Computational HEP AI/ML Initiative

HEP PI Meeting
May 21st, 2024

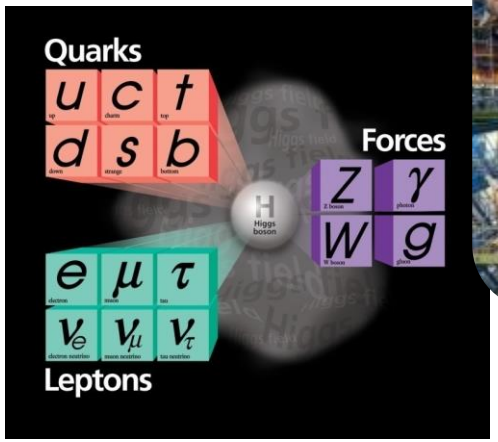
Jeremy Love & Eric Church
Computational HEP PM
Jeremy.Love@science.doe.gov Eric.Church@science.doe.gov

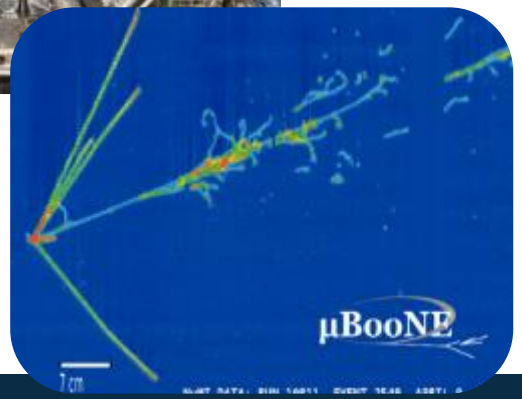
Office of Science Statement of Commitment & other Guidance

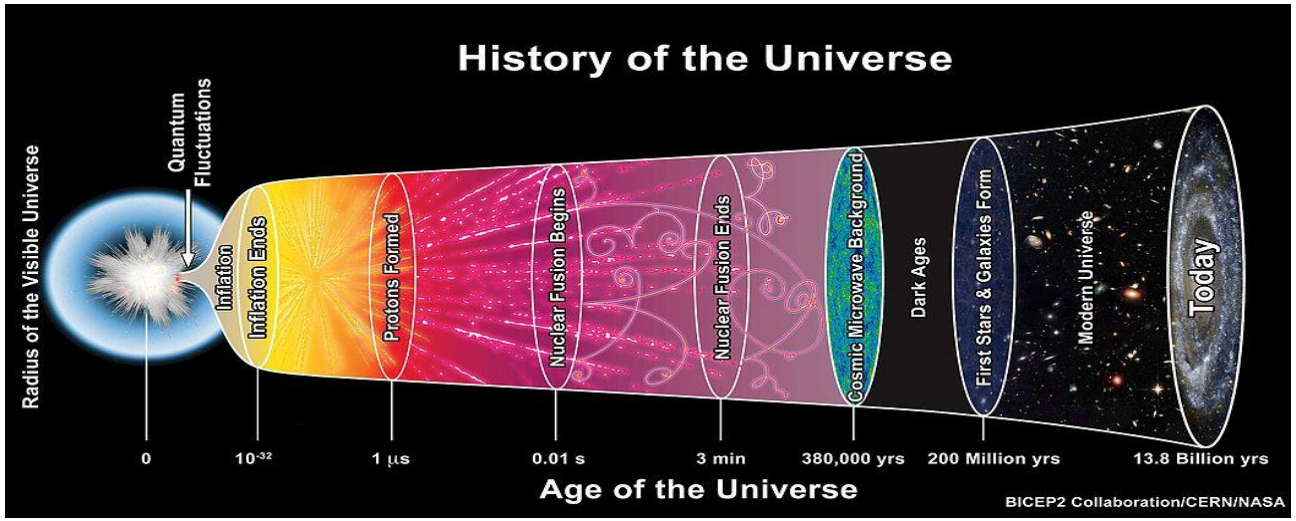
- **SC Statement of Commitment** – SC is fully and unconditionally committed to fostering safe, diverse, equitable, inclusive, and accessible work, research, and funding environments that value mutual respect and personal integrity. https://science.osti.gov/SW-DEI/SC-Statement-of-Commitment
- **Expectations for Professional Behaviors** –SC's expectations of all participants to positively contribute to a professional, inclusive meeting that fosters a safe and welcoming environment for conducting scientific business, as well as outlines behaviors that are unacceptable and potential ramifications for unprofessional behavior. https://science.osti.gov/SW-DEI/DOE-Diversity-Equity-and-Inclusion-Policies/Harassment
- **How to Address or Report Behaviors of Concern**– Process on how and who to report issues, including the distinction between reporting on unprofessional, disrespectful, or disruptive behaviors, and behaviors that constitute a violation of Federal civil rights statutes. https://science.osti.gov/SW-DEI/DOE-Diversity-Equity-and-Inclusion-Policies/How-to-Report-a-Complaint
- **Implicit Bias** – Be aware of implicit bias, understand its nature – everyone has them - and implicit bias if not mitigated can negatively impact the quality and inclusiveness of scientific discussions that contribute to a successful meeting. https://kirwaninstitute.osu.edu/article/understanding-implicit-bias

The Office of High Energy Physics (HEP) Program Mission

- DOE is a mission-oriented agency – mission includes **maintaining a vibrant U.S. effort in science and engineering as a cornerstone of our economic prosperity with clear leadership in strategic areas.**
- HEP's mission is to understand the universe at the most fundamental level:
  - **Discover** the elementary constituents of matter and energy
  - **Probe** interactions between them
  - **Explore** the basic nature of space and time
- DOE HEP supports ~85% of the U.S. HEP effort (in $) at Universities and National Labs

