

**US RESEARCH SOFTWARE
SUSTAINABILITY INSTITUTE
(URSSI)**

NSF S²I² CONCEPTUALIZATION

APRIL 10, 2018

VIPIN CHAUDHARY
RAJIV RAMNATH
PROGRAM DIRECTORS
CISE/OAC



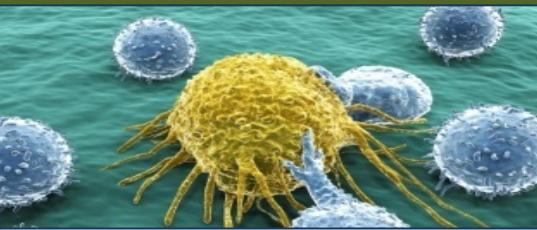
UNIVERSITY OF CALIFORNIA
BERKELEY, CA

URSSI: US Research Software Sustainability Institute

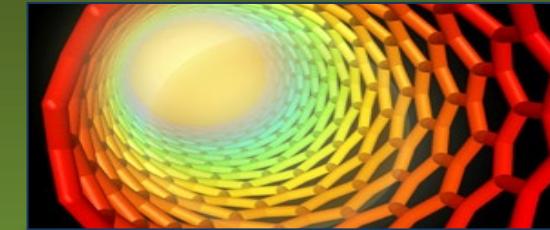
Program Directors' Perspectives



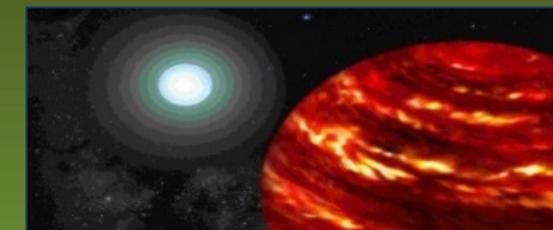
NSF Champions Research and Education Across All Fields of Science and Engineering



Biological Sciences



Engineering



Mathematical & Physical Sciences



Computer & Information Science & Engineering



Geosciences



Integrative Activities



Education & Human Resources



Social, Behavioral & Economic Sciences



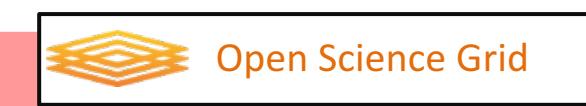
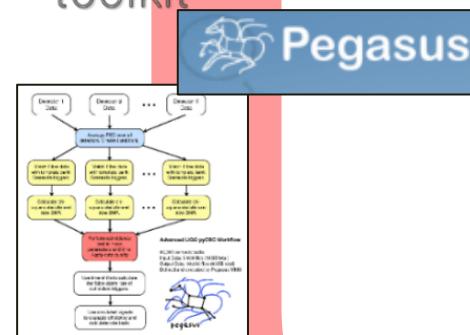
International Science and Engineering

CI enables Big Science

Gravitational wave detection enabled by NSF investments across the computational and data science workflow



einstein
toolkit



✓ Researcher access to sustained Advanced Computing resources

- New intensive simulations of relativity and magnetohydrodynamics. Massive, parallel event searches and validation (100,000 models).
- Advanced computing resources and services sponsored by NSF, DOE, and commercial cloud services.

✓ Interoperable Networking, Data Transfer, & Workflow Systems

- Pegasus, HTCondor, Globus workflow and data transfer management
- NSF funded 100 Gbps upgrades enabled huge throughput gains.

