Brookhaven National Laboratory

U.S. DEPARTMENT OF ENERGY

# Science Council

November 14th 2023

@BrookhavenLab

# Topics

1) LDRD update
2) Computing and data strategy: A plan for a plan
3) FY24 lab agenda
4) Future topics

# LDRD Update

1) **FY25 Type A**: Points of contacts working on presentations in 10 areas. Presentations will be week of 11/20 for down-selection. LD, DDST, DDO and L-1 Science managers to make down selection decision

2) **FY25 Type B**: Call is now out. Proposals due Jan 10th . Presently estimating the size of the “pot”

3) **FY24 Type C**: Umbrella language being developed for BHSO approval

# Data science

- All science is data science: We have to get this right to be successful!

- There are many aspects to getting this right:
  - Compute
  - Storage
  - Network
  - Computer science
  - Software development, production and maintenance
  - Workforce development
  - Cybersecurity
  - Culture

# Developing a strategic plan for data science: The vision

**Data Infrastructure Vision** (needs work!):
To have a seamless, robust and secure data infrastructure enabling large- and small- scale science with leading-edge AI-enabled operations and research

**Data Science Vision** (needs work!):
To be acknowledged leaders in autonomous experiments

# Process

**Phase I**: Self assessment: *Due Jan 1st 2024 (?)*
**Phase II**: Determine guiding parameters for strategic plan: *Due April 1st 2024 (?)*
**Phase III**: Write Strategic plan: *Due October 1st 2024*
(provide input to ALP in Spring 2024)

In parallel:
1) Reconstitute Brookhaven Data Infrastructure Steering Committee (BDISC)
   - Lab-wide body to provide input and help formulate the strategy
2) Form external Advisory Committee
   - Provides advice and feedback along the way

# Phase I: Self Assessment

1. Charge each of ITD, SDCC and CSI with documenting:
    a. What is your mission as you currently see it?
    b. SWOT for your organization in meeting that mission
    c. Benchmark your organization against equivalents at other national labs
    d. Barriers in achieving your mission
2. Charge EPS, NPP, EBNN, ATRO, CSI with: Facilities and Core Depts Separately
    a. What do you need from data science perspective from BNL?
    b. Were there opportunities lost because we did not have computing capabilities
    c. What are the key areas where data science will be needed by you in the next 5-10 years - what is needed to be competitive?
    d. SWOT analysis for Data Science in your organization
    e. In the area of scientific data and computing, for each of ITD, CSI and SDCC document:
        i. What is working
        ii. What are the barriers to working more effectively with that organization
        iii. What would be most helpful to you in enabling your science

# Phase II: Guiding parameters

Form internal committee (BDISC) to answer the following questions for the lab:

a. Building on analysis in phase I, determine areas of computer science R&D where BNL is leading, or poised to take a lead.
b. What are ASCR's priorities in the next 5 years?
c. Where can BNL have the biggest impact for ASCR?
d. What are the possible growth areas in Computer Science (if different from ASCR priorities) and how can we prepare for them?
e. Identify role of BNL in
    i. Developing AI/ML
    ii. Utilizing AI/ML
f. Identify unexploited synergies across existing lab efforts in computing.
g. What security posture is required for different use cases?
h. Determine areas of CS R&D required to support BNL mission-driven computing (development of new technologies or frameworks, as opposed to development required to adopt/adapt existing ones)
i. Identify unexploited synergies/engagements with Stonybrook CS (Institute for Advanced Computational Science) or other partners
j. How should BNL interact with the ASCR facilities including HPDF and LCFs
k. Analysis of on-prem vs off-prem for various use-cases
l. Identify areas of ITD, CSI and SDCC which need improvement in order to support the lab mission successfully

# Phase III: Strategic Plan

Strategic plan should include:

a. Vision for data science at BNL
b. Strategy to support to support small experiments and theory
c. Strategy to support large scale facilities
d. Strategy for computer science at BNL
e. Cross-cut strategies
    i. Role for HPDF and other LCFs
    ii. Role of cloud and on-prem compute and storage
    iii. Cyber security
    iv. AI/ML strategy
    v. Workforce development
    vi. Infrastructure development
        1. Compute
        2. Storage
        3. Network
        4. Policy
        5. ..
f. Identify the organizational structure that would best position the lab for achieving its data strategy.
g. Timeline with milestones to execute strategy

