Maximizing Returns on National Investments Through Cyberinfrastructure

Daniel E. Atkins

U. S. National Science Foundation

datkins@nsf.gov

www.nsf.gov/oci

Workshop on Advanced Collaborative Environments and CERN Large Hadron Collider Users

December 11-13, 2006
International Conference Centre
Geneva, Switzerland

The University of Michigan Upper Atmospheric Research Collaboratory (UARC)

The Initial Facility at Sondrestrom, Greenland

UARC Interface (circa 1988)
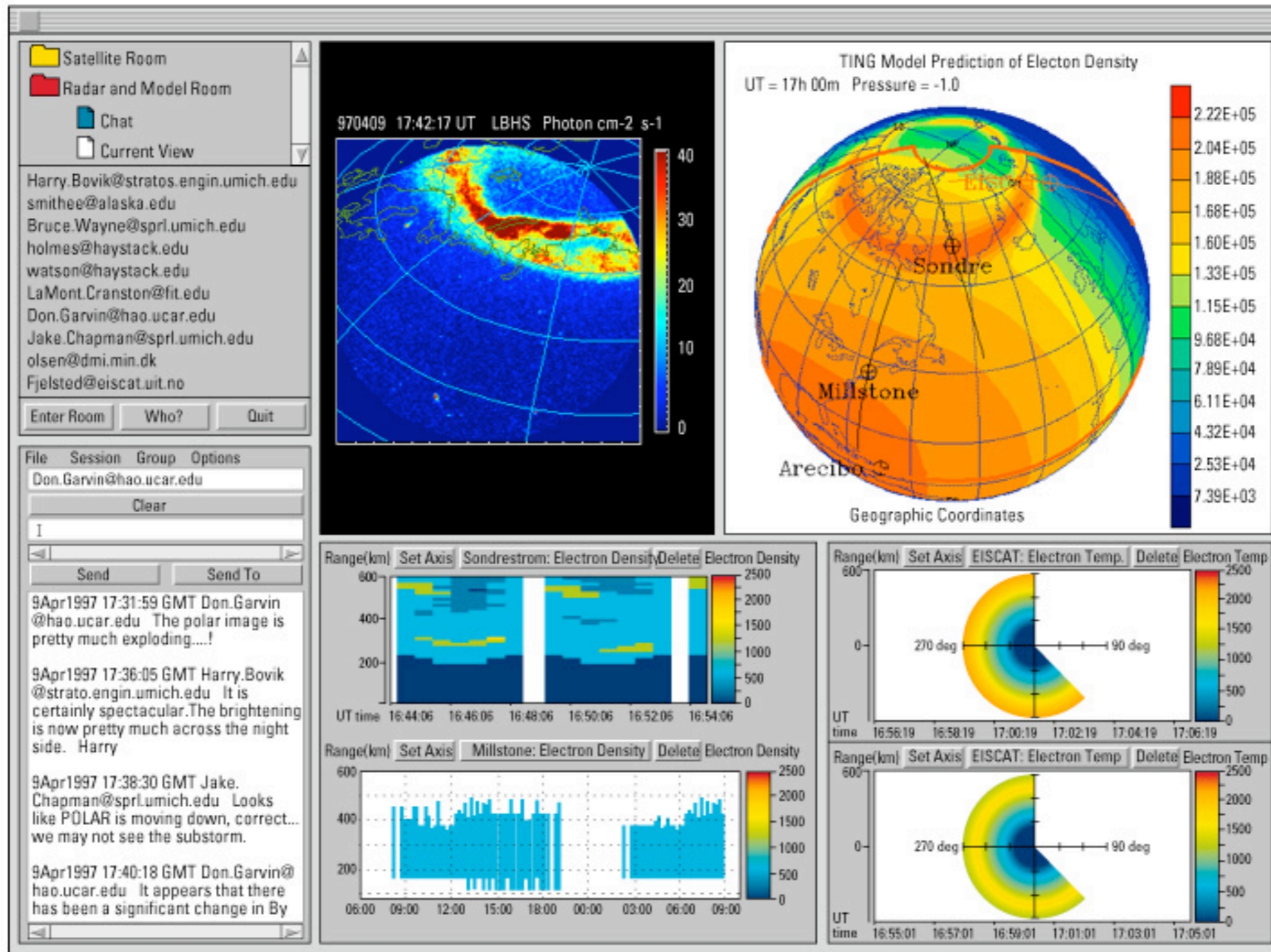
Real-time instruments   **computational models**

dynamic work rooms

annotation

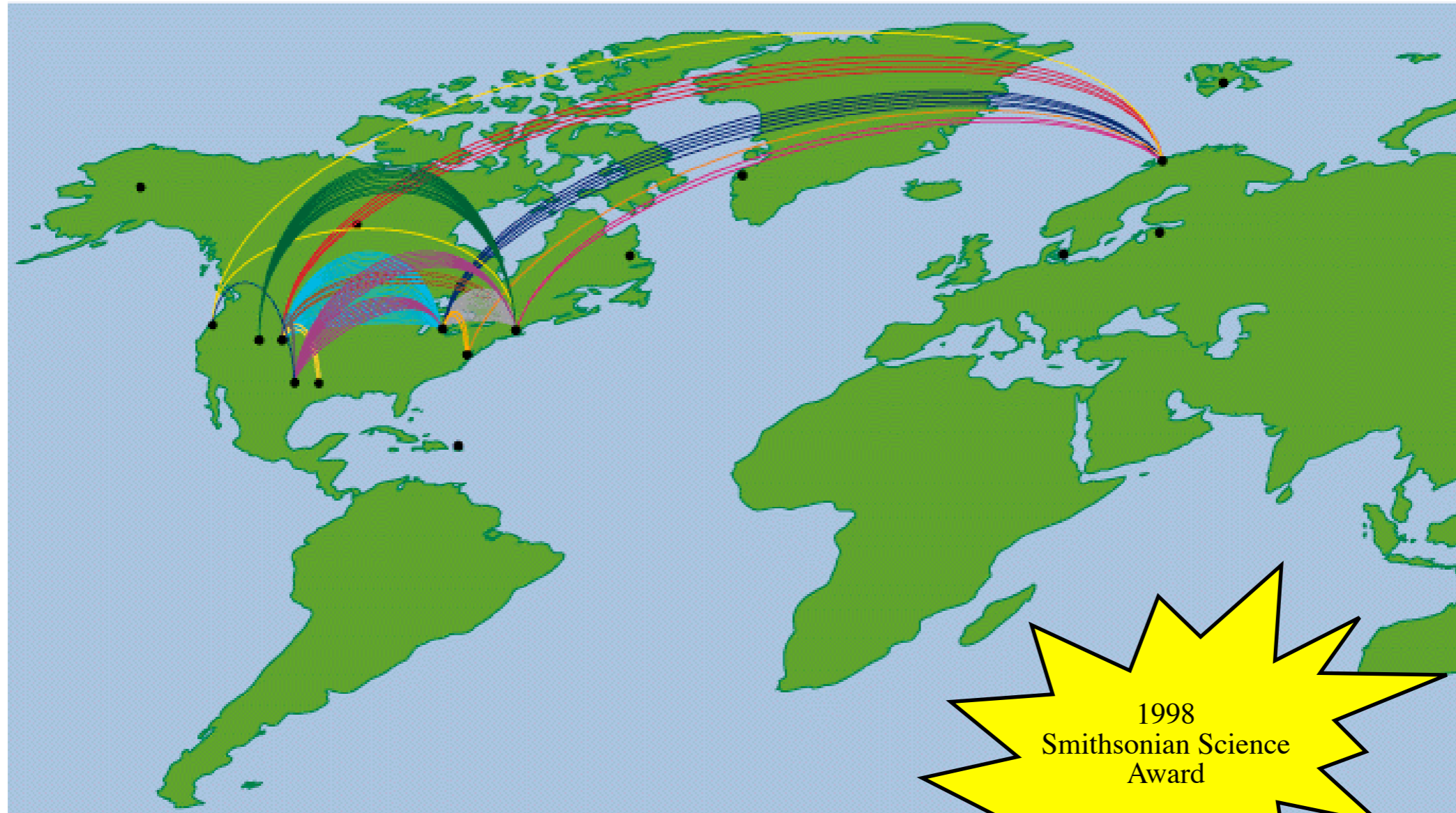**team chat**

Session replay

Archival data   **Journals**

Evolved into a Network of Instruments (one global instrument)

