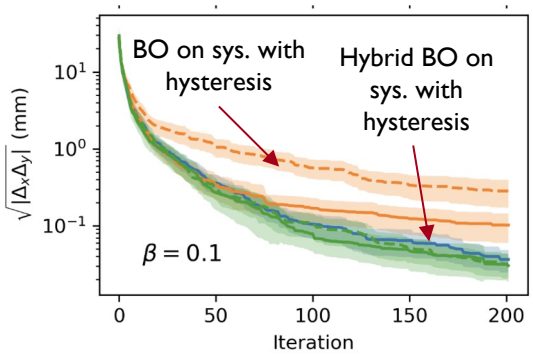
F. Ji, et al., Nat. Commun. 15, 4726 (2024)

Longitudinal phase space tuning



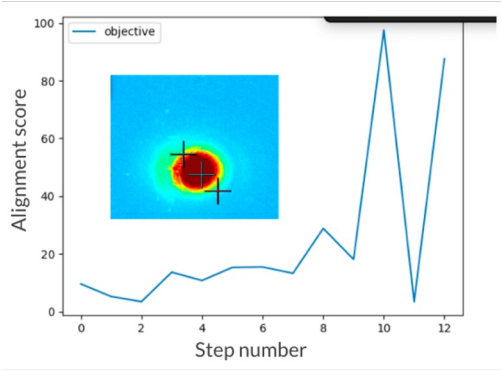
Hysteresis-aware magnet tuning





*Roussel et. al. PRL , **2022***

Photon beamline alignment at MFX



1990 1995 2000 2005 2010 2015 2020 2025

early exploration

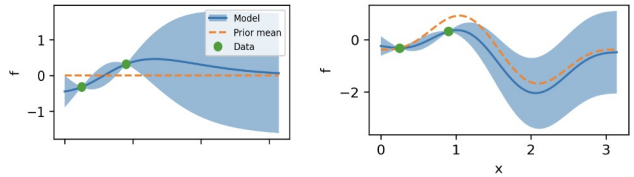
tentative re-exploration

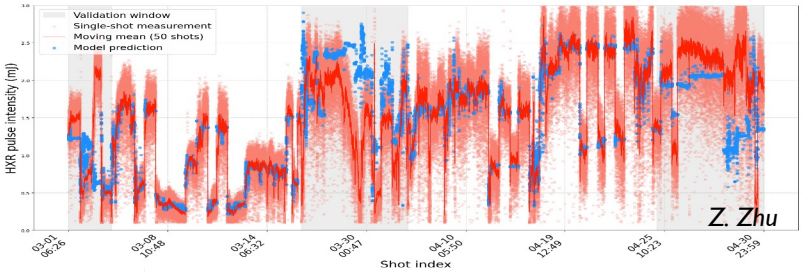
grassroots momentum

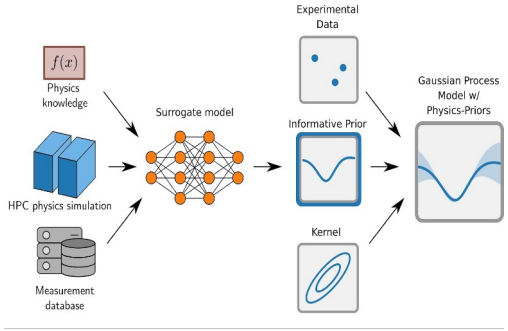
serious attention

Moving toward multi-component solutions …

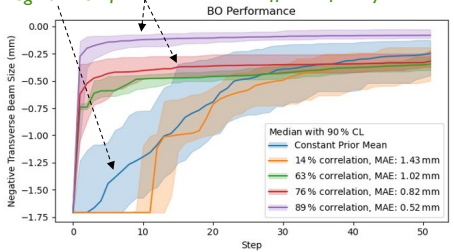
(e.g. adaptive digital twin → system model prior for BO)