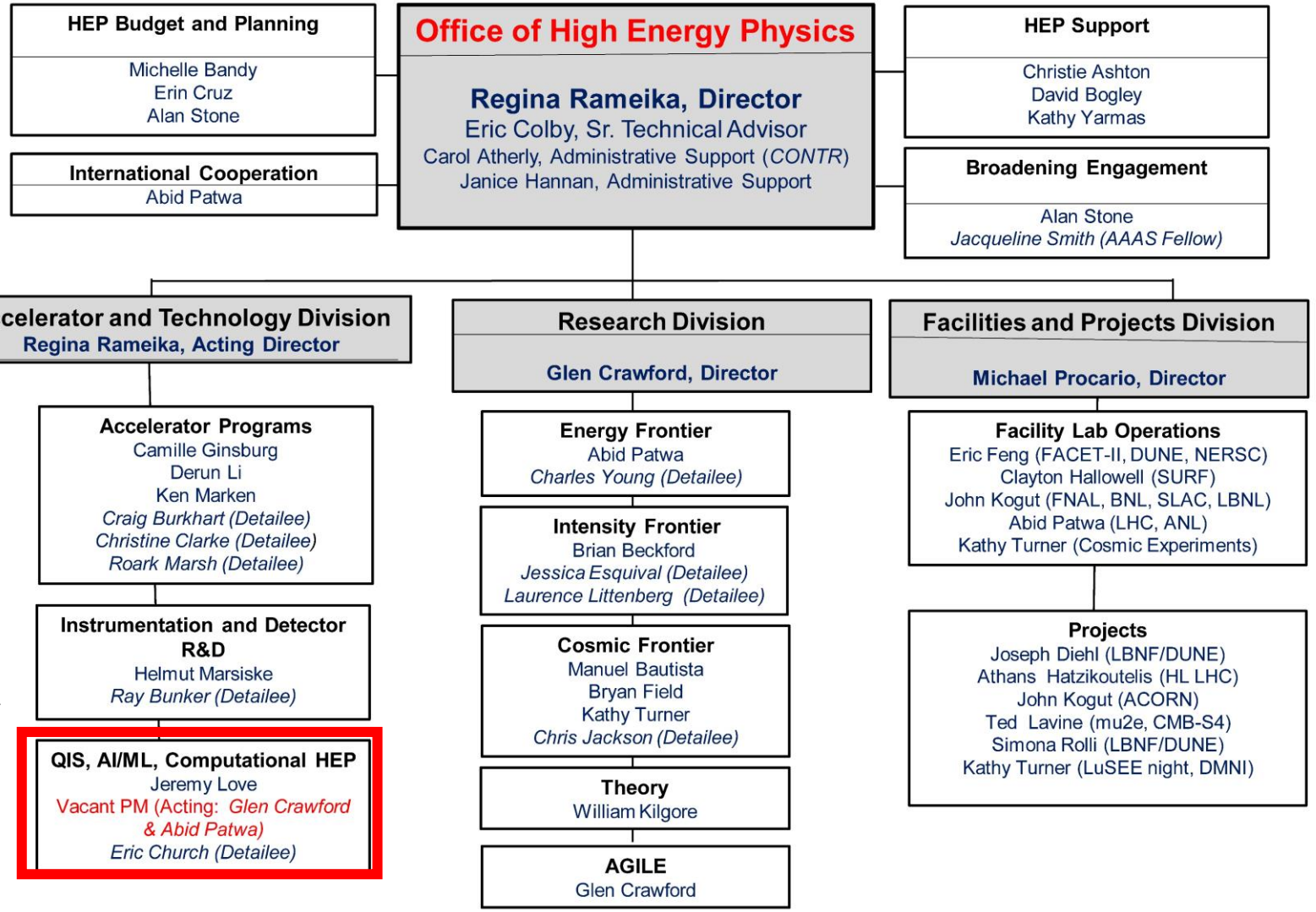






DOE HEP Organization

- The Office of High Energy Physics is organized into three divisions
  - Research, Facilities, and Technology
- All three are needed to achieve the field's long-term goals
  - Technology Division – supports Crosscutting HEP subprograms in enabling technologies: Accelerator Programs, Instrumentation and Detector R&D, Quantum Information Science, and Computational HEP and AI/ML Initiative
  - Work in new technologies allow for new facilities and experiments, more efficient Operations, and previously impossible discoveries
    - Broader impact on other science programs and society



Computational HEP

U.S. DEPARTMENT OF ENERGY | Office of Science

Energy.gov/science

Computational HEP Program

- **Computational HEP** – advanced computing research and development targeting challenges that are or have the capability of being broadly applicable to the increasingly complex HEP computing ecosystem
  - These challenges may include hardware-software co-design, development of collaborative software infrastructure, and research into high performance software and algorithms
  - Aspects of the program are supported by the ASCR Computational User Facilities
- Complimentary programs for research and development to ensure computational advances benefit the broader HEP community
    - **SciDAC** – Lab-led ambitious HPC research supported by ASCR SciDAC Institutes
    - **HEP-CCE** – Lab development of HPC applications for HEP experimental collaborations
    - **Computational HEP Traineeships** – University-led programs to train the computational experts needed to accomplish DOE HEP's mission of discovery science
    - **AI/ML Initiative** – **Programmatic AI/ML** research in the HEP subprograms and **Core AI** research in the Computational HEP program
      - **AI for HEP** – AI Research that furthers HEP priorities of pursuing the P5 science drivers
      - **HEP for AI** – Research that makes use of unique aspects of HEP (datasets, theory, etc.) to improve understanding of the theoretical capabilities and limitations of fundamental AI techniques
      - **HEP AI Ecosystem** – Production and development of open datasets, software ecosystems, or access to shared computing resources that enable broad democratic participation in AI research for HEP

HEP SciDAC-5

- Scientific Discovery through Advanced Computing (SciDAC)
  - Partner HEP and ASCR scientists to deliver high impact science that would not be possible without new innovative development to make use of DOE SC High Performance Computing User Facilities
    - Support for HEP domain scientists to work with scientists at the ASCR SciDAC institutes
  - FY22 was the 5th cycle of the ASCR-HEP SciDAC partnerships (five-year cadence)
    - Research topics of interest to both HEP and ASCR PIs
- Applications were sought to advance the P5 science drivers through HPC usage in three topic areas:
  - Integrated end-to-end simulation of conventional, hybrid, or "virtual" **particle accelerators**
  - Novel **detector simulation** and tracking models and data driven analysis techniques for HEP experiments
  - Innovative **theoretical**, computational, and simulation techniques to explore the unknown, including new particles, interactions and physical principles
- In total 5 awards were made across these three areas
  - Next Generation Precision for Neutrino and Collider Computations
  - Multiscale Acceleration: Powering future discoveries in High Energy Physics
  - Enabling Cosmic Discoveries in the Exascale Era
  - Collaboration for Advanced Modeling of Particle Accelerators (CAMPA)
  - Celeritas: GPU-accelerated particle transport for detector simulation in high energy physics experiments

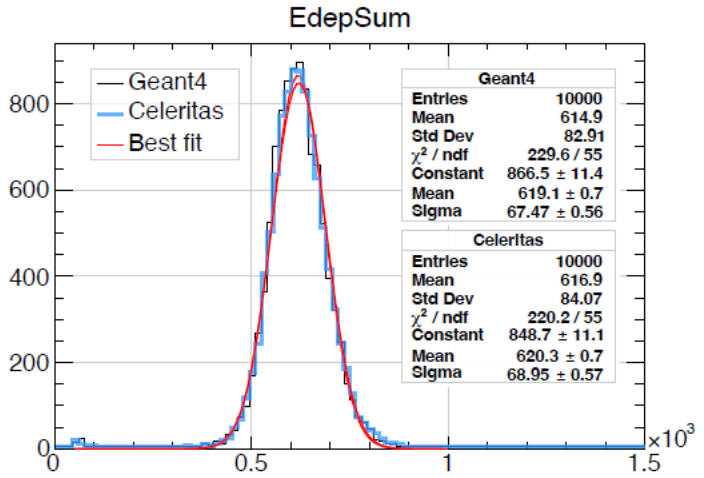
HEP-CCE

- High Energy Physics Center for Computational Excellence has been successful in connecting experiments to HPCs and experimental collaborations now routinely run on the DOE Computational User Facilities
  - Brought together experts from different HEP Frontiers to perform common R&D
- HEP-CCE is entering its second phase
  - Maintain focus on delivering research and development targeting HPCs outside the scope of experimental operations programs
    - HEP-CCE balances academic pursuits with the practical needs of experiments
    - Help bridge the gap between HPC R&D and Operations Programs
    - General R&D through joint work with specific focus
  - Phase-2 HEP-CCE Topics
    - **Optimizing Data Storage** – data management, reduction, and delivery
    - **Portable Applications to Portable Workflows** – HPC and hybrid CPU/GPU application support
    - **Scaling up HEP AI/ML Applications** – Large scale training and hyperparameter optimization on HPC systems
    - **Accelerating HEP Simulation** – Event generators for accelerated systems, GPU enabled simulation

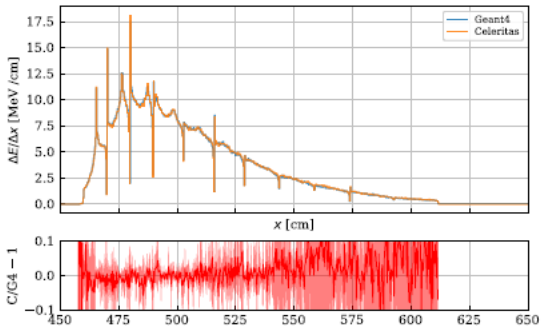
CompHEP Exciting Research

CELERITAS

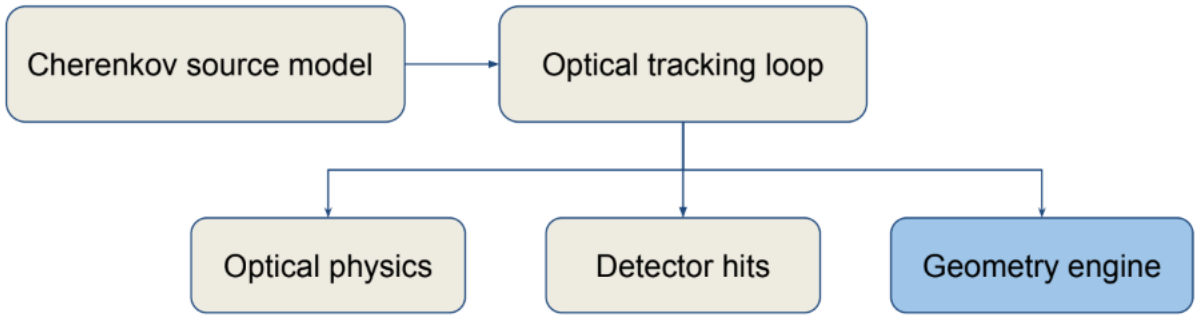
- Celeritas – Collaboration targeting exascale simulation of HEP detector modeling
  - Collaboration of HEP, ASCR, and NE scientists to parallelize and port particle transport codes to GPUs
  - Physics modeling of EM interactions offloaded to GPU shows promising speed-up
    - Integrated with GEANT4 speed-up of 42-256x for GPU:CPU core equivalence
- New directions development of Optical Photon simulation
  - Targeting simulation for LZ, DUNE, and CalVision