UARC Patterns of Communication

Pattern of Communication, UARC Campaign, April 9, 1997

1998
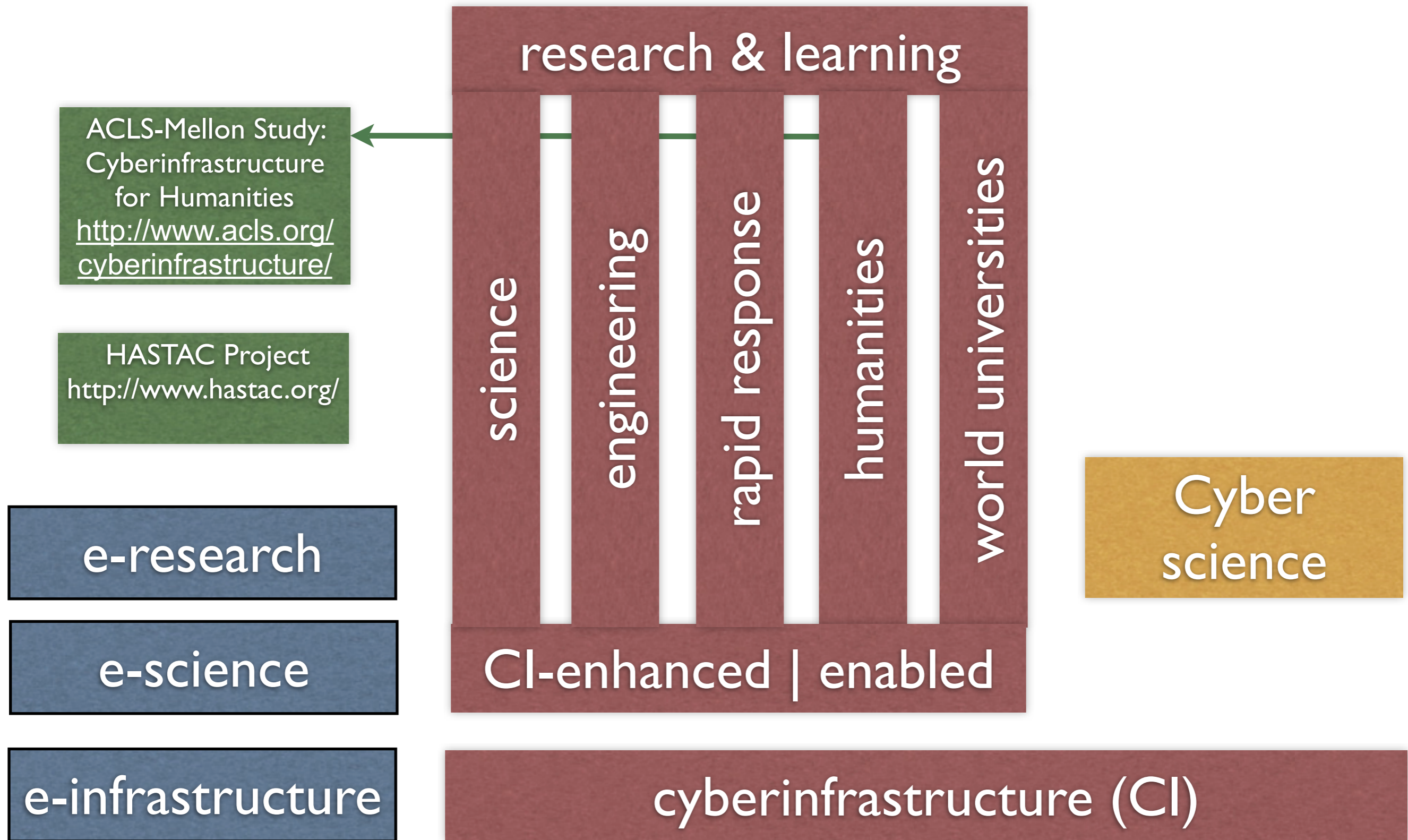Smithsonian Science
Award

Vignettes from UARC/SPARC

- Shared, tele-instruments, multiple eyes & complementary expertise.
- Rapid response, opportunistic campaigns.
- Isolated instruments became a global instrument chain.
- Cross-mentoring/training.
- New & earlier opportunities/exposure for grad students.
- Enhanced participation. Legitimate peripheral participation.

- Support for authentic, inquiry-based learning at UG and pre-college level.
- Distributed workshops for post-campaign data analysis.
- Session re-play for delayed participation and capture of process/provenance.
- Data-theory closure.
- A "living specification". Experiencing the new.

Current 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

e-research

e-science

e-infrastructure

Cyber science

cyberinfrastructure (CI)

Office of Cyberinfrastructure

e = electronic | enhanced | enabled

D. E. Atkins

Dualities

| cyberinfrastructure | → enables → ← | research & development |

CI is both an object and means for R&D

| collaboration | → enables → ← | CI environments |

Multi-stakeholder collaboration required to create, provision, and apply CI; CI supports collaborations across time and distance (geographic, disciplinary, institutional)

| learning | education | → enables → ← | CI environments |

Learning and workforce development initiatives required to create and use CI; CI enables/enhances learning/education

D. E. Atkins

NSF Blue Ribbon Advisory Panel on Cyberinfrastructure

"a new age has dawned in scientific and engineering research, pushed by continuing progress in computing, information, and communication technology, and pulled by the expanding complexity, scope, and scale of today's challenges. The capacity of this technology has crossed thresholds that now make possible a comprehensive Text "cyberinfrastructure" on which to build new types of scientific and engineering knowledge environments and organizations and to pursue research in new ways and with increased efficacy."

Revolutionizing Science and Engineering through Cyberinfrastructure:

Report of the National Science Foundation Advisory Panel on Cyberinfrastructure

February 3, 2003
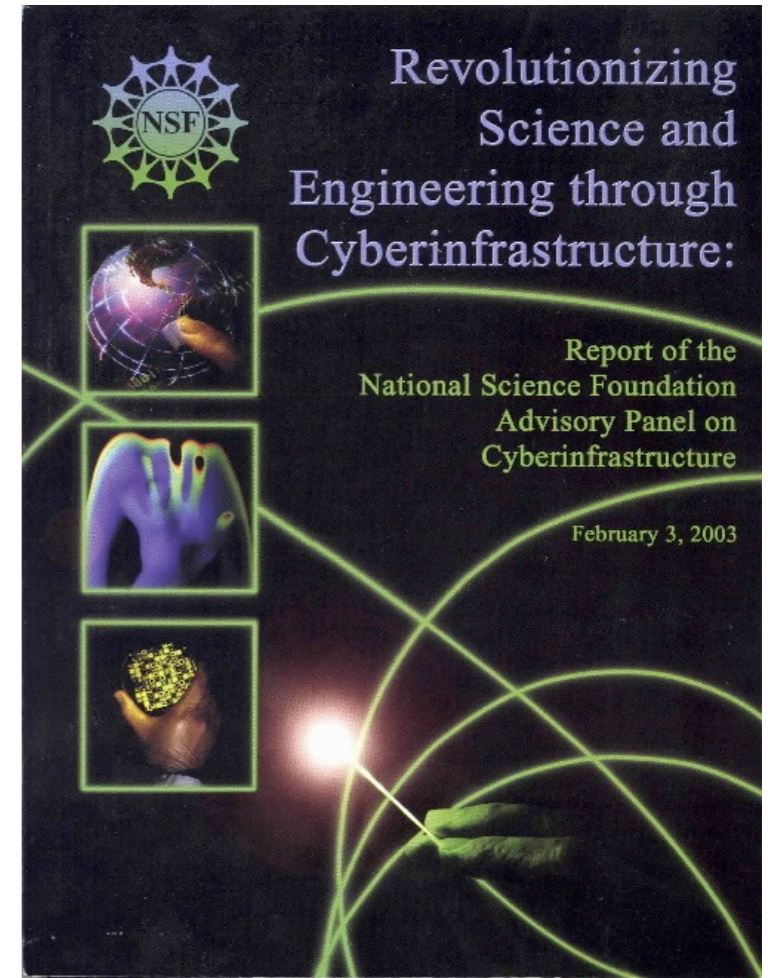
Daniel E. Atkins, Chair
University of Michigan
Kelvin K. Droegemeier
University of Oklahoma
Stuart I. Feldman
IBM
Hector Garcia-Molina
Stanford University
Michael L. Klein
University of Pennsylvania
David G. Messerschmitt
University of California at Berkeley
Paul Messina
California Institute of Technology
Jeremiah P. Ostriker
Princeton University
Margaret H. Wright
New York University

http://www.nsf.gov/od/oci/reports/toc.jsp

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

Revolutionizing Science and Engineering through Cyberinfrastructure:

