Connecting Science with the DOE's Integrated Research Infrastructure

NeRSC

HEP/HPC Strategy Meeting
30th Jan 2025

Debbie Bard
Department Head, Science Engagement
and Workflows, NERSC
Chair, IRI Leadership Group

The Integrated Research Infrastructure (IRI) is a new program encompassing the whole ASCR Facilities ecosystem (and potentially, the broader research computing ecosystem).

IRI is being designed to create better interfaces between DOE user facilities. It will include **frameworks for integration in software, policy and governance**.

It will layer on top of existing (and future) facility capabilities, to enhance integration and usability/user experience.

IRI is the means by which we will manage the ASCR ecosystem and the integration that is required to support Office of Science priorities.

IRI has been developing over the past 5 years. The program launched 1 year ago.

FY 2021 President's Budget Request includes **Integrated** Computation and Data Infrastructure Initiative

FY 2024 PBR advances IRI and the **High Performance Data Facility**

HPDF Selection

ASCR IRI Task Force launch

ASCR IRI Task Force report

SC IRI Blueprint Activity launch

IRI Blueprint Activity results

IRI Program Development

GO

Jan **2020** Jan **2021** Jan **2022** Jan **2023** Jan **2024**
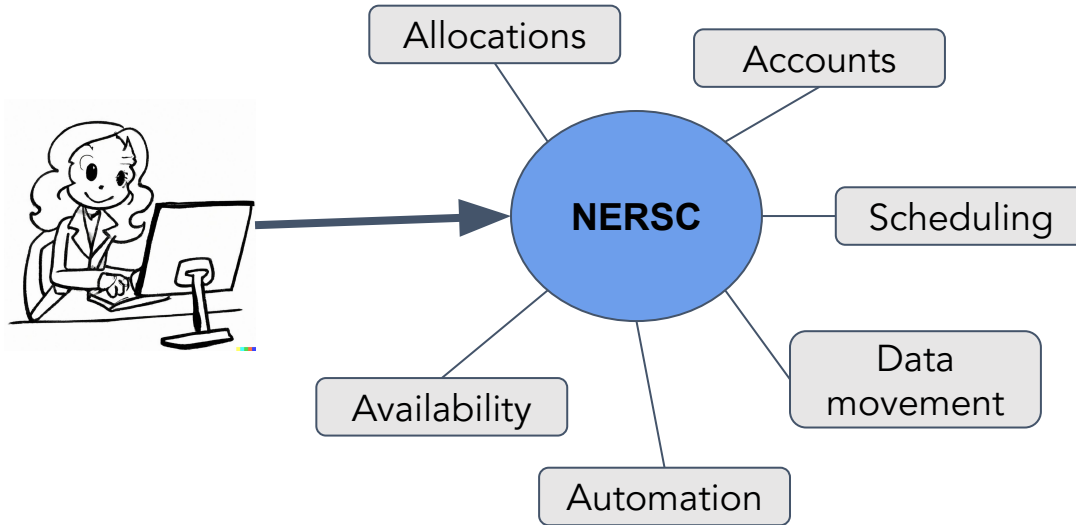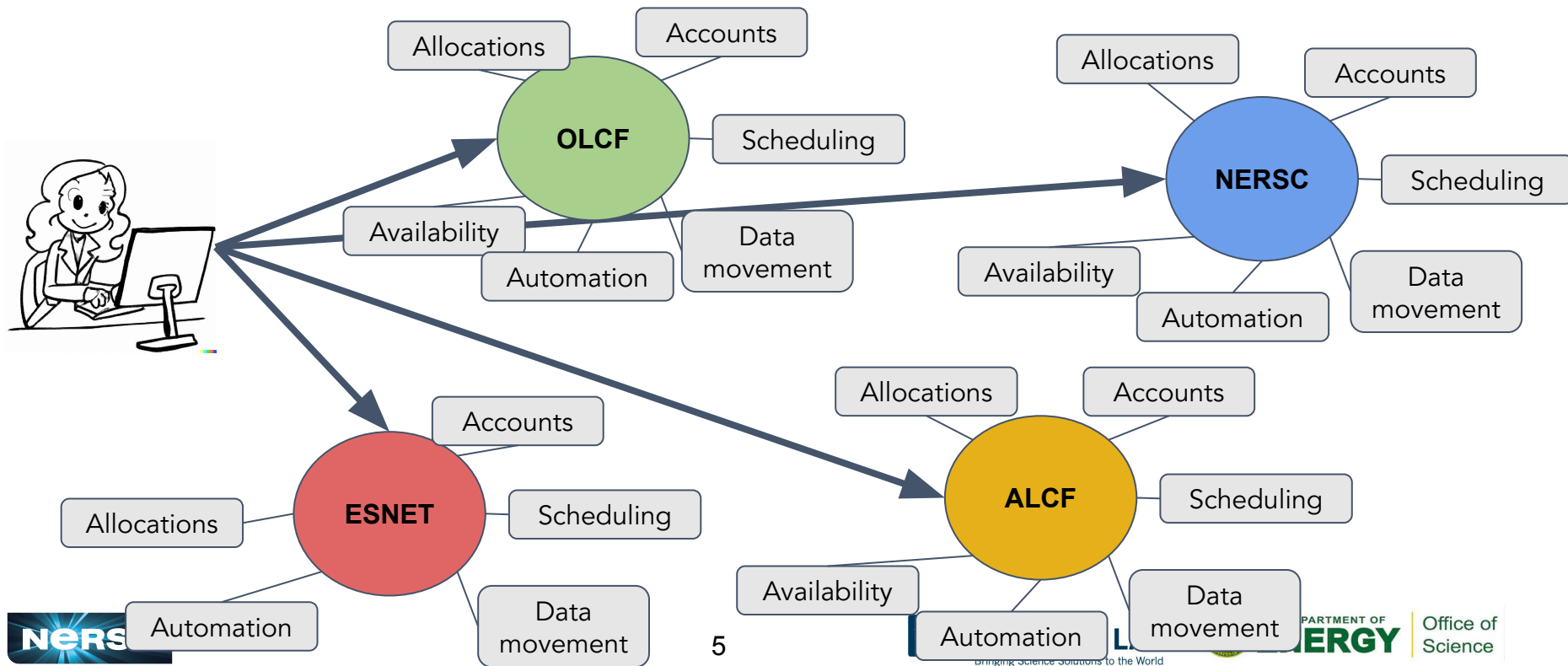
Vision **Strategy** **Implement**

