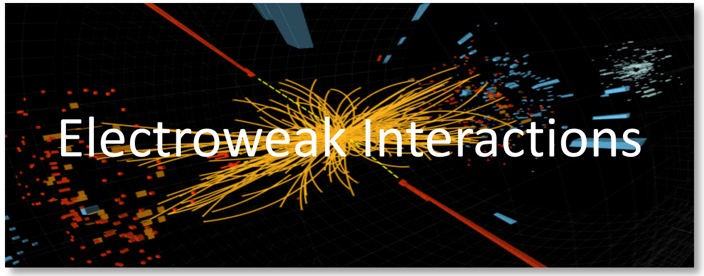
FUTURE TRENDS IN
NUCLEAR PHYSICS COMPUTING

Markus Diefenthaler (Jefferson Lab)

Further Exploration of the Standard Model

Dark matter searches

Electroweak symmetry breaking



Nuclear Physics

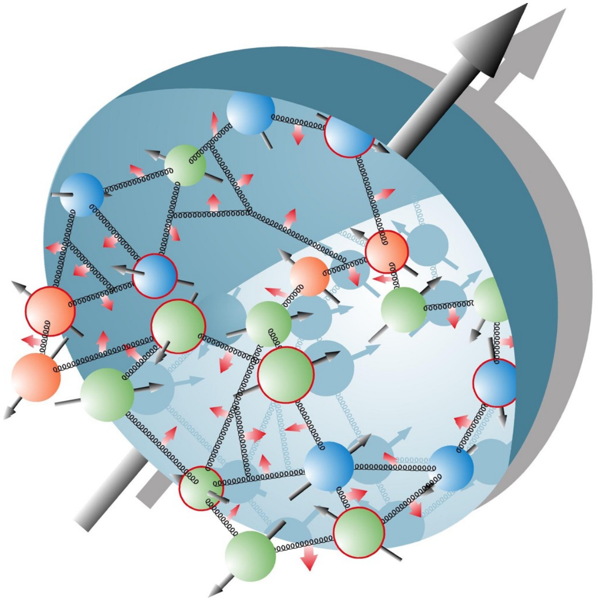
Deeper understanding of QCD

Mission of Nuclear Physics (NP)

Quest to understand the origin, evolution, and structure of the matter of the universe.

Exploring Nuclear Physics: Insights From the Conference Surroundings

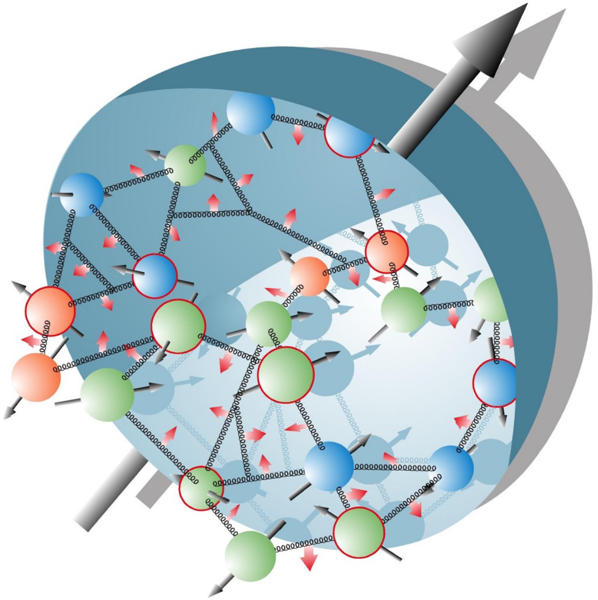
Visible mass in the universe largely from **protons** and heavier nuclei.



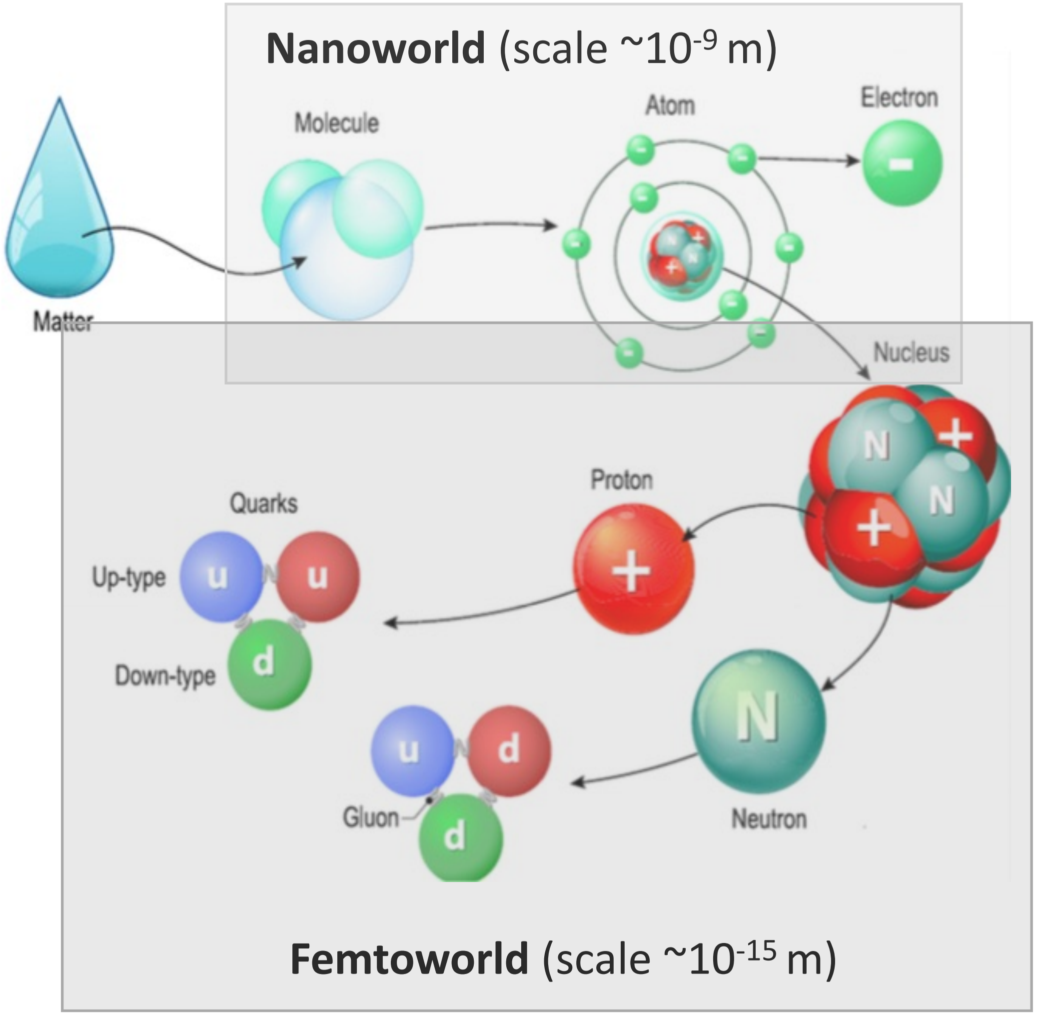
How do quark and gluons interact and combine to form the proton remains largely unknown?

Nuclear Matter is Unique

Molecular and atomic matter: Most known matter has localized mass and charge centers – **vast *open* space.**



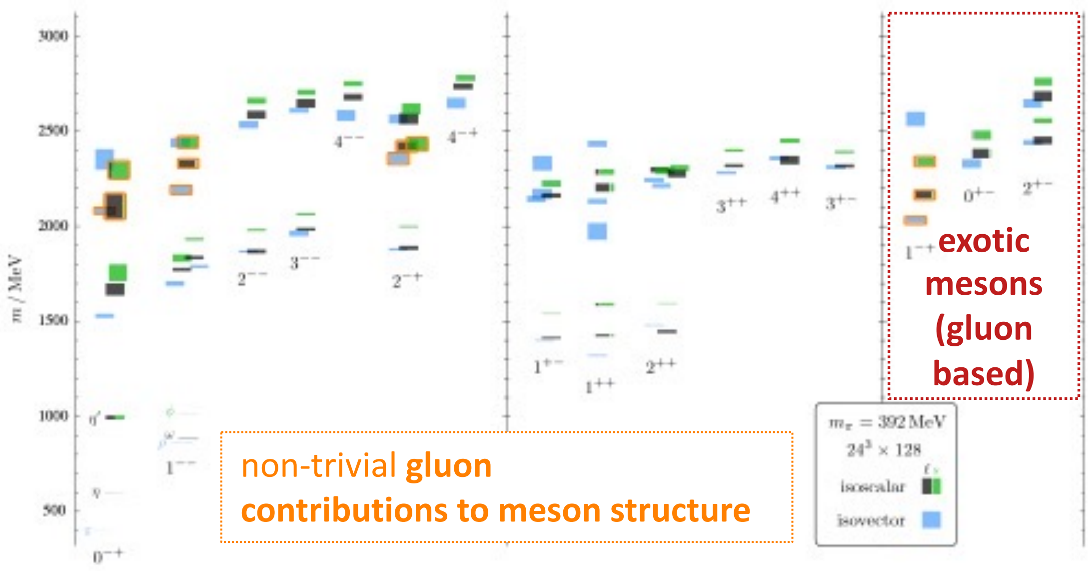
Not so in nuclear matter: **Interactions and structures are inextricably mixed up in protons** and other forms of nuclear matter.