regular BO prior means with different fidelity



T. Boltz https://www.nature.com/articles/s41598-025-95297-z

1990 1995 2000 2005 2010 2015 2020 2025

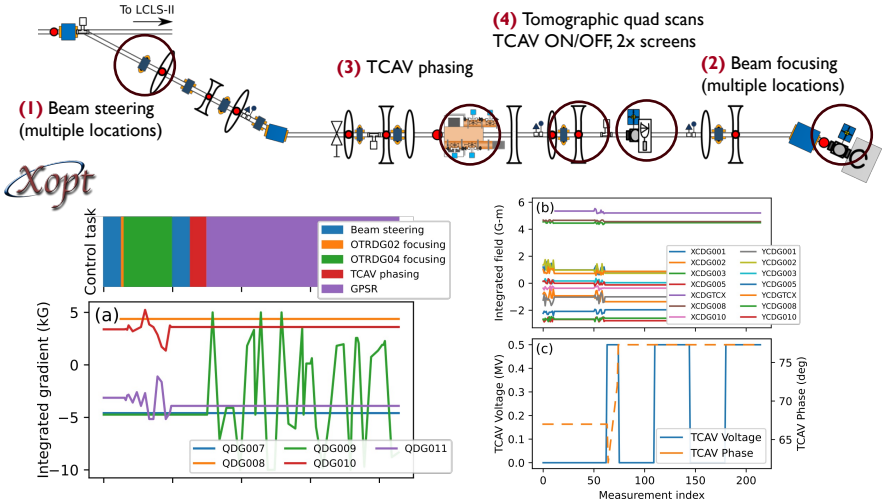
early exploration

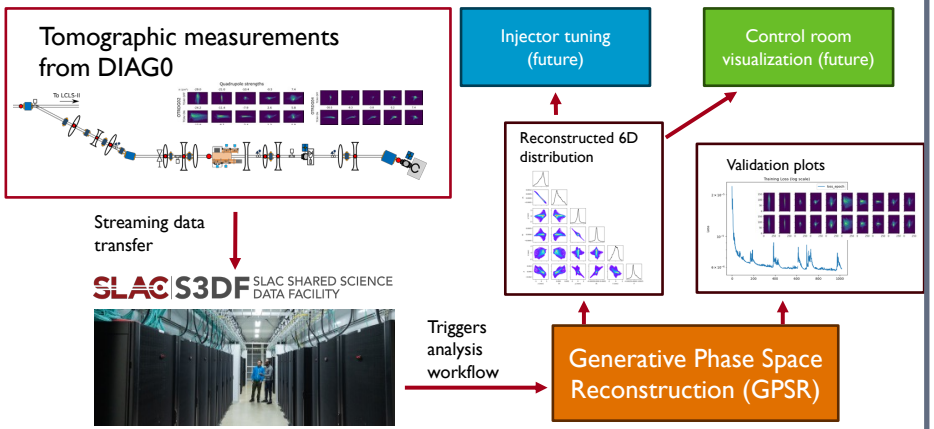
tentative re-exploration

grassroots momentum

serious attention

Moving from individual / isolated tasks toward autonomous chains of them …

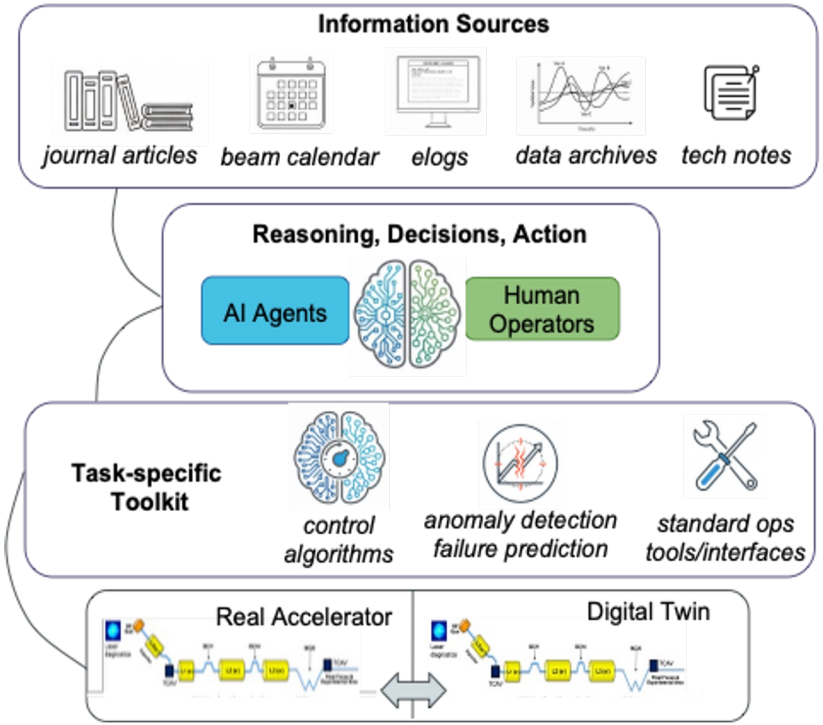




~5 minutes for entire process

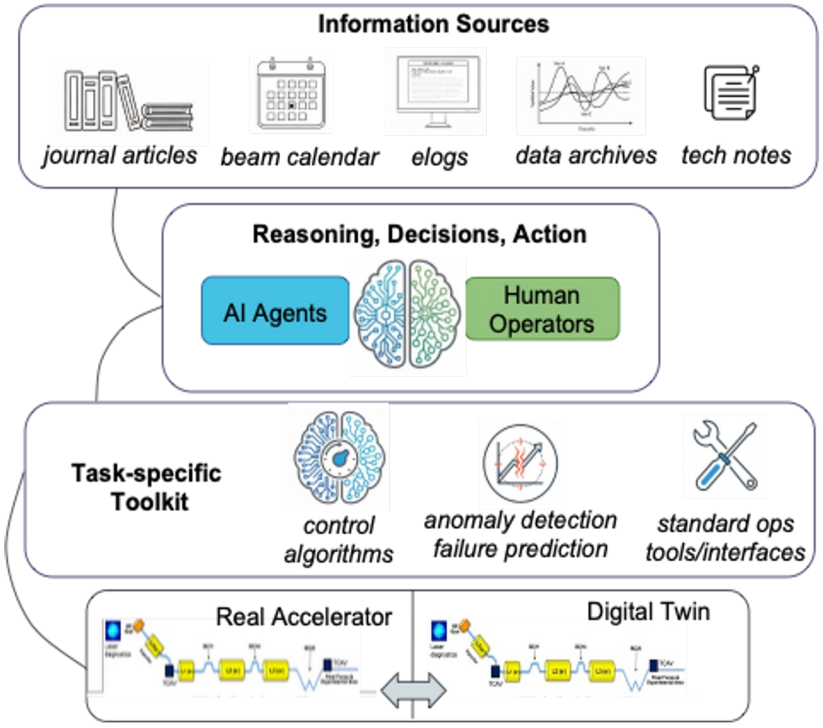
R. Roussel et al., https://arxiv.org/abs/2604.20125

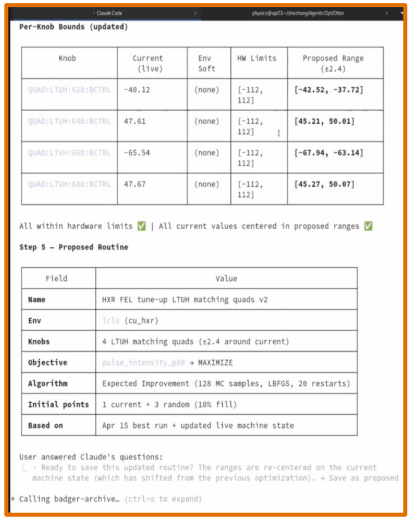
Agentic AI emerges …



Agentic AI emerges …

Osprey demo @LCLS





Per-Knob Bounds (updated)

Knob	Current (live)	Env Soft	HW Limits	Proposed Range (±2.4)
QUAD:LTUH:620:BCTRL	-40.12	(none)	[-112, 112]	[-42.52, -37.72]
QUAD:LTUH:640:BCTRL	47.61	(none)	[-112, 112]	[45.21, 50.01]
QUAD:LTUH:660:BCTRL	-65.54	(none)	[-112, 112]	[-67.94, -63.14]
QUAD:LTUH:680:BCTRL	47.67	(none)	[-112, 112]	[45.27, 50.07]

All within hardware limits ✅ | All current values centered in proposed ranges ✅

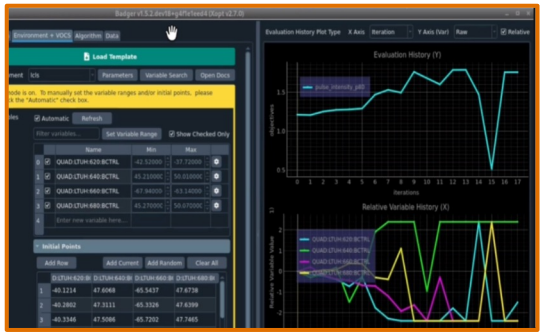
Step 5 — Proposed Routine

Field	Value
Name	HXR FEL tune-up LTUH matching quads v2
Env	lcls (cu_hxr)
Knobs	4 LTUH matching quads (±2.4 around current)
Objective	pulse_intensity_p80 → MAXIMIZE
Algorithm	Expected Improvement (128 MC samples, LBFGS, 20 restarts)
Initial points	1 current + 3 random (10% fill)
Based on	Apr 15 best run + updated live machine state

