Enabling Grids for E-sciencE

Sept. 25, 2006 Geneva

Collective Opportunities and Responsibilities for *Capitalizing on e- | cyber- infrastructures* in Service of International Discovery and Learning

Daniel E. Atkins
Director,
U.S. National Science Foundation
Office of Cyberinfrastructure
datkins@nsf.gov

Overview of Talk

- Appreciative of opportunity to learn and establish basis for better international collaboration.
  - UK e-science All Hands Meeting
  - CERN visit
  - EGEE06
- Background and Some Highlights of US NSF Cyberinfrastructure Activities
- Comments about International Opportunities, Challenges, and Priorities (and shared responsibilities)

Nomenclature

e-science

e = electronic | enhanced | enabled

Office of Cyberinfrastructure

D. E. Atkins

Nomenclature

e-science

e-infrastructure

Office of Cyberinfrastructure

e = electronic | enhanced | enabled

D. E. Atkins

Nomenclature

e-science

e-infrastructure

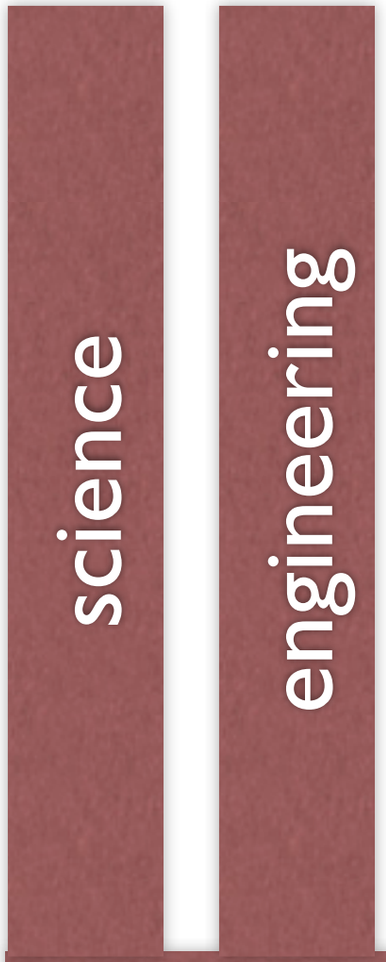
cyberinfrastructure (CI)

Office of Cyberinfrastructure

e = electronic | enhanced | enabled

D. E. Atkins

Nomenclature



e-science

CI-enhanced | enabled

e-infrastructure

cyberinfrastructure (CI)

Office of Cyberinfrastructure

e = electronic | enhanced | enabled

D. E. Atkins

Nomenclature

research & learning

science

engineering

e-science

CI-enhanced | enabled

e-infrastructure

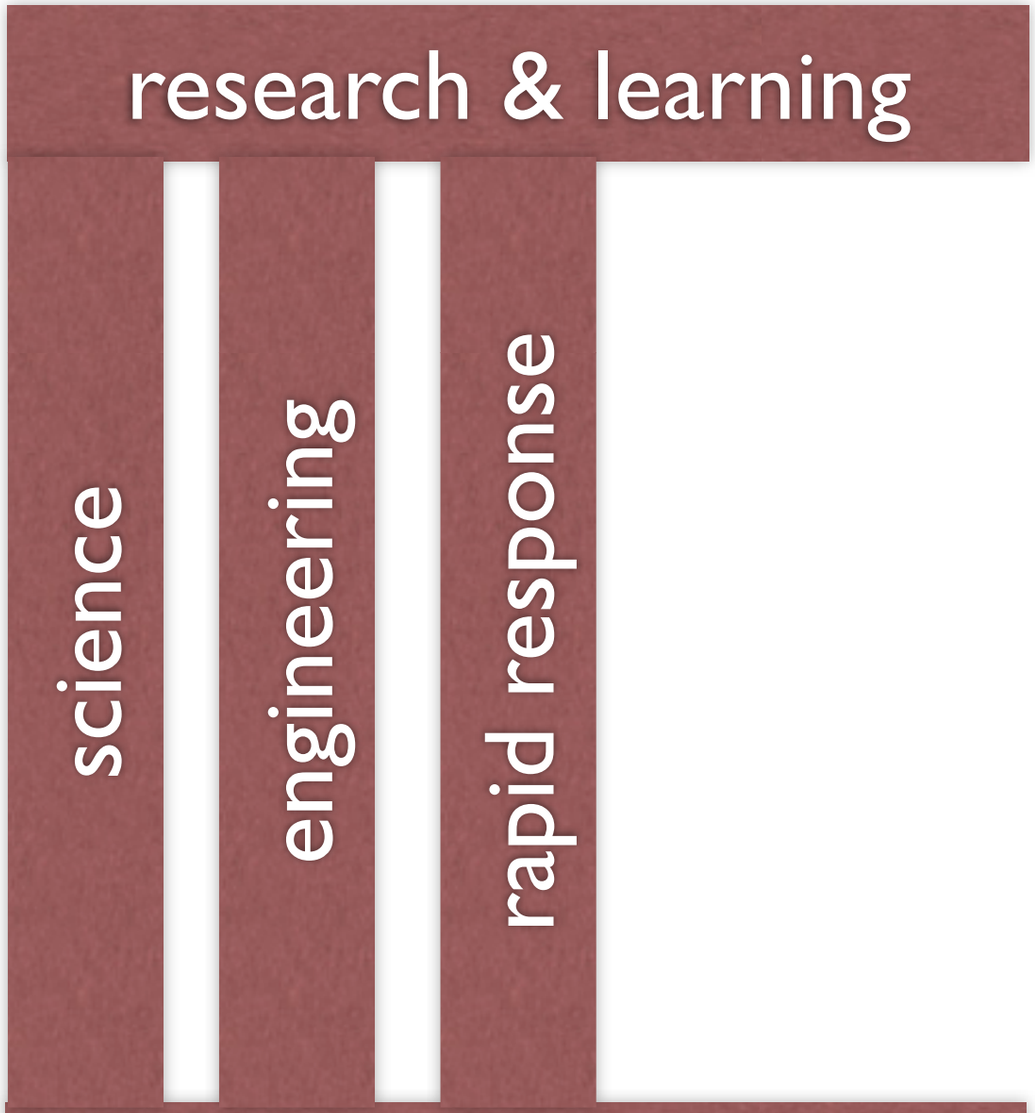
cyberinfrastructure (CI)

Office of Cyberinfrastructure

e = electronic | enhanced | enabled

D. E. Atkins

Nomenclature



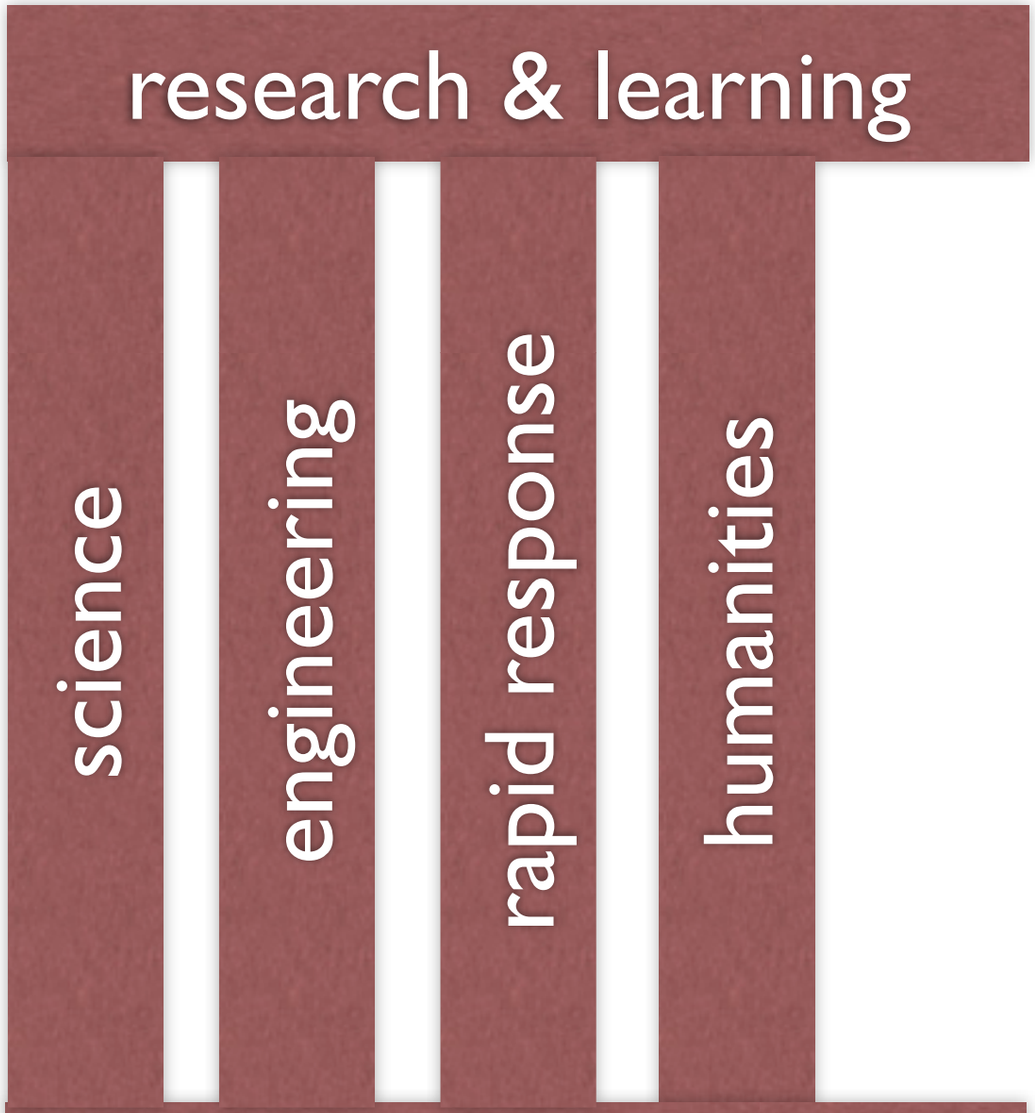
e-science

e-infrastructure

cyberinfrastructure (CI)

e = electronic | enhanced | enabled

Nomenclature



e-science

CI-enhanced | enabled

e-infrastructure

cyberinfrastructure (CI)