✓ Software Infrastructure

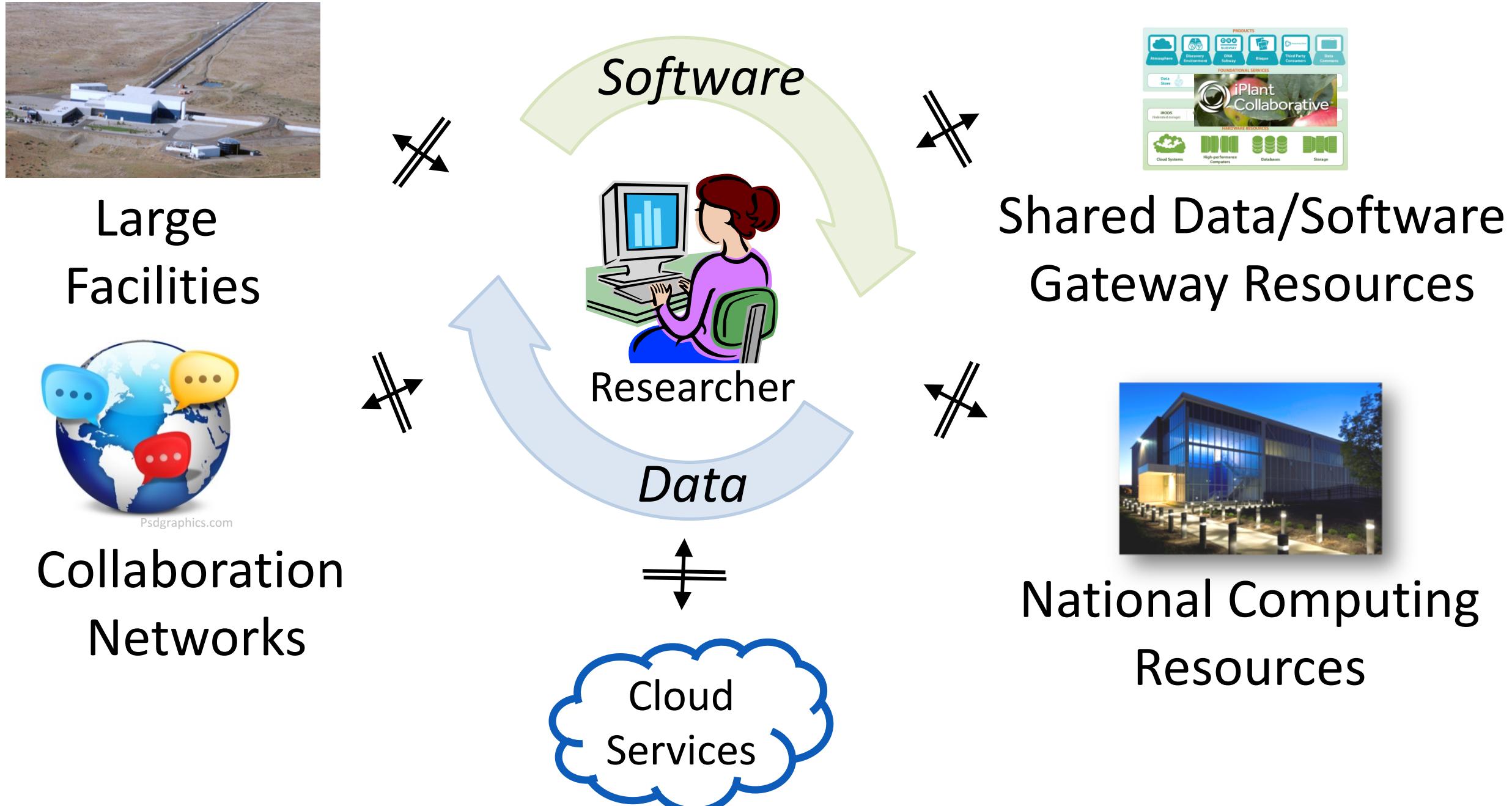
- Computational science advances embodied in Software Infrastructure, for simulations, visualizations, workflows and data flows

NSF programs: Data Building Blocks (DIBBs), Software Infrastructure (SI2), Campus Cyberinfrastructure Network Infrastructure and Engineering (CC*NIE, DNI), and others. OSG and Pegasus are also supported by the Dept of Energy.