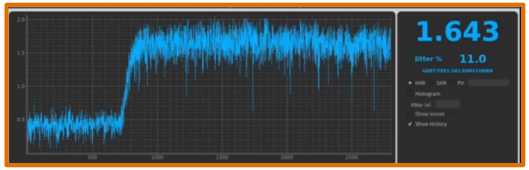
User answered Claude's questions:
· Ready to save this updated routine? The ranges are re-centered on the current machine state (which has shifted from the previous optimization). → Save as proposed

· Calling badger-archive… (ctrl+o to expand)

New optimization routine



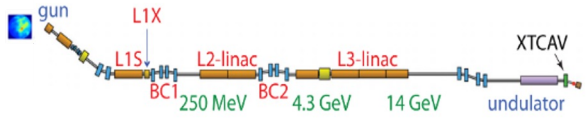
Automatic execution on machine via Badger tuning tool



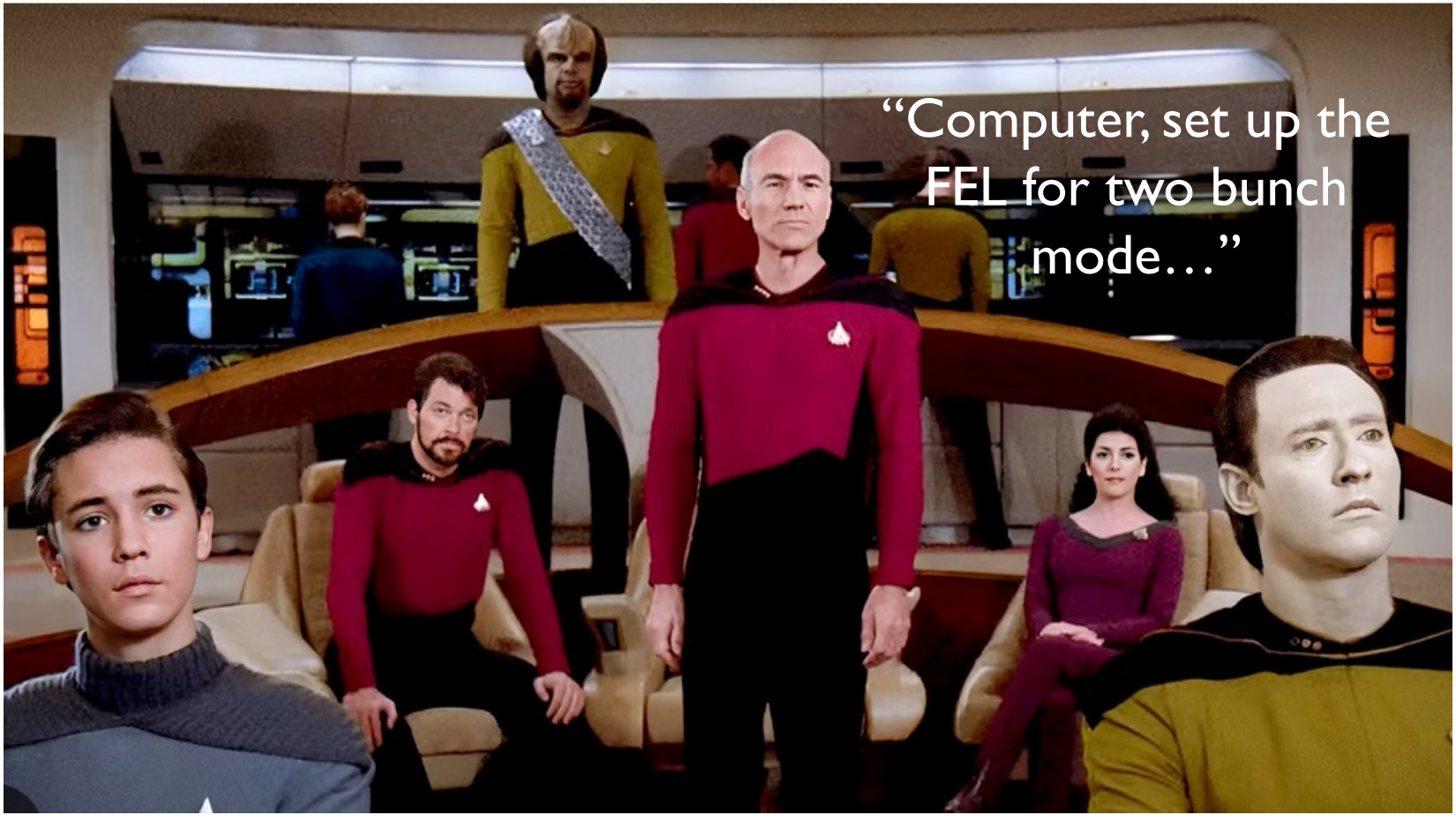
Rapid improvement in FEL performance

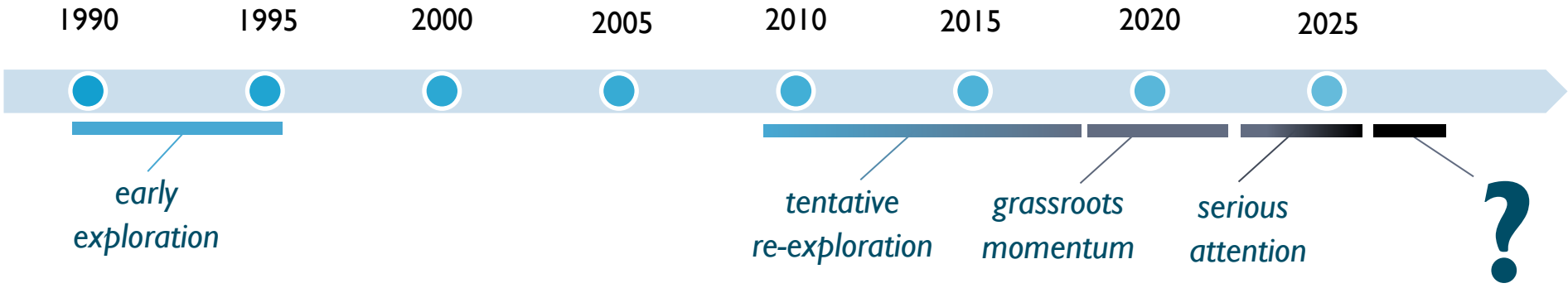
Badger run history

natural language description of tuning task



Prompt to completion: ~ 10 minutes


“Computer, set up the FEL for two bunch mode…”


1990
1995
2000
2005
2010
2015
2020
2025
early exploration
tentative re-exploration
grassroots momentum
serious attention
?