Jefferson Lab

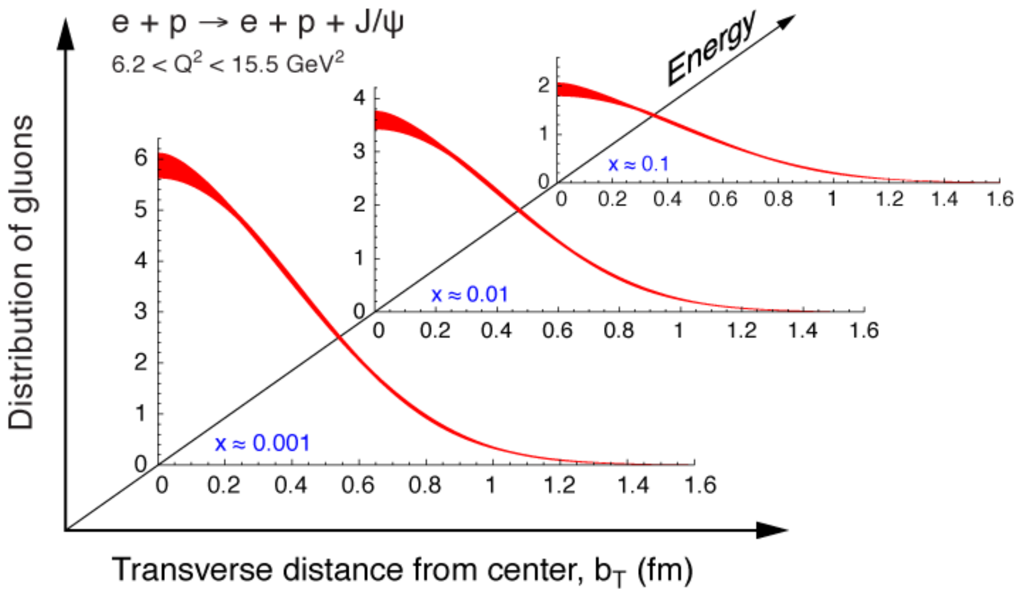
Two Examples of Studying Nuclear Matter

Multiple Channel Challenge,
e.g., discovery search of gluon-based exotic particles (partial wave analysis, 1000s of waves)



Strongly iterative analysis for reliable, model-independent analysis.

Multi-Dimensional Challenge,
e.g., 3D imaging of quarks and gluons in momentum or position space



High statistics in five or more strongly correlated kinematics and multiple particles.

Nuclear Physics is Diverse

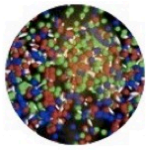
Diversity in the Research Program

The **Heavy Ions** program explores the high temperature frontier of QCD, aiming to recreate and study new forms of matter and phenomena that may exist in extremely hot and dense nuclear matter.

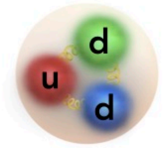
The **Medium Energy** program focuses on the low temperature frontier of QCD, aiming to understand how the properties of existing matter arise from the properties of QCD.

The **Nuclear Structure and Nuclear Astrophysics** program supports research in proton-rich and neutron-rich nuclei, as well as nuclear processes related to stellar nucleosynthesis, neutron stars, and Big Bang nucleosynthesis.

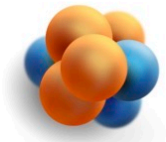
The **Fundamental Symmetries** program investigates the symmetries and forces governing the universe's history, seeking to answer questions such as why there is more matter than anti-matter, the neutrino's mass, and what new particles or forces remain to be discovered.



Hot and Dense Nuclear Matter

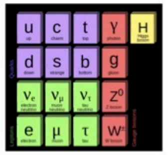


Hadrons



Atomic Nucleus

Nuceli in the Cosmos



Fundamental Interactions

Diversity in Facilities

Relativistic Heavy Ion Collider (**RHIC**) at BNL, Heavy Ion Program at LHC.

Continuous Electron Beam Accelerator Facility (**CEBAF**) at JLab, RHIC, Triangle Universities Nuclear Laboratory (TUNL), Fermilab.

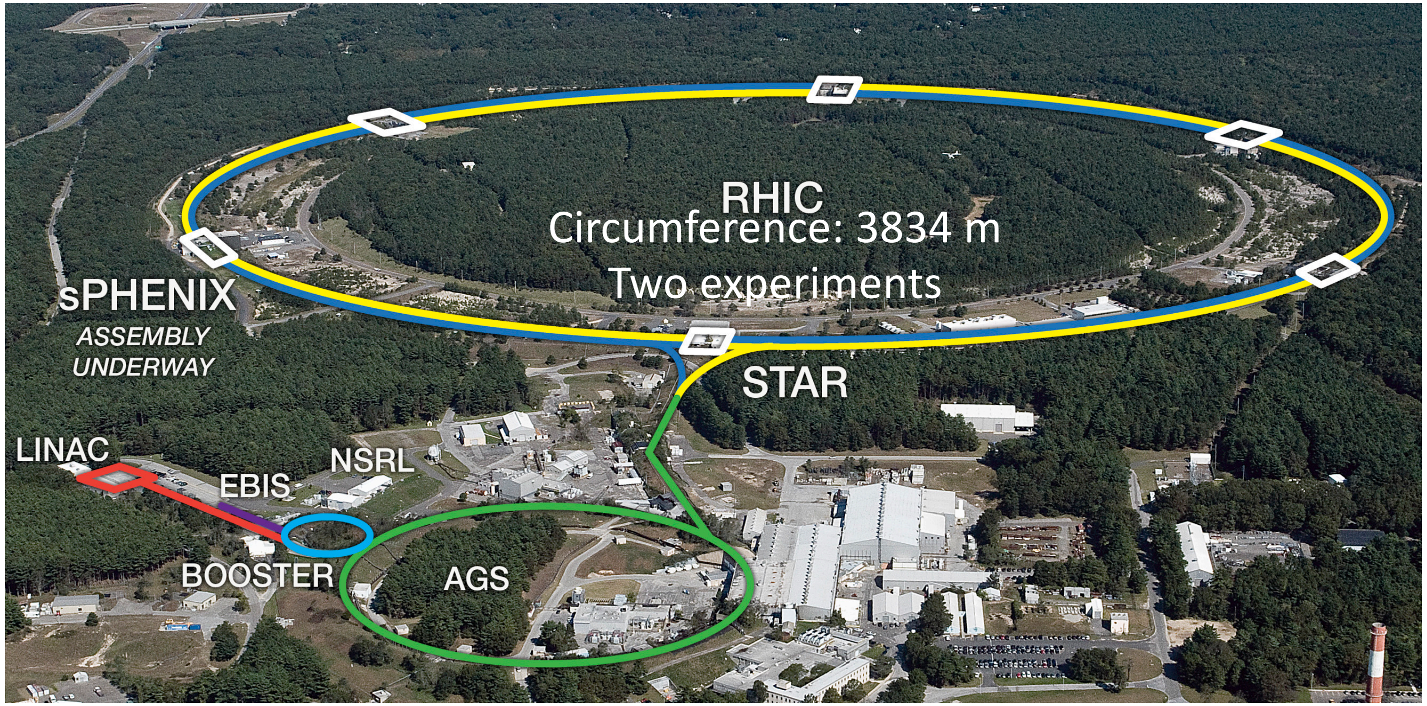
Argonne Tandem LINAC Accelerator System (ATLAS) at ANL, Facility for Rare Isotope Beams (**FRIB**) at MSU, TUNL, Texas A&M University Cyclotron Institute, 88-Inch Cyclotron at. LBNL.

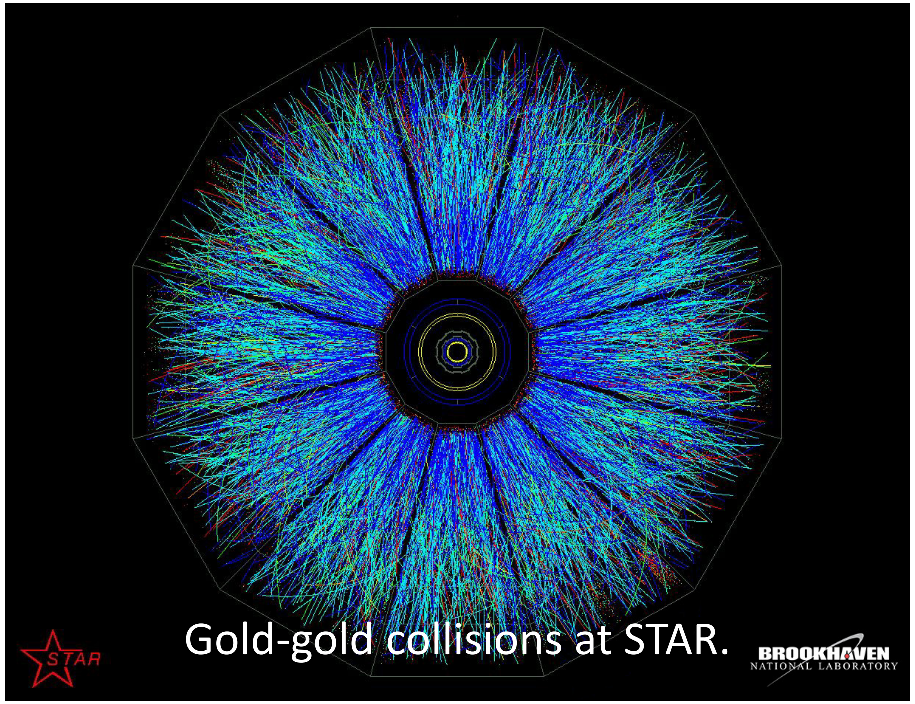
Deep underground labs, neutron facilities, and three university Centers of Excellence (CENPA, TUNL, and TAMU).

