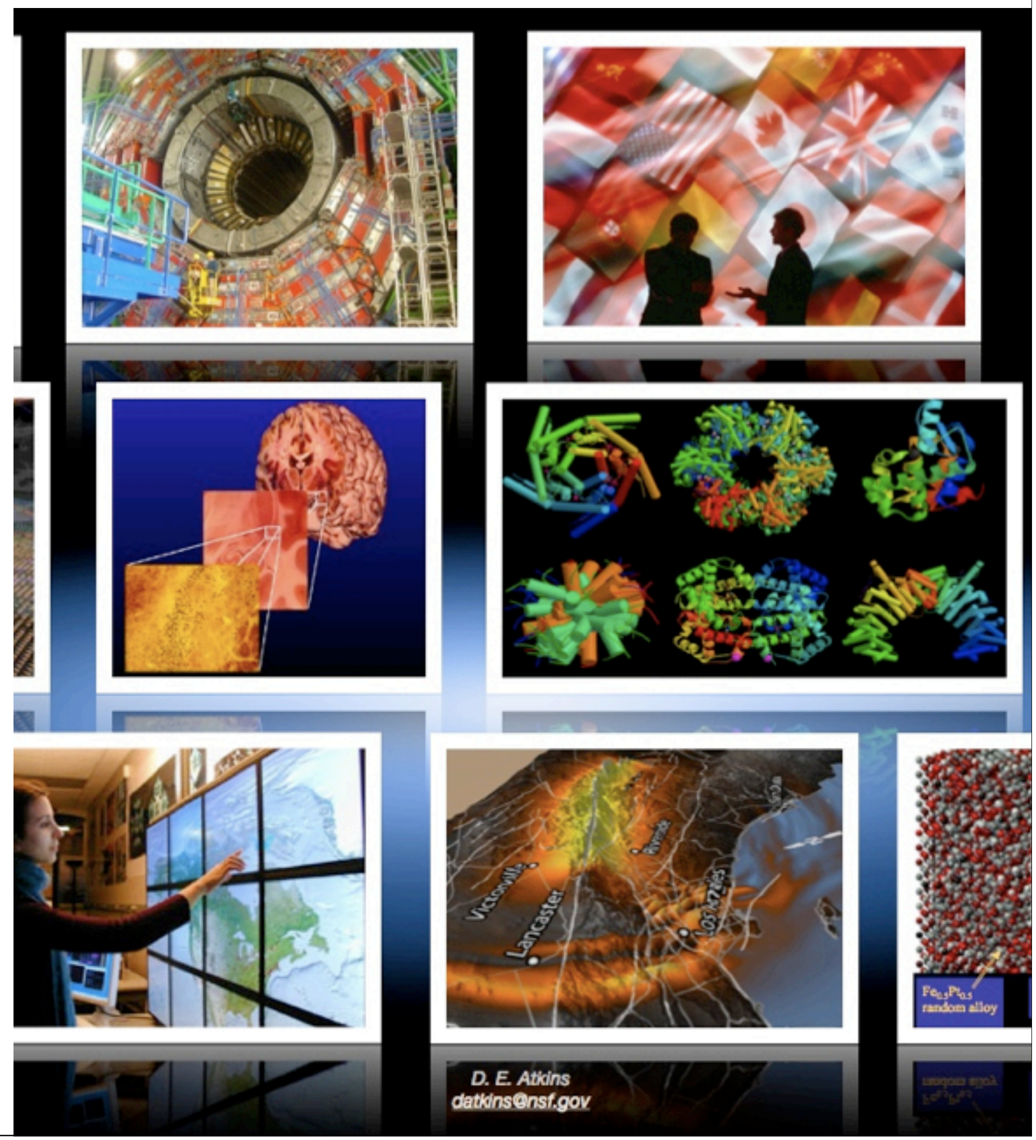
The International e-Science Movement: Status and Future

Daniel E. Atkins
W. K. Kellogg Professor of Community Informatics
Professor of EE and Computer Science
Associate VP for Research Cyberinfrastructure
University of Michigan, Ann Arbor, MI USA
atkins@umich.edu



Three Perspectives

- Funders
  - US initiatives: Cyberinfrastructure (CI) enabled science
  - UK initiatives: e-Science
- Researchers
- Universities