In an Integrated Research Infrastructure, simple things should be simple, and complex things should be possible

In an Integrated Research Infrastructure, simple things should be simple, and complex things should be possible
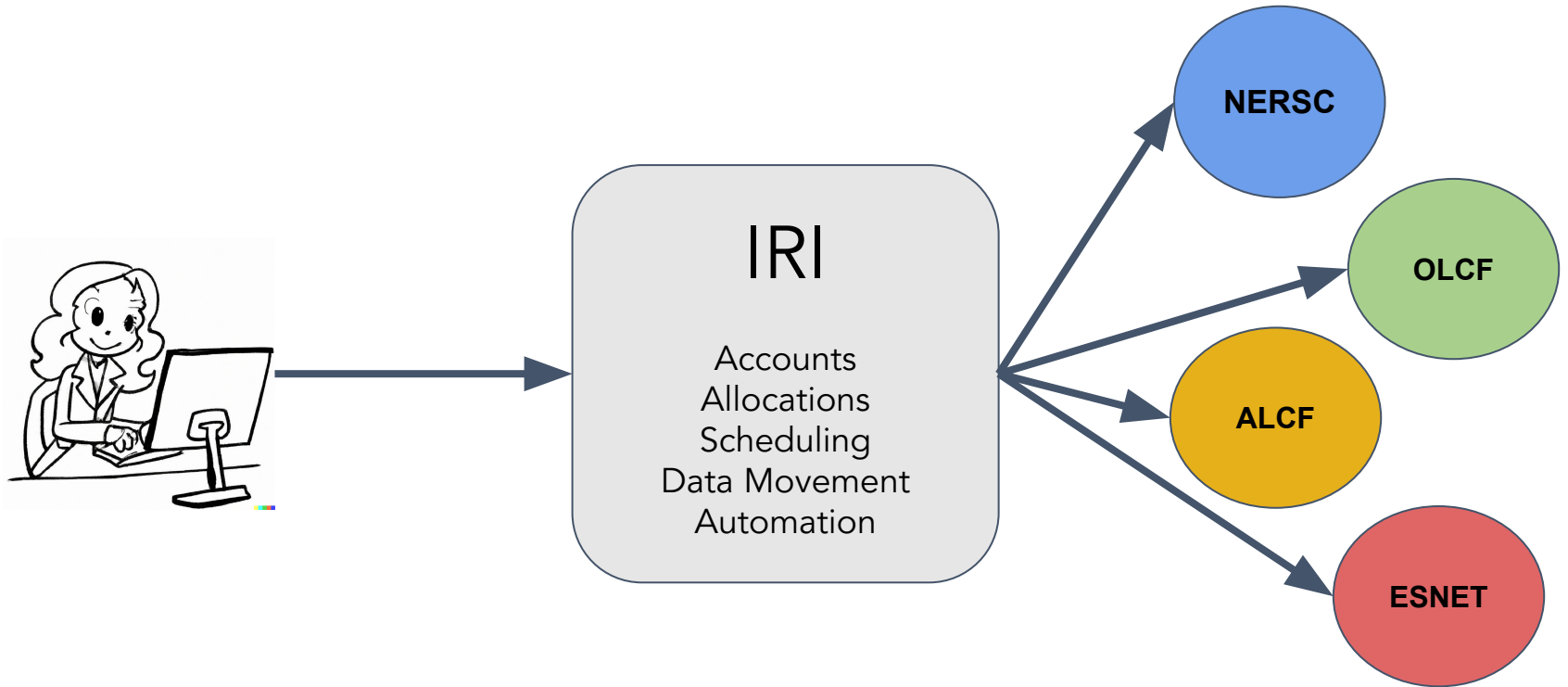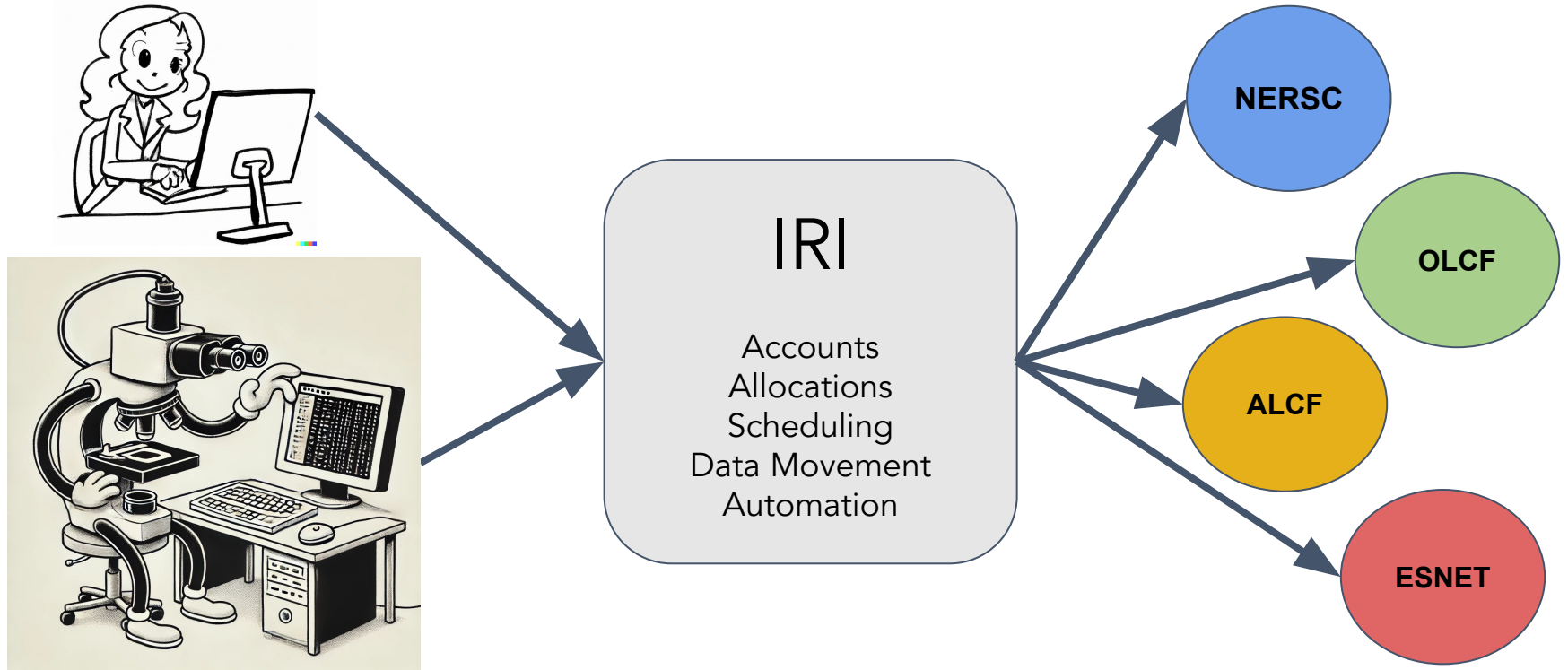
In an Integrated Research Infrastructure, simple things should be simple, and complex things should be possible

In an Integrated Research Infrastructure, simple things should be simple, and complex things should be possible

IRI is creating the OS for the DOE Ecosystem

IRI is an "operating system" that layers on top of existing facilities.