We are still holding ourselves back in serious ways…

A field fragmented at every level

Data **siloed** at facility, subsystem, or experiment level

Models developed for **single use cases** - rarely transferred

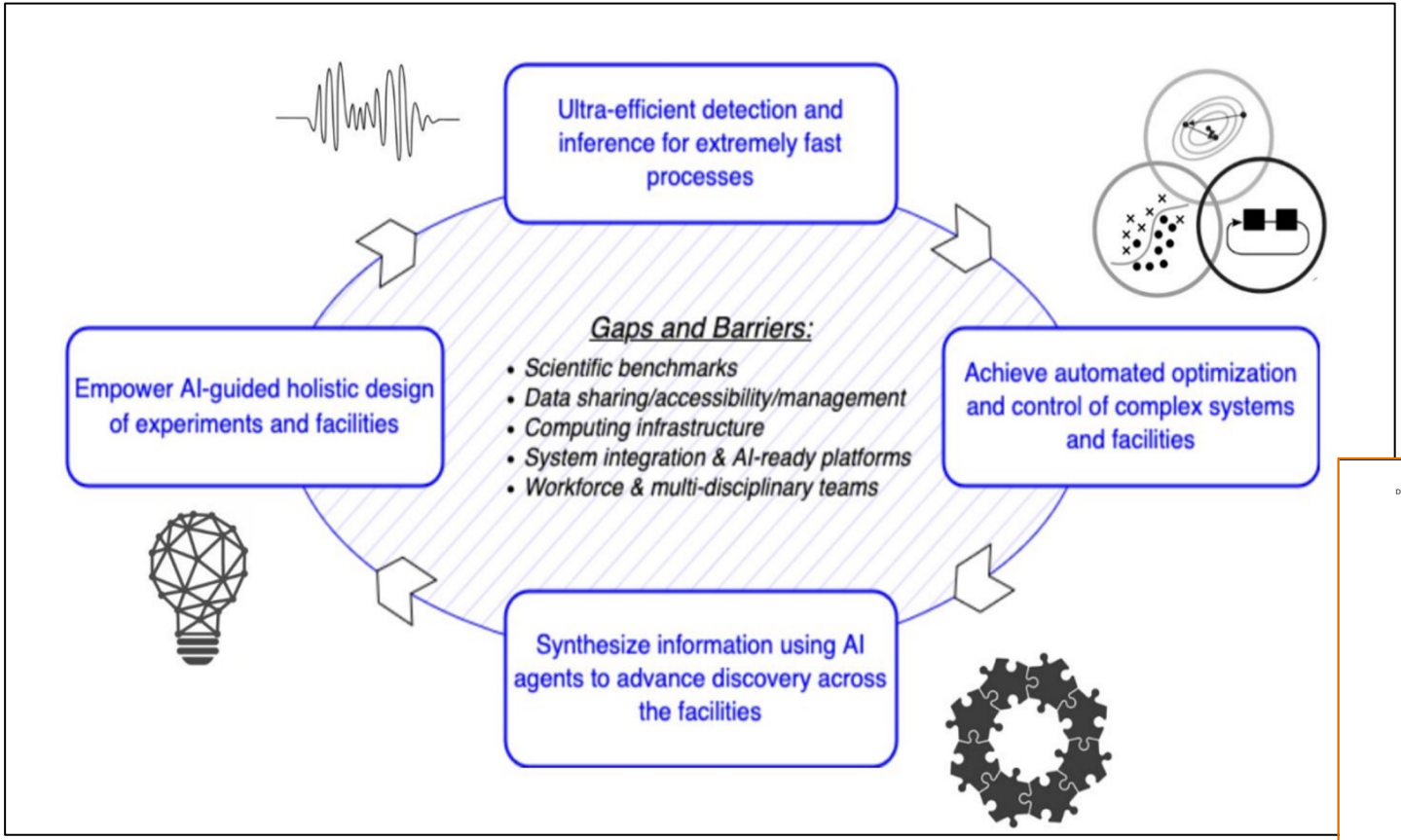
No shared **data standards**, no **benchmarks**

AI expertise in **tiny isolated groups**

➡ *These are the problems ARTIFACT was built to solve*

Projects end, **tools disappear**

A. Ghribi



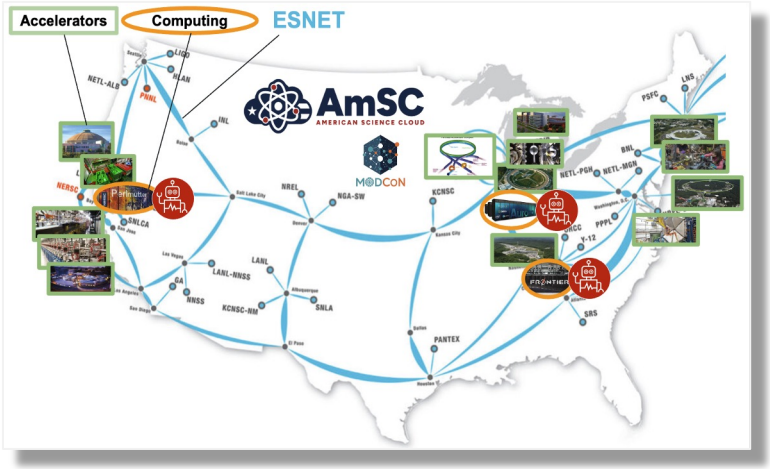
Draft

DOE-SC Roundtables:
Transformational Science Enabled by Artificial Intelligence: User Facility Science and Operations

March 19, 2025

DRAFT

DRAFT



MOAT

DOE-wide, highly-coordinated initiatives

- Ambitious vision to accelerate science and engineering
- Aim to leverage the full capabilities and talent of DOE