Office of Cyberinfrastructure

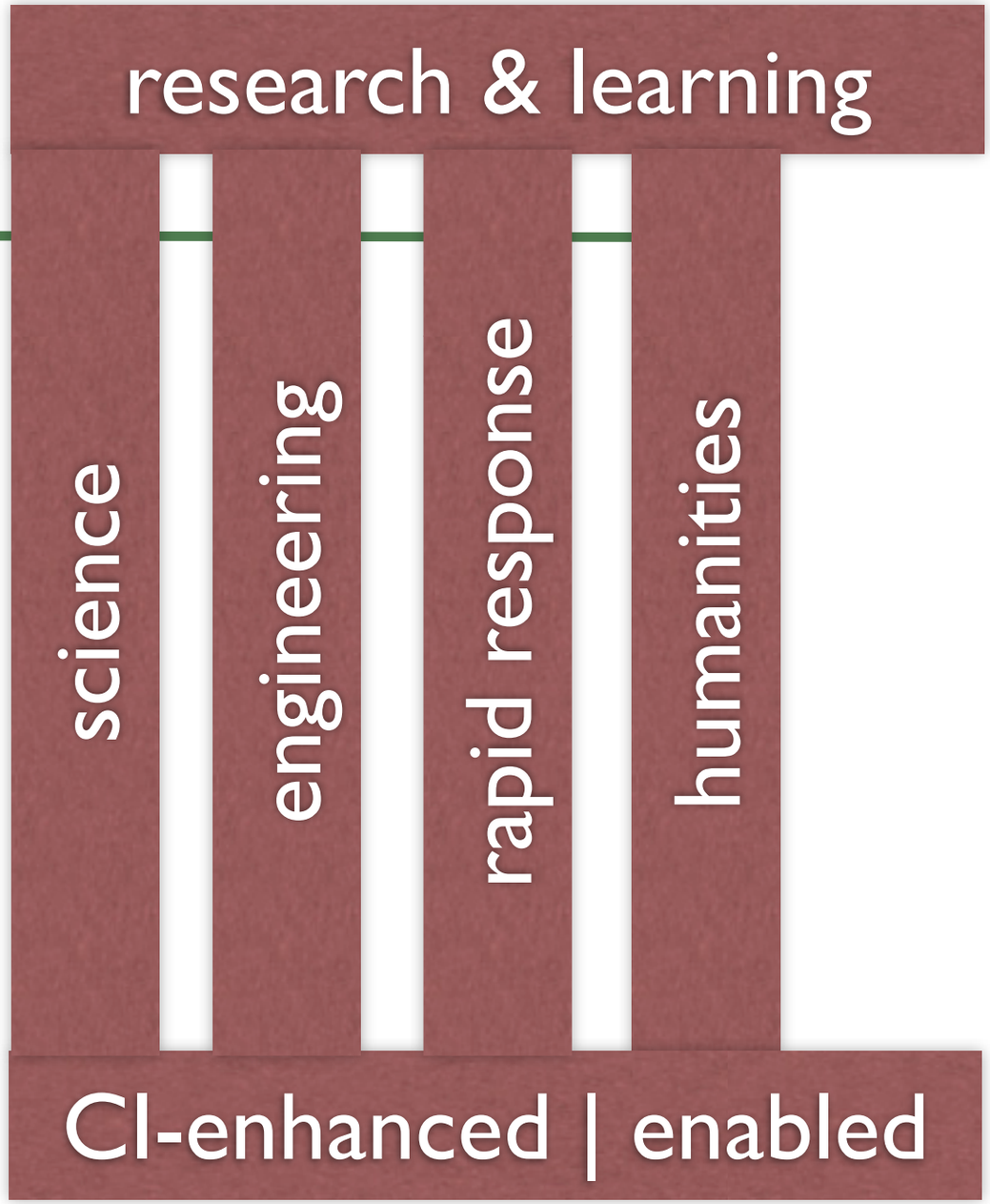
e = electronic | enhanced | enabled

D. E. Atkins

Nomenclature

ACLS-Mellon Study: Cyberinfrastructure for Humanities http://www.acls.org/cyberinfrastructure/

HASTAC Project http://www.hastac.org/



e-science

e-infrastructure

cyberinfrastructure (CI)

Office of Cyberinfrastructure

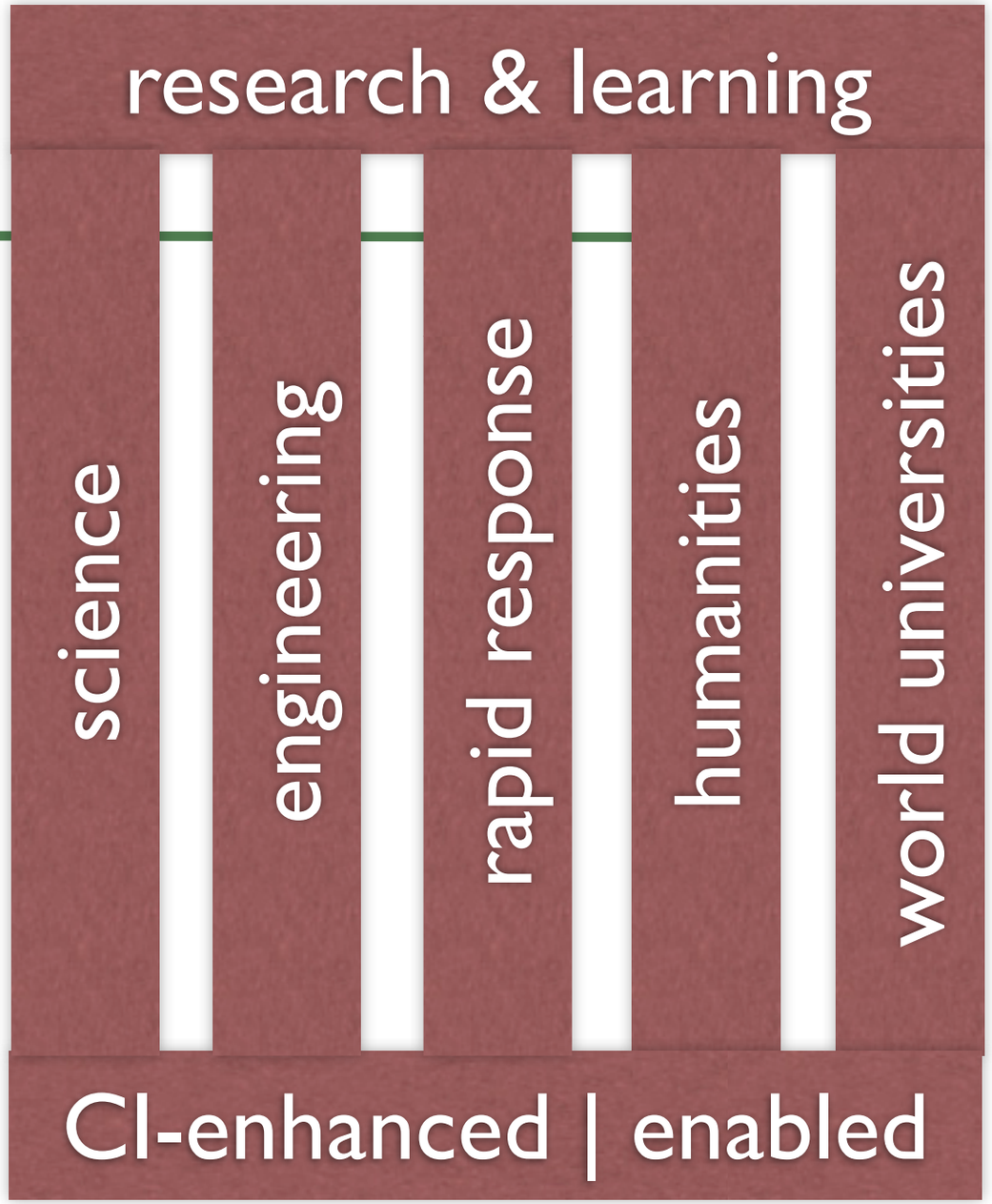
e = electronic | enhanced | enabled

D. E. Atkins

Nomenclature

ACLS-Mellon Study: Cyberinfrastructure for Humanities http://www.acls.org/cyberinfrastructure/

HASTAC Project http://www.hastac.org/



e-science

e-infrastructure

cyberinfrastructure (CI)