Facilities will need to make changes to support IRI software, policies and processes, but IRI does not dictate the physical hardware to be deployed at each site.

Facilities retain their unique mission and capabilities.

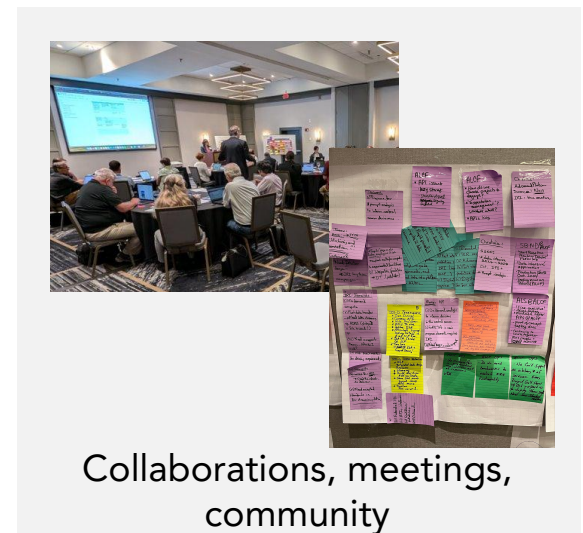| **Software**	**Policy Alignment**	**Governance**
Deployed across facilities to provide common interfaces and services	Security, allocation and other policies that enable and enhance integration	Transparent, open processes for developing and ratifying standards and practices

Coordinated Engagement
Engagement with facilities, projects and users to understand requirements, co-design solutions and develop best practices.