Heavy Ion Program

High temperature frontier of QCD.
Exploring new forms of matter and phenomena in dense, hot nuclear matter.

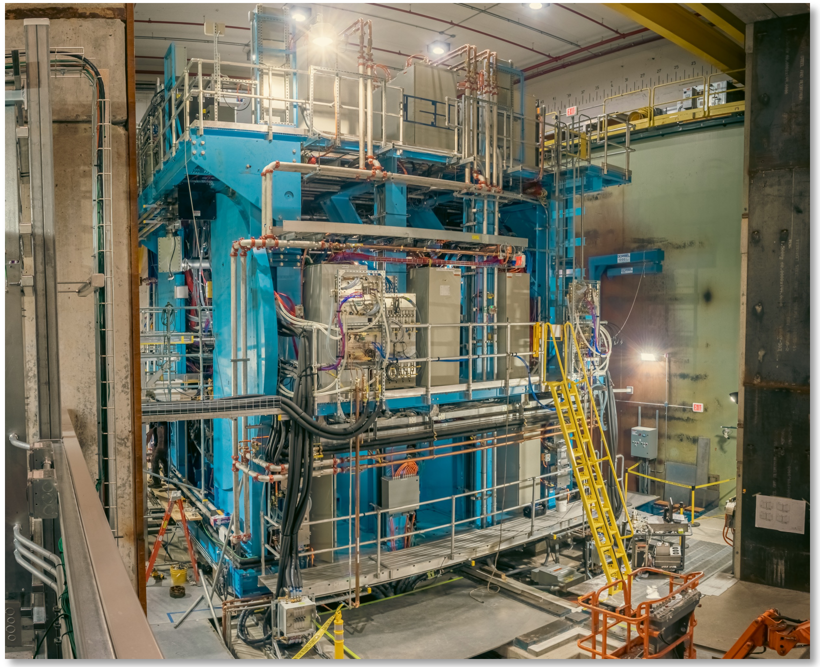
- **Relativistic Heavy Ion Collider** (RHIC) at BNL is the first heavy-ion collider worldwide (2000 – present).
- RHIC has collision energies that reach 100 GeV for gold ions and 250 GeV for protons:
- Study matter at densities that prevailed in the immediate aftermath of the Big Bang, particularly quark-gluon plasma.
- **RHIC-spin**: Only spin-polarized proton collider ever built, enables study of the gluon contribution to the proton spin and other proton structure measurements.
- The future Electron-Ion Collider will be built on the existing RHIC facility.





Gold-gold collisions at STAR.

sPHENIX (2023 – 2025)



Front End data
Front End data
Front End data

Offline data processing at SDCC (BNL). Distributed computing for simulations and physics analysis. Online calibration, and reconstruction at a rapid turnaround at SDCC.

Analysis data

Expected: 140 PB

Data format: ROOT (non-flat for reconstruction, flat for analysis)

2 Tb/s

40 Gb/s

135Gb/s after online data reduction (trigger throttling, compression)

- Collider experiment for high precision measurements of jets and heavy flavor observables (tracking, calorimetry).
- Study quark-gluon structure of strongly interacting quark-gluon plasma.

- **Software stack**: C++, Python for physics analysis, and ML.
- Triggered readout of calorimeter combined with streaming readout of tracking detectors.
- Aims to calibrate and reconstruct 100% of data in near real time. Will inform processing of streamed data for high data rates for other experiments.

- **AI/ML in production**: Fast ML for calibration and reconstruction; analysis.

Medium Energy Program

Low temperature frontier of QCD.
Aiming to understand how the properties of nuclear matter arise from QCD.

Two major accelerator facilities in the U.S.:

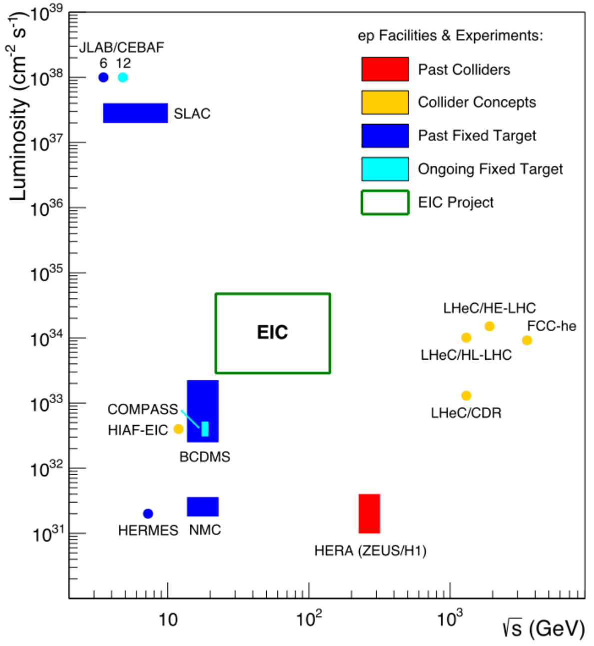
1. **RHIC-spin** (BNL), spin-polarized proton collisions to probe the spin structure of the proton.
2. **Continuous Electron Beam Accelerator Facility** (CEBAF) at Jefferson Lab (JLab).

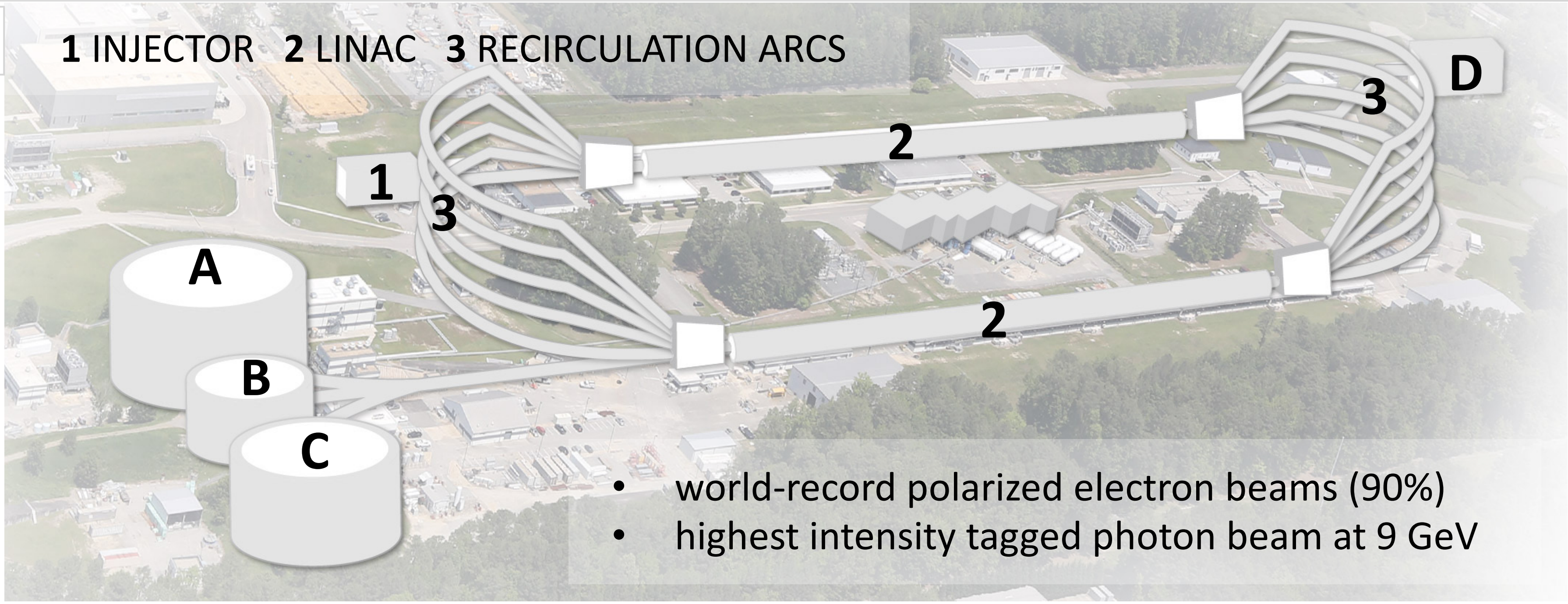
Next new accelerator facility, co-hosted by BNL and JLab: **Electron-Ion Collider** (EIC).

CEBAF12

Versatile: Deliver range of beam energies and currents to four experimental halls simultaneously.
62+ Experiments: Study of quark-gluon structure of nucleons and nuclei, and search for exotic mesons.

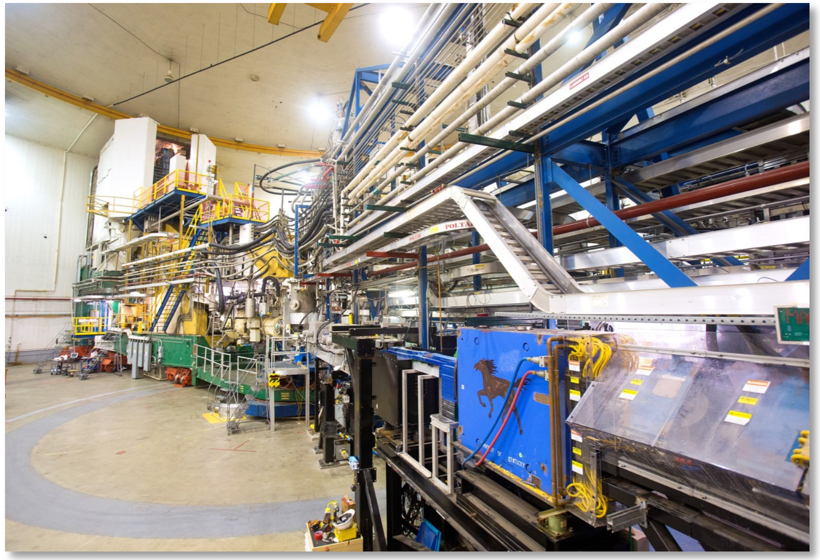
ultra-high luminosities





- world-record polarized electron beams (90%)
- highest intensity tagged photon beam at 9 GeV

Super Bigbite Spectrometer (SBS) in Hall A (2021 – 2024)



Front End data

Front End data

Front End data

20 Gb/s

Offline data processing at JLab

Online data processing for monitoring using 160 cores, detailed data quality checks as part of shift routine.