Office of Cyberinfrastructure

e = electronic | enhanced | enabled

D. E. Atkins

Nomenclature

ACLS-Mellon Study: Cyberinfrastructure for Humanities http://www.acls.org/cyberinfrastructure/

HASTAC Project http://www.hastac.org/

research & learning

science | engineering | rapid response | humanities | world universities

CI-enhanced | enabled

Cyber science

e-science

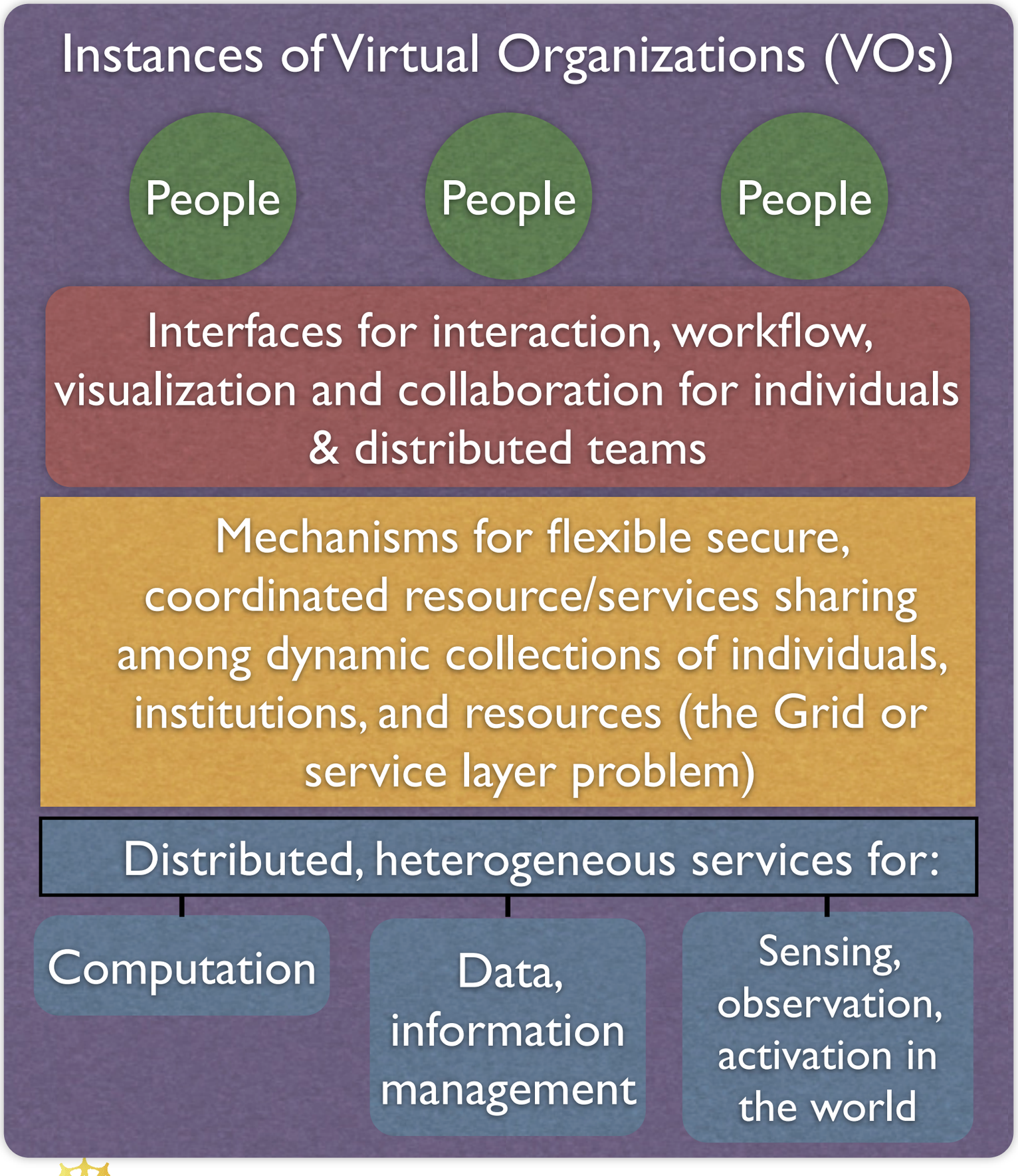
e-infrastructure

cyberinfrastructure (CI)

Office of Cyberinfrastructure

e = electronic | enhanced | enabled

D. E. Atkins



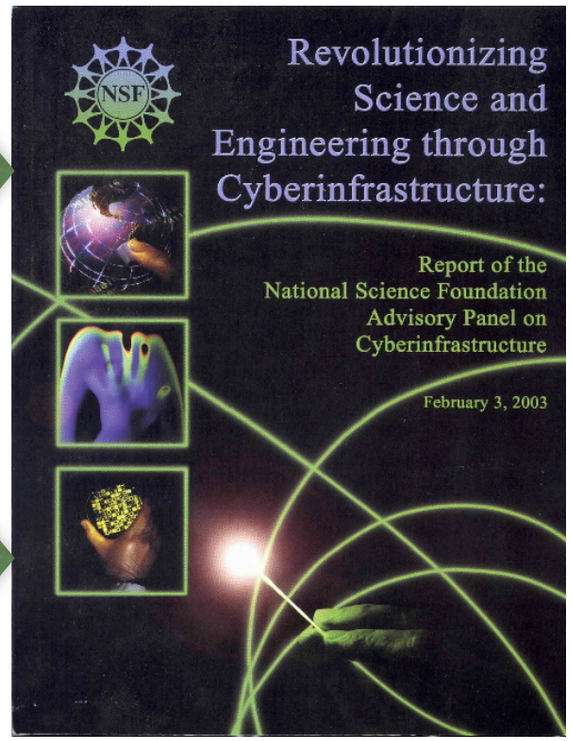
Alternate Names for Instances of VOs:

- Co-laboratory
- Collaboratory
- Grid (community)
- Network
- Portal
- Gateway
- Hub
- Virtual Research Environment (VRE)
- Other?

Vision and Activities Based on Broad and Diverse Community Engagement

Advances in components of CI-systems for S&E R&E

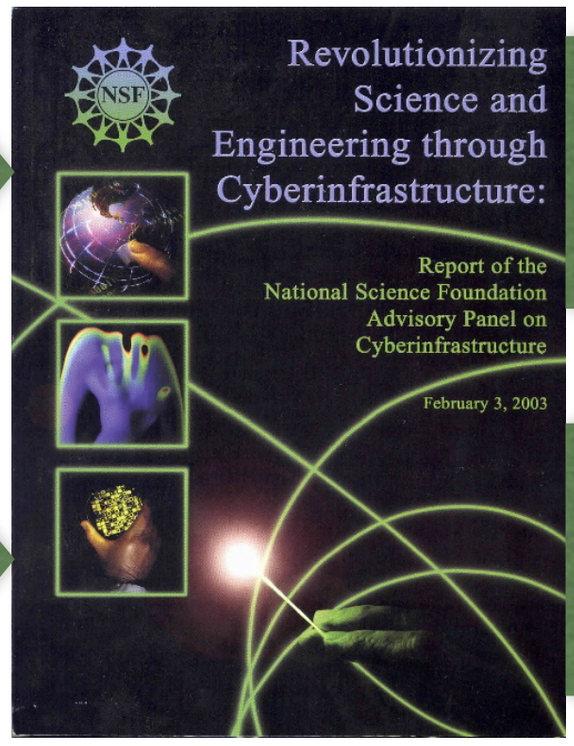
Complex, multi-scale, multidisciplinary S&E research challenges



Vision and Activities Based on Broad and Diverse Community Engagement

Advances in components of CI-systems for S&E R&E

Complex, multi-scale, multidisciplinary S&E research challenges



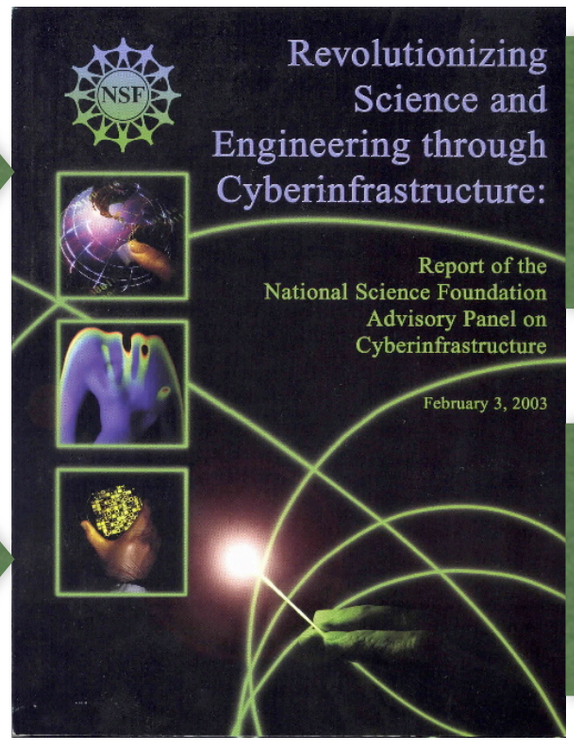
30+ disciplinary workshops on CI vision & impact