NeRSC

BERKELEY LAB
Bringing Science Solutions to the World

U.S. DEPARTMENT OF ENERGY | Office of Science

Pathfinder Projects

Architecture Blueprint Activity

Collaborations, meetings, community

IRI technical work is driven by requirements identified in the Architecture Blueprint Activity, partnership with Pathfinder projects, **and** many collaborations with science teams doing IRI-relevant work.

IRI technical work is coordinated via Technical Subcommittees, with membership across facilities

IRI Program Area	Status
IRI Allocations Program	Under development
Scheduling/Preemption Technical Subcommittee	Under development
TRUSTID Technical Subcommittee	Launched!
Outreach/Engagement Technical Subcommittee	Launched!
Interfaces Technical Subcommittee	Launched!
Data Movement Technical Subcommittee	Under development
Software Deployment and Portability	Under development

IRI technical work is managed via Technical Subcommittees

IRI Management Council

- IRI Executive Subcommittee
- IRI Leadership Group
- IRI Technical Subcommittees / IRI Outreach/ Engagement subcommittee

Their role is to identify goals and coordinate work in a particular technical area.

Outreach & Engagement Technical Subcommittee

Launched May 2024

- **Liaise with Pathfinder partners** to help demonstrate cross-facility workflows using early IRI frameworks
  - Coordinate work between Pathfinder teams and IRI
  - Identify and develop requirements and requirements strategy
- Develop and execute a plan for **broader community engagement**

Co-chairs: Eli Dart and Rafael Ferreira Da Silva. Members: Johannes Blaschke, Dale Carder, Jack Deslippe, Drew Paine, David Rogers, Katherine Riley, Tom Uram.

Trusted IRI Designs (TRUSTID) Technical Subcommittee

- Define <u>federated ID</u> **design patterns** to enable trusted interoperable cross-facility workflows within the IRI ecosystem
- Promote **interoperability** and use of existing standards
- Create documentation, best practices, and **reference implementations**

Convened by Carol Hawk. Co-chairs: Ryan Adamson and Tiffany Connors. Members: Bill Allcock, Carl Bai, Scott Campbell, Maxim Kowalski, Jeff Neel, John Riney, Adam Slagell, Veronica Vergara.

Interfaces Technical Subcommittee

- Design a **minimal functional API** and deploy it at multiple sites
  - Review existing API schemas
  - Propose IRI schema: endpoints, architecture, infrastructure,...
  - Implement MVP
- Explore how to align **Jupyter** across sites

Co-chairs: John MacAuley and Ryan Prout. Members: Bill Allcock, Ilya Baldin, Bjoern Enders, Paul Rich, Addi Thakur, Gabor Torok, Xi Yang

The ASCR Facilities are firmly embedded in the IRI ecosystem

IRI is permeating everything we do. **Multi-facility workflows** are an integral part of our major infrastructure design and strategic planning.
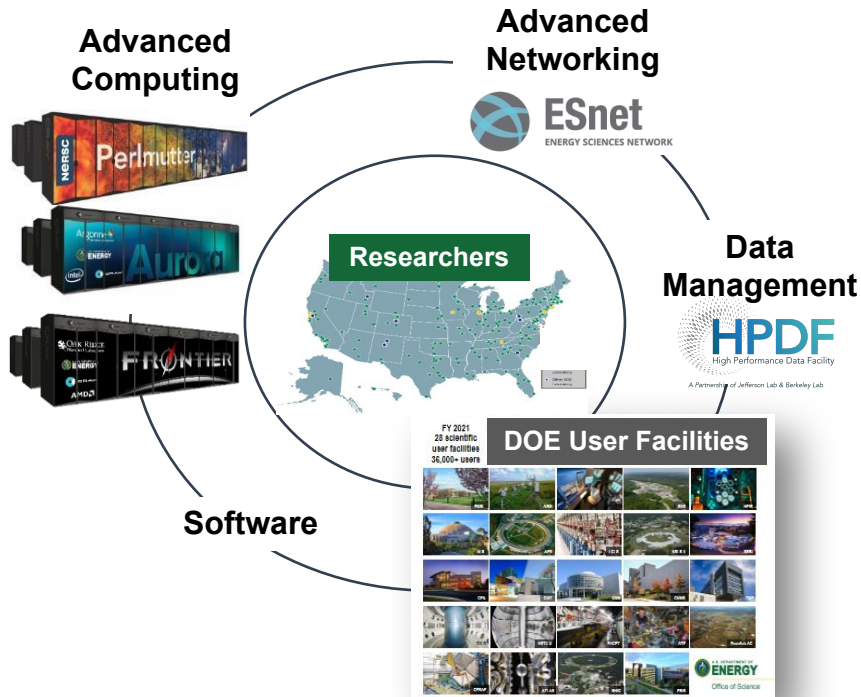
- ◆ ASCR facilities have made a lot of progress supporting complex workflows