Average energy deposition with pi⁺ test beam



Slab-integrated energy deposition



Computational Traineeships

- DOE seeks to support training of a domestic pool of technically competent computationally expert scientists and engineers needed to maintain U.S. leadership in activities supported by DOE SC
  - All science drivers identified by the 2014 Report of the Particle Physics Project Prioritization Panel (P5) require increasingly complex software and computing systems to deliver scientific discoveries
    - The lack of necessary computing expertise needed to realize these scientific drivers was viewed as a risk across the HEP program
- Support for graduate students working towards a Master's or PhD degree with a research thesis
  - Students' theses must include a Computational HEP research project carried out with a partner from DOE National Lab, that falls into one of three broad training categories:
    - **Hardware-Software co-design** – Software development that requires detailed knowledge and understanding of computing hardware systems
    - **Collaborative Software Infrastructure** – Collaborative software environments used to share tools and datasets in a coherent and efficient manner across heterogenous computing platforms for hundreds or thousands of scientific users.
    - **High Performance Software and Algorithms** – Software and algorithms that can take advantage of increasingly parallel computing platforms either synchronously or asynchronously.
  - The goal is to train scientists and engineers capable of designing, developing, deploying, and maintaining the software and computing infrastructure needed to deliver SC supported scientific discoveries
- Four programs:
  - TAC-HEP: Training to Advance Computational HEP in the Exascale Era (Univ. Wisconsin with U. Mass & Princeton)
  - WATCHEP: Western Advanced Training for Computational High-Energy Physics (UC Santa Cruz with Oregon State, UC Berkeley, UC Irvine, UC San Diego, and University of Washington)
  - C^2 the P^2: Chicagoland Computational Traineeship in High Energy Particle Physics (Northern Illinois with University of Illinois Chicago)
  - High Energy Physics Computing Traineeship for Lattice Gauge Theory (Michigan State with MIT, Univ. of Illinois Urbana-Champaign, UConn, Maryland, Univ. Colorado Boulder)
    - Started in FY23 with support from Theory Program

TAC-HEP

- Students/Fellows/Trainees: Physics PhD students (can get Computational minor, certificate etc) at the three universities, Research Software Engineers at Princeton
  - Each student will be a trainee for 2 years maximum after which they are expected to be funded as a Research Assistant (other grants)
- PhD theses on LHC (CMS/ATLAS), DUNE, and the Vera Rubin Observatory,
  - Three frontiers (Energy, Intensity, and Cosmic frontiers)
- Program started in Oct 2022
- 9 students currently involved:
  - https://tac-hep.org/trainees.html
  - Additional students joining at the start of summer

Future Directions and Opportunities in CompHEP

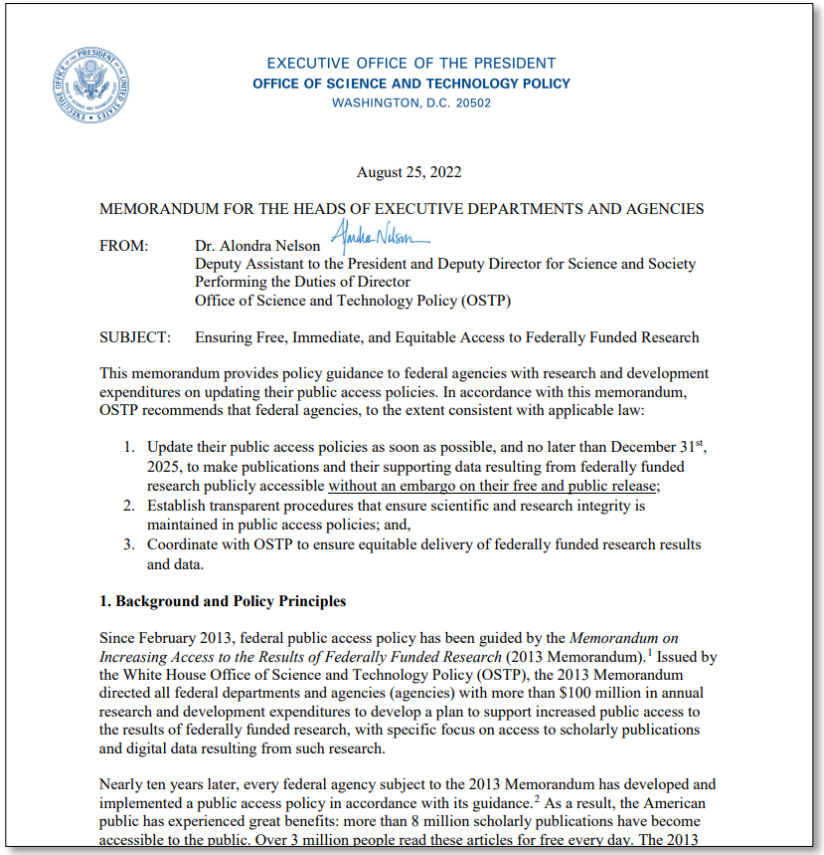
- Intention is to grow university participation in CompHEP program
  - To continue the pipeline started by Computational Traineeships and ensure a robust program
  - Research towards the entire software development cycle is sought in distinct categories:
    - **Seed:** proposals into forward looking blue-sky research and development into general HEP computing challenges
    - **Bridge**: development and support of mature software products for the broader HEP community
- CompHEP will participate in the FY25 Open Call Comparative Review program
  - Support for modest strategic efforts as budgets allow
    - Distinct CompHEP awards separate from HEP "umbrella grants"
    - Fractional support for researchers (PIs, students, postdocs) to maintain connection with physics program
    - Research aligned with the scope of existing CompHEP program is encouraged
      - SciDAC, HEP-CCE research thrusts, Traineeship topic areas, etc.
    - Priority support for research over hardware.
    - This program does not intend to support activities of benefit to a single project or experiment.
    - This program does not intend to support maintenance of legacy software packages.
      - Research and development into new features or functionality of widely used existing software is considered in scope.

Broadening Engagement

- Computational research has a lower barrier to entry compared with some other HEP subprograms
  - The CompHEP has an opportunity to be a leader in broadening engagement in HEP
  - Bringing together the broader community of researchers and engineers with relevant skills to HEP research is encouraged
    - Partnerships with CS, ECE, etc. researchers
    - Development of tools and datasets that facilitate participation and lower the barrier to entry
- PIs are encouraged to seek support through existing DOE Programs
  - Particularly programs partnering university researchers and students with lab staff:
    - Workforce Development for Teachers and Scientists: SCGSR, Visiting Faculty Program
    - SC Funding for Accelerated Inclusive Research (FAIR) & Established Program to Stimulate Competitive Research (EPSCoR)
- PIER Plans are required with all proposals (more details in backup)
  - They should be specific and unique to your institution and research group

2022 OSTP Public Access Memo

- 2022 OSTP Public Access Memo ("Nelson" Memo)
  - "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research"
    - Builds on the 2013 OSTP "Holdren" memo
    - "A federal public access policy consistent with our values of equal opportunity must allow for broad and expeditious sharing of federally funded research—and must allow all Americans to benefit from the returns on our research and development investments without delay."
  - Requirements include:
    - Removes 12-month embargo on access to scholarly publications – immediate access upon publication
    - Requires immediate access to data underlying publications and increased access to other data
    - Requires the use of persistent identifiers (PIDs) for research outputs (e.g., publications, data, software), researchers, and awards
- DOE SC has released an updated Public Access Plan and is in the process of updating the implementation and PI guidance
  - Additional details in May, 2024 HEPAP Talk
- CompHEP projects are expected to display leadership
  - FAIR – Findable, Accessible, Interoperable, Resusable



EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

August 25, 2022

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: Dr. Alondra Nelson
Deputy Assistant to the President and Deputy Director for Science and Society
Performing the Duties of Director
Office of Science and Technology Policy (OSTP)

SUBJECT: Ensuring Free, Immediate, and Equitable Access to Federally Funded Research

This memorandum provides policy guidance to federal agencies with research and development expenditures on updating their public access policies. In accordance with this memorandum, OSTP recommends that federal agencies, to the extent consistent with applicable law:

1. Update their public access policies as soon as possible, and no later than December 31st, 2025, to make publications and their supporting data resulting from federally funded research publicly accessible without an embargo on their free and public release;
2. Establish transparent procedures that ensure scientific and research integrity is maintained in public access policies; and,
3. Coordinate with OSTP to ensure equitable delivery of federally funded research results and data.

1. Background and Policy Principles

Since February 2013, federal public access policy has been guided by the *Memorandum on Increasing Access to the Results of Federally Funded Research* (2013 Memorandum).[1] Issued by the White House Office of Science and Technology Policy (OSTP), the 2013 Memorandum directed all federal departments and agencies (agencies) with more than $100 million in annual research and development expenditures to develop a plan to support increased public access to the results of federally funded research, with specific focus on access to scholarly publications and digital data resulting from such research.