NSF internal working groups

Vision and Activities Based on Broad and Diverse Community Engagement

Advances in components of CI-systems for S&E R&E

Complex, multi-scale, multidisciplinary S&E research challenges



30+ disciplinary workshops on CI vision & impact

NSF internal working groups

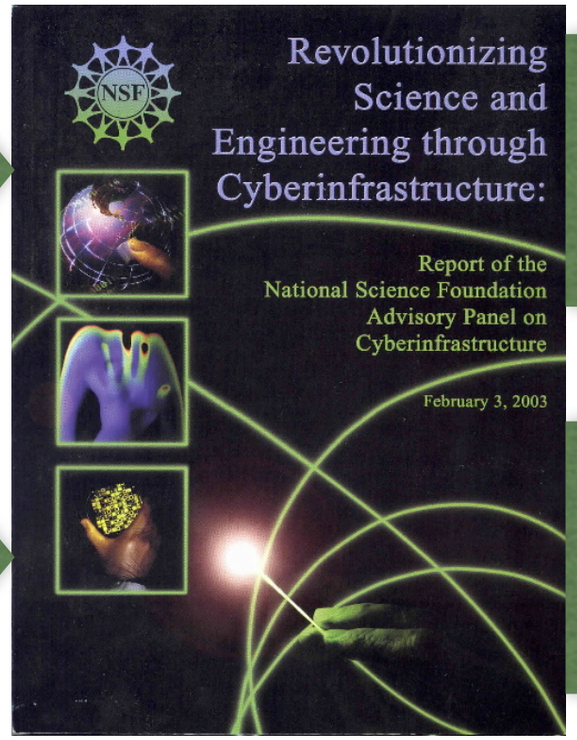
CI Council, Directorate/Office CI Activities, OCI, ACCI

NSB & Community Input

Vision and Activities Based on Broad and Diverse Community Engagement

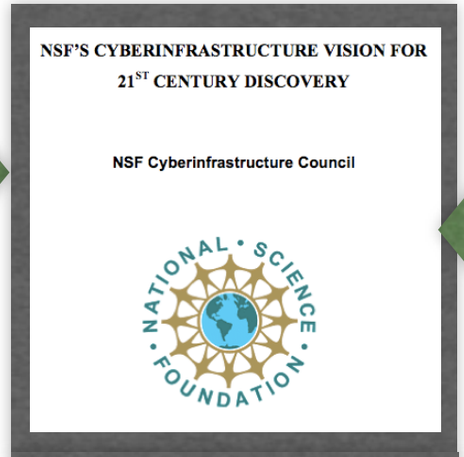
Advances in components of CI-systems for S&E R&E

Complex, multi-scale, multidisciplinary S&E research challenges



30+ disciplinary workshops on CI vision & impact

NSF internal working groups



CI Council, Directorate/Office CI Activities, OCI, ACCI

NSB & Community Input

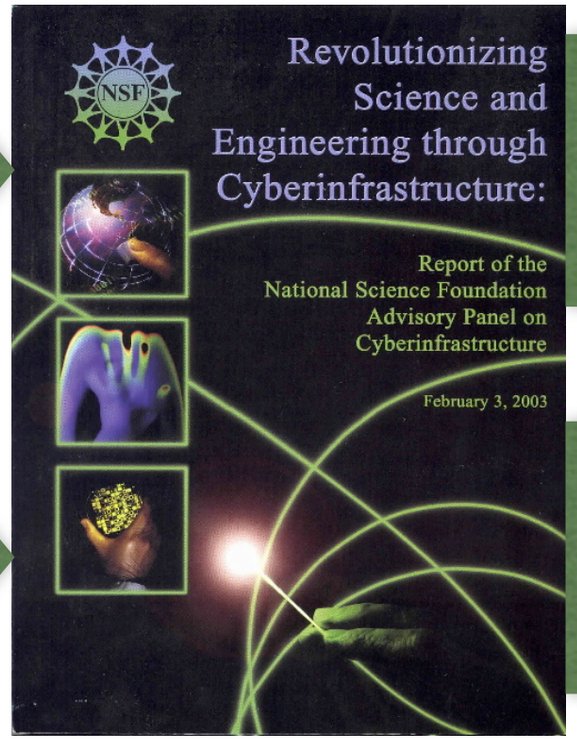
Vision Framework

- High Performance Computing
- Data, Data Analysis & Visualization
- Virtual Organizations
- Learning & Workforce Development

Vision and Activities Based on Broad and Diverse Community Engagement

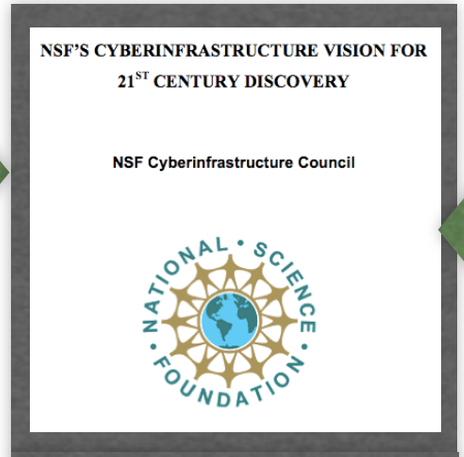
Advances in components of CI-systems for S&E R&E

Complex, multi-scale, multidisciplinary S&E research challenges



30+ disciplinary workshops on CI vision & impact

NSF internal working groups



CI Council, Directorate/Office CI Activities, OCI, ACCI

NSB & Community Input

Vision Framework

- High Performance Computing
- Data, Data Analysis & Visualization
- Virtual Organizations
- Learning & Workforce Development

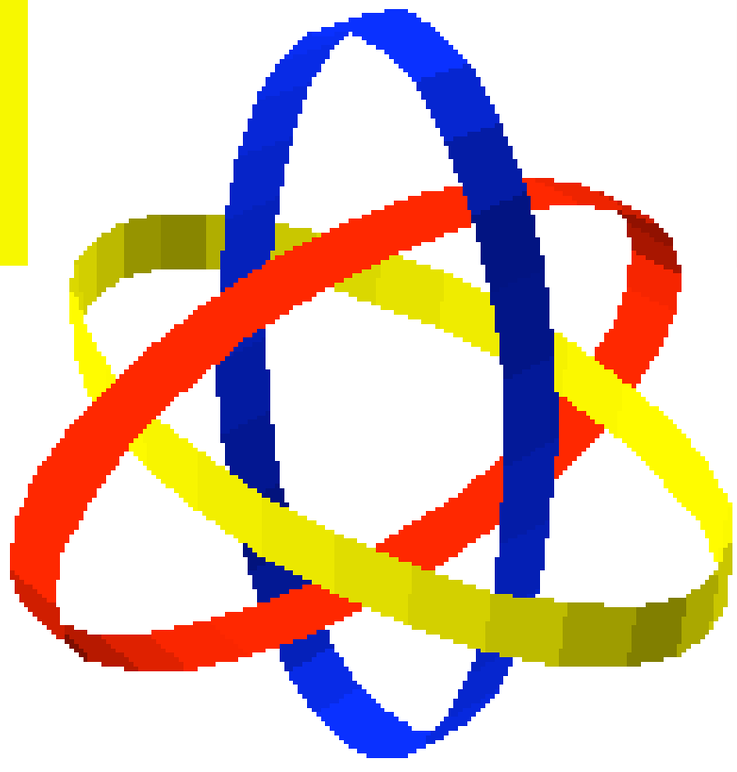
- *All directorates and offices support cyberinfrastructure.*
- *Science-driven partnerships between creation, provisioning and use of CI*
- *Supports integrated research and education and broadened access and participation.*

Some Science Drivers

- Inherent **complexity and multi-scale** nature of todays frontier science challenges.
- Requirement for **multi**-disciplinary, **multi**-investigator, **multi**-institutional approach (often international).
- High **data intensity** from simulations, digital instruments, sensor nets, observatories.
- Increased value of data and demand for data **curation & preservation** of access.
- Exploiting infrastructure **sharing** to achieve better stewardship of research funding.
- Strategic need for **engaging more students** in high quality, authentic science and engineering education.

Achieving the NSF CI (e-science) Vision requires synergy between 3 types of activities

Transformative Application - to enhance discovery & learning



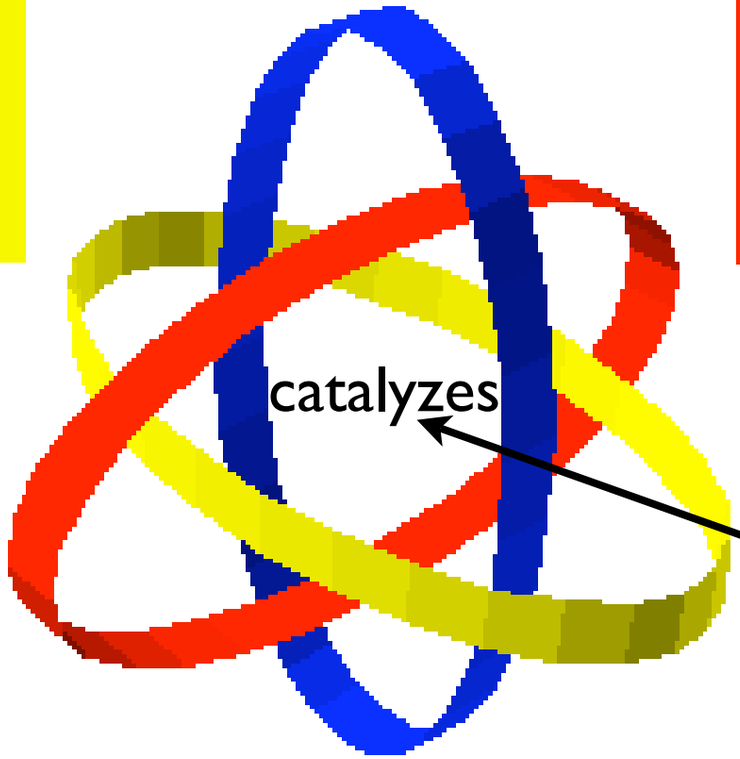
Provisioning - Creation, deployment and operation of advanced CI