CI Challenge: User-Centric Viewpoint

Revolution in the scientific workflow: many interfaces to shared services



CI-rich discovery pathways and communities demand diverse and dynamic approaches

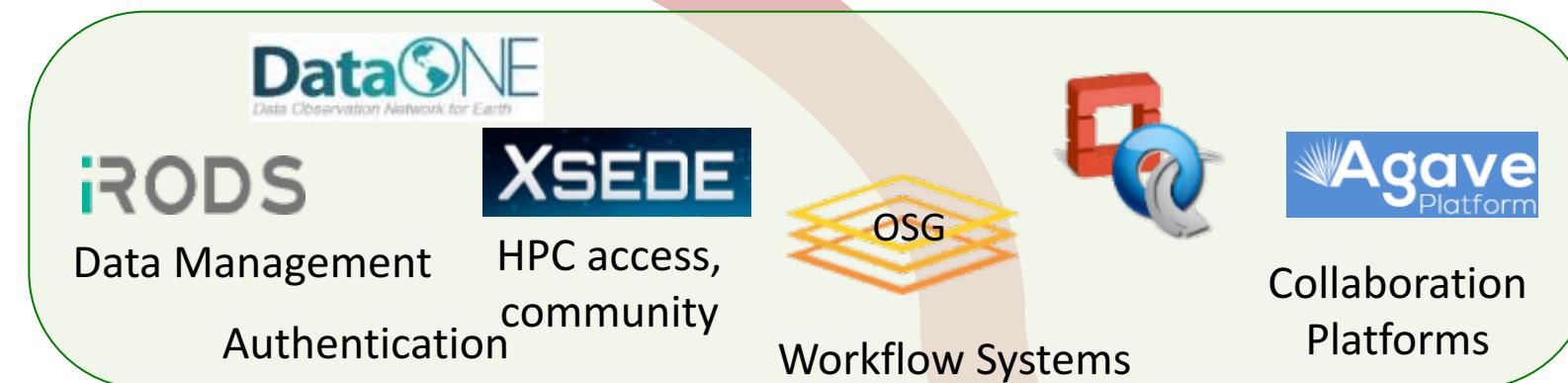
Measurement

Science
Portals

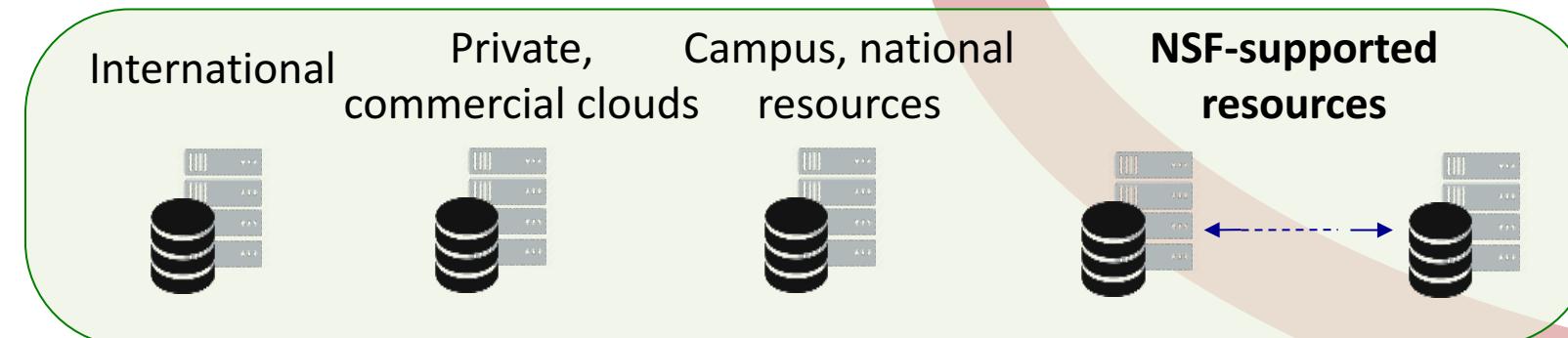
Applications,
Frameworks

Research
Facilities

**Discipline-specific
Environments**



**Integrative Services
("Middleware")**



**"Foundational"
CI Resources**

Discovery



Software: A primary modality for innovation and discovery permeating all layers of the cyberinfrastructure

Future Directions

The National Strategic Computing Initiative

1. Exascale computing system...
 - Foundational work on science, algorithms, programming environments, system software, architecture, and performance evaluation
2. Increasing coherence between ... simulation and data analytic...
 - Science and technology that use and enable applications involving both computational simulation and data analysis.
3. A viable path forward ... [in] the ‘post-Moore’s Law era
 - Foundational work on new device technology, fabrication methods, computer architectures, software techniques.
4. An enduring National HPC ecosystem...
 - Develop, integrate, and deploy building blocks of an HPC ecosystem.
 - Advance the organization, architecture applications of such a system,
 - Enhance user productivity, broaden participation, skilled workforce.
5. Public-private collaboration..."
 - Existing programs, such as GOALI, SBIR/SBTT, and IUCRC
 - Technology transition to and from practice
 - Advance the use of HPC technology in the commercial sector



NSF CI 2030 RFI - Snapshot of CI concerns and needs across all S&T fields...