Report of the National Science Foundation Advisory Panel on Cyberinfrastructure

February 3, 2003

30+ disciplinary workshops on CI vision & impact

NSF internal working groups

NSF'S CYBERINFRASTRUCTURE VISION FOR 21ST CENTURY DISCOVERY

NSF Cyberinfrastructure Council

NSB & Community Input

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

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.*

Cyberinfrastructure

www.nsf.gov/oci/

OCI to help the NSF community define and occupy the upper reaches of the "CI vision space"
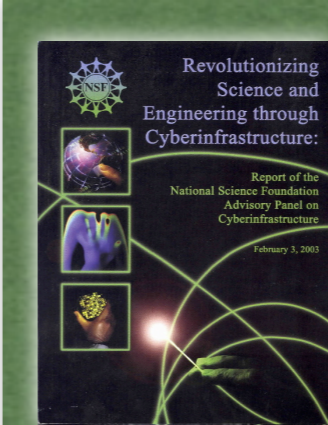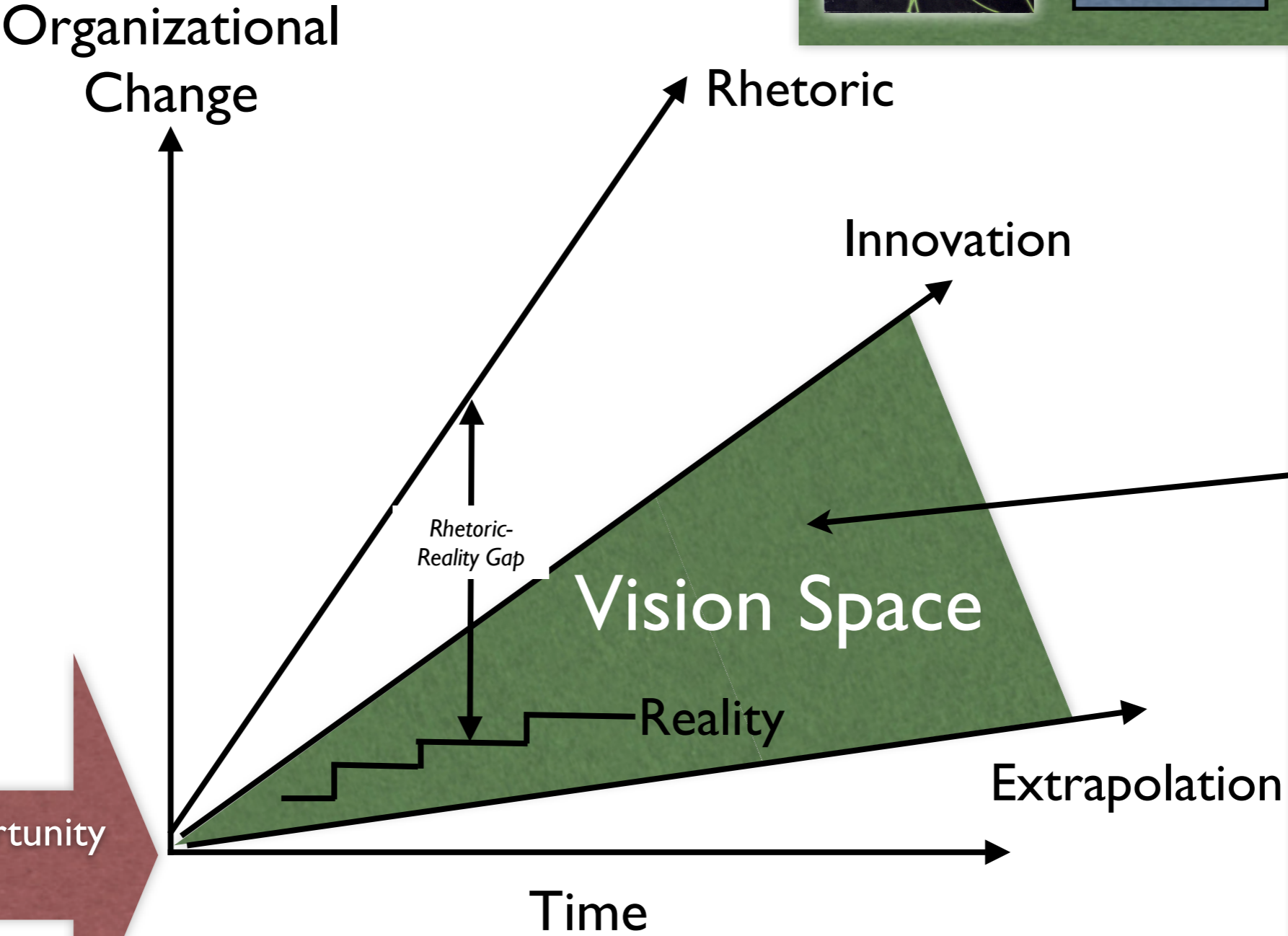
Adopted from John Anderson, formerly of EuroPARC.

D. E. Atkins

Office of Cyberinfrastructure
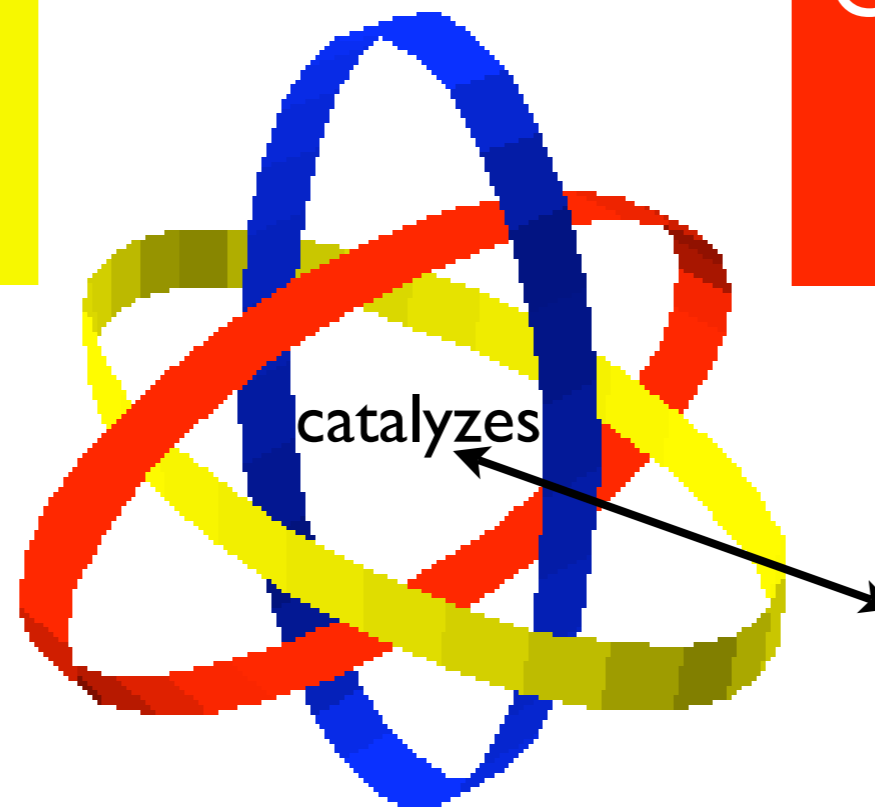
13

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.

D. E. Atkins

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/

catalyzes

provides shared and connecting CI

Office of Cyberinfrastructure

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

D. E. Atkins

High Performance Computing

increasingly important tool for understanding
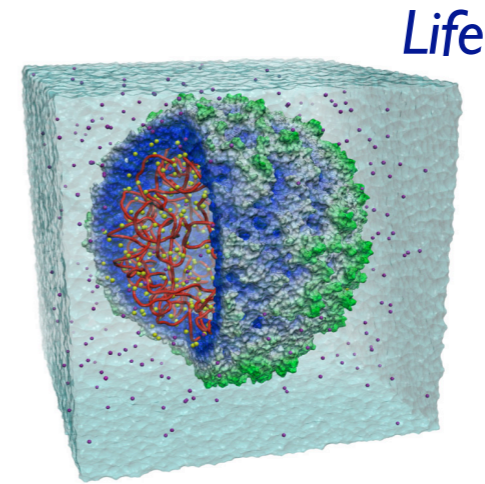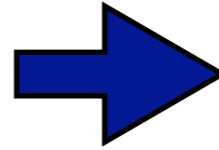
Life
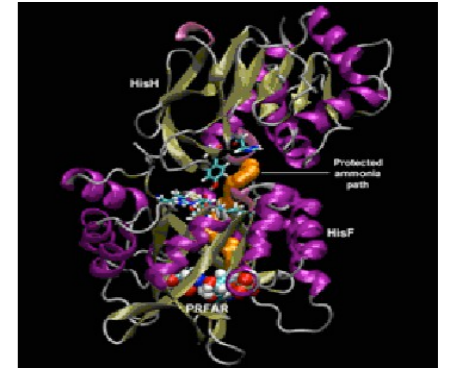
Satellite tobacco mosaic virus, P. Freddolino et al.

Aldehyde dehydrogenase, T. Wymore and S. Brown
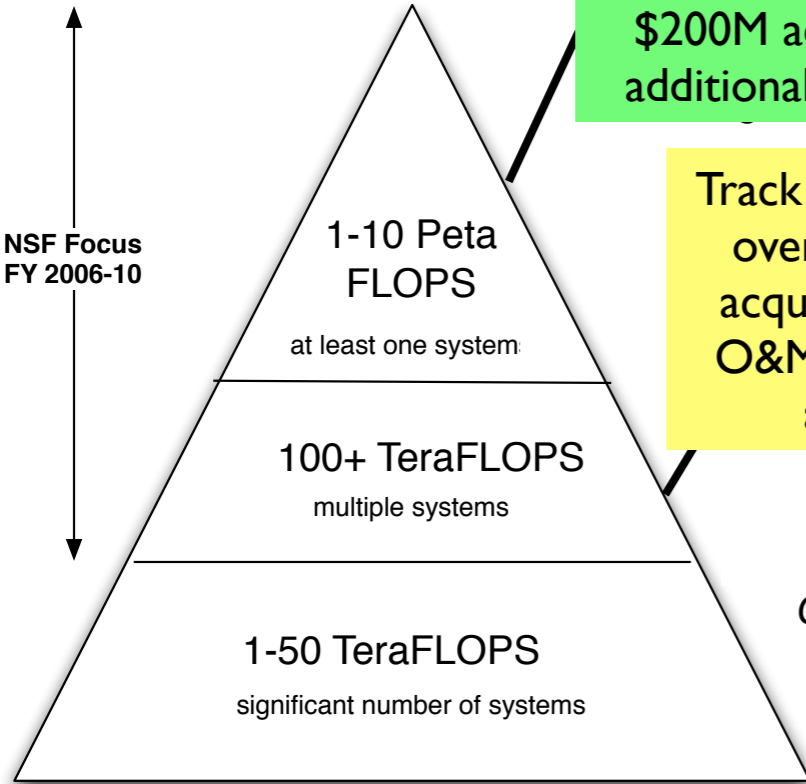
Track 1: One solicitation funded over 4 years: $200M acquisition + additional O&M cost.