Analysis data

O(2) Gb/s

Collected: 2.5 PB

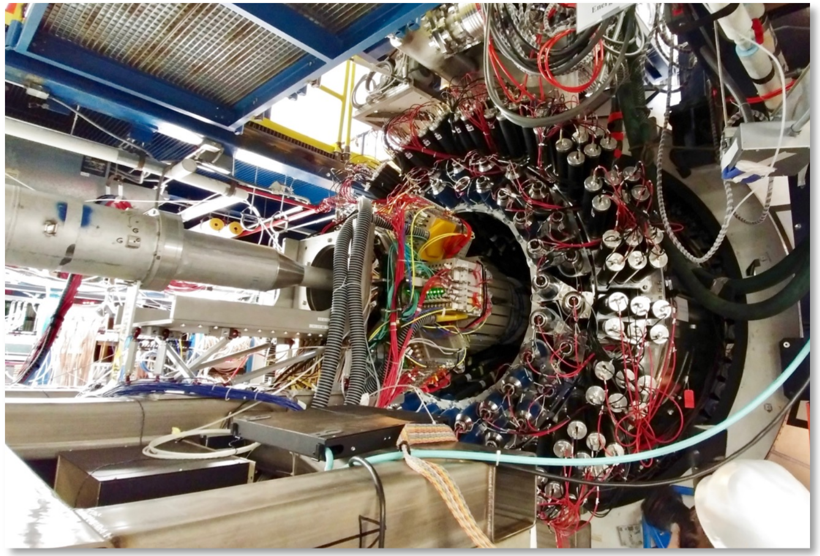
Data format: EVIO, ROOT (flat).

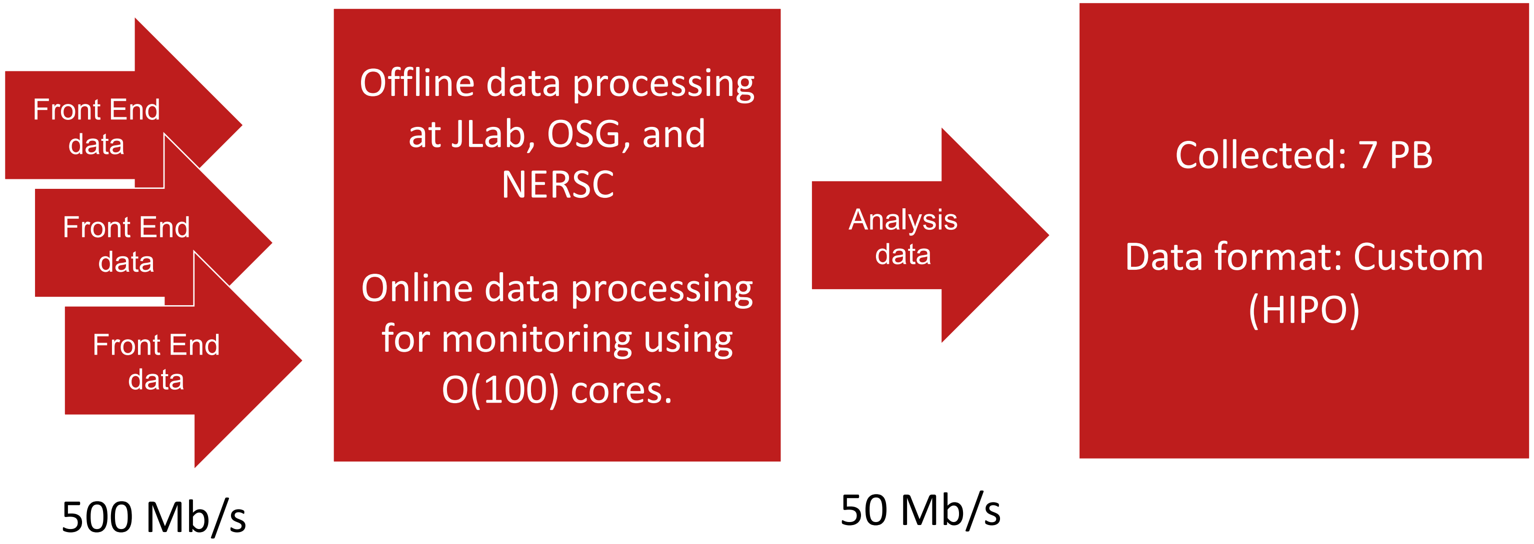
- Fixed-target experiment with high luminosity.
- Polarized electron beams off polarized targets.
- Two movable open-geometry medium-acceptance detector systems for coincidence measurements of multiple final states (tracking, calorimetry, particle identification).

- **Software stack**: C++.
- Hall A/C workflows have been standardized for over a decade, resulting in a highly trained workforce.
- Great success in preserving metadata in Git.

- **AI/ML in production**: None.
- **AI/ML in development**: Autonomous data quality monitoring (Hydra), data reduction for future experiments.

CLAS12 in Hall B (2018 – present)



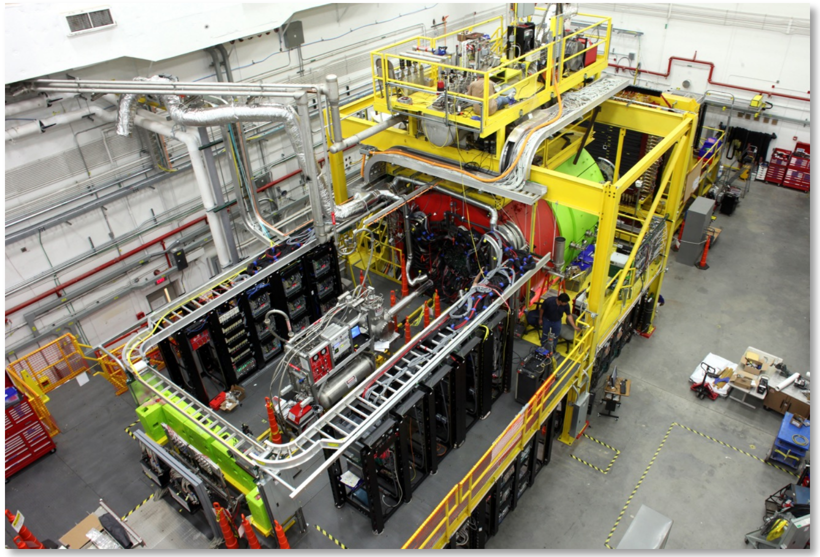


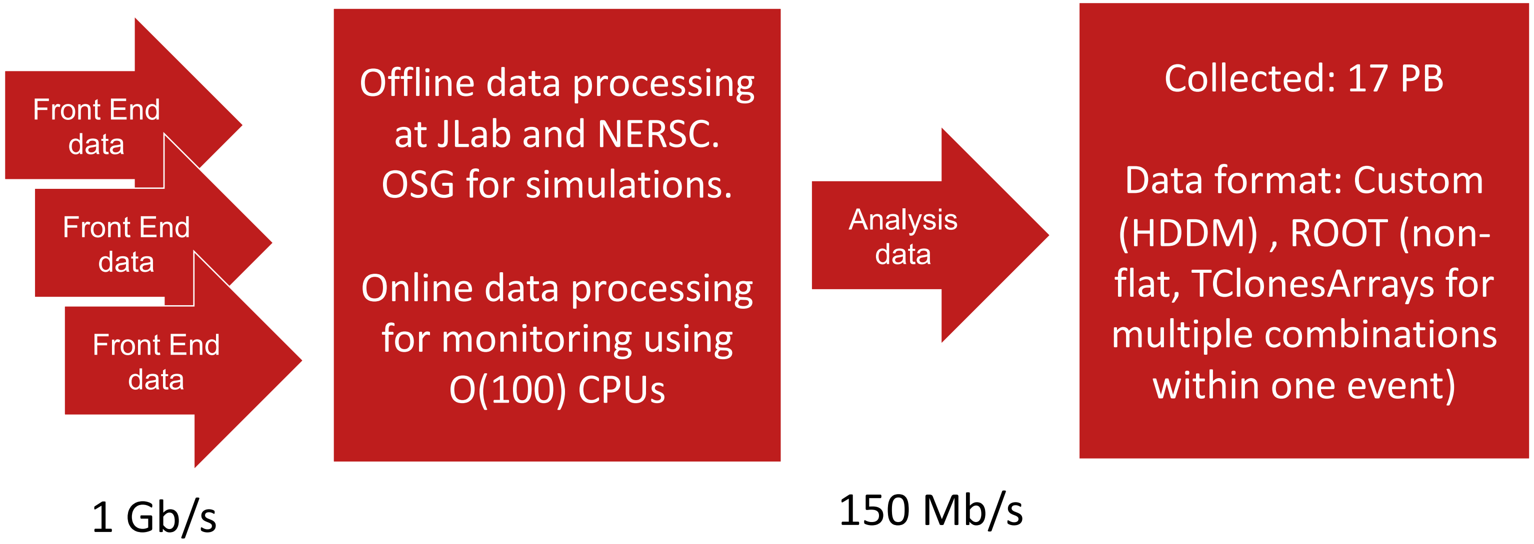
- Fixed-target experiment with medium luminosity.
- Polarized electron beams off a polarized targets.
- A large acceptance detector for the study of a multitude of final states (tracking, calorimetry, particle identification)

- **Software stack**: JAVA for almost all tasks, with C++ for detector simulations and Python and FORTRAN utilized for some physics analysis.
- Successes in standardized software workflows.