Nearly ten years later, every federal agency subject to the 2013 Memorandum has developed and implemented a public access policy in accordance with its guidance.[2] As a result, the American public has experienced great benefits: more than 8 million scholarly publications have become accessible to the public. Over 3 million people read these articles for free every day. The 2013

AI/ML Initiative

U.S. DEPARTMENT OF ENERGY | Office of Science

Energy.gov/science

Why Invest in HEP AI Research?

- Investment in AI/ML is a national strategic priority
  - Develop cutting edge tools and applications of AI/ML to maintain US expertise
  - Develop a technically capable workforce able to lead the economy of the future and speed-up societal benefits
- The White House AI Executive Order requested a report on "the potential role of AI, especially given recent developments in AI, in research aimed at tackling major societal and global challenges."
  - Delivered in April 2024 PCAST Report "Supercharging Research: Harnessing Artificial Intelligence to Meet Global Challenges"

"The cosmologists and particle physicists … **are some of the earliest adopters—and developers—of AI**, so an epoch of advanced AI is an epoch of exciting discoveries in fundamental physics and cosmology."

"Fundamental physics and cosmology are built on statistical analyses of data to test theory, so they require a deep understanding of the probabilities in the interpretation of data. **This requirement is driving the mathematical development of AI that can handle probabilistic rigor.** … Assessing uncertainties is crucial for fundamental physics, and **probabilistically rigorous AI would be a game changer for many other fields of science as well, in addition to being invaluable for applications beyond science.**"

- Our community is recognized as early adopters and developers of cutting-edge AI techniques, and a uniquely statistically rigorous and data driven field.
  - AI/ML techniques are embedded throughout our programs
- Investment in AI Initiative is an investment in HEP Research – not a zero-sum game
  - As early adopters HEP benefits from AI research by other sciences

AI Initiative in HEP

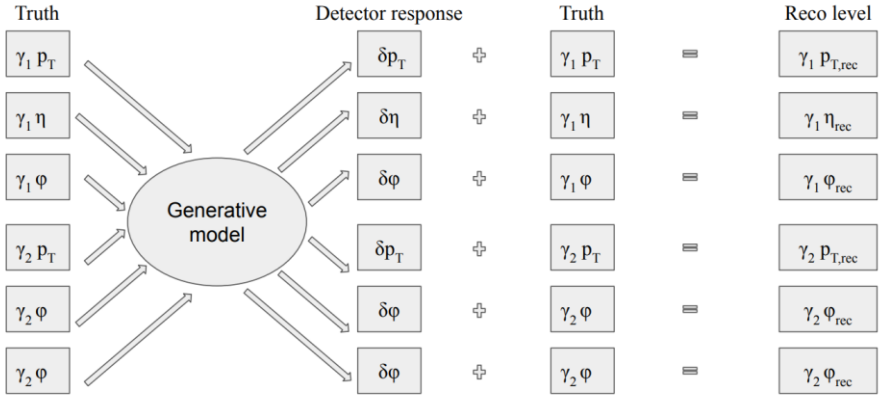
- For decades physics papers from all HEP subprograms have utilized and developed cutting edge AI/ML techniques
  - Pattern recognition, ML-assisted simulation, data classification, uncertainty quantification, real-time applications, etc.
- DOE HEP is pursuing research into AI/ML topics in two broad thrusts
  - **Programmatic AI/ML** – Furthers each subprograms pursuit of the P5 Science Drivers through integrated/embedded in the frontier programs.
    - Applications within a given frontier where primarily ML techniques are the best ones suited to improve physics results.
    - This is approximately 85% of our current AI/ML activities.
  - **Core AI/ML** – research into AI/ML topics from an HEP perspective and blue-sky R&D necessary to enable future HEP breakthroughs across frontiers
    - Go well beyond what is standard practice either through development of new methods, systems, or applications; or the study of fundamental AI techniques and their limitations

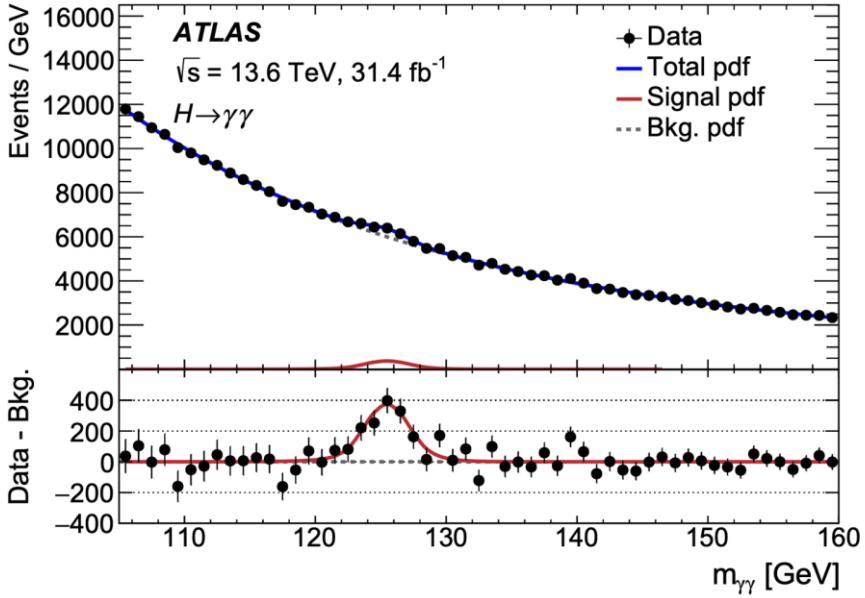
Programmatic AI/ML Research

- Research making use of existing AI/ML techniques in HEP context
  - Research that furthers the goals of the HEP subprograms
    - Administered by Subprogram PMs, contact with any questions
  - Value of the work should be subprogram science regardless of method
    - Programmatic AI/ML research is when the best techniques to improve a result happen to be ML
- Programmatic AI/ML Research is supported as a part of general HEP FOAs: Early Career, Open Call (Comparative Review), US/Japan, etc.
  - ML methods and techniques are expected to be described at an appropriate level just like any other technical work
    - Reviewers must be able to identify what work is being done, who is doing it, and what level of their time is needed to carryout the ML activities.

Generative Machine Learning for Detector Effect Modeling

- Normalizing flows as a generative ML model developed as surrogate for full detector simulatior
  - Capture detector response specific to object kinematics and event conditions, model non-Gaussian effects, reproduce correlations between measurements of different objects
    - JINST 19 (2024) 02, P02003
- This tool was used in the ATLAS Run-3 Higgs to diphoton measurement
  - Trained with ATLAS simulation samples to learn simulated detector response
  - Applied to *O*(100M) generator diphoton background events to obtain expected detector response
  - Used to determine the background PDF in the ATLAS paper
    - Eur. Phys. J. C 84 (2024) 78



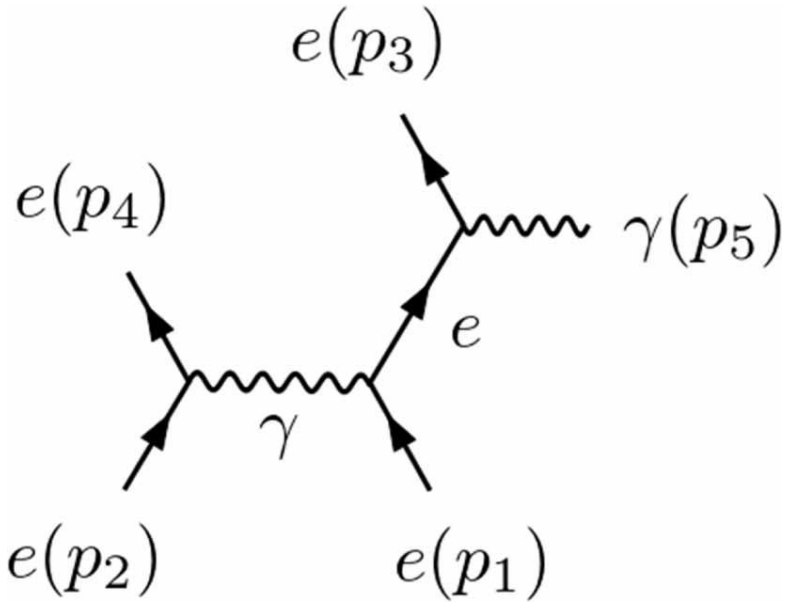


Core AI/ML Research

- Core AI/ML strategic focus:
  - **AI for HEP** – AI research that furthers HEP priorities of pursuing the P5 science drivers
    - Innovative applications of AI tools and techniques or demonstrations and development of new capabilities not currently available to HEP researchers and go well beyond **Programmatic AI/ML**
  - **HEP for AI** – AI research that makes use of unique aspects of the HEP (datasets, theory, etc.) to improve understanding of the theoretical capabilities and limitations of fundamental AI techniques
  - **HEP AI Ecosystem** – Production of open datasets, software ecosystems, or access to shared computing resources that enable broad democratic participation in AI research for HEP
    - Including democratic participation from historically underserved communities
- Supported through dedicated FOAs and strategic projects
  - Exa.TrkX – using exascale computing to develop Graph Neural Nets for LArTPC and tracking detector reconstruction
  - Data Science and Machine Learning for Scientific User Facilities FY20
  - Artificial Intelligence Research in High Energy Physics FY22
  - Hardware-Aware AI for HEP FY24/FY25