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

NSF Focus FY 2006-10

1-10 Peta FLOPS
at least one systems

100+ TeraFLOPS
multiple systems

1-50 TeraFLOPS
significant number of systems

Campus Level
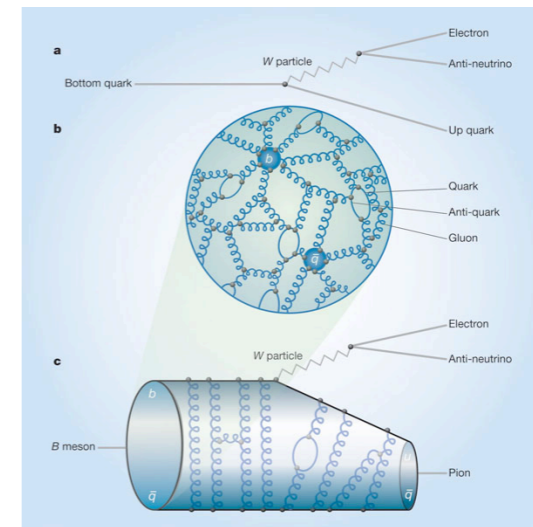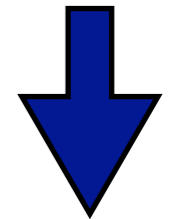
Matter
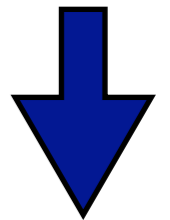
I. Shipsey

The Environment

Society

John Q Public
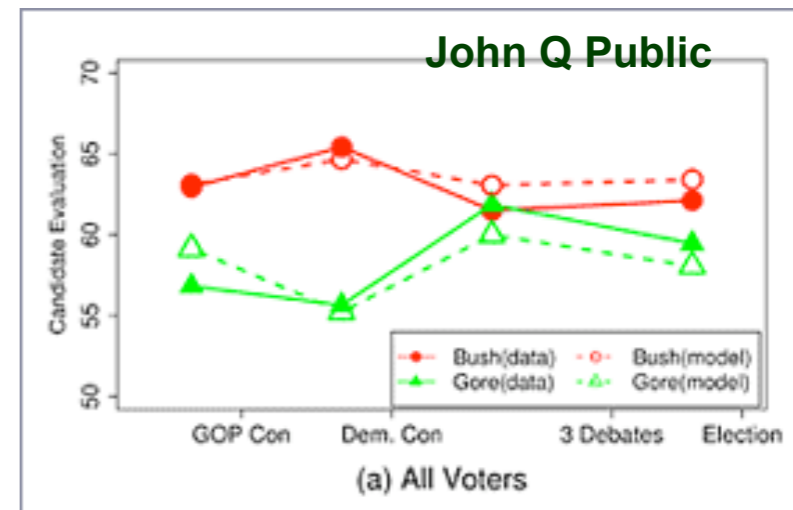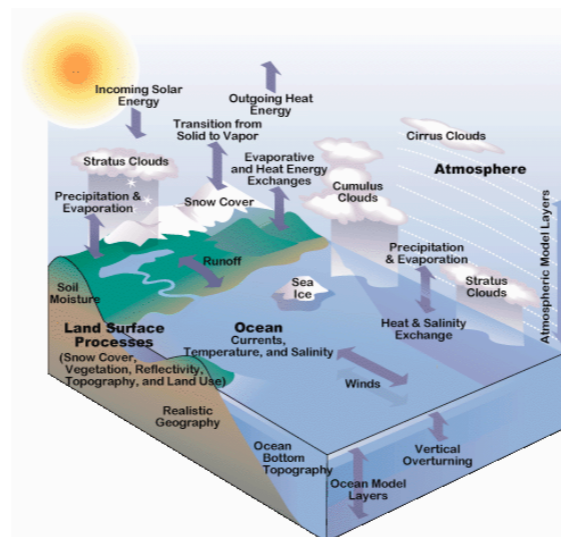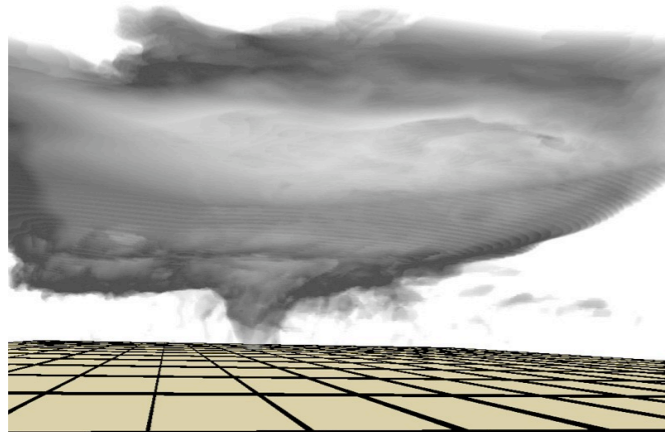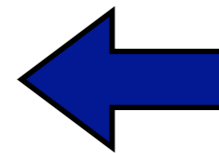
S.-Y. Kim, M. Lodge, C. Taber.

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
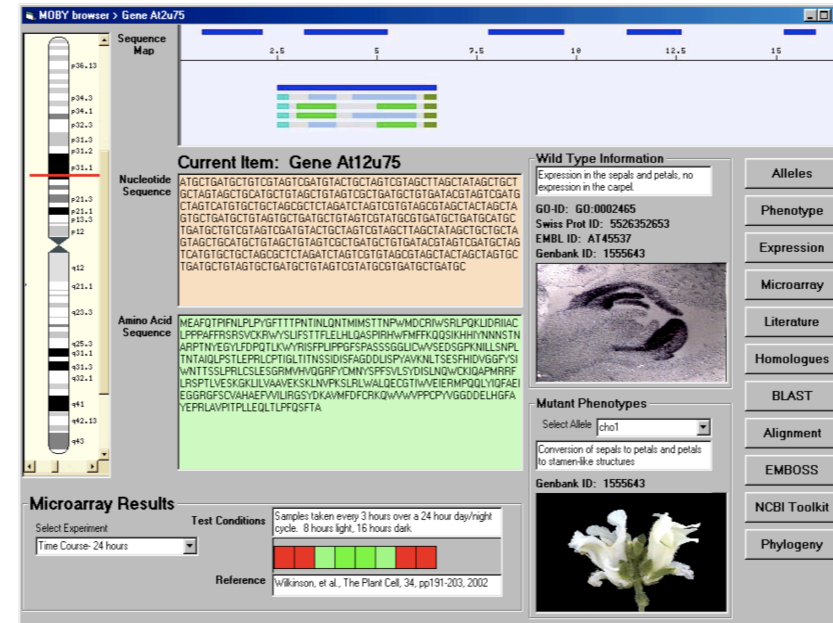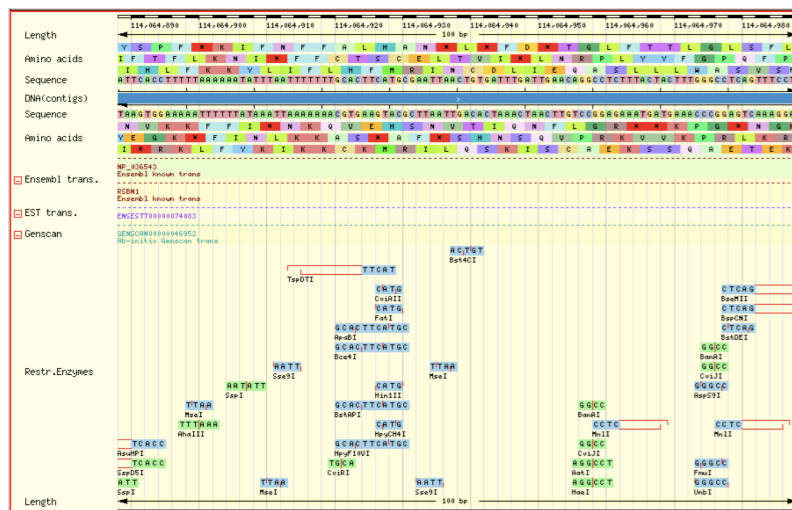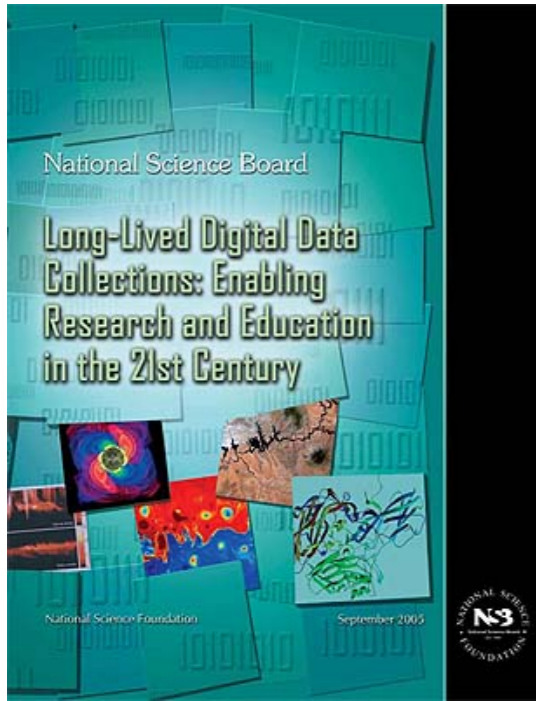
New Report: *To Stand the Test of Time*
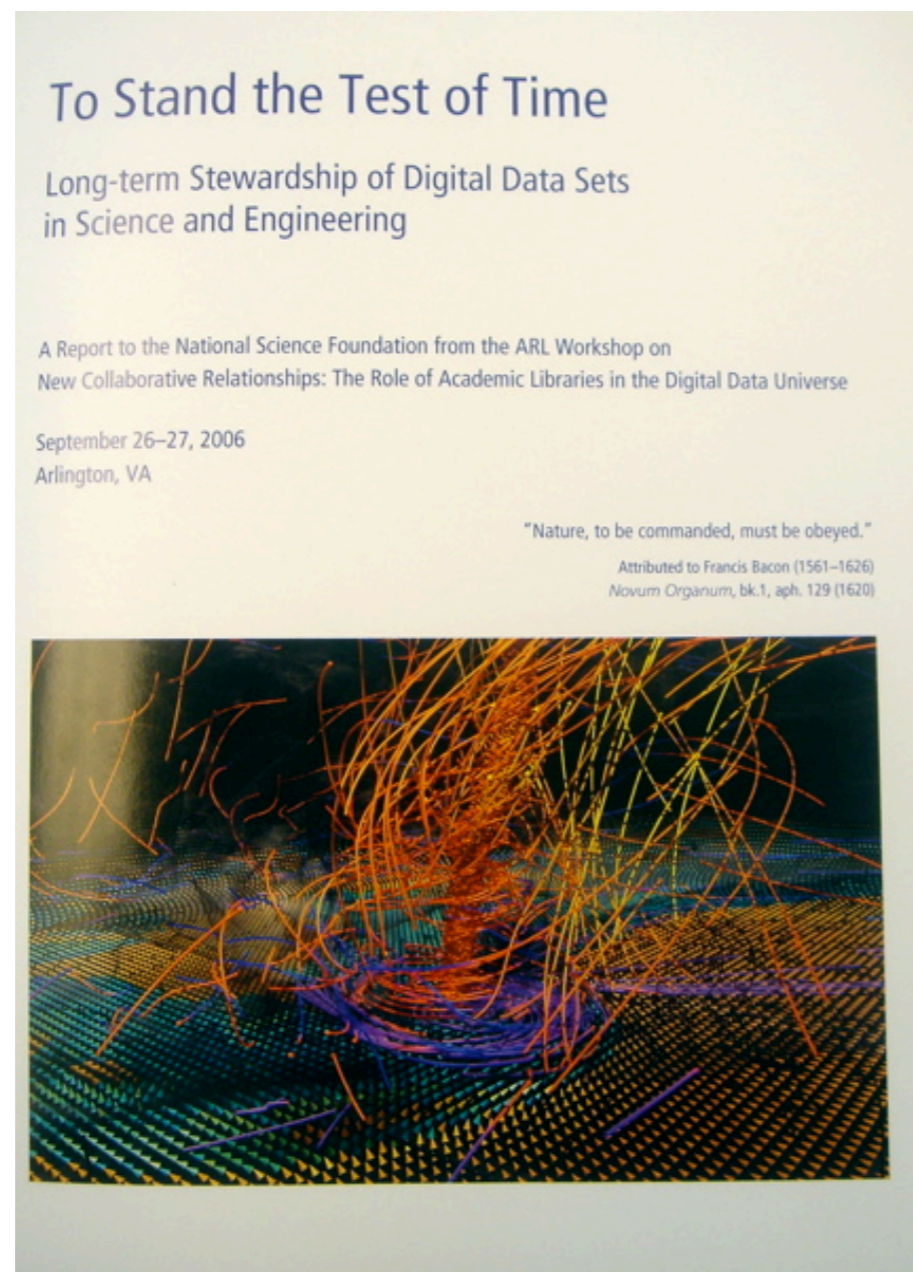
Available online at http://www.arl.org/info/events/digdatarpt.pdf