- **AI/ML in production**: Noise reduction for tracking and track finding, autonomous data quality monitoring.
- **AI/ML in development**: L3 trigger system for higher luminosity running, particle identification.

GlueX in Hall D (2017 – present)





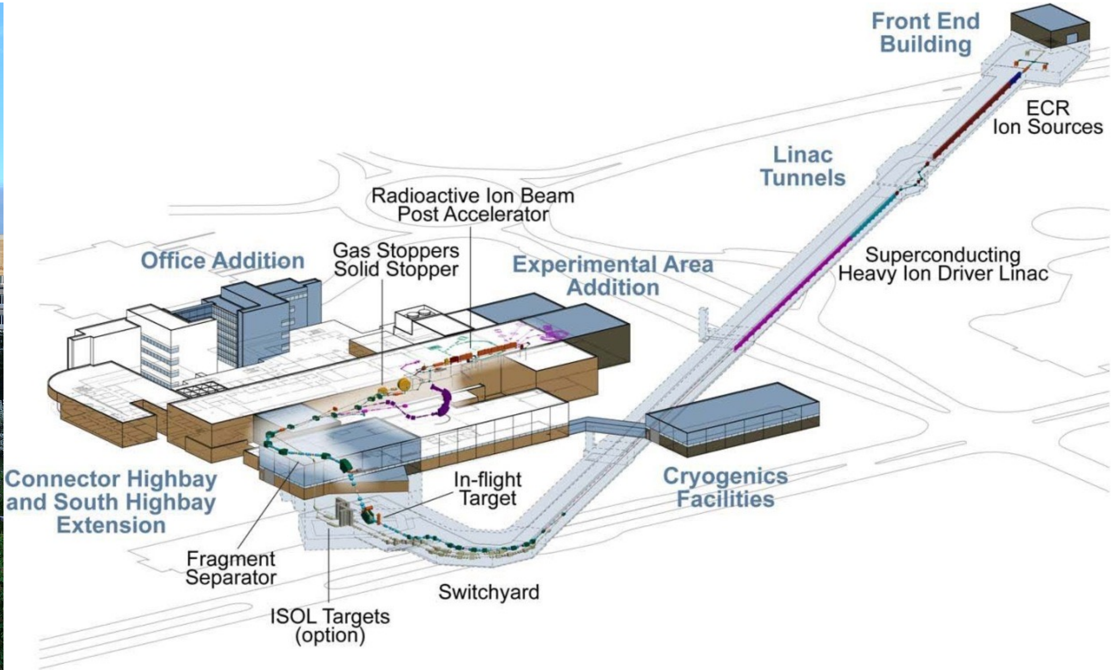
- Fixed-target experiment with high luminosity.
- Linearly polarized photon beam of 9 GeV off liquid hydrogen target.
- Spectrometer with solenoidal magnet designed for the search of light hybrid mesons with high statistical accuracy (tracking, calorimetry, particle identification)

- **Software stack**: C++.
- Parallelization and multi-threading with JANA framework were essential for efficient data processing and analysis.

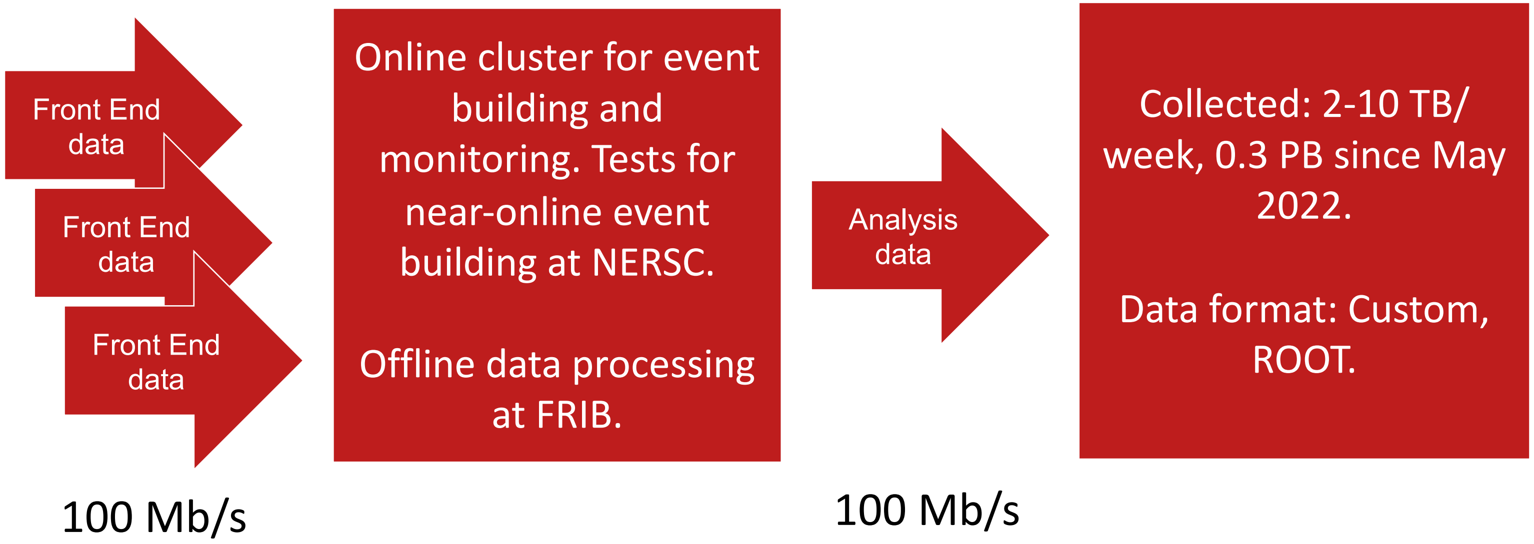
- **AI/ML in production**: Autonomous control (AIEIC) and data quality monitoring (Hydra) particle identification.
- **AI/ML in development**: Autonomous calibration, PWA.

Nuclear Structure and Nuclear Astrophysics

- **Facility for Rare Isotope Beams** (**FRIB**) is a major scientific user facility for the Nuclear Structure and Nuclear Astrophysics program.
- FRIB is located at Michigan State University and was completed in 2022.
- One of the most powerful rare isotope facilities globally, producing **isotopes with unique properties** for **research in NP, astrophysics, and medical fields**.
- Research focus on **nuclear structure of a rare isotopes**, measurement of nuclear reaction rates that are crucial to **understanding the behavior of matter in extreme astrophysical environments**, and to **test nuclear models** that accurately describe the behavior of such environments.



FRIB Experiments (2022 – present)

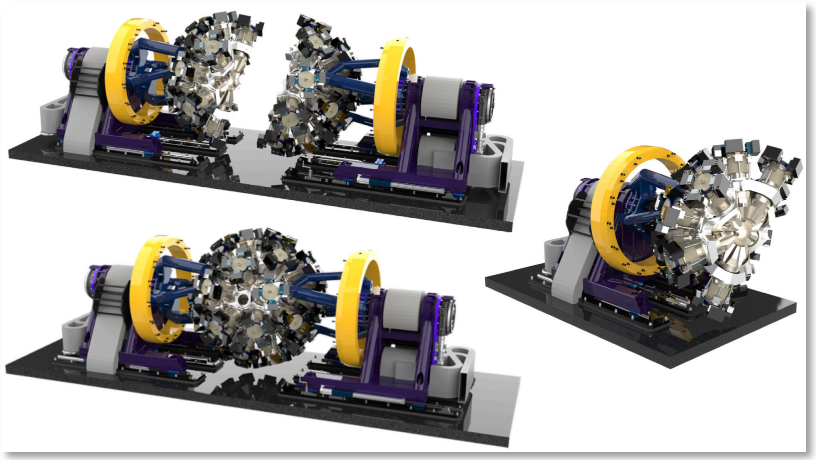


- Data rates vary from experiment to experiments.
- Support for customizable detector configurations with auxiliary detectors from users.

- **Software stack**: C++, Python for slow control configuration, online monitoring tools, and physics analysis.
- Streaming readout to bypass issues as event pile-up or overlapping signals from different events.

- **AI/ML in production**: Autonomous control, event clustering and classification, physics analysis.

GRETA (planned for 2025)



Front End data

Front End data

Front End data

Mid-scale GPU cluster for near real time data analysis (< 10s) of the entire data stream. Remote processing on HPC being explored.

Analysis data

Expected: a few 100 TB /year.

Data format: Custom, ROOT.

8 Tb/s

4 Gb/s

32Gb/s FPGA farm for near real time data analysis

- GRETA is a large-scale gamma-ray tracking spectrometer to be used for nuclear structure and nuclear astrophysics studies at FRIB.

- **Software stack**: Go for high performance network aggregation components and data pipeline control plane, C++ for online data analysis component, Python and Javascript for UI.
- Working with Interdisciplinary software teams for the development of a modern detector computing system using streaming readout.

- **AI/ML in development**: None.