- ◆ Ongoing projects, testbeds, and major acquisitions are targeting IRI (ESnet6, NERSC-10, OLCF-6, ALCF-4 and HPDF)
  - • KPPs and technical requirements explicitly mention IRI

- ◆ IRI requires a growing amount of collaboration across ASCR facilities - it's changing how we work together

IRI will provide a foundational infrastructure for next-generational experimental science, post-exascale computing, AI/FASST and other initiatives.

In FY25, we will take the first steps towards a fully-featured IRI framework

- Develop the first IRI frameworks: guidelines for federated ID and a prototype API
- Iterate via test and development cycles between Technical Subcommittees (TSs) and Pathfinder partners
- IRI Pathfinders use prototypes on testbeds and in a limited way in production
- Expand the IRI science community via an IRI Community Group
- Resources permitting, expand participation in all aspects of IRI work.

By the end of FY25, IRI will have demonstrated the first capabilities at multiple ASCR facilities, and laid the groundwork for future IRI work.

Visit iri.science to keep up to date with our work!

DOE's Integrated Research Infrastructure (IRI) will seamlessly link DOE's experimental and observational scientific user facilities, data assets, and advanced computing resources so that researchers can combine these tools in novel ways that radically accelerate discovery
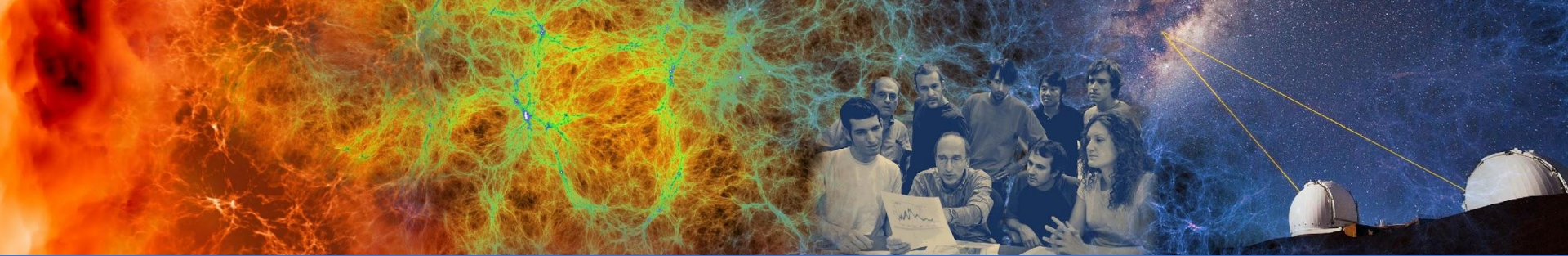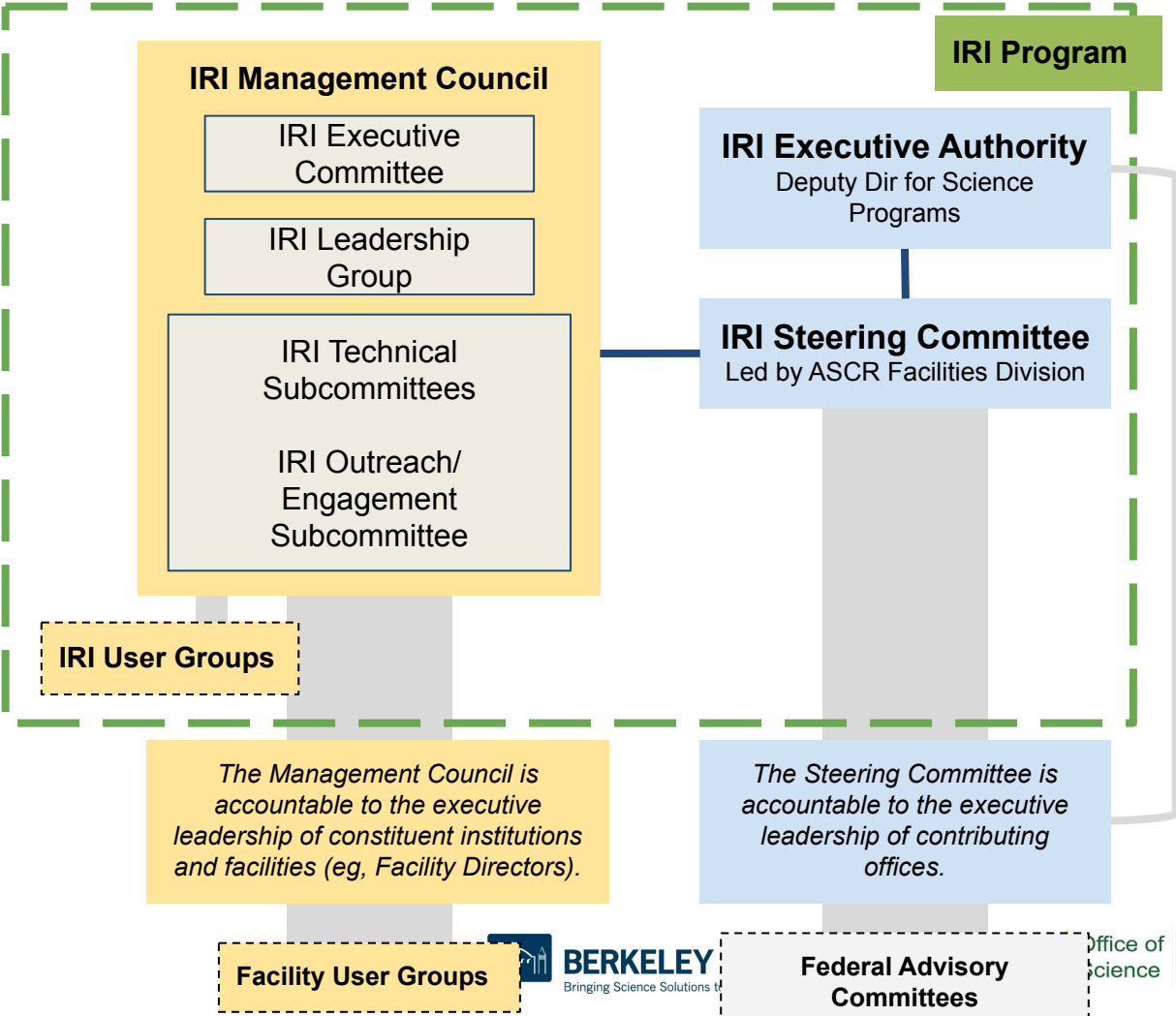
Extra slides