Follows years of community-driven info. gathering and strategizing

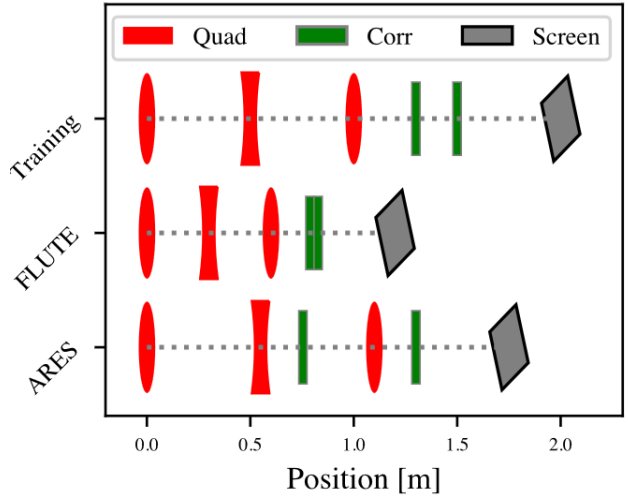
Genesis Mission

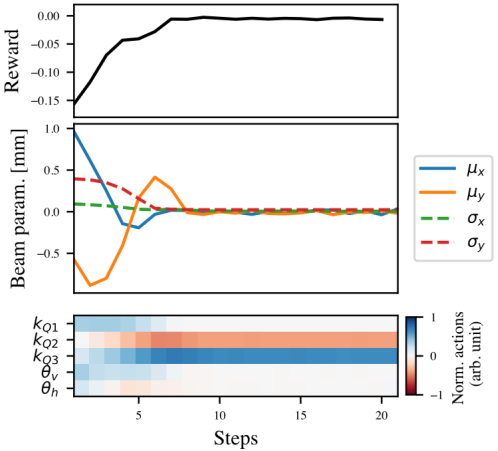
Direct cross-facility algorithm use of trained models/controllers is not as "crazy" as it sounds

14th International Particle Accelerator Conference,Venice, Italy JACoW Publishing
ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL029

BEAM TRAJECTORY CONTROL WITH LATTICE-AGNOSTIC REINFORCEMENT LEARNING

C. Xu*, E. Bründermann, A.-S. Müller, A. Santamaria Garcia
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
J. Kaiser, A. Eichler
Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany

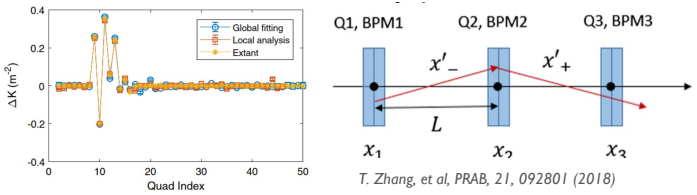




C. Xu et al, IPAC23 https://inspirehep.net/files/f031fdaacc0a5131f8bc83c9d9586caf

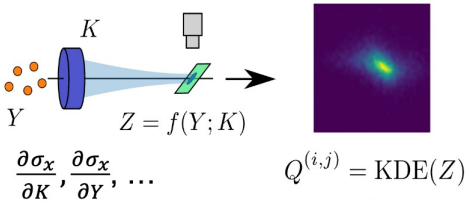
Things are not as different as they seem: ML as another accelerator physicist's tool

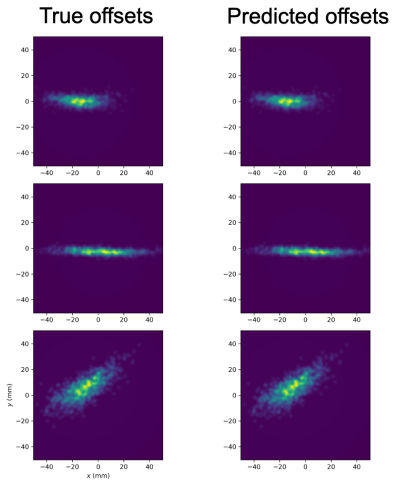
Classical Analytic and Brute Force Methods



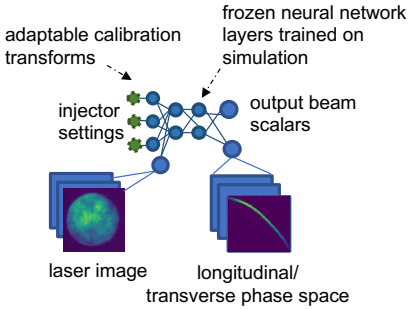
T. Zhang, et al, PRAB, 21, 092801 (2018)

Differentiable Physics



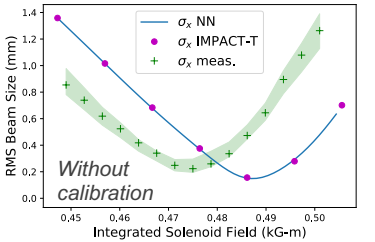


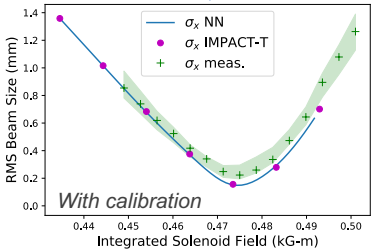
ML model trained on simulation
→ transfer learning to measured data

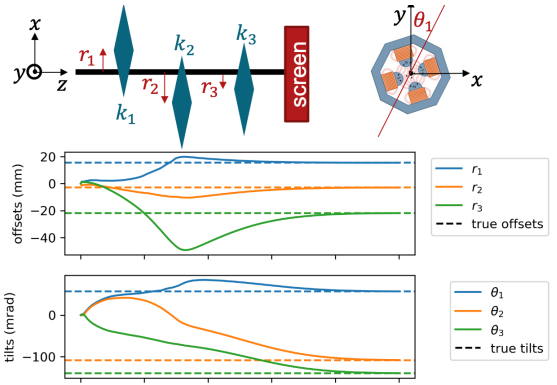


Inputs
Laser radius
Laser spot sizes
Pulse length
Charge
Solenoid
L0A phase
L0B phase
SQ quad
CQ quad
6 matching quads