(A simple wordle of all text of all responses)



NSF CI 2030 RFI - common emerging needs across diverse domains

- **Advanced computing.** Growing need for on-demand computing for steering simulations, rapid data processing, and experiments. Also, comparing simulations and data.
- **Data Science and management.** Big Data and Machine Learning cited in ~30% of RFI responses as critical. Automated mining, analytics, visualization, provenance tracking and credit, Open Data/Science, discoverability, accessibility, reproducibility, security, privacy.
- **Multi-source streaming data.** Dealing with data from myriad sensors, devices (Internet of Things, IOT), and cyber-physical systems, at all scales – human, community, built environments, ecosystems, biodiversity...
- **Networking, communication, workflows.** Technologies and approaches that scale with increasing performance demands. Processing and integrating streaming data. Long-tail and big data workflows, storage, identity management, cybersecurity.
- **Software.** Porting, accelerating, validating algorithms and community codes. Software quality, reliability, validity, practices.
- **Training and workforce development.** For researchers and computing professionals, diversity and inclusion. CS/CI experts that collaborate closely with domain researchers.



NSF, OAC, and URSSI

- NSF Priorities: NSCI, Big Ideas, Convergent Research
- OAC Priorities: CI 2030, National CI ecosystem
- URSSI: Effort to address the above priorities
 - Where does URSSI fit in current ecosystem?



OAC Integrated Software/Data Investments

- The Cyberinfrastructure for Sustained Scientific Innovation (CSSI) umbrella program encompasses

- Data Infrastructure Building Blocks (DIBBs)
- Software Infrastructure for Sustained Innovation (SI2) programs

to enable funding opportunities that are flexible and responsive to the evolving and emerging needs in data and software cyberinfrastructure.

Elements: Small groups - create & deploy robust capabilities for demonstrated need to advance science & engineering.

Framework Implementations: Larger teams organized around the development and application of common infrastructure aimed at solving common research problems, resulting in a sustainable community framework serving a diverse community or communities.

Planning Grants for Community Cyberinfrastructure: Focus on long-term capabilities in cyberinfrastructure to serve a research community of substantial size and disciplinary breadth.

Community Cyberinfrastructure Implementations: Focus on long-term hubs of excellence in cyberinfrastructure and technologies, to serve a research community of substantial size and disciplinary breadth.

Planning Grants for Community Cyberinfrastructure

- Aimed at
 - Organizing an interdisciplinary community
 - For a vision for a hub of excellence
 - that will enable transformative new science
 - by providing expertise, software, and other cyberinfrastructure resources
 - to serve a research community of substantial size and disciplinary breadth.
- Requirements and challenges for establishing robust and sustainable cyberinfrastructure should be presented.

Community Engaged Planning Activities

- Refining science drivers and utilization scenarios.
- Identifying the major stakeholders and target user base.
- Definition of a technical conceptualization, scientific requirements, utilization scenarios, major functional features, architectural strategies, envisaged cyberinfrastructure lifecycle duration.
- An initial strategy for design and implementation.
- Initial cost estimates
- Sustainability strategies.
- Prototyping activities as part of the scope of the project, as appropriate.
- A strategic, timeline- and measurement- driven approach.

Solicitation specific review criteria

Science-driven: Promotes science and engineering excellence, enabling fundamentally new scientific and engineering advances; benefits science and engineering communities beyond initial targets;

Innovative: Emphasizes unique NSF contributions; builds a national CI ecosystem; considers both human and technical aspects of the CI; Embeds research and innovation within the project activities

Collaborative: Fosters partnerships and community development; actively engages CI experts, specialists, and scientists working in concert with domain scientists who are users of CI;

Leveraged: Builds on existing, recognized capabilities;

Strategic: Encourages measurement of progress and sharing of results; and

Sustained: Provides benefits beyond the participants and the lifetime of the award.



Evaluation of Software Institutes

Performance of Software Institute will be evaluated based on the following criteria:

- Is the Institute enabling new science and broader impacts?
- Is the Institute demonstrating success in addressing the issues it identified to motivate its formation?
- Has the Institute responded to changes in the issues of concern from the community? Is the Institute achieving a consensus among the software developers in the Institute's topic area?
- Are the software products available in the Institute's topic area being integrated or made interoperable, including, but not limited to the products from NSF-supported teams working in its area?
- Is the Institute achieving broad buy-in among the software user community in the Institute's topic area?
- Has the Institute created and begun to implement a sustainability plan that is realistic and likely to be successful?
- Is the Institute succeeding in terms of its success metrics?



Thank You!