Lead NSF Program Officer for Data Initiatives, Chris Greer is at this meeting.

Virtual Organizations

NVO

LEAD

iVDgL

- To catalyze the development, implementation and evolution of a national cyberinfrastructure that integrates both physical and cyberinfrastructure assets and services.

- To promote and support the establishment of world-class VOs that are secure, efficient, reliable, accessible, usable, pervasive, persistent and interoperable, and that are able to exploit the full range of research and education tools available at any given time

- To support the development of common cyberinfrastructure resources, services, and tools that enable the effective, efficient creation and operation of end-to-end cyberinfrastructure systems for and across all science and engineering fields, nationally and internationally.

eGee Enabling Grids for E-sciencE

TeraGrid

Open Science Grid

NEES
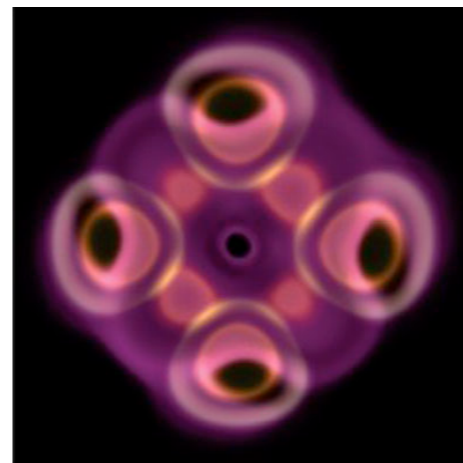
NanoHub
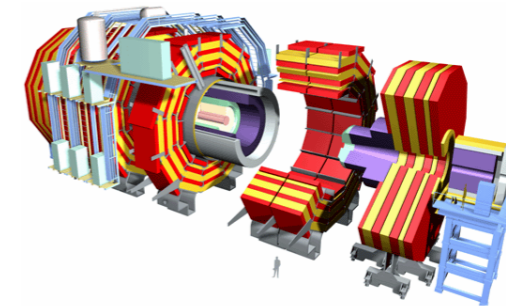
CMS
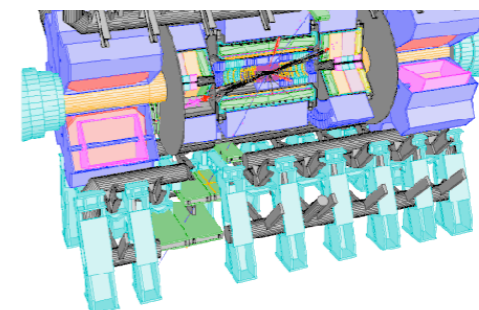
ATLAS

Instances of Virtual Organizations (VOs)

People* People* People*

Interfaces for interaction, workflow, visualization and collaboration for distributed teams in domain/project specific and potentially functionally-complete VOs.

Mechanisms for flexible secure, coordinated resource/services sharing among dynamic collections of individuals, institutions, and resources (the Grid or service layer problem)

Distributed, heterogeneous services for:

Computation

Data, information management

Sensing, observation, activation in the world

Alternate Names for Instances of VOs:
• Co-laboratory
• Collaboratory
• Grid (community)
• Network
• Portal
• Gateway
• Hub
• Virtual Research Environment (VRE)
• Cyberinfrastructure Collaborative

* People engaged in discovery and learning as individuals and in teams.

D. E. Atkins

CI/VO Enabled Science
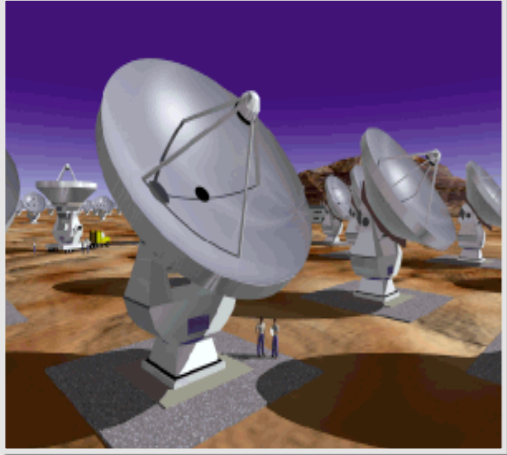
NVO and ALMA
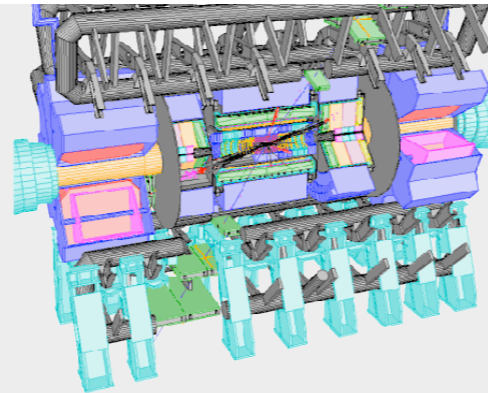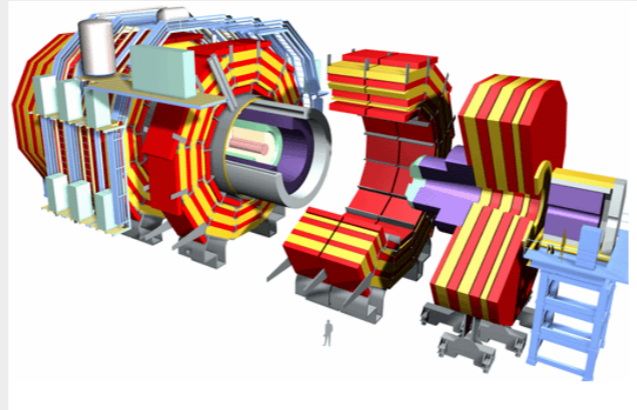
ATLAS and CMS
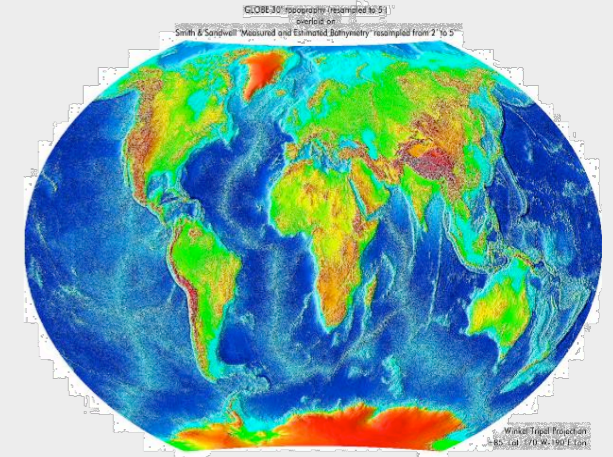
Climate Change