SYMBA: Symbolic calculation of squared amplitudes in high energy physics with machine learning

- Develop natural language processing models to calculate matrix elements from Feynman Diagrams
  - Automate calculation of scattering processes with higher order corrections in collider experiments that can not be done by hand
- For QED processes the model accurately predicts 99% of terms from the diagram
  - QCD 73% accuracy



A. Alnuqaydan et al 2023 Mach. Learn.: Sci. Technol. 4 015007

- **The amplitude** $(e\ e \to e\ e\ \gamma)$:

$$i\mathcal{M} = \frac{\frac{1}{2} ie^3 \left(p_{3\rho}\,\gamma^{\rho}_{\epsilon}\,\gamma_{\rho\eta}\,A^{\rho *}_{j}(p_5)\,\mathbf{e}^{*}_{i\eta}(p_4)\,\mathbf{e}^{*}_{l\epsilon}(p_3)\,\mathbf{e}_{k\delta}(p_2)\,\mathbf{e}_{i\delta}(p_1) - \frac{1}{2}\,p_{5\sigma}\,\gamma_{\rho\epsilon}\,\gamma^{\rho}_{\epsilon}\,\gamma_{\rho\eta}\,\gamma^{\sigma}_{\epsilon}\,A^{\rho *}_{j}(p_5)\mathbf{e}^{*}_{i\eta}(p_4)\,\mathbf{e}^{*}_{l\epsilon}(p_3)\,\mathbf{e}_{k\delta}(p_2)\,\mathbf{e}_{i\delta}(p_1)\right)}{((m_e^2 - \vec{p_2}.\vec{p_4}) * \vec{p_3}.\vec{p_5})}$$

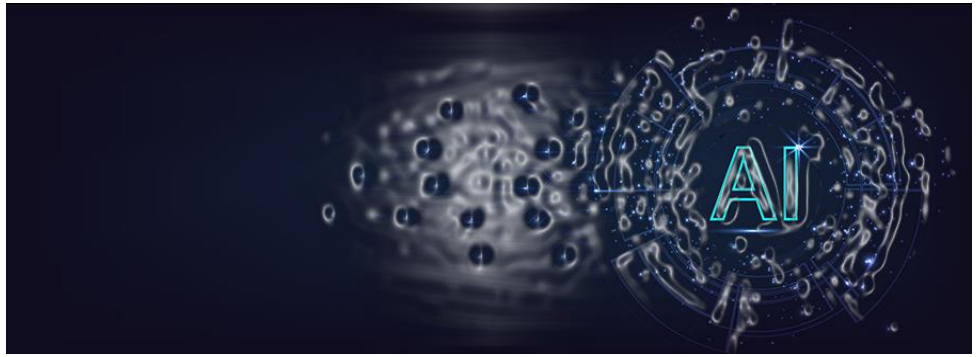
- **The squared amplitude** $(e\ e \to e\ e\ \gamma)$:

$$|\mathcal{M}|^2 = -\frac{e^6}{((\vec{p_3}.\vec{p_5})^2 * (m_e^2 - \vec{p_2}.\vec{p_4})^2)}(2m_e^6 + m_e^4 * (-\vec{p_1}.\vec{p_3} - \vec{p_1}.\vec{p_5} - \vec{p_2}.\vec{p_4} + 2\vec{p_3}.\vec{p_5}) + m_e^2 * (\vec{p_1}.\vec{p_2} * \vec{p_3}.\vec{p_4} +$$
$$\vec{p_1}.\vec{p_2} * \vec{p_4}.\vec{p_5} + \vec{p_1}.\vec{p_4} * \vec{p_2}.\vec{p_3} + \vec{p_1}.\vec{p_4} * \vec{p_2}.\vec{p_5} + \vec{p_1}.\vec{p_5} * \vec{p_3}.\vec{p_4} - \vec{p_2}.\vec{p_4} * \vec{p_3}.\vec{p_5}) - \vec{p_1}.\vec{p_2} * \vec{p_3}.\vec{p_5} * \vec{p_4}.\vec{p_5} - \vec{p_1}.\vec{p_4} * \vec{p_2}.\vec{p_5} * \vec{p_3}.\vec{p_5})$$

Hardware-Aware AI for HEP LAB 24-3305

HEP POC: Jeremy Love

Lab Call Issued : May 1



- University PIs may submit proposals in this area for a Research review through the SC Open Call FOA-3317

Pre-applications/LOIs due: June 26

Full applications due: July 24

Webinar date : May 29. Registration required.

https://science.osti.gov/hep/Research/Artificial-Intelligence-AI

- DOE HEP intends to hold a review for new ambitious research projects where detailed knowledge of HEP hardware systems informs the AI techniques and methods required for implementation
- Applications are sought in two broad categories
  - **Smart Detectors –** Intelligence on detector in readout and control electronics
  - **AI for Operations** – AI/ML for improved experiment and facility operations and control
- Multi-institution team applications are allowed. See lab call for details.
- Limited to two applications per lab (as lead institution). No restriction on number of submissions as subawardee. No restrictions on number of submissions per PI.

Application Track	Award Floor (Annual)	Award Ceiling (Annual)	Award Duration
Lab Application	$350,000	$3,000,000	3 years
University Review	$100,000	$350,000	3 years

Hardware-Aware AI for HEP

- AI for this opportunity are computational systems that respond to data and take action, without human intervention, to achieve a goal and the development of those systems. It includes validation and interpretation work beyond dataset cultivation, training, and algorithm optimization to allow trustworthy autonomous action by the computational system.
- HEP Hardware for this opportunity are an HEP specific detector or sensor technology deployed in HEP experiments and facilities or under development for future HEP applications including Application Specific Integrated Circuits (ASICs) and readout electronics that provide real-time operation of facilities, experiments, and observatories.
- **Smart Detectors** – Informed by the Advanced Detector BRN seek applications that move intelligence as close to the sense element as possible.
  - This may include applications for real time readout or control in ASICs or FPGAs
- **AI for Operations** – Continuation and broadening of the Data Science and Machine Learning for Scientific User Facilities
  - All HEP experiments, experimental facilities, test facilities, observatories, user facilities, etc. are considered HEP Hardware
  - Proposals for AI applications up to a human shifter or operator of the HEP Hardware system are considered in scope
- Topics considered out of scope:
  - Widely available computational hardware and accelerators such as CPUs, GPUs, FPGAs, or Quantum Processors as well as emulations of those systems are not considered Hardware systems as defined for this call.
    - These computing platforms may be used but details of these computational hardware alone is insufficient
  - Proposals for applications that rely on detailed knowledge of HEP Hardware for offline or post-shifter/operator applications
  - Proposals for ML-based triggering algorithms in traditional hardware architectures are not encouraged
    - Applications for non-centralized read-out decision making are encouraged

Hardware-Aware AI for HEP

- Like the AI for HEP FOA this opportunity is expected to be extremely competitive
  - Institutions are encouraged to prioritize quality of proposals over quantity
- Partnerships with non-traditional HEP institutions and PIs are encouraged
  - Especially when those individuals or institutions bring unique and necessary expertise to the project
- Applications targeting an experiment, facility, observatory, etc. are encouraged to obtain letters of collaboration from an appropriate representative to assure reviewers that the stated impact of the project will be realized and there is a willingness to provide the necessary and appropriate support for the system after the project period
- The intention of this opportunity is to support research activities
  - Support for substantial hardware investment may be appropriate if it would realize a fully developed system that would be operations ready
  - Support for fabricating prototypes is expected to be modest, and any request should make clear the specific questions a physical development system will answer that simulation alone cannot.

Thank you.