Borromean Ring: The three rings taken together are inseparable, but remove any one ring and the other two fall apart. See www.liv.ac.uk/~spmr02/rings/

R&D to enhance technical and *social* effectiveness of future CI environments

Achieving the NSF CI (e-science) Vision requires synergy between 3 types of activities

Transformative Application - to enhance discovery & learning



Provisioning - Creation, deployment and operation of advanced CI

provides shared and connecting CI

Office of Cyberinfrastructure

Borromean Ring: The three rings taken together are inseparable, but remove any one ring and the other two fall apart. See www.liv.ac.uk/~spmr02/rings/

R&D to enhance technical and *social* effectiveness of future CI environments

NSF CI FY07 Budget Request Total of $600M in CI Funding with $182M in OCI

Cyberinfrastructure Funding

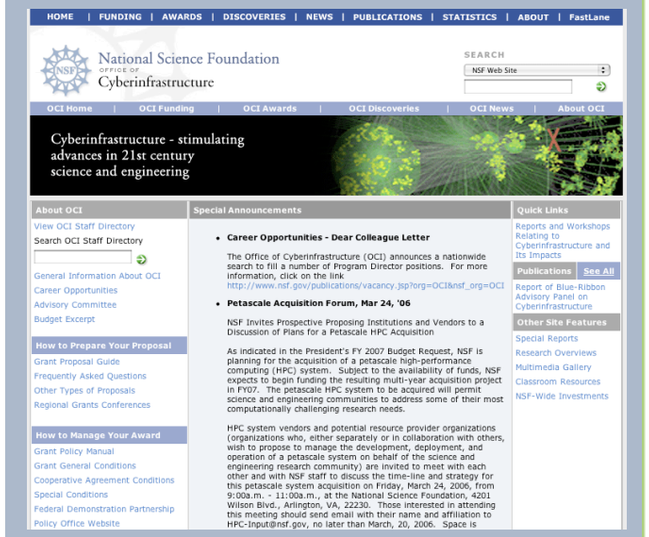
(Dollars in Millions)

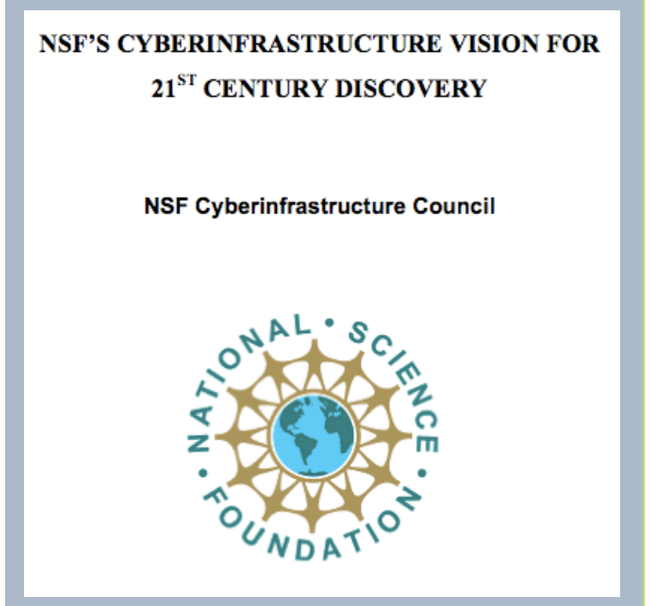
	FY 2005 Actuals	FY 2006 Current Plan	FY 2007 Request	Change over FY 2006 Amount	Change over FY 2006 Percent
Biological Sciences	$77.00	$84.00	$90.50	$6.50	7.7%
Computer and Information Science and Engineering	45.32	63.00	68.00	5.00	7.9%
Engineering	52.00	52.00	54.00	2.00	3.8%
Geosciences	71.35	71.35	75.00	3.65	5.1%
Mathematical and Physical Sciences	56.52	59.30	63.56	4.26	7.2%
Social, Behavioral and Economic Sciences	20.39	20.54	20.54	-	-
Office of Cyberinfrastructure	123.28	127.12	182.42	55.30	43.5%
Office of International Science and Engineering	0.22	1.00	1.05	0.05	5.0%
Office of Polar Programs	25.38	26.24	26.24	-	-
Subtotal, Research and Related Activities	471.47	504.55	581.31	76.76	15.2%
Education and Human Resources	20.27	15.02	15.52	0.50	3.3%
Total, Cyberinfrastructure Funding	$491.74	$519.57	$596.83	$77.26	14.9%

Totals may not add due to rounding.

www.nsf.gov/oci/

OCI Website - Visit often and provide feedback on the Vision document.





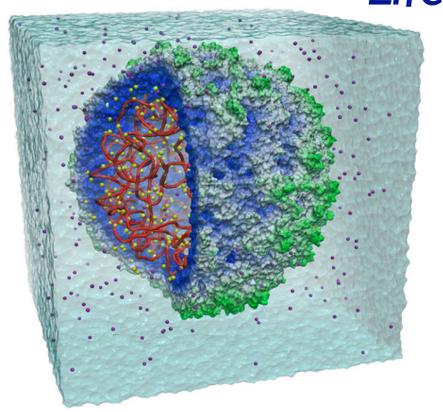
www.nsf.gov/od/oci/ci-v7.pdf

National Science Foundation
www.nsf.gov

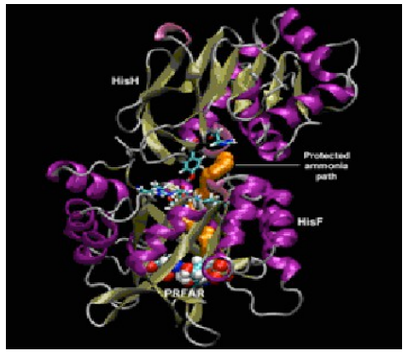
High Performance Computing

increasingly important tool for understanding

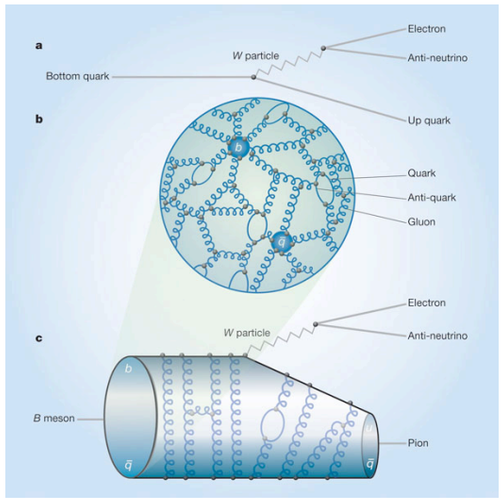
Life



Satellite tobacco mosaic virus, P. Freddolino et al.



Aldehyde dehydrogenase, T. Wymore and S. Brown

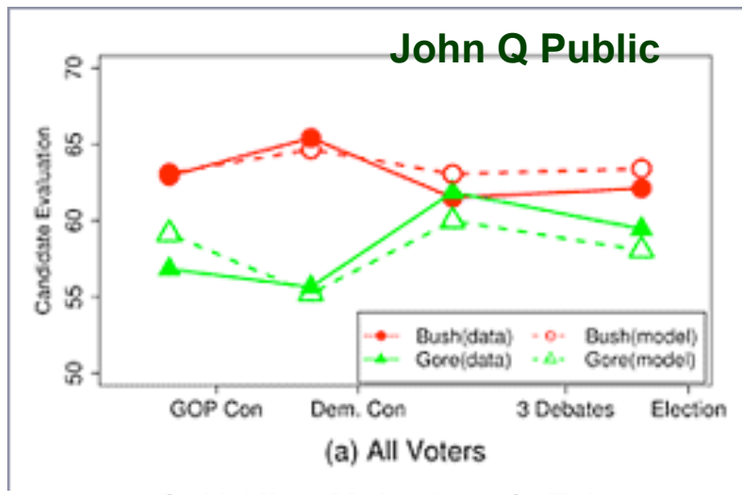


I. Shipsey

Matter

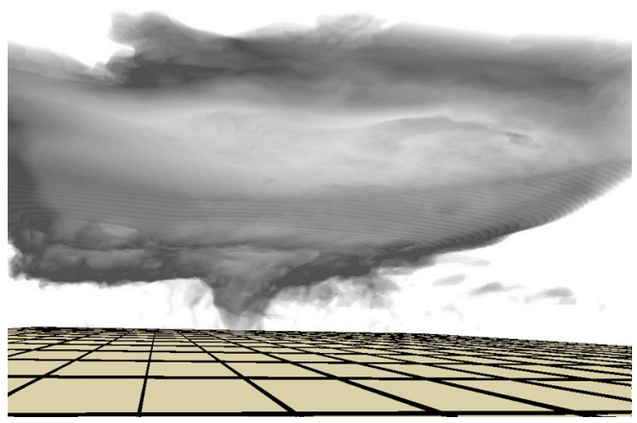
Society

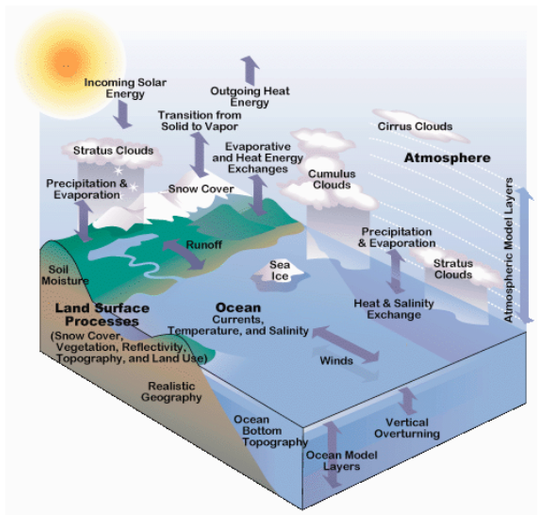
John Q Public



S.-Y. Kim, M. Lodge, C. Taber.