LIGO
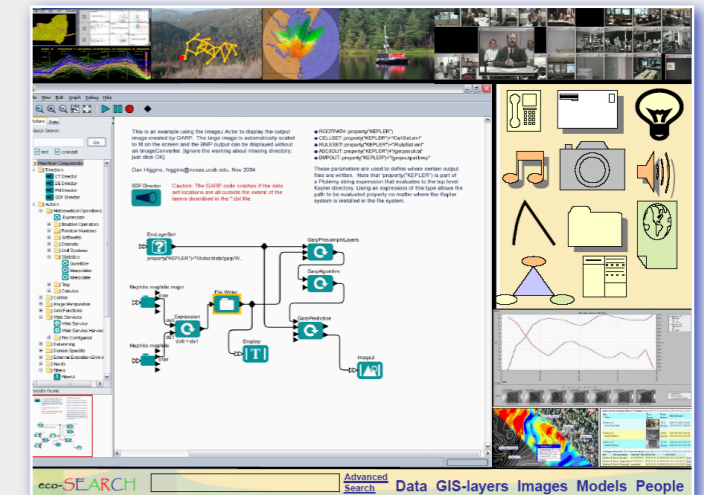
NEON

The number of international-scale projects is growing rapidly!

Alternate Generic Names for *Virtual Organizations*

- Distributed Knowledge Communities

- Virtually-augmented Communities

- Participation Environments built upon Participation Architectures

- Learning and Discovery Ecosystems

D. E. Atkins

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

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

Geographic Place

Physical + Virtual, Not Physical vs. Virtual

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.

Optimizing CI Investment

The Grid Movement

Let's look at a few real example Grid Science Gateways

(about a dozen … many more exist!)

These example slides courtesy of D. Gannon

NEESGrid

Realtime access to earthquake
Shake table experiments at remote sites.

BIRN – Biomedical Information

Mesoscale Meteorology

NSF LEAD project - making the tools that are needed to make accurate predictions of tornados and hurricanes.
- Data exploration and Grid workflow

The LEAD Vision: Adaptive Cyberinfrastructure
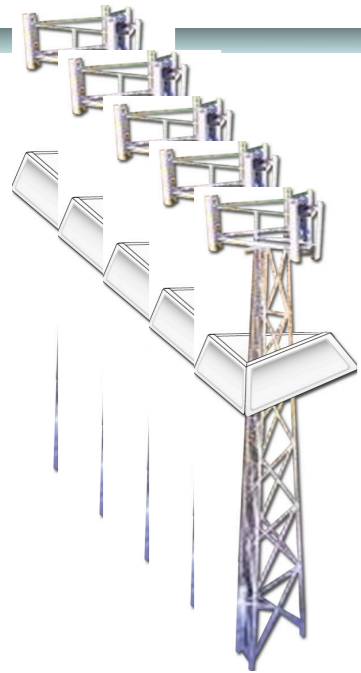
DYNAMIC OBSERVATIONS
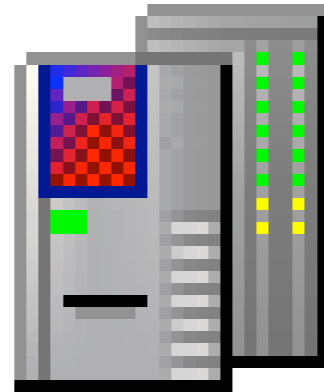
Analysis/Assimilation

Quality Control
Retrieval of Unobserved Quantities
Creation of Gridded Fields
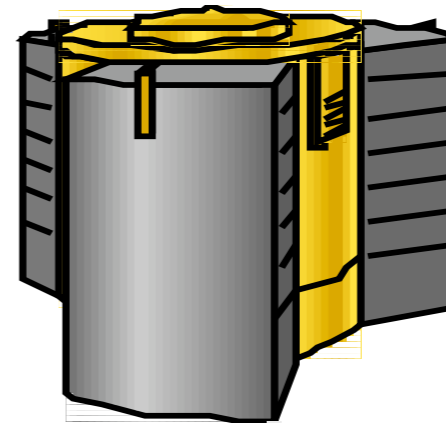
Prediction/Detection

PCs to Teraflop Systems

Product Generation, Display, Dissemination

Models and Algorithms Driving Sensors

End Users

NWS
Private Companies
Students

The CS challenge: Build cyberinfrastructure services that provide adaptability, scalability, availability, useability, and real-time response.

From D. Gannon

Nanohub - nanotechnology

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

Engaging People in Cyberinfrastructure

EPIC

MARIACHI \mär-ē-ˈach-ē\ n: The Study of Ultra High Energy Cosmic Rays.
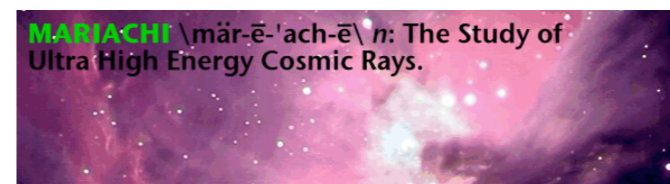
MARIACHI

cyberbridges

CyberBridges

To be a genuinely competitive knowledge economy, Europe must be better

- **in producing knowledge through research**
- **in diffusing it through education**
- **in applying it through innovation**
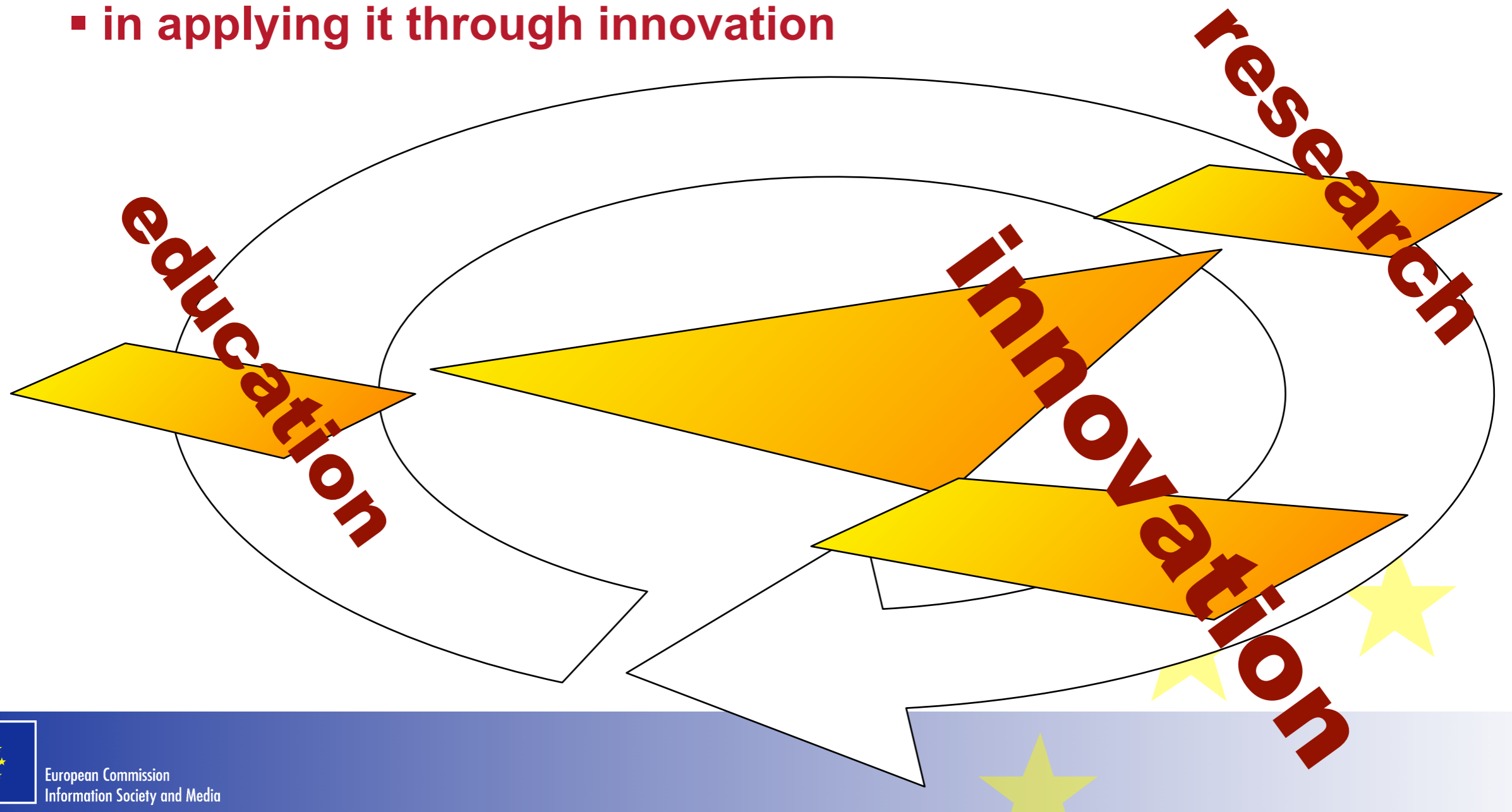
Bringing the best brains together