Additional Material

Promoting Inclusive and Equitable Research (PIER)

- The Office of Science introduced a new Merit Criterion to all FOAs in 2023:

5. Quality and Efficacy of the Promoting Inclusive and Equitable Research (PIER) Plan

- Is the proposed Promoting Inclusive and Equitable Research (PIER) Plan suitable for the size and complexity of the proposed project and an integral component of the proposed project?
- To what extent is the PIER plan likely to lead to participation of individuals from diverse backgrounds, including individuals historically underrepresented in the research community?
- What aspects of the PIER plan are likely to contribute to the goal of creating and maintaining an equitable, inclusive, encouraging, and professional training and research environment and supporting a sense of belonging among project personnel?
- How does the proposed plan include intentional mentorship and are the associated mentoring resources reasonable and appropriate?
- For renewal applications only: How does the proposed plan build or expand upon strategies to promote diversity, equity, accessibility, and inclusion of the currently supported research?
- Are any plans proposed for recruiting additional scientific and/or technical personnel including new senior staff, students, and postdocs reasonable, justified, and appropriate?

SC expects to receive a wide range of ideas and approaches in applicants' PIER Plans; these questions do not represent boxes that must all be checked! Some questions may not apply to every PIER Plan.

A large fraction of HEP Research funding is devoted to salary support for faculty, postdocs, and students. It is appropriate that we consider the effectiveness of an applicant's record and plans for conducting research in an inclusive and equitable manner and for recruitment and mentoring in the allocation of research funds.

SC Commitment to Diversity, Equity, Inclusion, and Accessibility

As a steward of public funding, the Office of Science has a responsibility to ensure that we are serving the public.

SC is deeply committed to:

- Supporting diverse, equitable, inclusive, and accessible work, research, and funding environments that value mutual respect and personal integrity;
- promoting people of all backgrounds, including individuals from groups and communities historically underrepresented and minoritized in STEM fields;
- Advancing scientific discovery by harnessing a diverse range of views, expertise, and experiences to drive scientific and technological innovation.

The FY 2023 new proposal requirements are a reflection of this responsibility and of this commitment.

Promoting Inclusive and Equitable Research (PIER) Plans

At-a-glance:

- Should describe the activities and strategies proposed by the Principal Investigator (PI)/project team to promote equity and inclusion integral to the research project;
- Are between 1-3 pages long, and included as an appendix to the research proposal narrative;
- Will be evaluated as part of the merit review process used to inform funding decisions;
- Are required for all research proposals submitted to SC through FOAs, Laboratory Announcements, and invitational proposals from DOE Labs;
- Are not required for existing awardees unless they are submitting a renewal proposal starting in FY 2023;
- Are not required for applications for supplemental funding on existing awards;
- Are not required for applications requesting funding to support conferences (*but there are new conference proposal requirements for FY 2023*)
- Are not required for proposals submitted to SBIR/STTR Programs announcements. A requirement will be phased in at a later date.

PIER Plans: General Guidance Language

APPENDIX [#]: Promoting Inclusive and Equitable Research Plan

"All applications must provide a Promoting Inclusive and Equitable Research (PIER) Plan as an appendix to the research proposal narrative. The PIER plan should describe the activities and strategies of the applicant to promote equity and inclusion as an intrinsic element to advancing scientific excellence in the research project within the context of the proposing institution and any associated research group(s). Plans may include, but are not limited to: strategies of your institution (and collaborating institutions, if applicable) for enhanced recruitment of undergraduate students, graduate students, and early-stage investigators (postdoctoral researchers, and others), including individuals from diverse backgrounds and groups historically underrepresented in the research community; strategies for creating and sustaining a positive, inclusive, safe, and professional research and training environment that fosters a sense of belonging among all research personnel; and/or training, mentoring, and professional development opportunities. Plans may incorporate or build upon existing diversity, equity, accessibility, and inclusion efforts of the project key personnel or applicant institution(s), but should not be a re-statement of standard institutional policies or broad principles. The complexity and detail of a PIER Plan is expected to increase with the size of the research team and the number of personnel to be supported.
See also Section V for information on the Merit Review Criteria associated with this section.

- Do not attach a separate file.
- This response should not exceed three (3) pages. This appendix will not count in the project narrative page limitation."

Merit Review Criteria

DOE SC's standard merit review criteria are set forth by 10 CFR Part 605.10 and may include additional criteria relevant to the scope and objectives of the solicitation. ***Unless otherwise tailored in the solicitation*** (Funding Opportunity Announcement or DOE Laboratory Call), the merit review criteria for the evaluation of applications are as follows, in descending order of importance:

- Scientific and/or Technical Merit of the Project;
- Appropriateness of the Proposed Method or Approach;
- Competency of Applicant's Personnel and Adequacy of Proposed Resources;
- Reasonableness and Appropriateness of the Proposed Budget; and
- **Quality and Efficacy of the Plan for Promoting Inclusive and Equitable Research.**

The sponsoring SC Program Office may elect to modify this order at the time the solicitation is developed, as appropriate for the scope and objectives of the solicitation.

Guiding Reviewer Questions for PIER Plan Criterion

Quality and Efficacy of the Plan for Promoting Inclusive and Equitable Research

- Is the proposed Promoting Inclusive and Equitable Research (PIER) Plan suitable for the size and complexity of the proposed project and an integral component of the proposed project?
- To what extent is the PIER Plan likely to lead to participation of individuals from diverse backgrounds, including individuals historically underrepresented in the research community?
- What aspects of the PIER Plan are likely to contribute to the goal of creating and maintaining an equitable, inclusive, encouraging, and professional training and research environment and supporting a sense of belonging among project personnel?
- How does the proposed Plan include intentional mentorship and are the associated mentoring resources reasonable and appropriate?

Additional reviewer questions may be included in the solicitation if applicable to the scope of the solicitation and history of the research efforts.

Examples of insufficient PIER Plans

- A copy of a departmental or institutional DEIA plan or listing standard institutional policies and procedures.
- A proposal for STEM K-12 or community outreach that **is not** integral to the proposed research.
- A DEIA activity that is already being carried out by the applicant or partners and is not related to or relevant to the proposed research.

PIER Plans should be integral to the proposed research.

More on PIER Plans

- PIER is not meant to be a general-purpose exercise in Diversity, Equity, and Inclusion (DEI), nor does it ask for participation in unrelated outreach efforts. **PIER** is **P**romoting **I**nclusive and **E**quitable **R**esearch and a PIER Plan should describe how inclusivity and equity are to be expressed in the research being proposed, and how senior investigators on the proposal are involved in the effort. A PIER Plan can leverage institutional DEI plans and resources, but it is not enough to simply describe those programs and resources; the PIER Plan must discuss how they are to be implemented in the proposed research. Please look at the information available at: https://science.osti.gov/grants/Applicant-and-Awardee-Resources/PIER-Plans.

C^2-THE-P^2: Chicagoland Computational Traineeship in High Energy Particle Physics)

(Pronounced “CP squared”)