Outputs
Beam size (x,y)
Emittance (x,y)
Bunch length

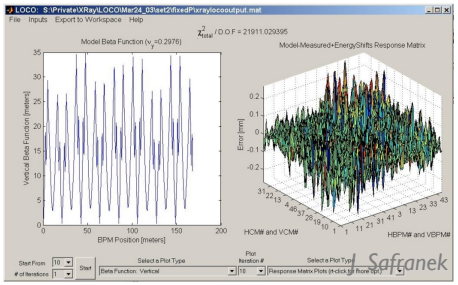




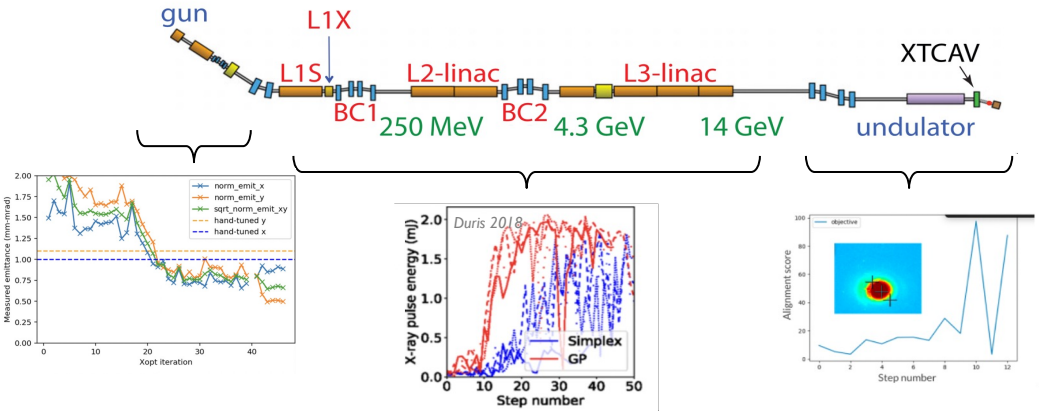


JP Gonzalez-Aguilera DOI: 10.18429/JACoW-IPAC2023-WEPA065

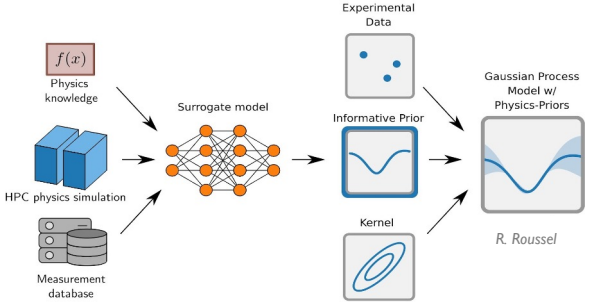
Things are not as different as they seem: ML as another accelerator physicist's tool



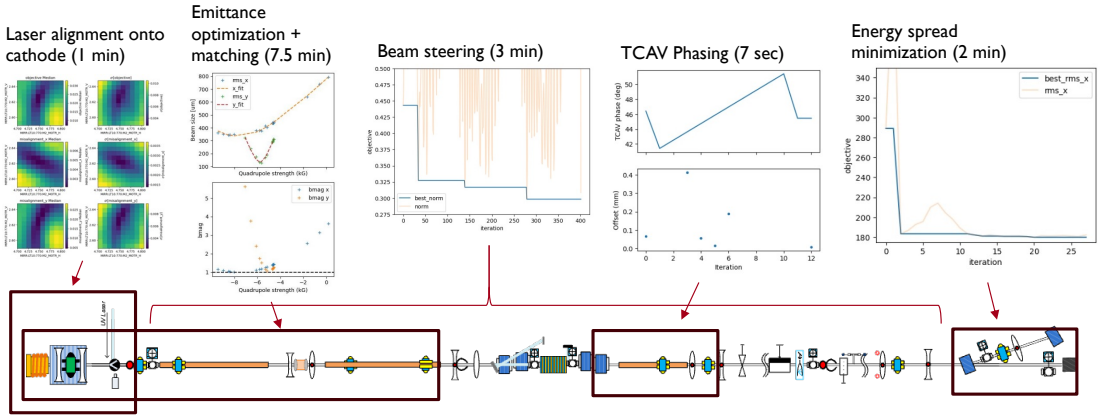
classical optics fitting + feedback



sub-system optimization/control with on-the-fly learning in BO

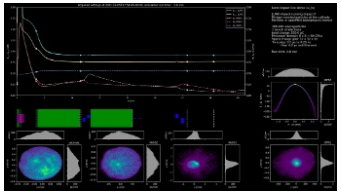


stacking algorithms:
system model + ML-based tuning



chaining together tasks: sequencer vs. agentic workflow

Things are not as different as they seem: ML as another accelerator physicist's tool



Typical Control Room Apps

Additional Visualization

High Level Output (e.g. EPICS PV)

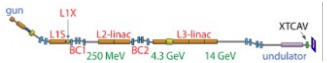
Physics Process Emulators (phys. sims, ML)

beam dynamics, cooling systems, user beamlines, etc
range of fidelities and comprehensiveness

Low-level Controls

e.g. PID loops on magnet settings

Hardware I/O

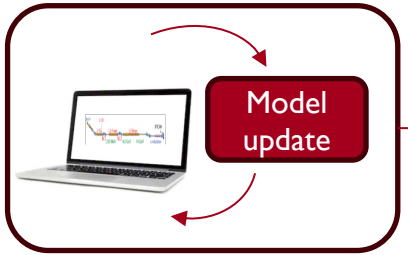


NATIONAL ACADEMIES Sciences Engineering Medicine

NATIONAL ACADEMIES PRESS Washington, DC

Foundational Research Gaps and Future Directions for Digital Twins

https://nap.nationalacademies.org/catalog/26894/foundational-research-gaps-and-future-directions-for-digital-twins

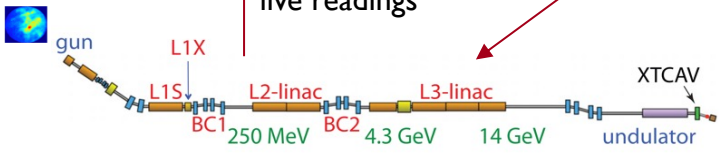


Model update

Decision support
Automated control
Visualization

live readings

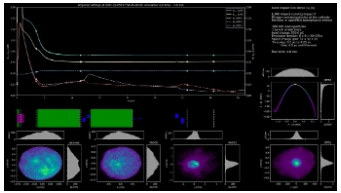
intervention



Digital twin – adaptive online model

Virtual accelerator – comprehensive offline model

Things are not as different as they seem: ML as another accelerator physicist's tool



Typical Control Room Apps

Additional Visualization

High Level Output (e.g. EPICS PV)

Physics Process Emulators (phys. sims, ML)