Unique Computing Challenges for Nuclear Physics

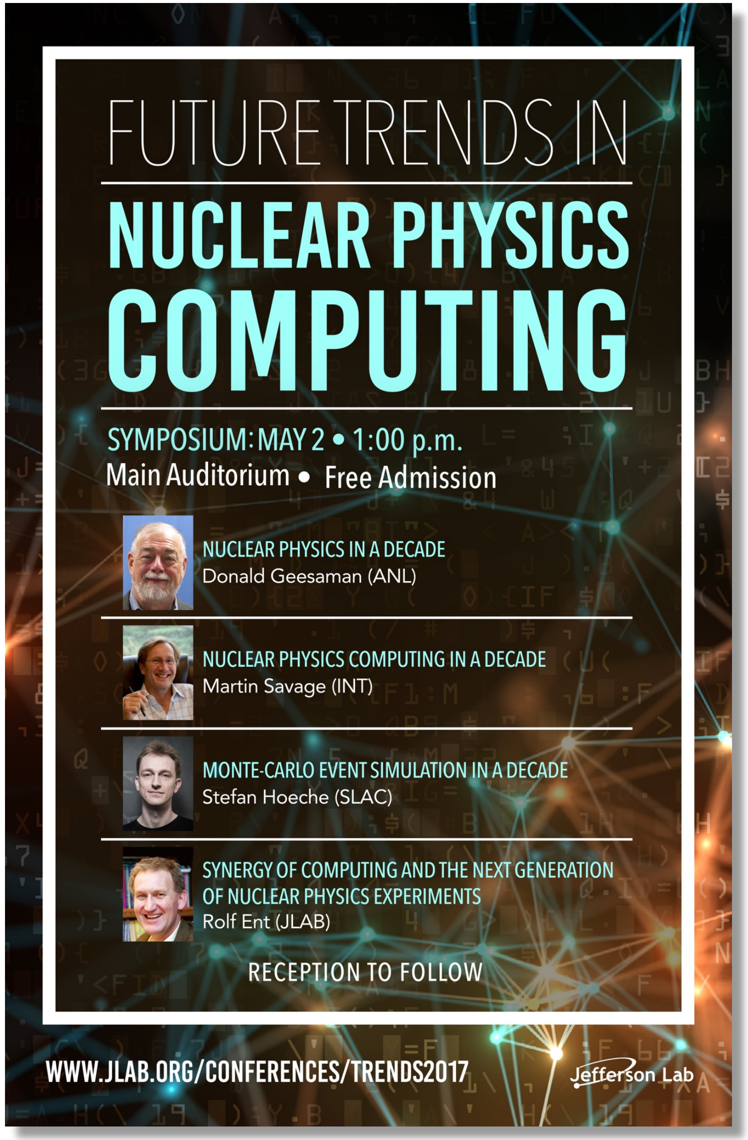
Nuclear Physics is diverse:

- Broad research program with many facilities and experiments.
- Research program extends across a broad range of collaborative scales, in average smaller than HEP.
  - Relatively smaller size of experiments goes along with shorter experimental life cycles and faster changes in scientific goals.

Software and computing efforts are diverse:

- Vary according to collaborative scale, from pragmatic do-it-yourself approaches among a few, to substantial organized software and computing activities within large experiments.
- Relatively smaller group size requires careful planning and design of the software effort:
  - Need to find right balance between in-house development and adoption of common software packages and data management practices.
  - Balancing maintenance and improvement of original software simultaneously with development and incorporation of new tools requires continual attention.
  - Data and analysis preservation for re-producing, re-using, and re-interpreting analyses major challenge.
- New experiments and increasing data volumes drive the need for new approaches to data processing and analysis:
  - Even at small experiments due to rapidly increasing data volumes and processing demands.

The Role of Advanced Computing in Nuclear Physics



Future Trends in Nuclear Physics Computing

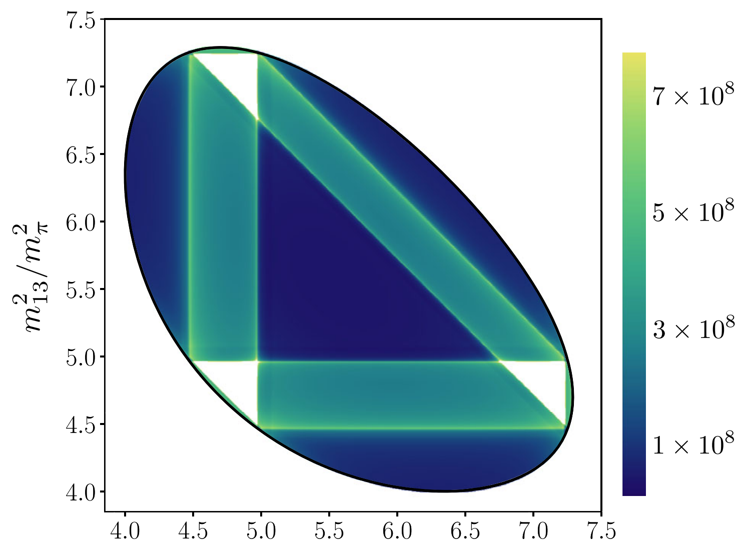
- **Recent years** Discussion about the next generation of data processing and analysis workflows that will maximize the science output.
- **One context for this discussion**
  - Workshop series on Future Trends in Nuclear Physics Computing

Donald Geesaman (ANL, former NSAC Chair) "*It will be **joint progress of theory and experiment** that moves us forward, not in one side alone*"

Martin Savage (INT) "*The next decade will be looked back upon as a **truly astonishing period in NP** and in our understanding of fundamental aspects of nature. This will be **made possible by advances in scientific computing** and in how the Nuclear Physics community organizes and collaborates, and how DOE and NSF supports this, to take full advantage of these advances.*"

- LQCD develops **theoretical, algorithmic, and software tools for lattice QCD**, using cutting-edge HPC systems and exploring the role of AI/ML for lattice QCD.
- It enhances our understanding of heavy-ion measurements at RHIC; nuclear structure studies at JLab, RHIC-spin, and the upcoming EIC; and the search for excited and exotic mesons at JLab.

First nonperturbative QCD calculation of an three-hadron scattering amplitude (π^+, π^+, π^+)



Scattering rate of the resulting amplitude.

Lattice QCD and Data-Intensive Challenges

- Generating snapshots of the QCD vacuum on 4-d lattice:
  - O(1000) configurations of the gluon fields (O(10GB)).
- Solving 4D Dirac equation in each of these configurations:
  - O(1M) individual files (O(1PB)).
- Contractions results in additional O(1M) files.

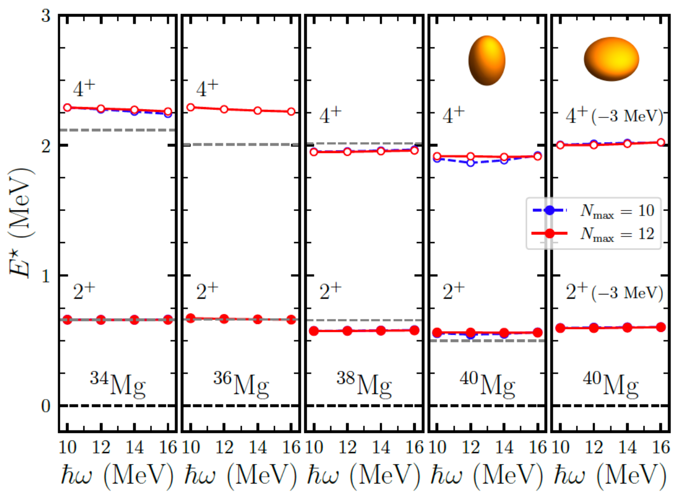
- Online disk in high demand: Reading O(100TB) can saturate the IB to disk if O(100s) jobs are running.

- O(1PB) are available online.
- Rest kept on tape, currently 14PB.

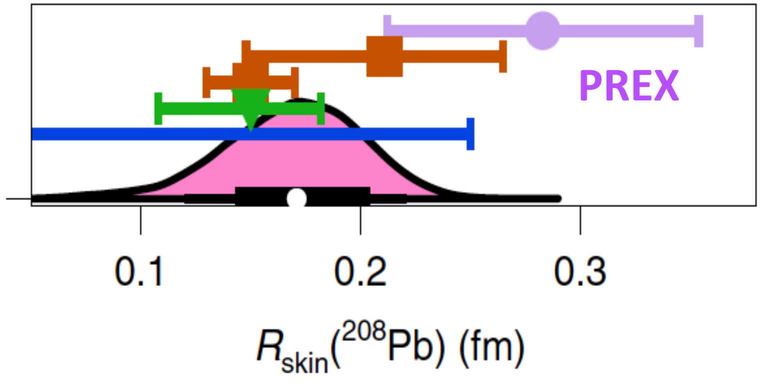
SciDAC-5 project

- **Goal**: Improve theoretical predictions for low-energy NP by:
  - Advanced computation of accurate and precise nuclear interactions and currents using HPC,
  - More sophisticated quantification of uncertainties using data science.
- **Research**: Computational low-energy NP and applied math/computer science.
- Relevant to experimental facilities FRIB, ATLAS at ANL, and JLab and to future 1-ton scale neutrino experiments such as LEGEND.