https://www.c2thep2.org/

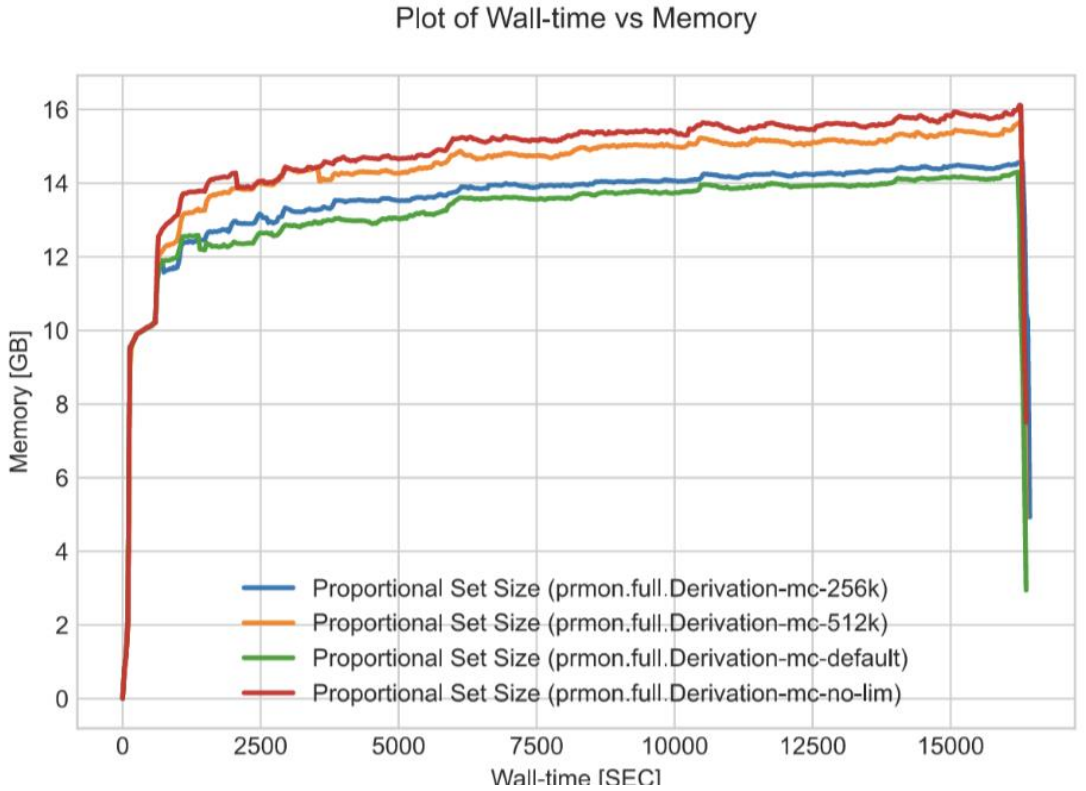
C^2-THE-P^2 program

- A joint program between NIU, UIC, ANL and FNAL to train both physics (NIU+UIC) and computer science (UIC) graduate students in computational skills necessary to advance the particle physics field
- Students obtain a MS degree as part of the two-year traineeship, and then can optionally stay in their department to obtain a PhD
    - Allows students to take the MS and apply for PhDs or jobs elsewhere, or to continue with the PhD
- Focus on Chicagoland area, allows for students to easily work in person with lab mentors and to collaborate with and learn from one another
- Monthly technical seminars on computational topics of interest to HEP
- Eligible students need only mention interest in the traineeship in their personal statements when applying to the respective program

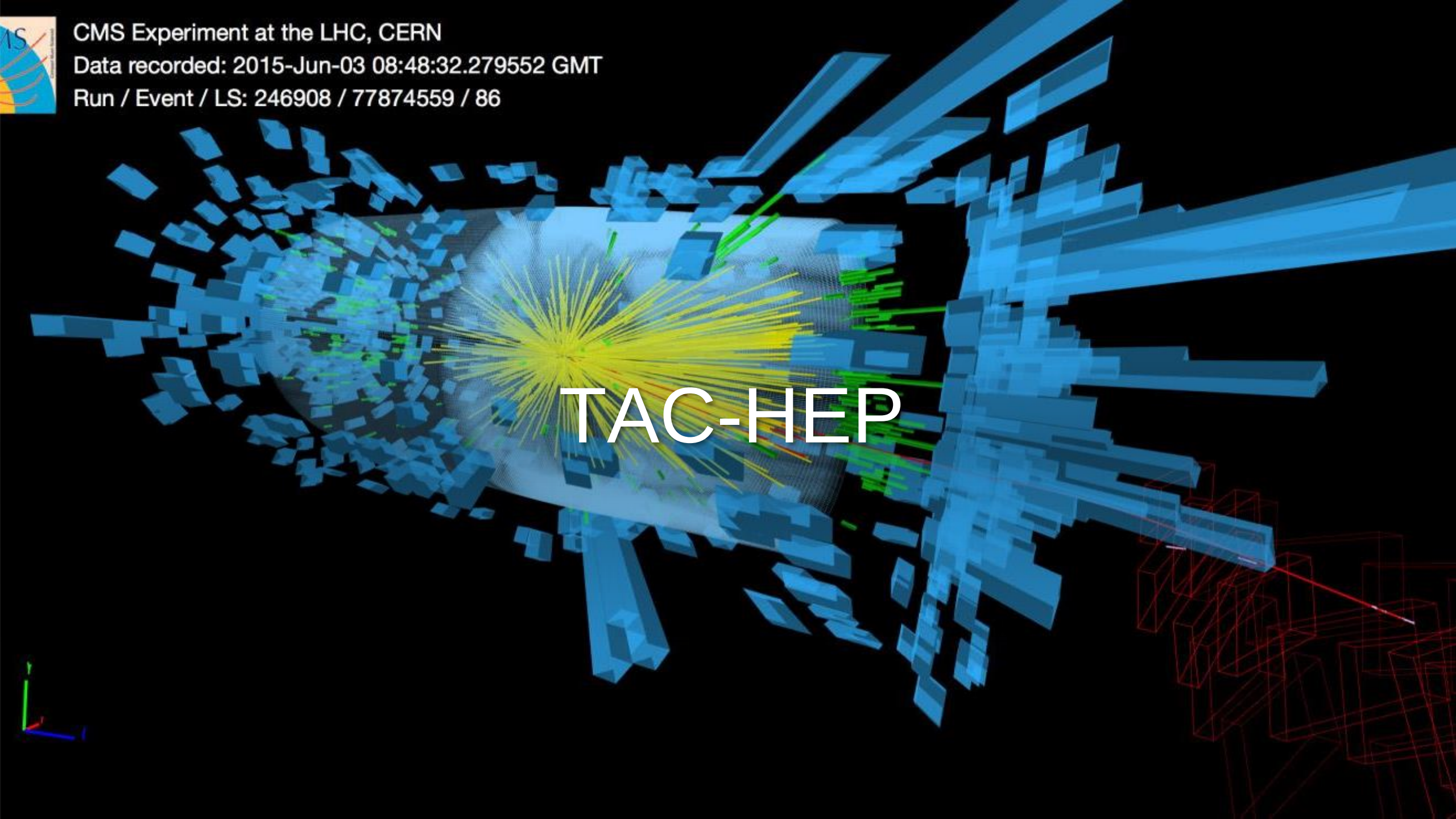
C^2-THE-P^2 student research

- Students take courses at their home institution and work closely with their faculty mentor, but research is guided by a mentor at either ANL or FNAL
- Current students working in varied topics including: Data storage I/O, management and performance, HLS development for triggers and optimization of novel analysis software
- Projects on ATLAS, CMS and DUNE

Effects on memory usage in ATLAS derivation jobs when varying ROOT basket buffer limits



C^2-THE-P^2 2023 annual retreat at ANL


CMS Experiment at the LHC, CERN
Data recorded: 2015-Jun-03 08:48:32.279552 GMT
Run / Event / LS: 246908 / 77874559 / 86
TAC-HEP

TAC-HEP

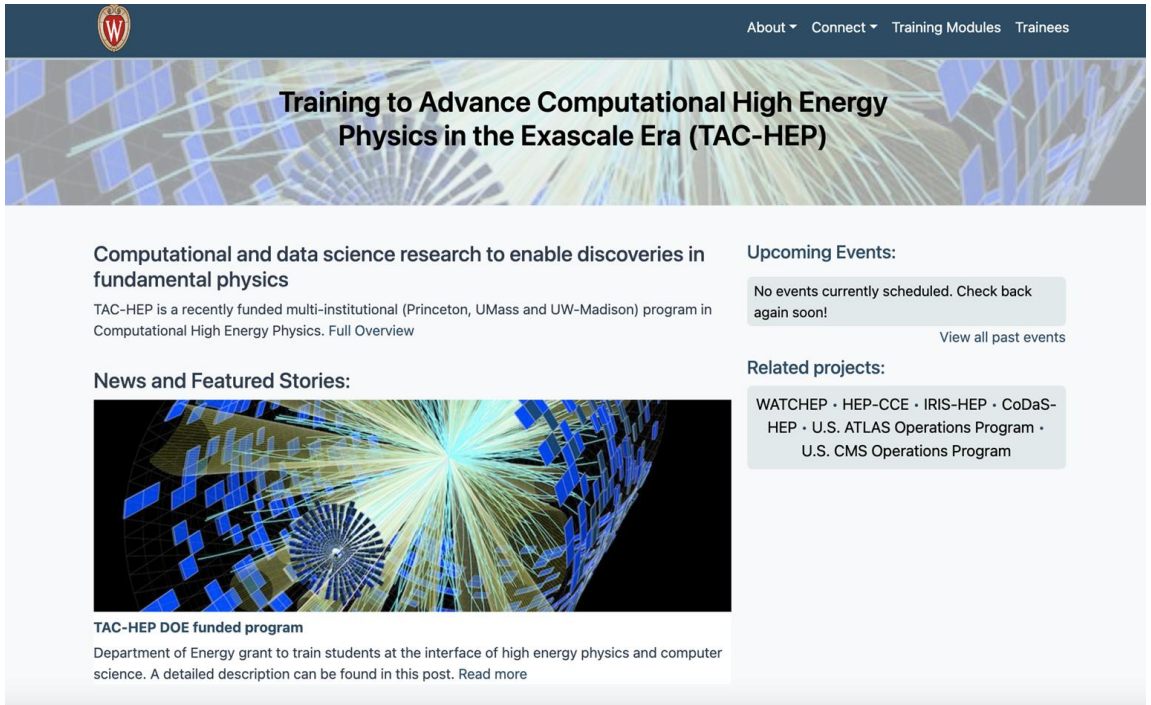
- **Training to Advance Computational High Energy Physics in the Exascale Era (TAC-HEP)**