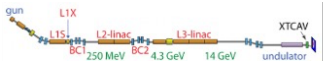
beam dynamics, cooling systems, user beamlines, etc

range of fidelities and comprehensiveness

Low-level Controls

e.g. PID loops on magnet settings

Hardware I/O



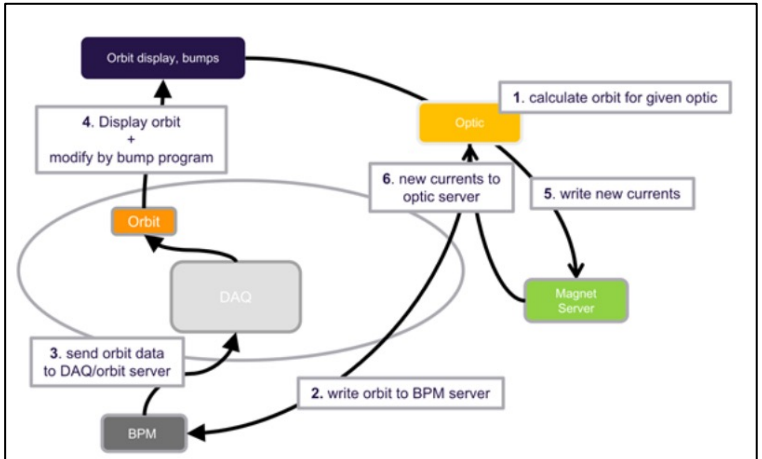


Figure 2: Simulated closed loop data flow for the case of the beam position monitors.

TUD3O04 Proceedings of ICALEPCS2015, Melbourne, Australia

THE VIRTUAL EUROPEAN XFEL ACCELERATOR

R. Kammering, W. Decking, L. Froehlich, O. Hensler, T. Limberg, S. Meykopff, K. Rehlich, V. Rybnikov, J. Wilgen, T. Wilksen, DESY, Hamburg, Germany

Digital twin – adaptive online model

Virtual accelerator – comprehensive offline model

What's new: ML allows a level of speed/fidelity not previously achievable and easier sim2real + continuous calibration

Current status

- Substantial gains in modeling, control, and design are proven
  → *improvement in speed / solution quality range from several factors to orders of magnitude*
- Many solutions in routine use → *mostly for subsystems / individual tasks*
- Infrastructure has been a major limiter *(software, computing, data handling)*

Moving toward more comprehensive treatment within facilities

- Accelerator and user side, chaining together multiple tasks, linking subsystems

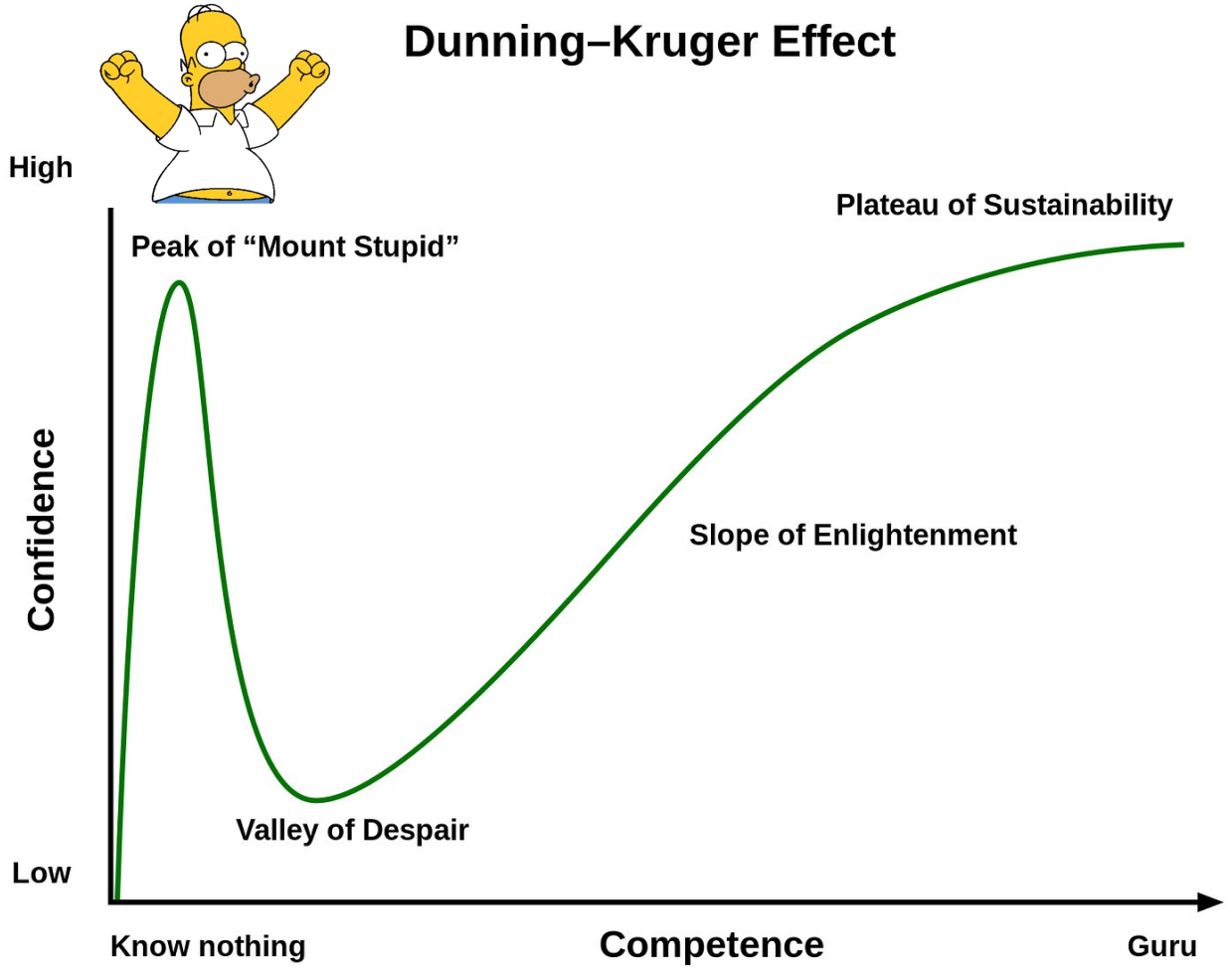
Continuing momentum toward shared development and cross-facility AI/ML

- Community standards / co-development, cross-laboratory collaborative projects
- Much can transfer directly/indirectly → faster progress and increased reliability/robustness

Agents / frontier models clearly powerful and transformative *(when used wisely)*