The IRI Governance model

IRI Program

IRI Management Council

- IRI Executive Committee
- IRI Leadership Group
- IRI Technical Subcommittees
- IRI Outreach/ Engagement Subcommittee

IRI Executive Authority
Deputy Dir for Science Programs

IRI Steering Committee
Led by ASCR Facilities Division

IRI User Groups

The Management Council is accountable to the executive leadership of constituent institutions and facilities (eg, Facility Directors).

The Steering Committee is accountable to the executive leadership of contributing offices.

Facility User Groups

Federal Advisory Committees

BERKELEY
Bringing Science Solutions t

Office of Science

NERSC

IRI has been developing over the past 5 years. The program launched this year.

Identify the problem.
What's needed?
What's been tried,
what can we learn?
What's in it for me?

Community engagement

Useful

Translate IRI concepts into real software frameworks and policies.

Iterate with partners

Useable

Deploy IRI frameworks on production systems with a long-term support model.

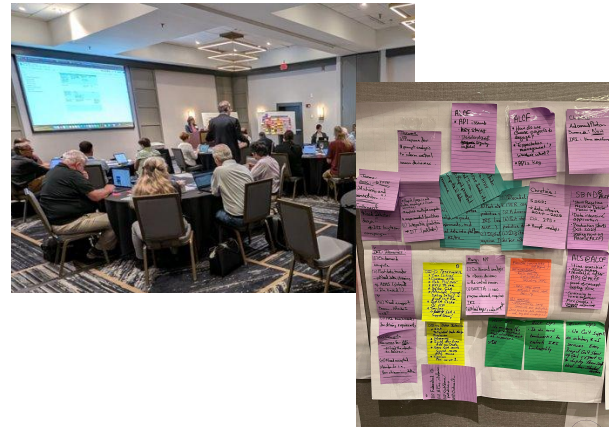*Follow-up and follow-through with community*

Used

We have held multiple rounds of requirements gathering, prioritization and review with the DOE community

Identify the problem.
What's needed?
What's been tried,
what can we learn?
What's in it for me?