# Advisory Committees

**Internal**: Brookhaven Data infrastructure Steering Committee (BDISC)
Needs to be reconstituted with a new charge.
Current membership:

1.Stuart Campbell (NSLS-II, Current Chair)
2.Jamie Dunlop (past Chair)
3.Kevin Brown
4.Srini Rajagopalan
5.Alejandro Sonzogini
6.Eric Lancon
7. Kerstin Kleese van Dam
8. Tom Schlagel David Salbego)
9. Sioan Zohar (now left the lab)

**External:** Yet to be formed. Bi-annual? Needs a charge.
Some suggested names:

1. Jeff Nichols (ORNL retired)
2.Frank Alexander
3.Jeanette Wong (Columbia VPR)
4. Alun Ashton (PSI)
5. Gulzow Volker (DESY) Or new person
6. Amedeo Perazzo (SLAC)
7. Georgia Tourassi (ALD for CS, ORNL)

# BNL AI Policy Working Group

Kerstin Kleese van Dam (Chair)

Kevin Yager, Daniel Hornback, David Asner, David Salbego, Jessica Wilke, Dan Forino, Stuart Wilkins, Michiko Minty, Torre Wenaus, Ian Ballantyne, Noel Blackburn, Thomas Schlagel

**Charge:**

1. Carry out a brief survey of AI policies from other national labs (to the extent that the exist). Could be broadened to look at university examples too, where that was deemed useful.
2. Develop a simple “lite” policy for BNL. Expected to be short, on the order of 1 page. It should:
   a. Define what we mean by AI to be covered by the policy
   b. Guiding principles for AI at BNL
   c. Guidance on use of generative AI such as ChatGPT
3. Identify issues that should be addressed in the future as we develop more detailed policies.

**Due: November 21st 2023**

# AI: White House Executive Order


October 30th 2023

Very long order, with many charges. One is:

Within 270 days of the order, the Secretary of Commerce, in coordination with the Secretary of Energy and Secretary of Homeland security will
develop guidelines, standards and best practices for AI safety and security to help ensure the development of safe, secure and trustworthy AI systems

# FY24 Lab Agenda

- Previous lab agenda's followed ALP and were supposed to be tactical – what we would do in the coming year(s) to execute our strategy. It was also supposed to be trackable and to feed the operations business plans.
- But there was no ALP in FY23 and lab agenda had become cumbersome and not so useful
- In FY24 we will therefore do a truncated version. Each science organization will write an agenda that is useful to *them*.
- One page per initiative. Org's define their own initiatives. Some will be lab initiatives. No fixed number.
- Audience for agenda is internal: Science organization's management and staff, other science organizations, lab management, operations organizations
- **Due**: Dec 22nd and then iterated on

# FY24 Lab agenda template

1 page per initiative:

1. Description of initiative:
2. Strategic Statement:
3. Metric: how do we measure progress on this task
4. Potential: what is the impact expected on the laboratory if successful? This might include science/technology impact in addition to laboratory program impact. What sponsor(s) might this appeal to?
5. Scale of effort:
6. Point of contact:
7. Objectives for FY24:
8. Stakeholder Engagement tasks:
9. Lab support needed:
10. Team Needed:

# Planning process (towards FY25 Strategic Plan)

Activities currently underway:

- FY24 Agendas
- FY25 LDRD Type A down-selection process
- Computing and data strategy
- Accelerator strategic planning exercise

Each will be input for CY24 strategic retreats and ultimately FY25 ALP and beyond (next slide)

# Future Science Council Topics (CFN list)

1. Institutional Computing
   - Institutional level infrastructure for data/meta-data databases, archiving
   - Policy towards open data (inc user agreements)
   - Open source policy and culture
2. AI/ML strategy
   - Resolution of bureaucracy (reqs for cloud access)
3. Strategic Planning
   - Want to be more proactive, planning across the lab for future strategic science directions
4. Strategic partnerships (i.e at the institutional level)
5. Post-docs
   - Costs
   - Hiring
   - Recruitment (“soft” incentives)

# Backup