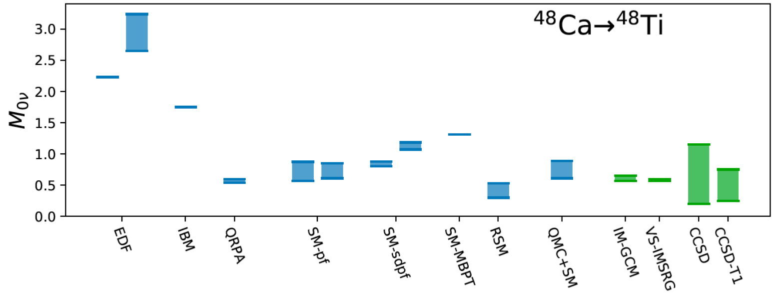
Neutron-rich Mg isotopes are deformed



Neutron *skin* thickness of Pb



Nuclear matrix elements for neutrinoless double beta decay



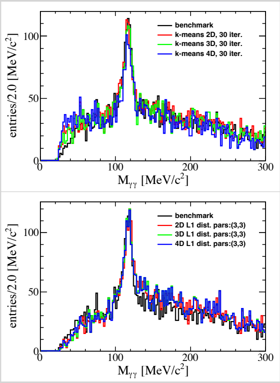
Compute-Detector Integration to Maximize Science

- **Problem** Data for physics analyses and the resulting publications available after O(1year) due to complexity of NP experiments (and their organization).
    - Alignment and calibration of detector as well as reconstruction and validation of events time-consuming.
- **Goal** Rapid turnaround of data for physics analyses.
- **Solution** Compute-detector integration using:
    - AI/ML for autonomous alignment and calibration as well as reconstruction in near real time,
    - Streaming readout for continuous data flow and heterogeneous computing for acceleration.

On-Beam Validation of Streaming Readout at Jefferson Lab (*Eur.Phys.J.Plus* 137 (2022) 8, 958)

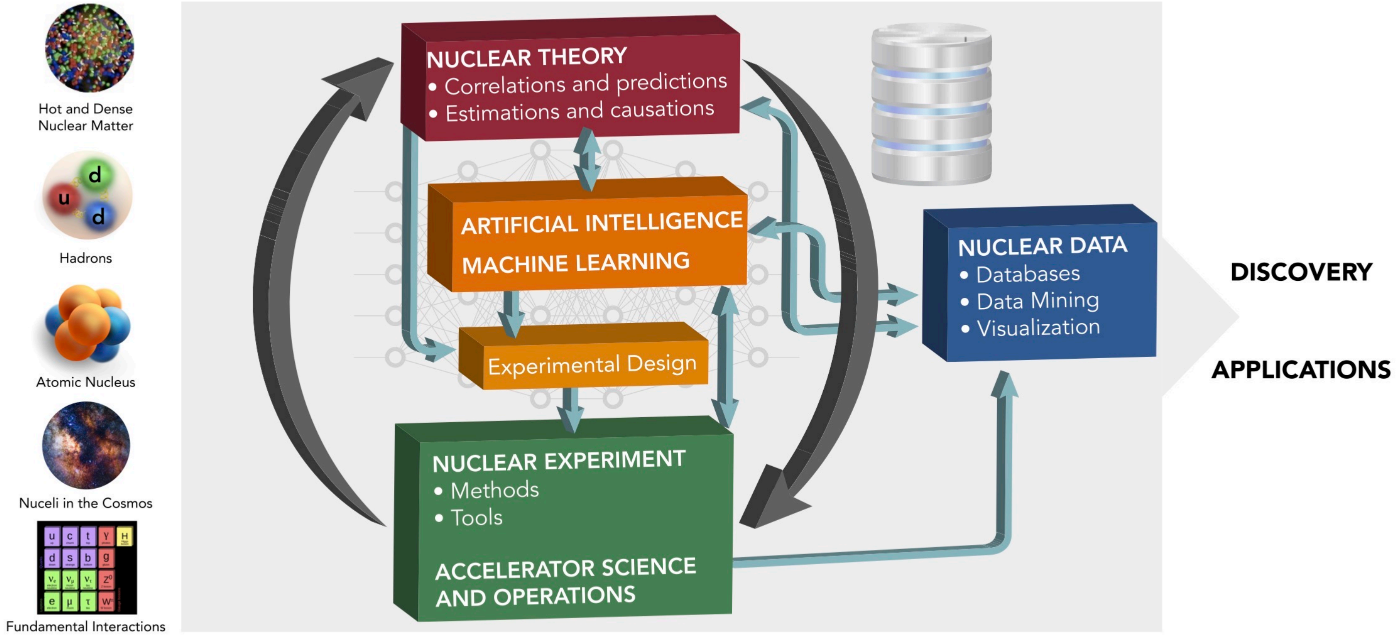
Tests at CLAS12 and GlueX included **AI-supported real-time tagging and selection** algorithms

- Standard operation of **Hall-B CLAS12** with high-intensity electron-beam
    - Streaming readout of forward tagger calorimeter and hodoscope
    - Measurement of inclusive π^0 hadronproduction
- Prototype of EIC PbWO4 crystal EMCAL in **Hall-D Pair Spectrometer**
    - Calorimeter energy resolution of SRQ compatible with triggered DAQ.



Lessons will be learned from the streaming readout at sPHENIX: High-rate processing of streamed data.

AI/ML in Nuclear Physics

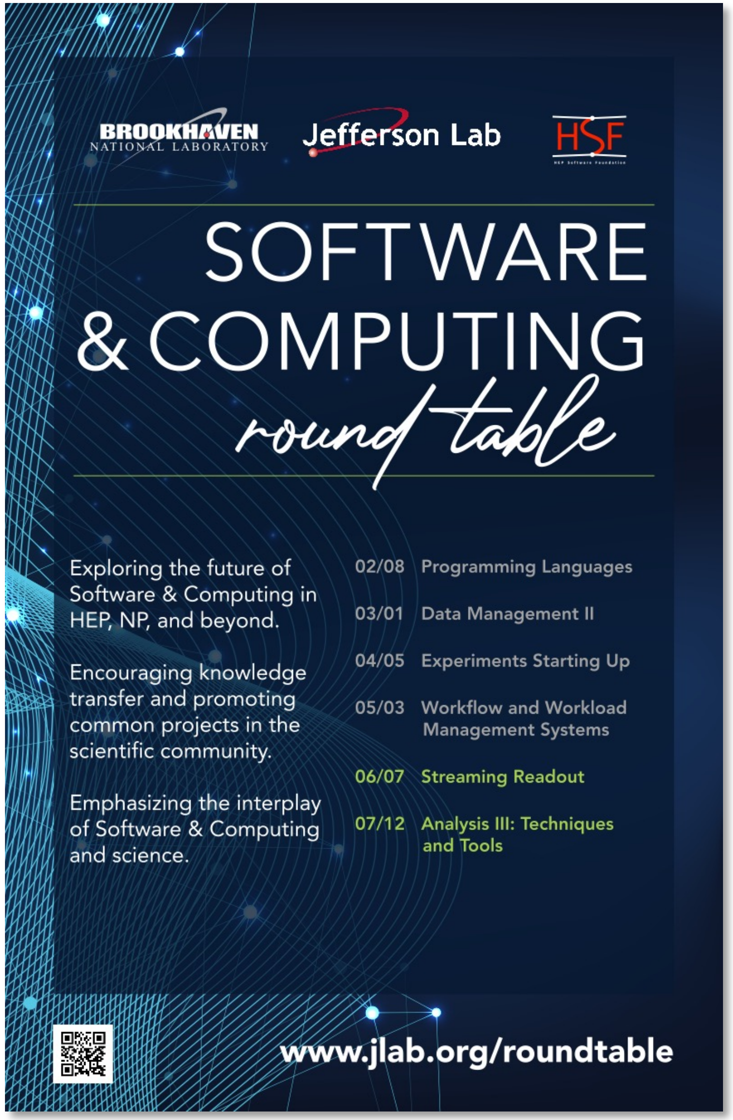


NP is a highly distributed scientific field, utilizing various data types across different scales, making it **ideal for AI/ML applications** (Colloquium Article).

Tremendous interest and activity in AI/ML in NP:

- NP researchers already have the talent and many of the tools required for the AI/ML revolution.
- NP addresses challenges that are not addressed in current technologies.
- NP presents data sets that expose limitations of cutting edge methods.
- **Cross collaboration**: To solve the many complex programs in the field and facilitate discoveries strong collaborations between NP, data science, and industry would be beneficial for all parties.
- **Education** is key to increase the level of AI-literacy – research programs and curricula in data science can help to attract students.

Common Scientific Software



Nuclear Physics Software Community Building

- **Software & Computing Round Table** (BNL, HSF, JLab) explores interplay of computing and science and aims to promote for knowledge transfer and encourage common projects.
- **Participation in HEP Software Foundation** Involvement in MC event generators, frameworks, reconstruction and software triggers, training.

Common Scientific Software – Lessons learned from ACTS and Rucio

- **The team is the most important**: Do not separate development and operations.
- **The project**: Clear, focused short-term goals should align with a sustainable long-term plan that accommodates external collaborators.
- **The management**: Manage expectation to allow the team enough time to achieve success.