Community engagement

Useful

THE DOE OFFICE OF SCIENCE

Integrated Research Infrastructure Architecture Blueprint Activity

FINAL REPORT
2023

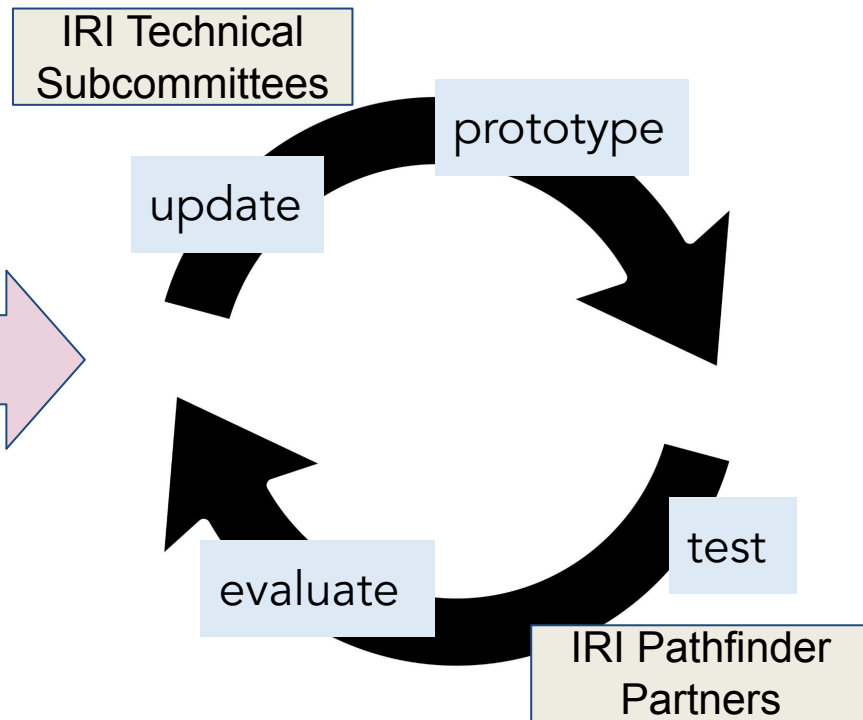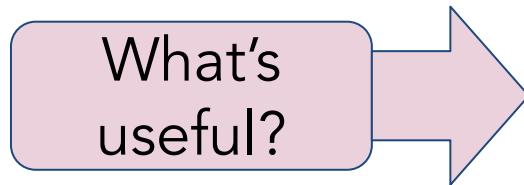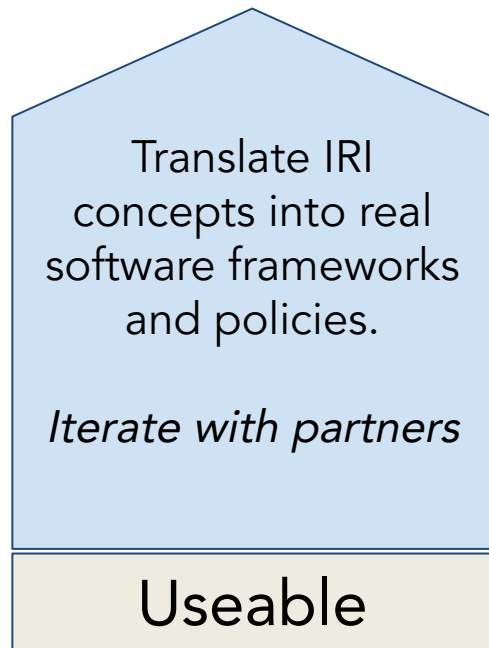ASCR Integrated Research Infrastructure Task Force   March 8, 2021

Toward a Seamless Integration of Computing, Experimental, and Observational Science Facilities: A Blueprint to Accelerate Discovery
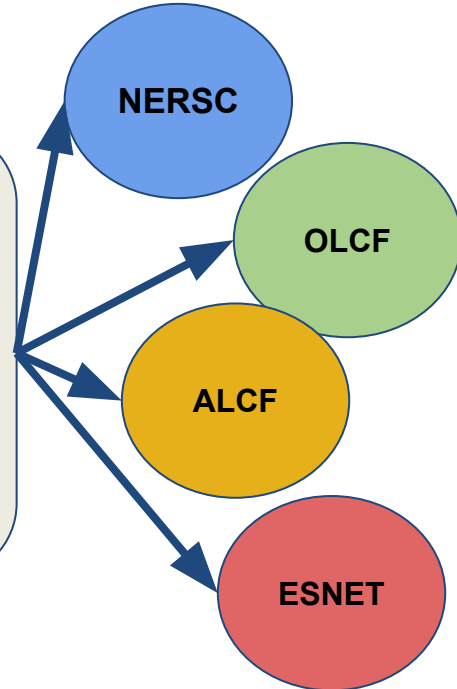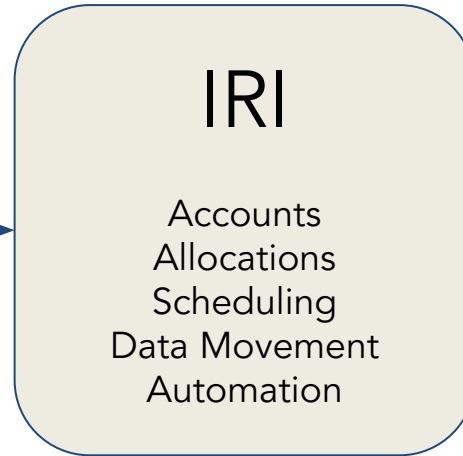
About the ASCR Integrated Research Infrastructure Task Force

There is growing, broad recognition that integration of computational, data management, and experimental research infrastructure holds enormous potential to facilitate research and accelerate discovery.[1] The complexity of data-intensive scientific research—whether modeling/simulation or experimental/observational—poses scientific opportunities and resource challenges to the research community writ large.