Sharing the best scientific resources

Weather Forecast VO

Biomedics VO

Astrophysics VO

e-Infrastructure

Scientific Data

Sharing Scientific Resources

Communication Network

. . . .
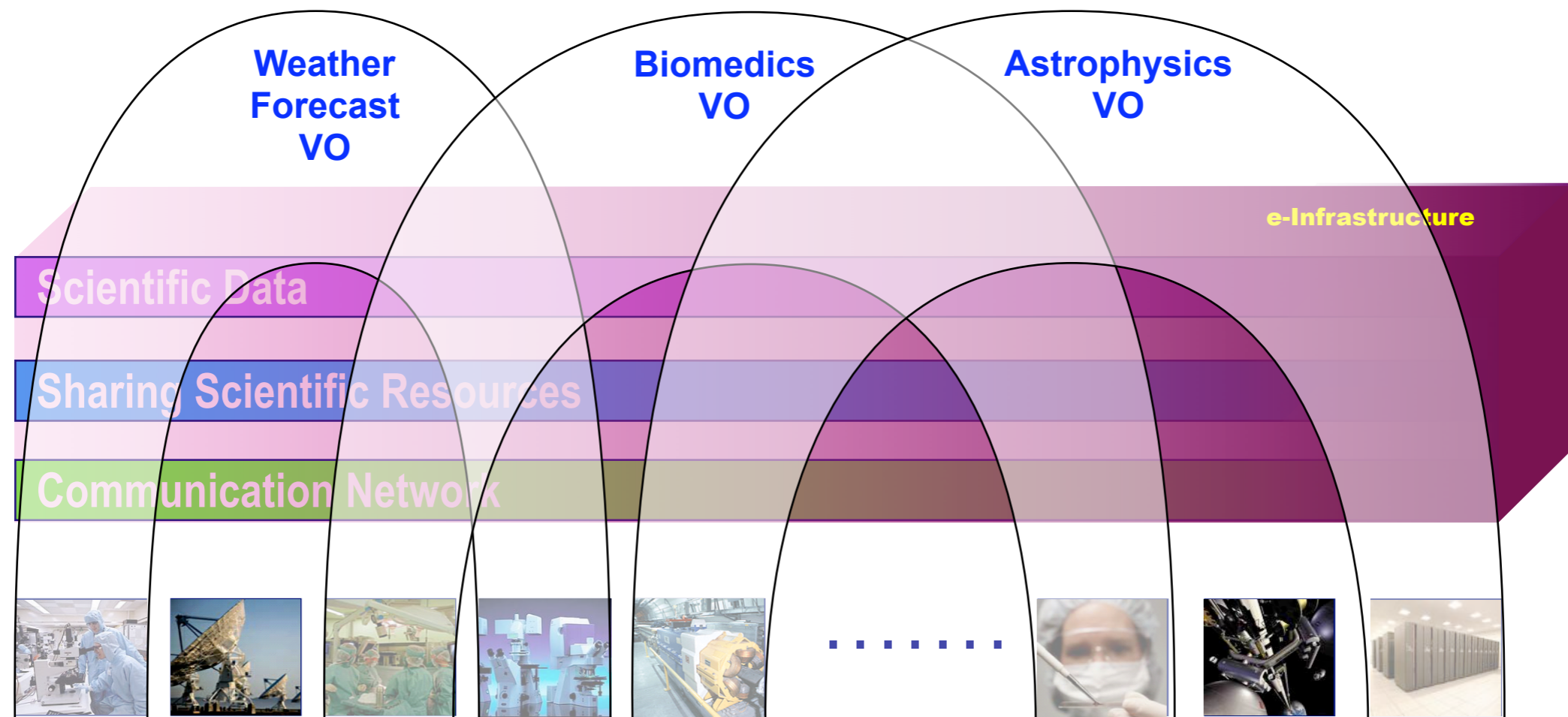
Producing the best science

And while we are at it...

- can we create CI platforms in support of research, learning, and societal engagement in ways that exploit complementarity between them?

  - *"Better than being there."*
  - *Pasteur's Quadrant research*
  - *Ubiquitous learning ecologies*

teaching learning

service engagement

Knowledge Communities Enhanced by Cyberinfrastructure

research scholarship
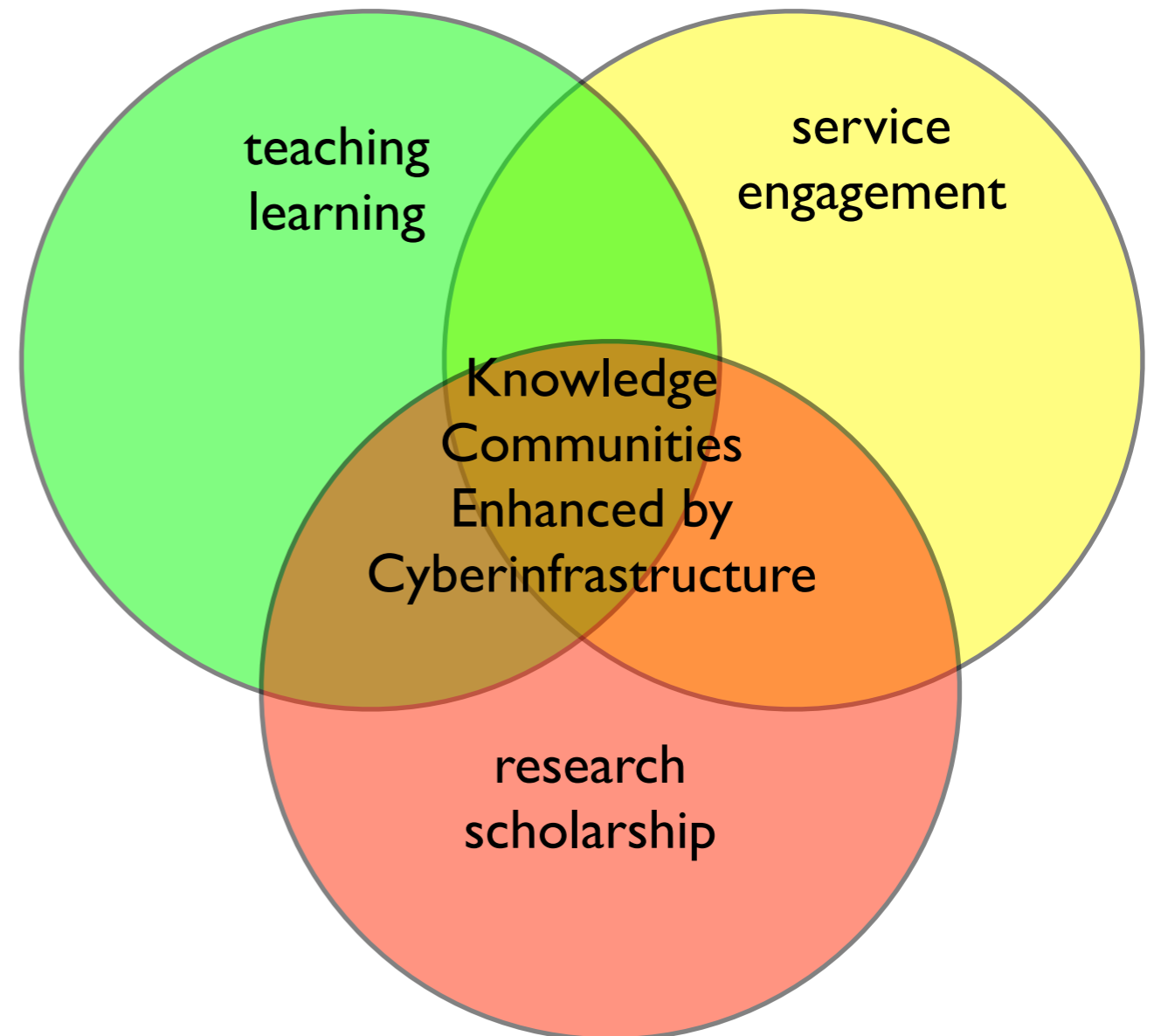
The opportunity is more than distributed computing and data storage. It is about using ICT/CI as an enhancement platform for knowledge-based communities to learn, discover, and engage in more "barrier free" ways.

D. E. Atkins
atkins@umich.edu

The Meta University
A Personal View

What we are observing is the early emergence of a *Meta University* -- a transcendent, accessible, empowering, dynamic, communally-constructed framework of open materials and platforms on which much of higher education worldwide can be constructed or enhanced.

From Charles M. Vest, President Emeritus of MIT   37

The Meta University

- Will enable -- not replace -- residential campuses
- Will bring cost efficiencies to institutions through shared development;
- Will be adaptable -- not prescriptive;
- Will serve both teachers and learners;
- Will speed the propagation of high-quality education and scholarship;
- Will build capacity for economic development;
- Will build bridges across cultures and political boundaries; and
- Will be particularly important to the developing world.

From Charles M. Vest, President Emeritus of MIT   38

Something to ponder.....

Global Needs

Half of the world's population is under 20 years old.

Today, there are over 30 million people who are fully qualified to enter a university, but there is no place available. This number will grow to over 100 million during the next decade.

To meet the staggering global demand for advanced education, a major university would need to be created every week.

"In most of the world, higher education is mired in a crisis of access, cost, and flexibility. The dominant forms of higher education in developed nations–campus based, high cost, limited use of technology–seem ill-suited to addressing global education needs of the billions of young people who will require it in the decades ahead."