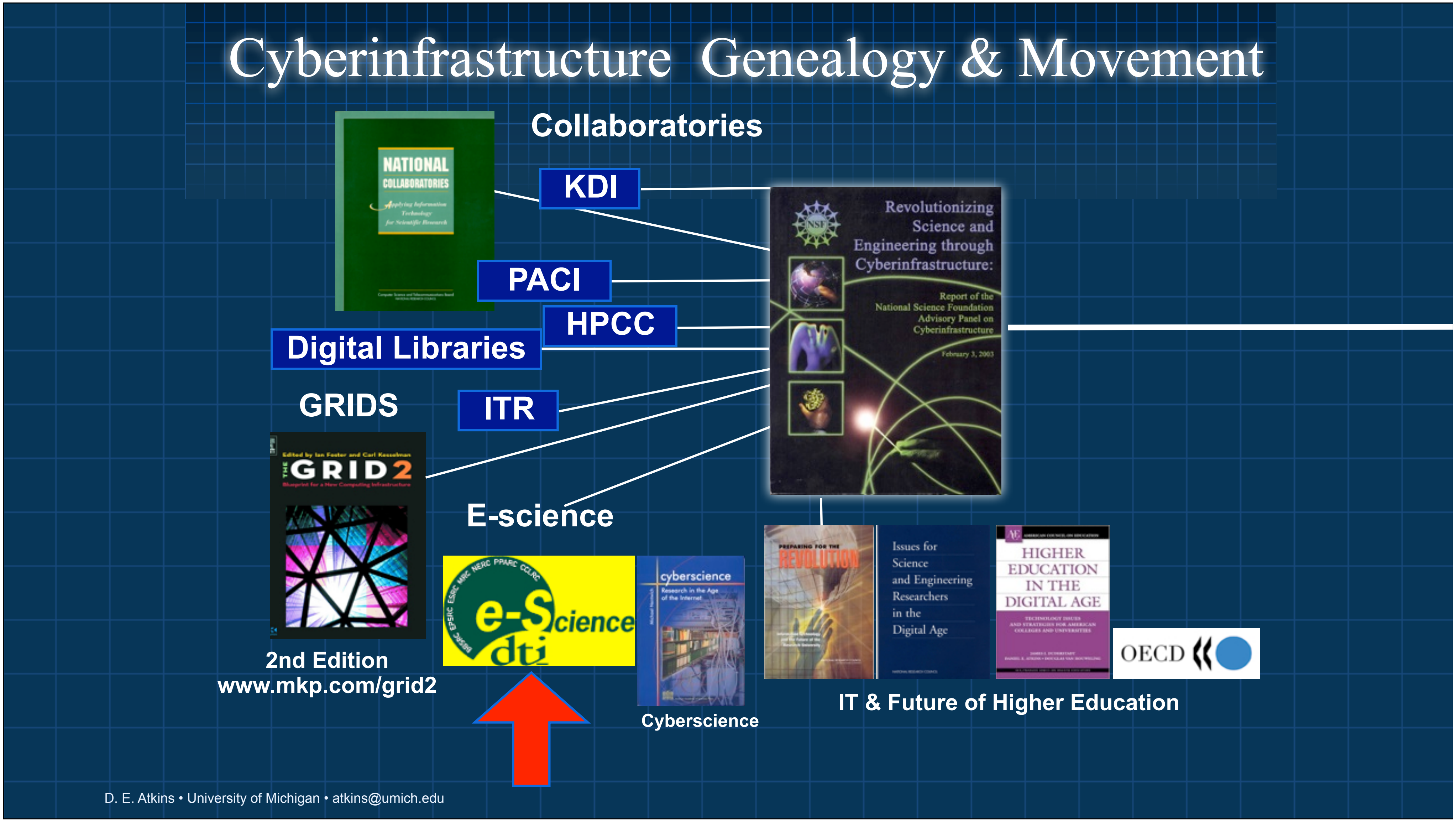
"Matrices of relationships, approaches and sets of goals"