Many applications closely tie accelerator physics and ML

A word of caution on agents / LLMs



https://commons.wikimedia.org/w/index.php?curid=86929278

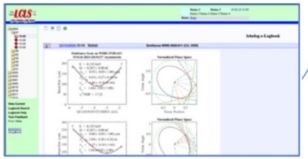
The field is steadily moving toward the shared vision of a fully AIML integrated particle accelerator

Human-computer interaction

Agentic AI

Language modeling / multi-modal data

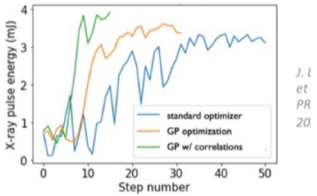
(e.g. electronic logbook)

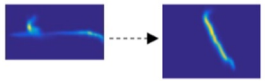


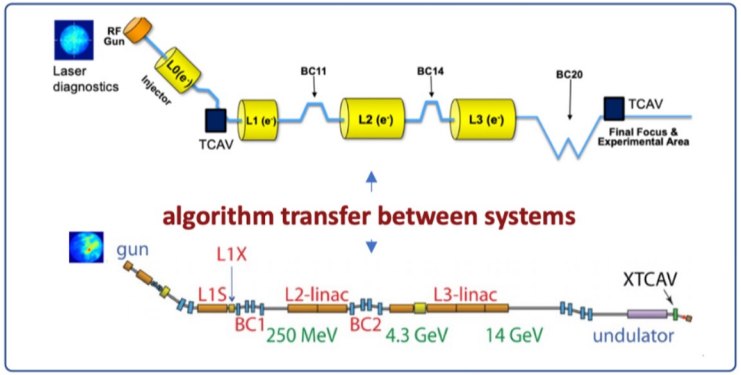
Data reduction/rejection *(kHz/MHz data streams)*

Event triggering

Automated control + optimization

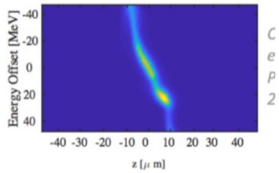






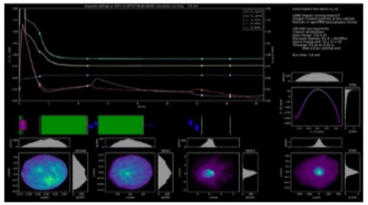
ML-enhanced diagnostics

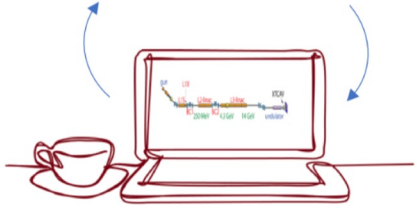
(provide insight at faster rate, at higher resolution, non-invasively)



Anomaly detection failure prediction

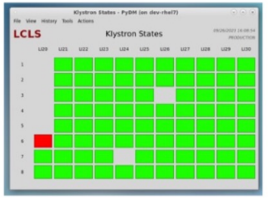
(plan maintenance; alert to changes in machine; alert to interesting science)





Extract unknown relationships + correlations

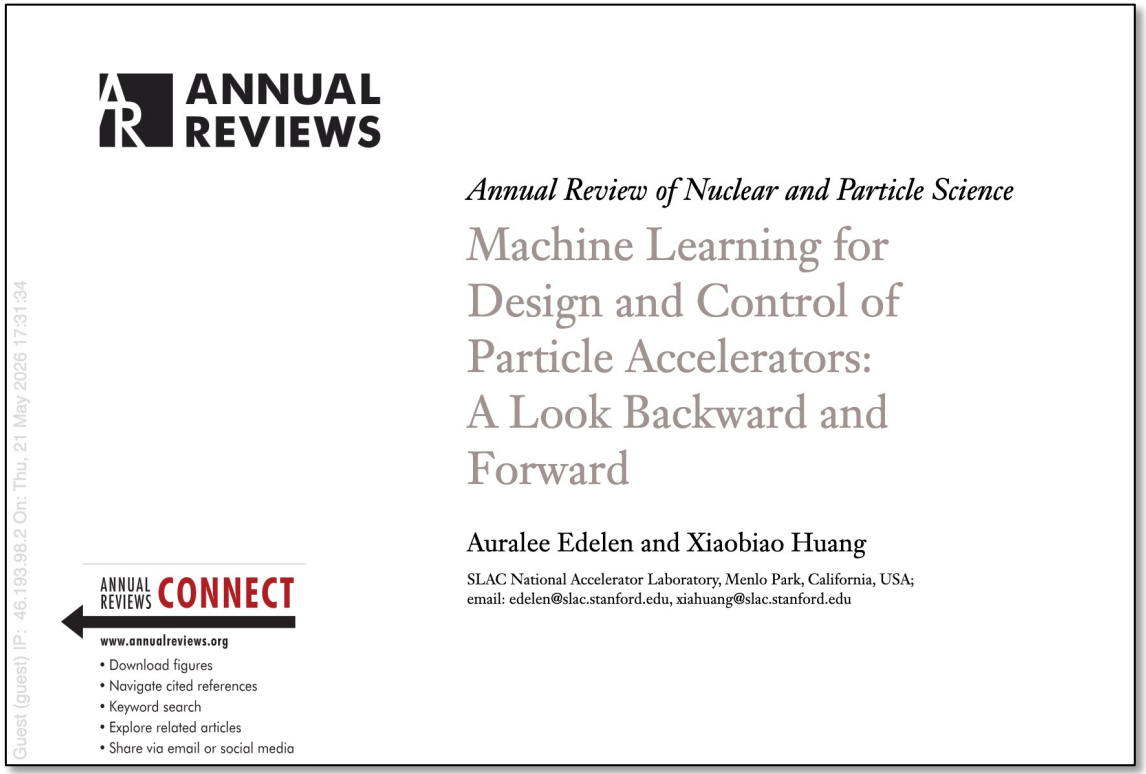
(feed into future control / design)



Digital twins + online modeling

(fast sims, differentiable sims, model calibration, model adaptation)

+ need uncertainty quantification for all
+ can incorporate physics information in all



ANNUAL REVIEWS

Annual Review of Nuclear and Particle Science

Machine Learning for Design and Control of Particle Accelerators: A Look Backward and Forward

Auralee Edelen and Xiaobiao Huang

SLAC National Accelerator Laboratory, Menlo Park, California, USA;
email: edelen@slac.stanford.edu, xiahuang@slac.stanford.edu

Thank you!

Annu. Rev. Nucl. Part. Sci. 2024. 74:557–81

https://www.annualreviews.org/docserver/fulltext/nucl/74/1/annurev-nucl-121423-100719.pdf