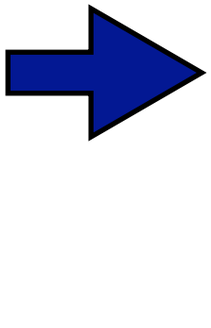
The Environment



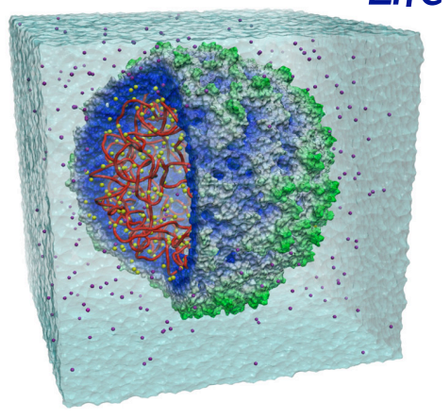


High Performance Computing

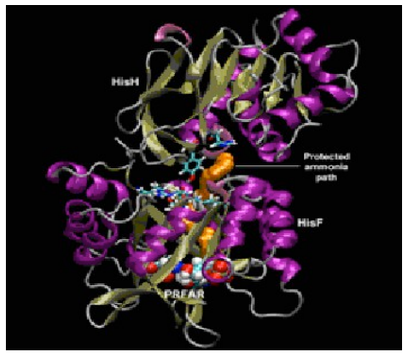
increasingly important tool for understanding



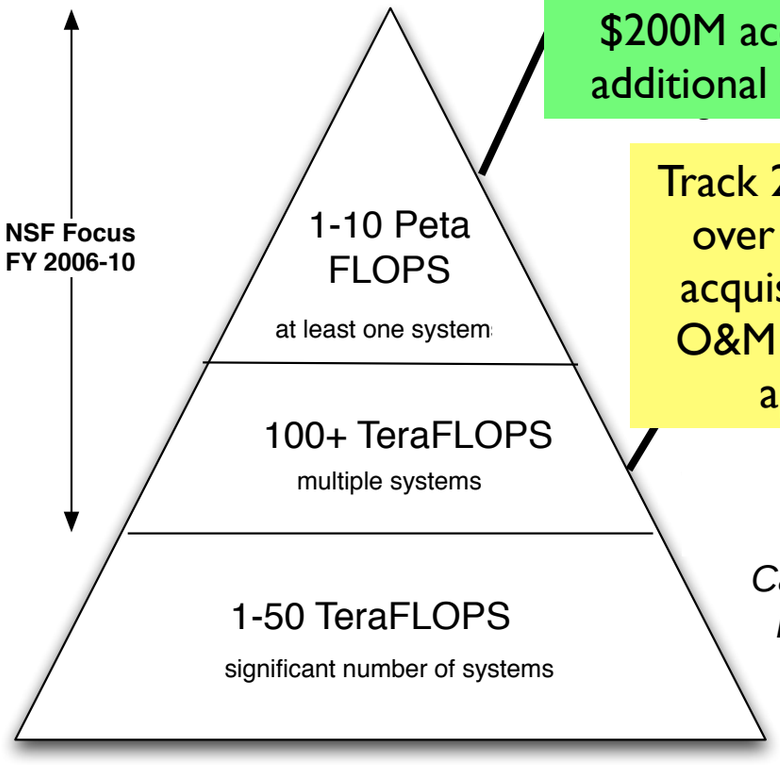
Life



Satellite tobacco mosaic virus, P. Freddolino et al.



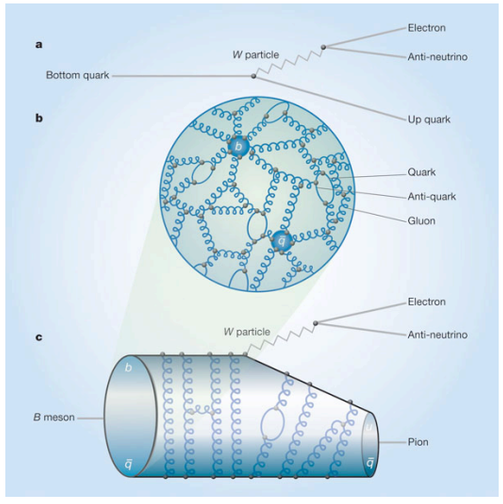
Aldehyde dehydrogenase, T. Wymore and S. Brown



Track 1: One solicitation funded over 4 years: $200M acquisition + additional O&M cost.

Track 2: Four solicitations over 4 years: $30M/yr acquisition + additional O&M cost. First track 1 approved 8-07

Campus Level

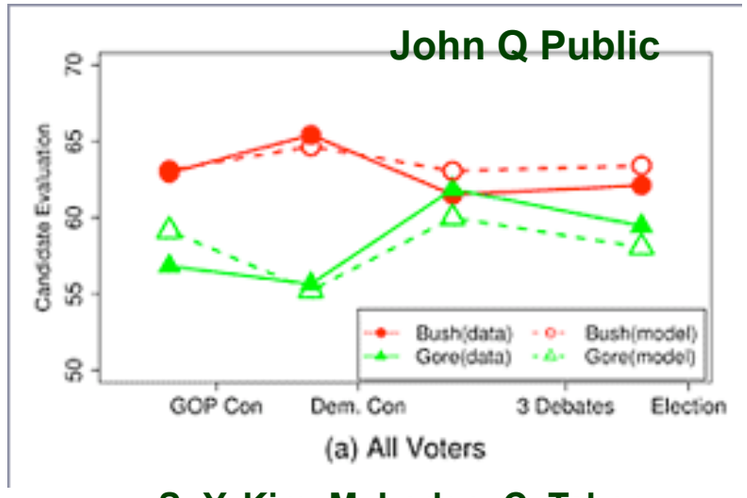


I. Shipsey

Matter

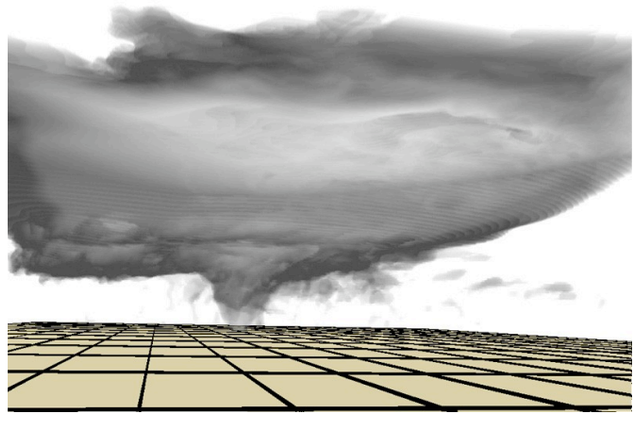
Society

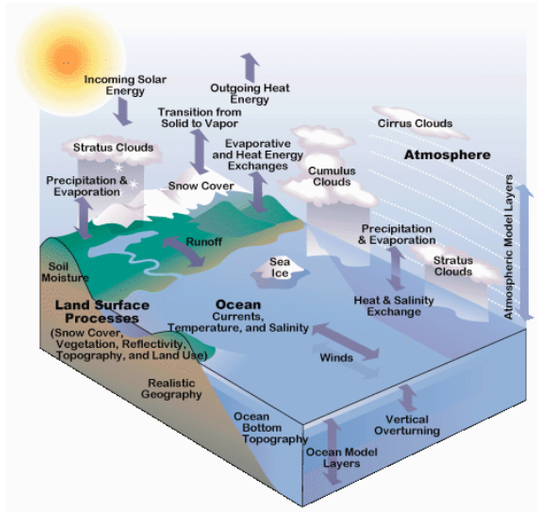
John Q Public



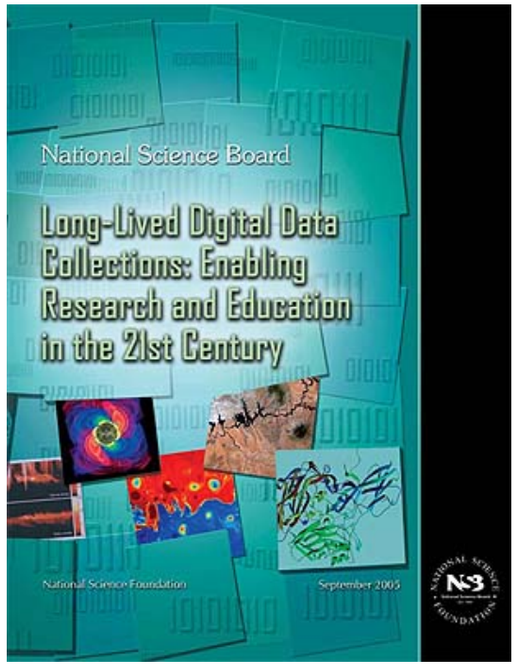
S.-Y. Kim, M. Lodge, C. Taber.