Future Focus

- Computational Discovery, Computational Thinking for All at Various Service Levels
  - Modeling, Simulation, Prediction, Exacting Knowledge from Data (of all kinds) UK initiatives: e-Science, “Fourth Paradigm”
- ICT/Cyberinfrastructure Rationalization, Multi-Use Cyberinfrastructure. Focus on Efficiency and Mission Effectiveness, Consolidate Demand and Sourcing of Services/
- End-to End Performance: Portfolio of Options, Campus and “Above the Campus”, International
- HPC: High Performance Computing --> High Performance *Collaboration*
- *Managing Data in Difficult Times: policies, strategies, technologies and infrastructure to manage research and teaching data in a fast changing technological and economic environment. Who has the responsibility?*

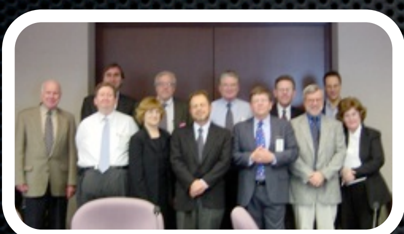
Definition should likely now be made more agnostic wrt technology

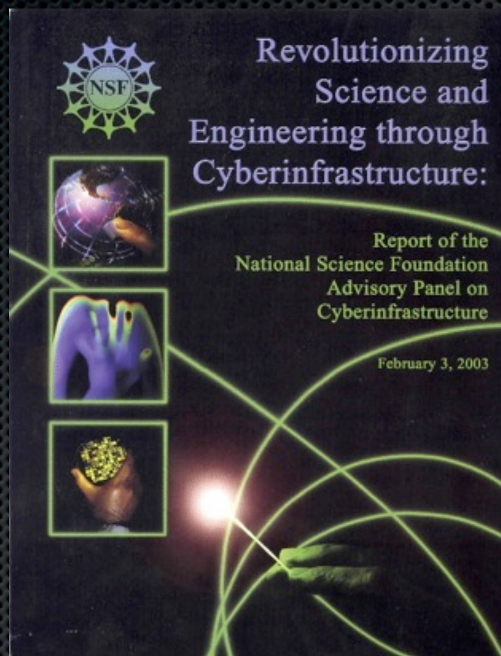
Original UK Definition of e-science by Sir John Taylor


Cyberinfrastructure Genealogy & Movement
Collaboratories
NATIONAL COLLABORATORIES
Applying Information Technology for Scientific Research
KDI
PACI
HPCC
Digital Libraries
ITR
GRIDS
GRID 2
2nd Edition
www.mkp.com/grid2
E-science
e-Science
dti
cyberscience
Cyberscience
Revolutionizing Science and Engineering through Cyberinfrastructure:
Report of the National Science Foundation Advisory Panel on Cyberinfrastructure
February 3, 2003
PREPARING FOR THE REVOLUTION
Issues for Science and Engineering Researchers in the Digital Age
HIGHER EDUCATION IN THE DIGITAL AGE
OECD
IT & Future of Higher Education
D. E. Atkins • University of Michigan • atkins@umich.edu

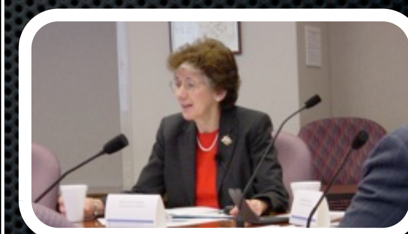
NSF Blue Ribbon Advisory Panel on Cyberinfrastructure

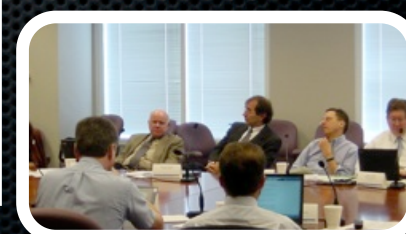
Daniel E. Atkins, Chair
University of Michigan
Kelvin K. Droegemeier
University of Oklahoma
Stuart I. Feldman
IBM
Hector Garcia-Molina
Stanford University
Michael L. Klein
University of Pennsylvania
David G. Messerschmitt
University of California at Berkeley
Paul Messina
California Institute of Technology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





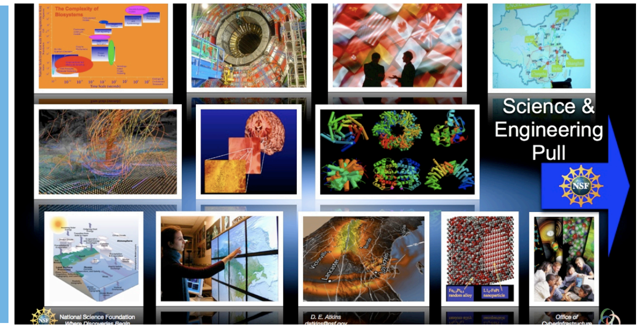
"a new age has dawned in scientific and engineering research, pushed by continuing progress in computing, information, and communication technology, and pulled by the expanding complexity, scope, and scale of today's challenges. The capacity of this technology has crossed thresholds that now make possible a comprehensive "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."