We are developing useable IRI infrastructure via iterative co-design

Translate IRI concepts into real software frameworks and policies.

Iterate with partners

Useable

What's useful?

IRI Technical Subcommittees

prototype

update

test

evaluate

IRI Pathfinder Partners

We aim to deploy the first IRI software and policy frameworks in the next year, but we're planning for long-term, sustainable support

Deploy IRI frameworks on production systems with a long-term support model.

Follow-up and follow-through with community

Used

IRI

Accounts
Allocations
Scheduling
Data Movement
Automation

NERSC

OLCF

ALCF

ESNET
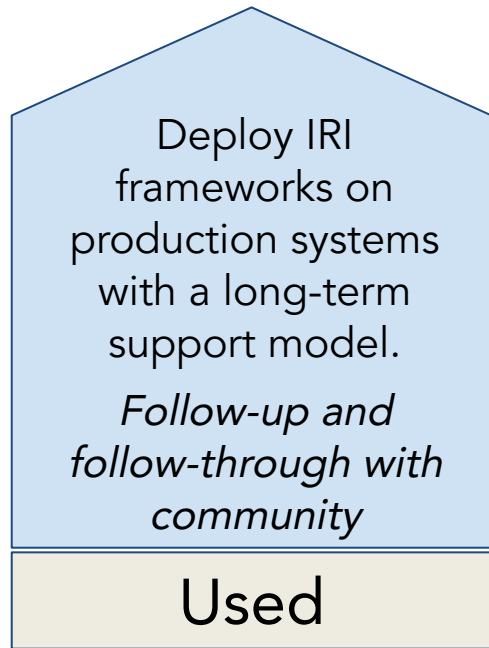
Pathfinder Projects - "It takes a village …."

The IRI program has identified "pathfinder projects", nominated by their program managers as ready to partner with us in the first stages of developing IRI technologies. Pathfinders will partner closely with IRI to:

- Facilitate collaborative co-design to leverage expertise across multiple disciplines
- Demonstrate feasibility to validate initial ideas and technical capabilities
- Promote knowledge sharing to accelerate paths to solutions

DIII-D

ESGF

BES Light Sources

In December 2023, the SC Director charged the Office of Science to conduct an assessment of major projects within the Office

The Chairs of the SC Federal Advisory Committees were asked to "look toward the scientific horizon and identify what new or upgraded facilities will best serve our needs in the next ten years (2024-2034)."

This charge was only the third time the Office of Science (SC) conducted an SC-wide projects assessment.

Department of Energy
Office of Science
Washington, DC 20585

Office of the Director

December 1, 2023

To: CHAIRS OF THE OFFICE OF SCIENCE FEDERAL ADVISORY COMMITTEES:
  Advanced Scientific Computing Advisory Committee
  Basic Energy Sciences Advisory Committee
  Biological and Environmental Research Advisory Committee
  Fusion Energy Sciences Advisory Committee
  High Energy Physics Advisory Panel
  Nuclear Science Advisory Committee

The Department of Energy's Office of Science (SC) has envisioned, designed, constructed, and operated many of the premiere scientific research facilities in the world. More than 38,000 researchers from universities, other government agencies, and private industry use SC User Facilities each year—and this number continues to grow.

Stewarding these facilities for the benefit of science is at the core of our mission and is part of our unique contribution to our Nation's scientific strength. It is important that we continue to do what we do best: build facilities that create institutional capacity for strengthening multidisciplinary science, provide world class research tools that attract the best minds, create new capabilities for exploring the frontiers of the natural and physical sciences, and stimulate scientific discovery through computer simulation of complex systems.

The ASCAC Facilities charge report emphasized the importance of an integrated facility ecosystem for DOE science

Recommendation 2: Science demands integration. We advocate viewing ASCR facilities not as isolated entities, but as integral components of a single, larger integrated computational *ecosystem...*, with a single governance model.

... Further, this integrated ecosystem is required for programs of other agencies, and industry. Its critical role in bolstering national scientific and technological capabilities, as well as its status as a model internationally, cannot be overstated.

IRI has achieved a lot in 2024!

- Stood up the **management structures** to design and coordinate IRI work
- Stood up **3 Technical Subcommittees**, who have made real progress in designing the first IRI capabilities
- Started collaborating with our **Pathfinder Project** partners
- Held a successful **kick-off meeting** in DC in July, getting feedback from scientists and Program Managers across the Office of Science
- Strong **IRI presence** at multiple meetings, conferences and workshops, including a very successful SC!
- Made the case for a more integrated infrastructure, and **gained support** for IRI within DOE…