Sir John Daniels

From James J. Duderstadt, President Emeritus, U. of Michigan

International e-science (cyberinfrastructure) movement

Office of Cyberinfrastructure

D. E. Atkins

Openness Movement

Openness Movement
Culture of Contribution

Open Courseware Consortium

Creative Commons Licensing

Open Repositories, Digital Libraries, Archives

Open Software: middleware, toolkits, apps, services

Open scholarly communication

Web 2.0 Collaborative contribution environments
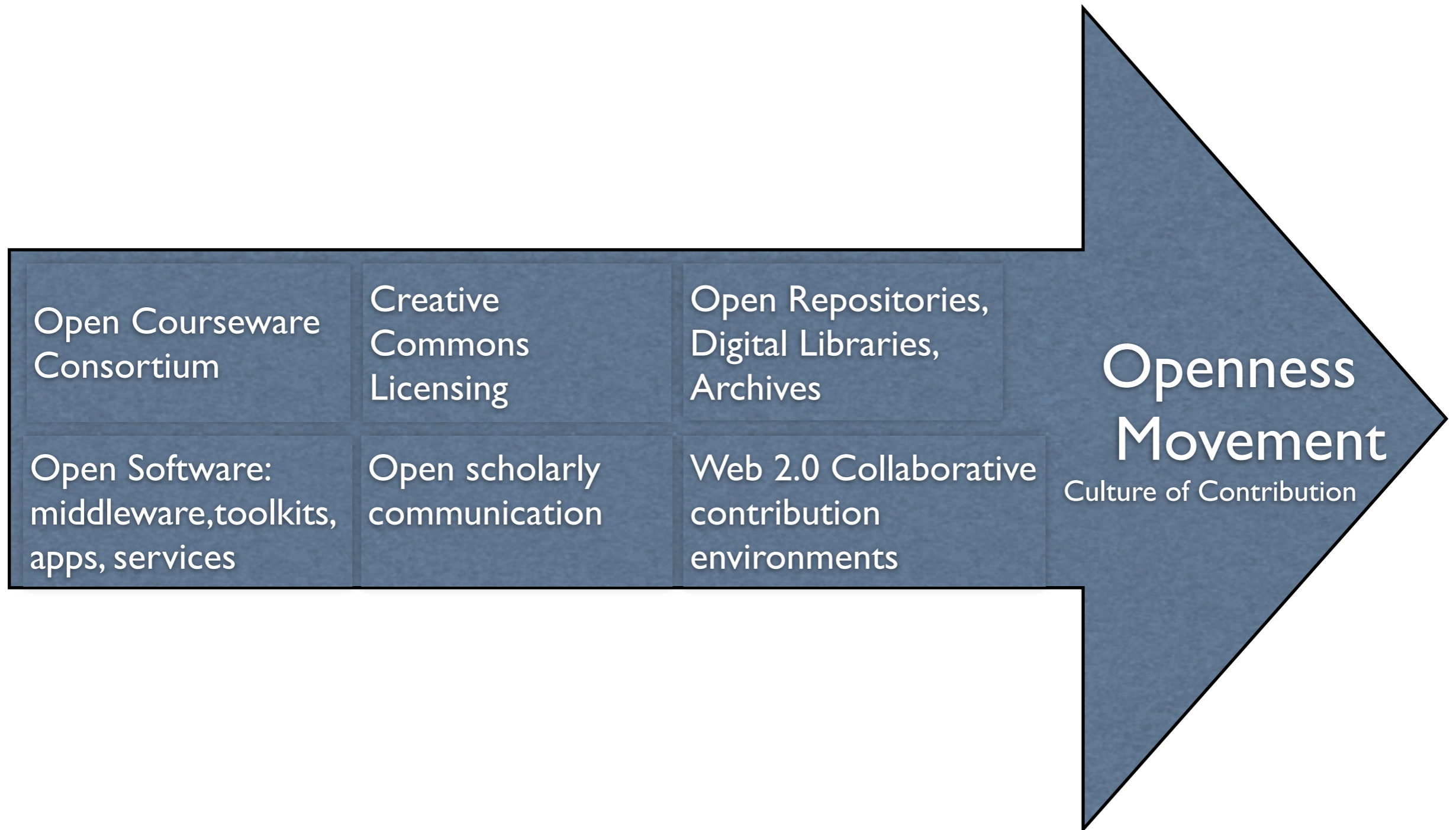
Concepts, Visions of the Future of Learning & Discovery

Meta-University, Universities in and of the World	National Academies ITFRU Studies	OECD and other international visioning activities
JSB Culture of Learning	NSF CI Vision, ACLS CI & Humanities	OTHER?

Studies, Concepts, Visions

D. E. Atkins

Multiple Stakeholders

Federal and State Government	Philanthropy	Industry
Academia	Civil Society	International version of all these

Stakeholders
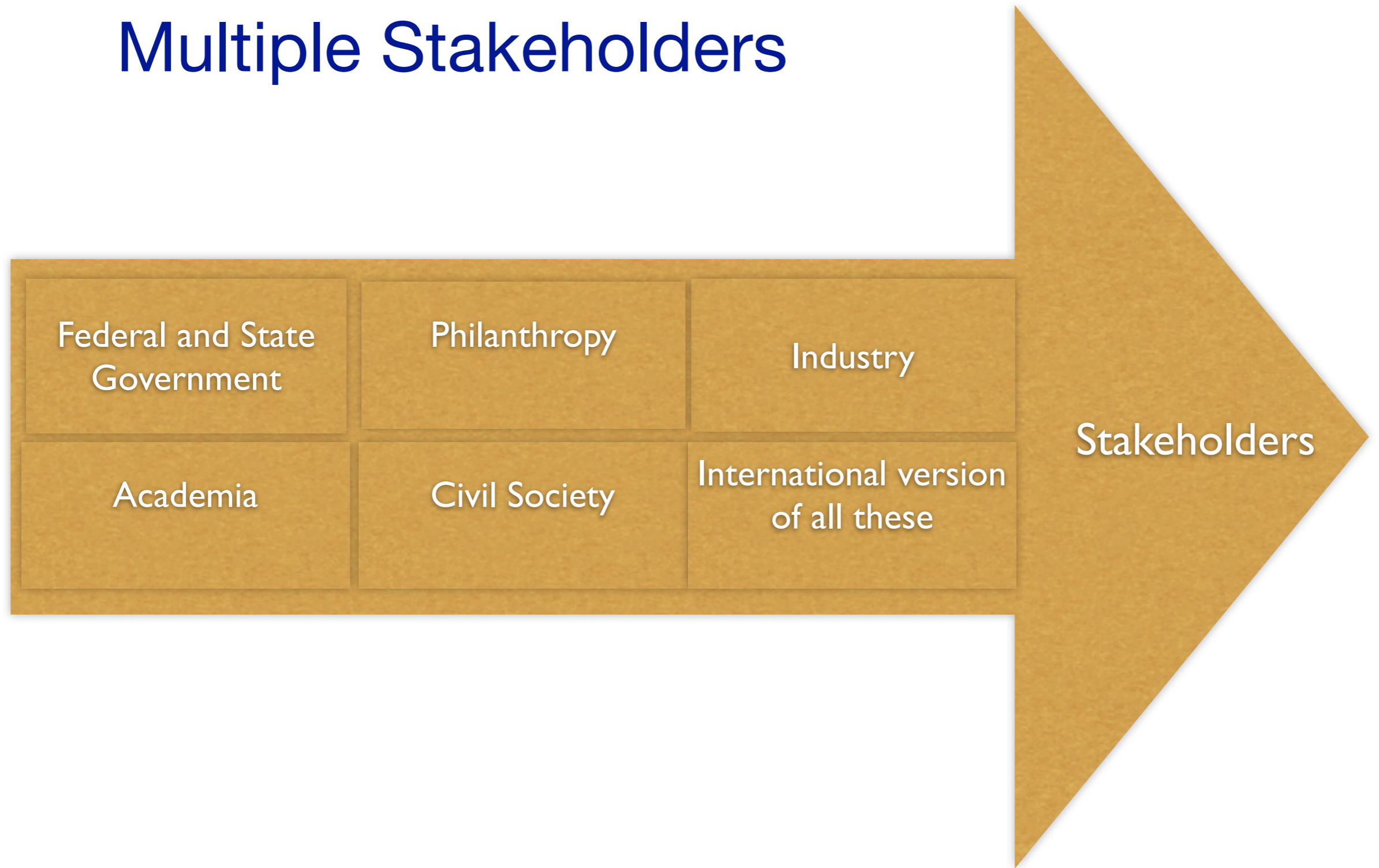
Gaming, collaboration, virtual, immersive and augmented reality

CI Movement

Openness Movement

Visions, Frameworks

Stakeholders

Gaming,...

Opportunity for Innovation and Leadership in Learning and Discovery

Critical Societal Needs

D. E. Atkins

Questions

- The LHC community has made massive investments in high-capacity, globally-distributed grid computing to support the data analysis mission. But to what it extent is it leveraging this investment to reduce the barriers of geography for the thousands of people involved in LHC projects?

  - Being at CERN when on your campus. Being on your campus when at CERN.

  - Broadened participation.

  - Research and learning, including participation by students, citizen scientists, general public.

D. E. Atkins

- Could at least a subset (microcosm?) of the LHC community develop a prototype of a "global meta university;" an open ecosystem of discovery and learning.

  - Primary goal: more discovery and learning in the field of physics.

  - Secondary goal - a large scale prototype of an open, global scale, "meta-university"?

  - Possibility of broadened set of sponsors.

- Why not leadership in new CI-enabled platforms for participation in learning and discovery from the particle physics community?

D. E. Atkins

- Perhaps such global-scale CI-enabled platforms for participation will be **essential** for success with the next generation linear accelerator.

- Furthermore perhaps eventually it can mitigate against the growing concern about the expense of creation, operation, and sustaining very expensive research infrastructure, and the attendant erosion of research funding for the use of such facilities.

Questions & Discussion