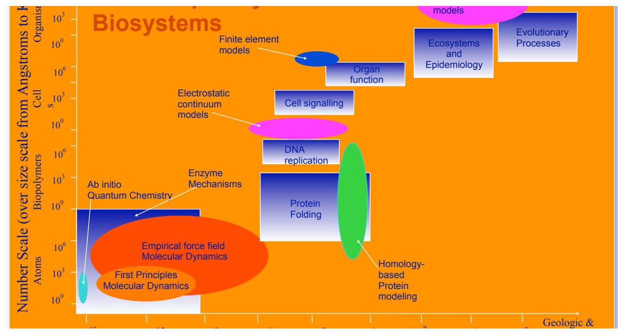


http://www.nsf.gov/oci

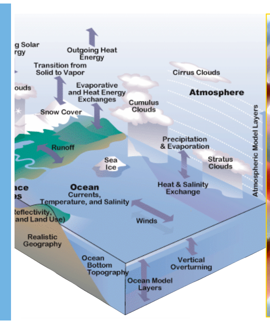
e-science (research enabled by e-infrastructure/ICT) is increasingly essential for meeting 21st century challenges in scientific discovery and learning



The inherent complexity, multi-scale, and multi-science nature of today's frontier science challenges.

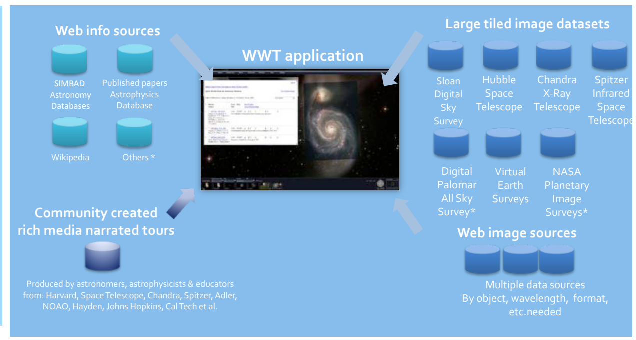


The accompanying requirement for multi-disciplinary, multi-investigator, multi-institutional approach (often international in scope).

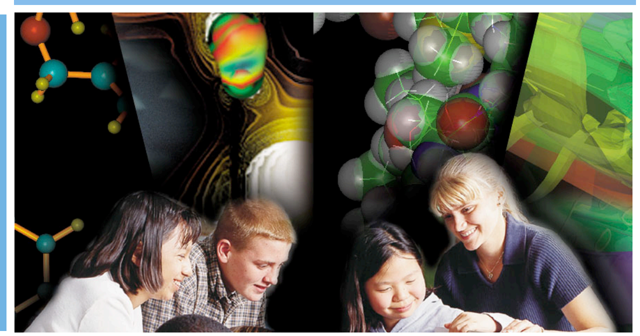


The high data intensity and heterogeneity from simulations, digital instruments, sensor nets, and observatories.

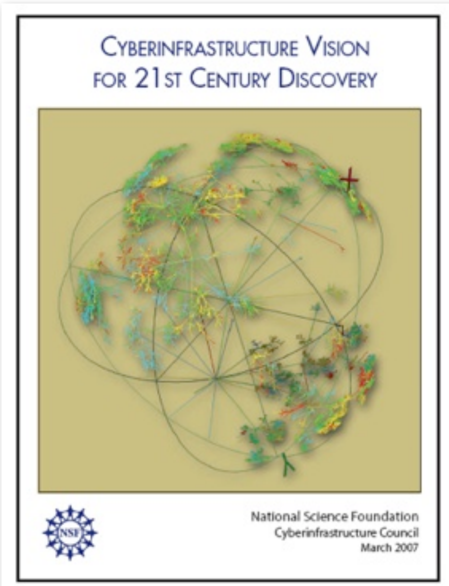
The increased scale and value of data and demand for semantic federation, active curation and long-term preservation of access.