- DOE funded multi-institutional program with University of Wisconsin-Madison, Princeton University and the University of Massachusetts-Amherst
  - UW-Madison: Tulika Bose, Kevin Black, Kyle Cranmer, Sridhara Dasu, Brian Rebel, Keith Bechtol
  - UMass: Stephane Willocq, Verena Martinez Outschoorn, Rafael Coelho Lopes de Sa
  - Princeton: Peter Elmer, Isobel Ojalvo, Jim Olsen, assessment expert

- Collaboration with two national labs
  - Fermi National Accelerator Laboratory
  - Brookhaven National Laboratory

- Leverage IRIS-HEP and HSF

- https://tac-hep.org/



TAC-HEP

- Students/Fellows/Trainees: Physics PhD students (can get Computational minor, certificate etc) at the three universities, Research Software Engineers at Princeton
  - Each student will be a trainee for 2 years maximum after which they are expected to be funded as a Research Assistant (other grants)
- PhD theses on LHC (CMS/ATLAS), DUNE, and the Vera Rubin Observatory,
  - Three frontiers (Energy, Intensity, and Cosmic frontiers)
- Program started in Oct 2022
- 9 students currently involved:
  - https://tac-hep.org/trainees.html
  - Additional students joining at the start of summer

TAC-HEP Trainees

Elise Chavez
University of Wisconsin-Madison

Miranda Gorsuch
University of Wisconsin-Madison

Matthew Harris
University of Massachusetts Amherst

Stephanie Kwan
Princeton University

Matthew Maroun
University of Massachusetts Amherst

Trevor Nelson
University of Wisconsin-Madison

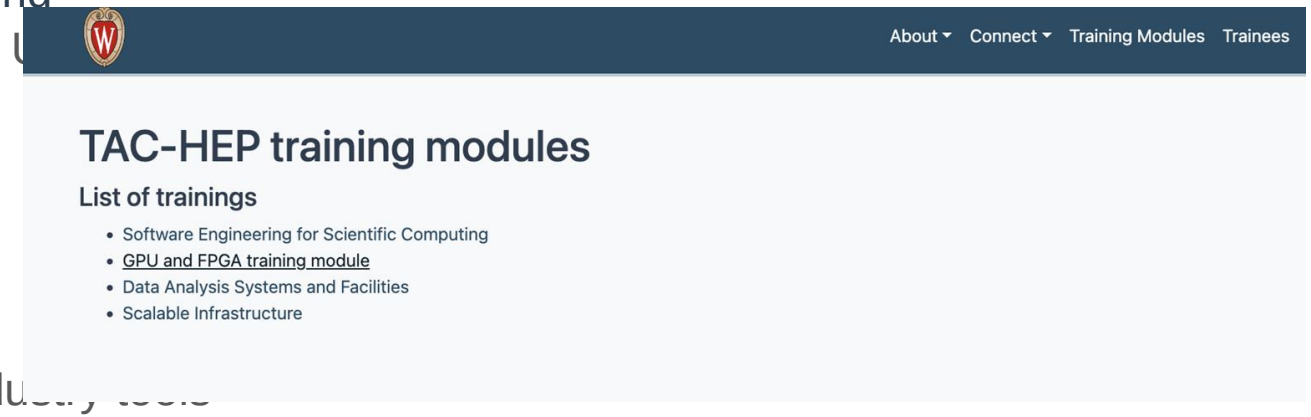
Ryan Simeon
University of Wisconsin - Madison

Andrew Toler
University of Massachusetts Amherst

Lael Verace
University of Wisconsin-Madison

Program components

- Coursework
  - e.g. Software Engineering for Scientific Computing
    - Zoom/hybrid modality: Taught by Princeton U
- Dedicated Training modules
  - Parallel programming
    - GPUs, FPGAs
  - Analysis systems
    - Based on Python ecosystem and using indu
  - And more…
- R&D Projects with lab partners
  - Reconstruction, algorithm development, cyberinfrastructure, analysis systems, simulation….
  - Developed together with faculty advisor and lab mentor
- Summer program where everybody gets together for training (incl. professional development, mentoring etc.)
- Regular (independent) assessment to improve the program



Summer Program

- 2-week event this summer (hosted by Princeton University)
  - 1 week of HEP “school” – CoDAS-HEP
    - https://codas-hep.org
  - 1 week of programming together with students from other DOE computational training groups (from across the country)
    - https://indico.cern.ch/event/1293313/timetable/
- Planning ongoing for this year’s summer program

Summer School Program

- **Hands-on Demo Sessions**
  - Hands on exercises focusing on data analysis demonstrations
- **Big Picture R&D Talks**
  - Presentations on current R&D topics in software & computing
- **Workshop on Communication**
  - How to improve your communication skills from an expert
- **Pitching Your Project - A Workshop on Proposal Writing**
  - How to improve your proposal preparation skills
- **Careers in Physics**
  - Presentation from APS Careers & a panel including research scientists at labs and universities, faculty at universities and a senior data scientist at SFL Scientific, a Deloitte Business

This is an informal & interactive event please ask questions!!!

Upcoming Events: Lightning Round & Coding Jam

- **Lightning Round Presentations Tomorrow**
  - **Tuesday after lunch - 5 minute presentations from each participant**
  - Please prepare no more than 1-2 slides
  - Introduction & background - eg current graduate student year, school, program, etc
  - Interest in software & computing
  - If you are already doing the project part of the program, please mention what your project is!
- **Coding Jam Group Exercises**
  - **A few hours each day except for Thursday**
  - Group work and an opportunity to practice what you are learning in the hands-on demos

Trainees with some faculty, instructors, mentors at Princeton Univ…

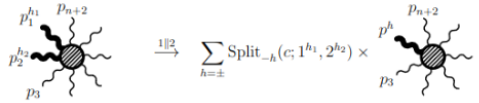
Recruitment

- Students are encouraged to reach out to individual faculty with their interest/questions and mention interest in the traineeship program in their graduate school applications
- TAC-HEP faculty & students regularly participate in Graduate Student “Open Houses” and other recruitment events and advertise the program
- Faculty advertise the traineeship program during research seminars organized for 1st and 2nd year graduate students
- Department graduate student coordinators are familiar with the program and its eligibility requirements and help direct potential candidates
- We encourage students participating in external programs (e.g. IRIS-HEP fellowships) to consider applying to the program

Core AI Exciting Research

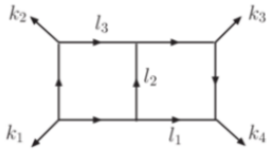
- HEP for AI Team award (SLAC with U Wisconsin & Meta) – use attention transformers to learn the grammar and language of analytic amplitude calculation
  - Train LLM on N=4 Super-Yang-Mills calculations where higher order calculations can be analytically computed to learn how to calculate QCD amplitudes where the answer is not known

Limits & Known Symmetries



Planar N=4 Super-Yang-Mills Multiple Polylog Integrals

$$G(a_1,a_2,\ldots,a_n,x)=\int_0^x \frac{dt}{t-a_1}G(a_2,\ldots,a_n,t),$$



"Bootstrapped" Symbol Form of Integrals

$$\begin{aligned}
\mathcal{S}[F_3^{(1)}] &= (-1)\Big[b\otimes d+c\otimes e+a\otimes f+b\otimes f+c\otimes d+a\otimes e\Big],\\
\mathcal{S}[F_3^{(2)}] &= 4\Big[b\otimes d\otimes d\otimes d+c\otimes e\otimes e\otimes e+a\otimes f\otimes f\otimes f\\
&\quad +b\otimes f\otimes f\otimes f+c\otimes d\otimes d\otimes d+a\otimes e\otimes e\otimes e\Big]\\
&\quad +2\Big[b\otimes b\otimes b\otimes d+c\otimes c\otimes c\otimes e+a\otimes a\otimes a\otimes f\\
&\quad +b\otimes b\otimes b\otimes f+c\otimes c\otimes c\otimes d+a\otimes a\otimes a\otimes e\Big].
\end{aligned}$$

Some "Words" & Coefficients for Training

Some "Words" & Coefficients for Validation/Test

Transformer NLP Model

Undiscovered Symmetries and Recurrences in "Symbol Words"

Q: can we reduce the need for costly limit computations in the bootstrap?

PI L. Dixon