Scientific software careers need support

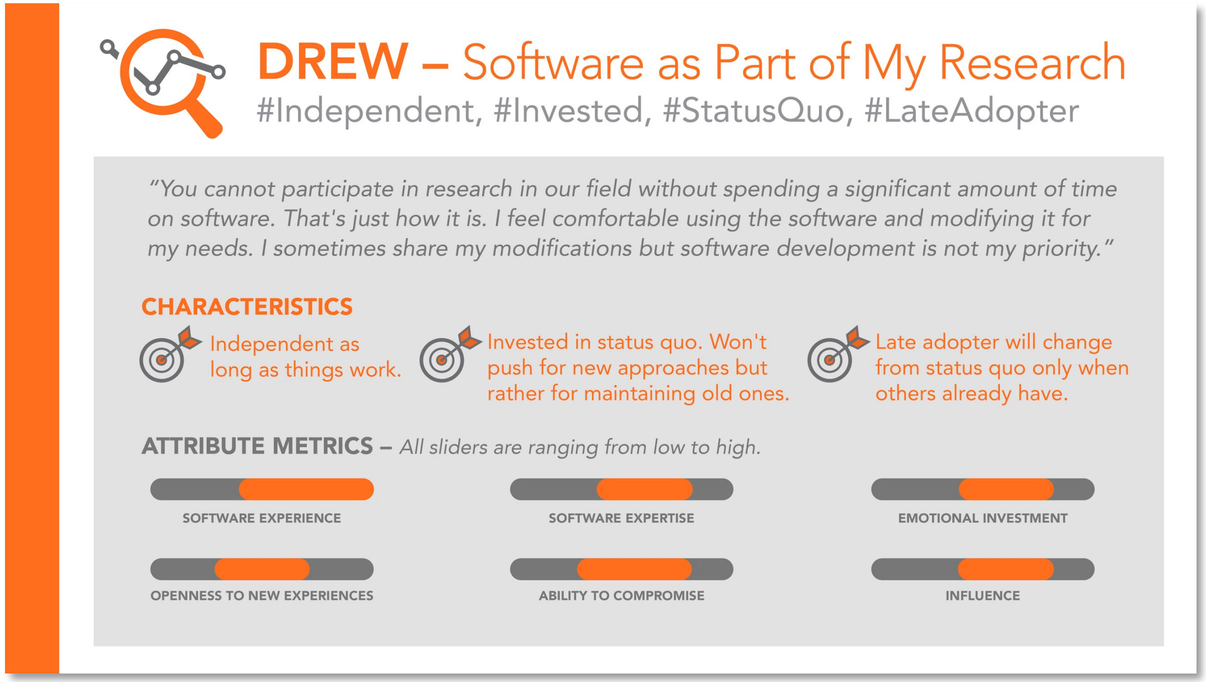
- Support for education and training in software development.
- Provide career paths and funding that allow for and value software development.

Gaining Insight Into the Community

- **Goal**: Enable active participation in physics analysis, regardless of career stage, beyond just students and postdocs.
- **Survey**: On average, 78% of students' and postdocs' research time is devoted to software and computing.

→ **User-Centered Design**: Engage community in development. Listen to users, then develop software.

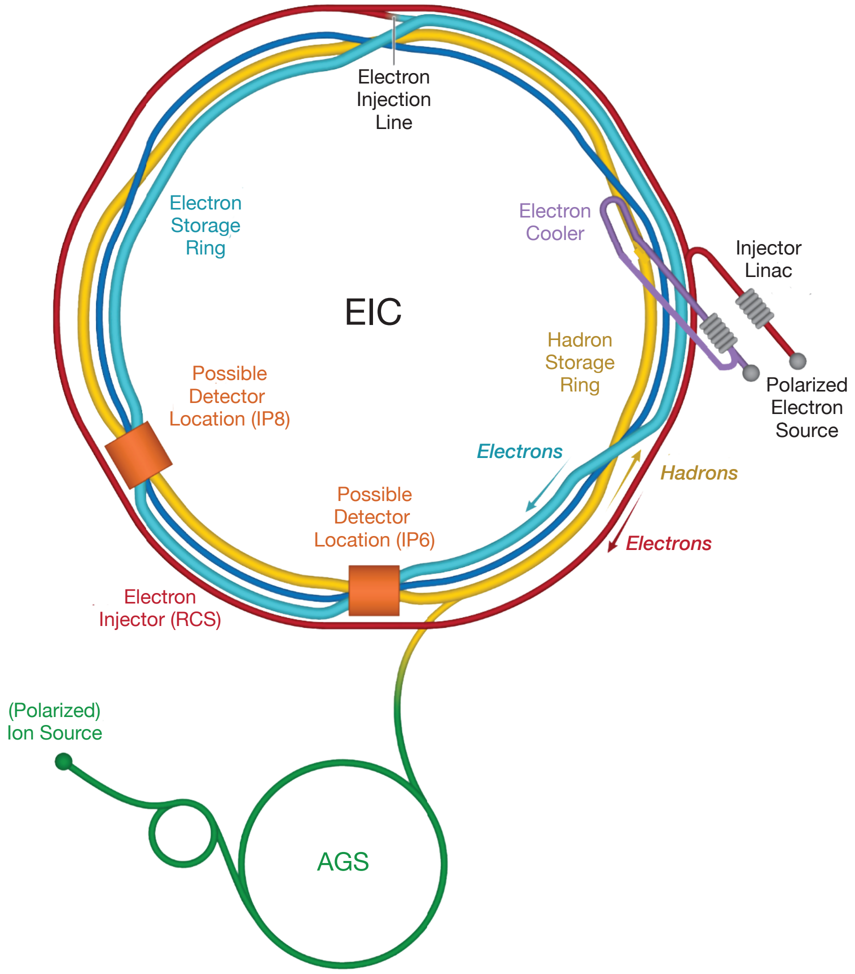
- User archetypes developed on feedback from focus group discussions.
- Input to software developers as to which users they are writing software for:



User Archetypes

Software is not my strong suit.
Software as a necessary tool.
Software as part of my research.
Software is a social activity.
Software emperors.

The Electron-Ion Collider (EIC)



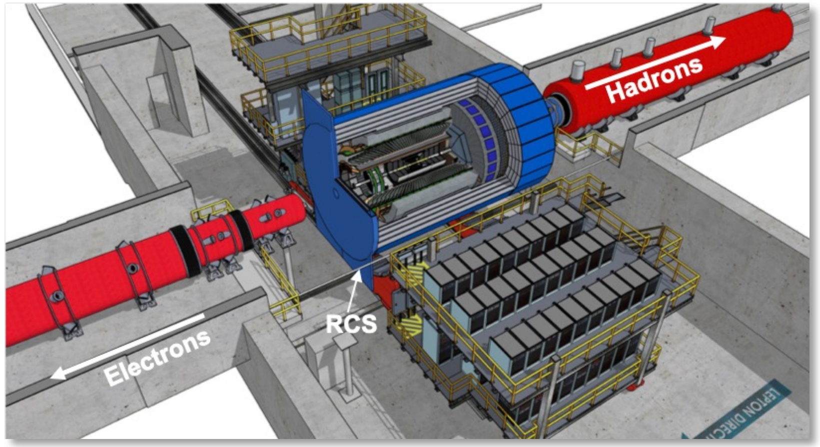
Frontier accelerator facility in the U.S.

- **World's first collider of**:
  - Polarized electrons and polarized protons,
  - Polarized electrons and light ions (d, ^{3}He),
  - Electrons and heavy ions (up to Uranium).

- The EIC will enable us to embark on a **precision study of the nucleon and the nucleus at the scale of sea quarks and gluons**, over all of the kinematic range that is relevant.
- The **EIC Yellow Report** (Nucl.Phys.A 1026 (2022) 122447) describes the physics case, the resulting detector requirements, and the evolving detector concepts for the experimental program at the EIC.

- BNL and Jefferson Lab will be host laboratories for the EIC Experimental Program. Leadership roles in the EIC project are shared.
- EIC operations will start in about a decade.

ePIC (EIC Project Detector)



Front End data → Front End data → Front End data →

Common software stack for online and offline software. Rapid, near-real-time data processing using streaming readout, AI/ML, and heterogeneous computing.

Analysis data →

Expected: 20PB /year

Data format: ROOT (flat)

10 Tb/s

100 Gb/s

- Collider experiment with high luminosity.
- Polarized electron beams off polarized light ions or unpolarized heavy ions.
- Integrated interaction and detector region of ~90m to get ~100% acceptance for all final state particles, and measure them with good resolution (tracking, calorimetry, particle identification).

- **Software stack**: Modular simulation, reconstruction, and analysis toolkit using tools from the NP-HEP community (Geant4 and DD4hep, JANA, EDM4eic and podio, ACTS). C++ and Python.
- **Software design** based lessons learned in the worldwide NP and HEP community, including statement of software principles.

- **AI/ML in production**: N/A.
- **AI/ML in development**: AI-assisted detector design; autonomous control and experimentation.; fast detector simulations integrated in Geant4; reconstruction using holistic detector information.

FUTURE TRENDS IN

NUCLEAR PHYSICS COMPUTING

- **Software & Computing** play an ever-growing role in modern science, including NP, HEP, and related fields.
- As new experiments commence and data volumes rapidly increase, the NP community is exploring the **next generation of data processing and analysis workflows to optimize scientific output**:
  - This includes **streaming readout**, **AI/ML**, and **common scientific software**.
- The next decade promises to be exciting for NP, with diverse scientific programs ongoing at facilities such as CEBAF, FRIB, RHIC, the upcoming EIC, and many others.
- To achieve our goals for next-generation software and computing for NP, we must **work together globally** and **across various fields**.

Thanks for preparing the NP summary: Alexander Austregesilo, Nathan Baltzell, Giordano Cerizza , Mario Cromaz, Robert Edwards, Ole Hansen, Tanja Horn, Jin Huang, Jeff Landgraf, Witold Nazarewicz, Thomas Papenbrock, Jianwei Qiu, Brad Sawatzky, and Brad Sherrill.
Special thanks also to: Amber Boehnlein, Andrea Bressan, Wouter Deconinck, Rolf Ent, Jin Huang, Sylvester Joosten, David Lawrence, Graeme Stewart, Torre Wenaus, and Rik Yoshida.