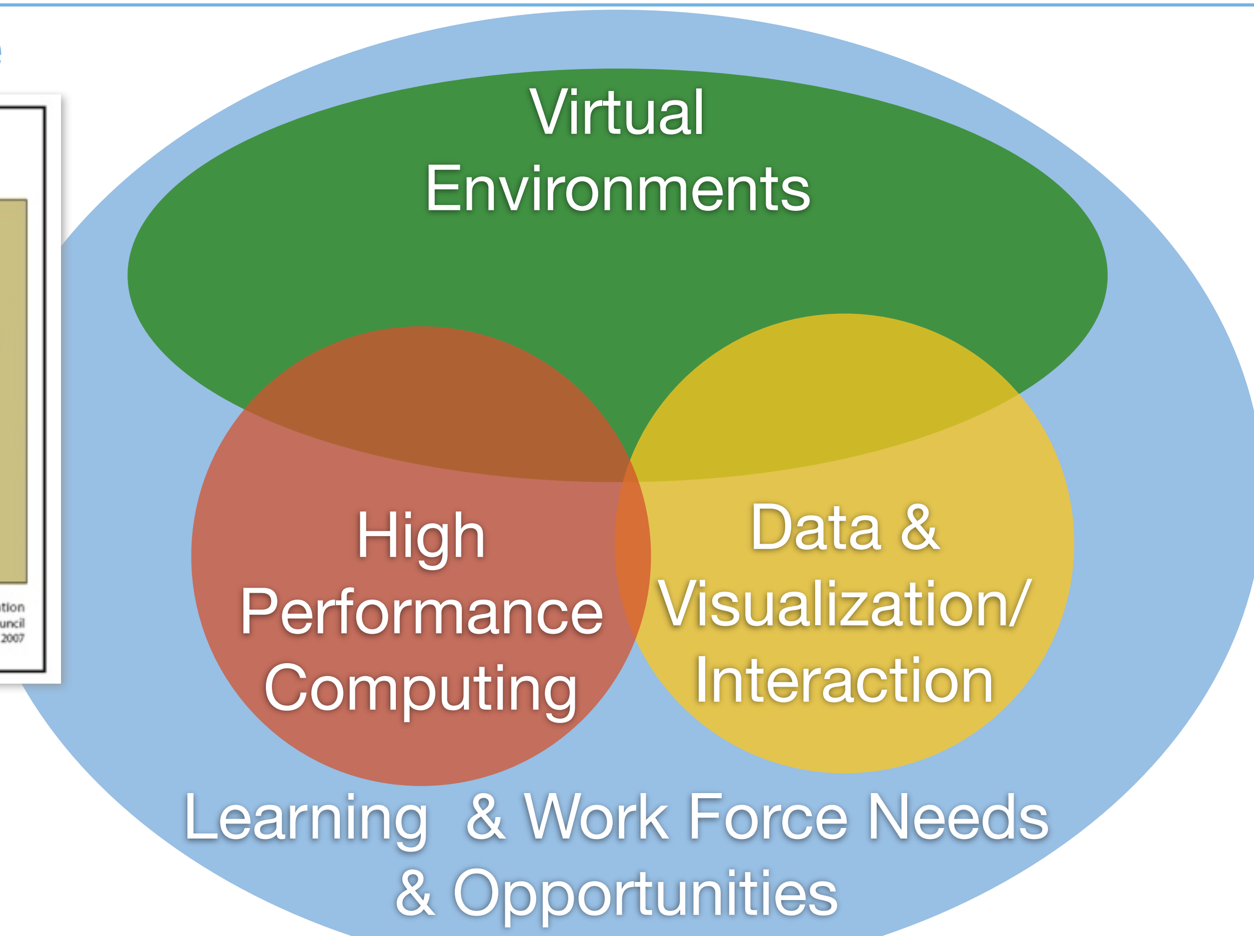
And the need to engage more students in high quality, authentic, passion-building science and engineering education.