The Environment

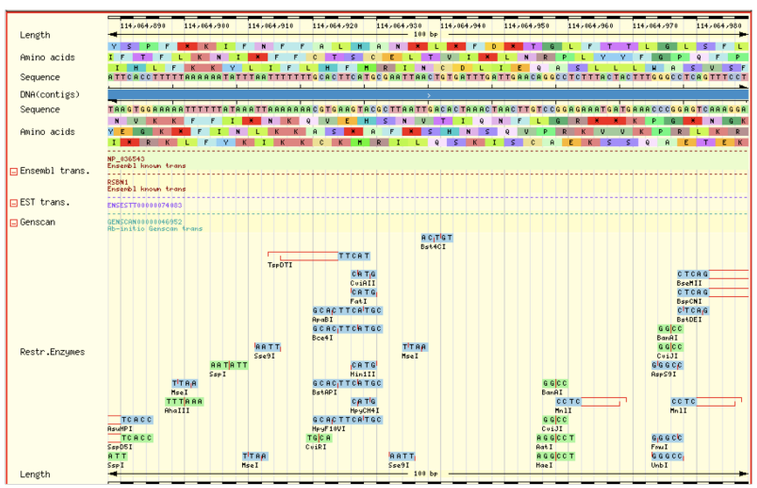


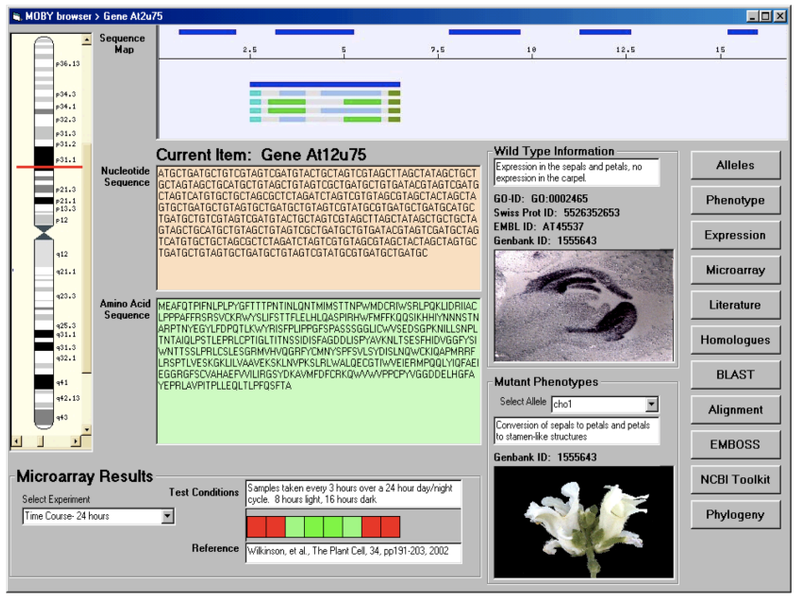


Data, Data Analysis & Visualization



- Challenges: **increased scale, heterogeneity, and re-use value** of digital scientific information and data. Inadequate digital preservation strategy of long-lived data.
- Taking initial steps to **catalyze the development** of a federated, global system of science and engineering data collections that is open, extensible, evolvable, (and appropriately curated and long-lived.)
- Complemented by a **new generation of tools** and services to facilitate data mining, integration, analysis, visualization essential to transforming data into knowledge.
- NSF Leadership for OSTP/Interagency Working Group on Digital Data





Virtual Organizations

NVO

LEAD

iVDgL

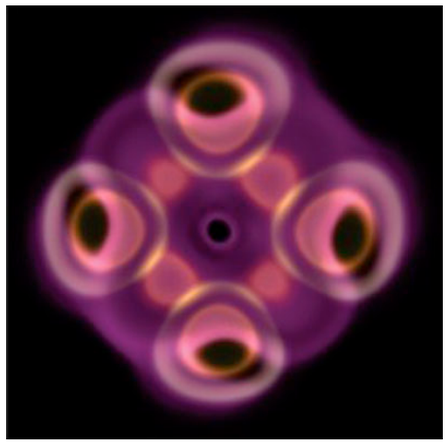
TeraGrid

Open Science Grid

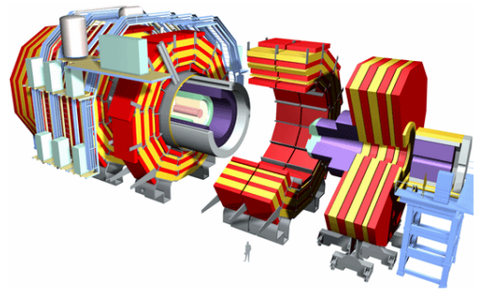
eGee Enabling Grids for E-sciencE

- Distributed virtual organizations are **based upon CI** that provides flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions, and resources.
- **Resources and services** include HPC, data/information management, sensor-nets/observatories, linked through global networking and middleware, and accessed by people through web portals and workflow environments.
- Increasing numbers of **virtual organizations are required** by S&E research and education communities. Referred to by many names, e.g. *collaboratory, co-laboratory, grid, gateway, portal,. hub, ...*
- **Challenges** being address include tools for more rapid building and ease of use, interoperability/middleware, high performance, end-to-end networking, and dynamic reconfiguration, social issues, assessment of impact, and economic and technical sustainability.

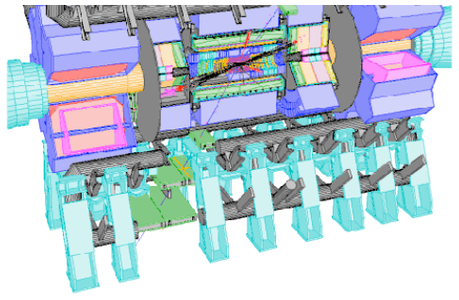
NEES



NanoHub



CMS



ATLAS

Virtual Organizations offer additional modes of interaction between People, Information, and Facilities

Geographic Place	Time	Same (synchronous)	Different (asynchronous)
	Same	**ST-SP** **P**: Physical mtgs **I**: Print-on-paper books, journals **F**: Physical labs, studios, shops	**DT-SP** **P**: Shared notebook **I**: Library reserves **F**: Time-shared physical labs, ...
	Different	**ST-DP** **P**: AV conference **I**: Web search **F**: Online instruments	**DT-DP** **P**: Email **I**: Knowbots **F**: Autonomous observatories

P: people, **I**: information, **F**: facilities, instruments

Virtual Organizations offer additional modes of interaction between People, Information, and Facilities

Time

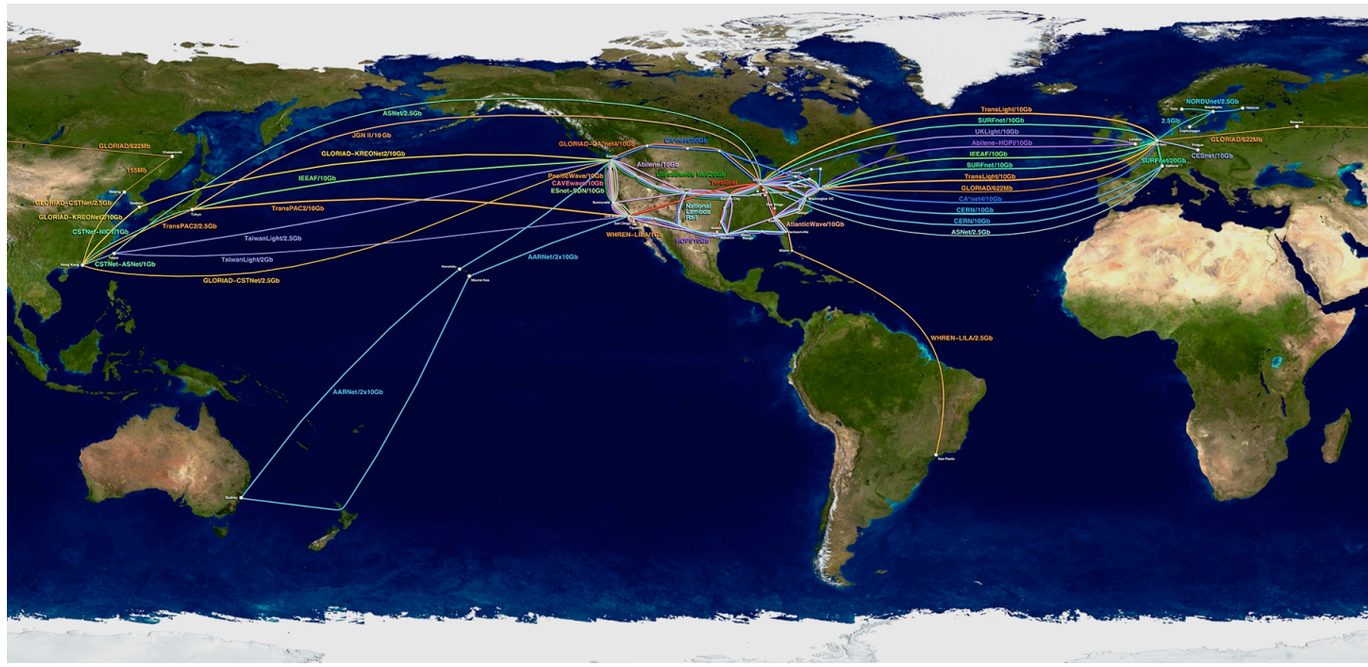
Geographic Place	Same (synchronous)	Different (asynchronous)
Same	**ST-SP** **P**: Physical mtgs **I**: Print-on-paper books, journals **F**: Physical labs, studios, shops	**DT-SP** **P**: Shared notebook **I**: Library reserves **F**: Time-shared physical labs, ...
Different	**ST-DP** **P**: AV conference **I**: Web search **F**: Online instruments	**DT-DP** **P**: Email **I**: Knowbots **F**: Autonomous observatories

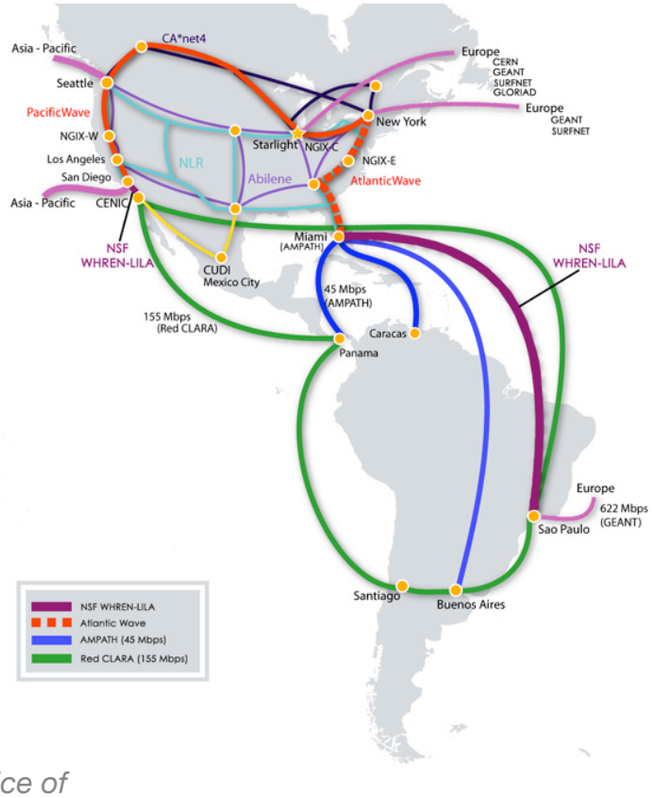
P: people, **I**: information, **F**: facilities, instruments

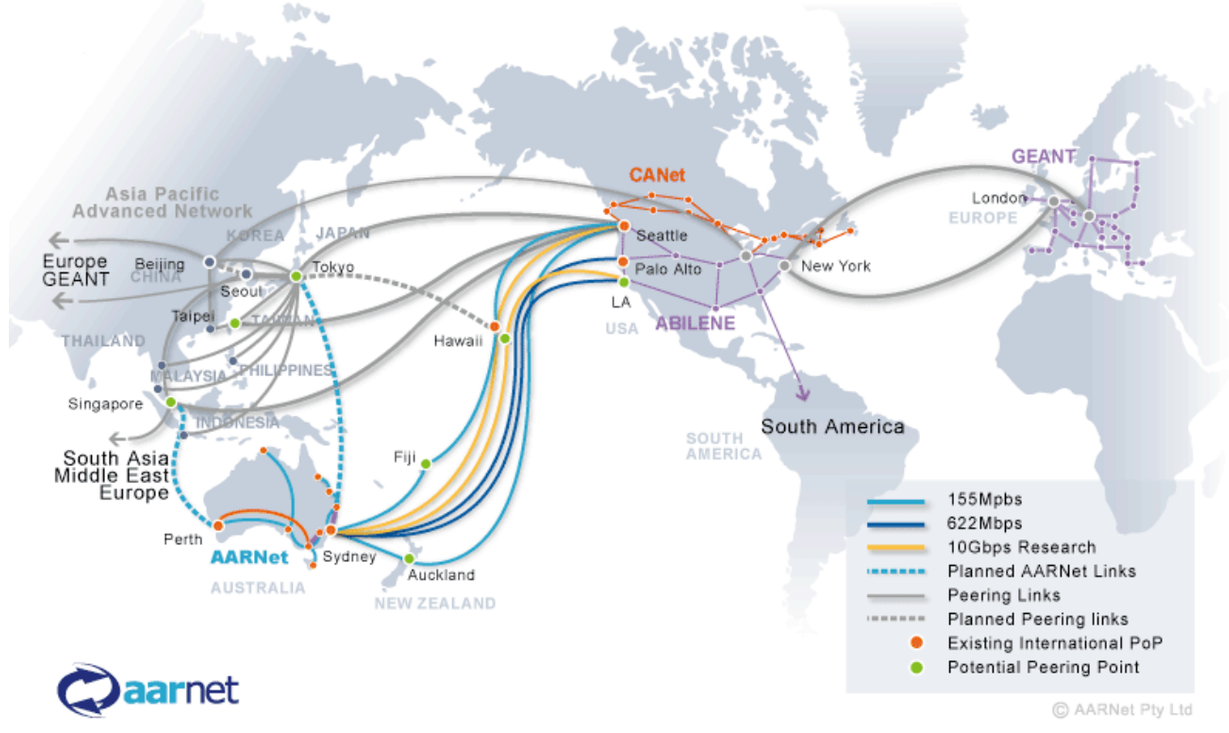
Need a socio-technical approach: Realizing the potential of e- | cyber science to support effective VOs requires attention to *institutional infrastructure*

- Need more than good ICT systems and tools for individuals and organizations. (Technology determinism alone is not enough).
- No less important is the *institutional contexts* (i.e. norms of practice and rules) to facilitate collaboration within science and technical research communities.
- The institutional and organizational environment of e-science encompasses a wide and diverse array of interrelated social, economic, and legal factors that
  - create incentives for, and constraints upon individual and collective action; and
  - thereby shape the production, utilization, consumption, and governance of e-science capabilities and products.

VO-substrate: International R&E Networking







Learning & Workforce Development

- Learning **supported by** CI. (cyber-enabled learning).
- Workforce development **to create and use** CI for S&E research and education.
- **Broadened participation**: Exploit the new opportunities that cyberinfrastructure brings for … people who, because of physical capabilities, location, or history, have been excluded from the frontiers of scientific and engineering research and education.
- Explore CI support for **integrated research and education.**

BIOINFORMATICS CI INSTITUTE

EPIC

MARIACHI

CyberBridges

Priorities (Joint?)

- **Complete, robust**, and **sustainable middleware** for sharing resources to enable virtual organizations for many specific research domains.
- International grid interoperability. EGEE, OSG, TeraGrid, ... More common components; more joint development??
- Grids as platforms for **functionally complete collaboratories** (computation, data/information, instruments, people) in a growing number of domains.
- Commitment to **openness**: standards, software, content, participation.
- **Multi-purpose** Grid-based VOs built to support wide spread use for research, education, public awareness. Integration with campus environments.

Some Existing & Potential Interactions

International Science Projects, e.g. ATLAS, CMS

Funding & science collaboration

U.S. Investments

- U.S. part of international science/engineering research projects.

- Open Science Grid (OSG)
- TeraGrid & Science Gateways
- Grid Interoperabilty Now (GIN)
- GLOBUS
- Condor Technologies
- Virtual Data Toolkit (VDT)
- NMI Build and Test
- Shibolleth
- GridShib
- Other?

EGEE

- gLITE
- Experience with large, production, international Grid operation
- Other?

staff R&D interactions

use of components

shared development

Other National/ Regional Grid Projects

New Opportunities

"Cyberinfrastructure-enhanced knowledge communities offer the potential for enabling a new wave of global-scale collaboration across multiple disciplines, geography, and institutions. It could empower a revolution in ***what*** *science explores,* ***how*** *it is done, and* ***who*** *participates.*

Realizing this potential will, however, also required ***a new wave of commitment to collaboration*** *between the complex array of stakeholders necessary to create, deploy, sustain, and apply cyberinfrastructure in transformative ways.*

***Cyberinfrastructure both enables and requires a new wave of collaboration.*"**

D. E. Atkins, Keynote for EDUCAUSE Australasia, Auckland, NZ, April 5-8, 2005

New Opportunities

"Cyberinfrastructure-enhanced knowledge communities offer the potential for enabling a new wave of global-scale collaboration across multiple disciplines, geography, and institutions. It could empower a revolution in ***what*** *science explores,* ***how*** *it is done, and* ***who*** *participates.*

Realizing this potential will, however, also required ***a new wave of commitment to collaboration*** *between the complex array of stakeholders necessary to create, deploy, sustain, and apply cyberinfrastructure in transformative ways.*

Cyberinfrastructure both enables and requires a new wave of collaboration*."*

To the EGEE: *Are we doing enough of the right things together??*

D. E. Atkins, Keynote for EDUCAUSE Australasia, Auckland, NZ, April 5-8, 2005