e-science



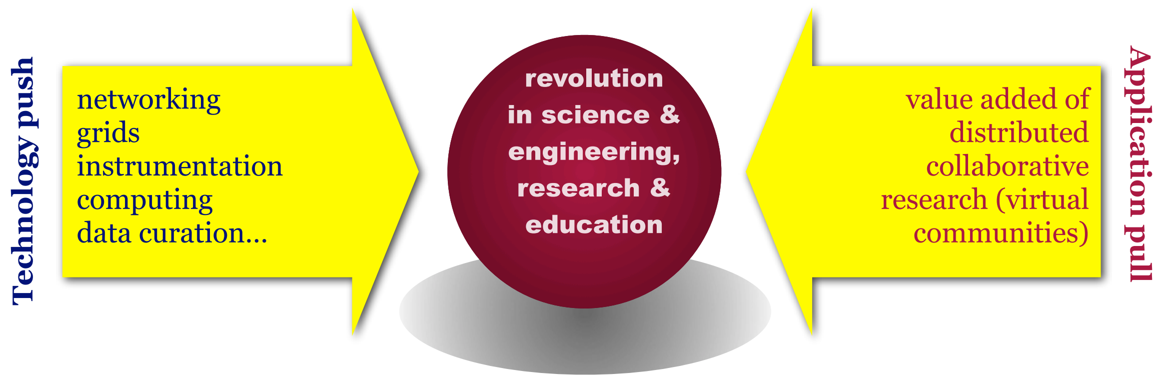
CI-enabled science



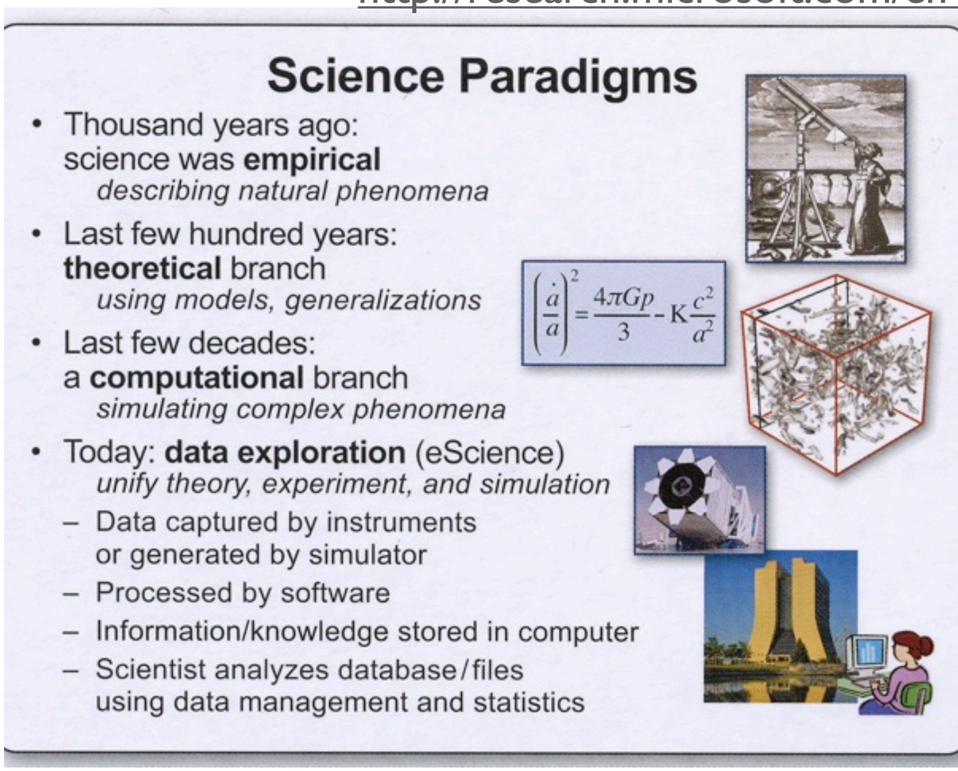
Note: Many other reports on discipline-specific visions of and drivers for e-Science are available at www.nsf.gov/oci

a new vision for Science

- Global challenges with high societal impact
- Data deluge… wet-labs versus virtual-labs
- Improved scientific process
- Cross-disciplinarity
- Virtual Research Communities



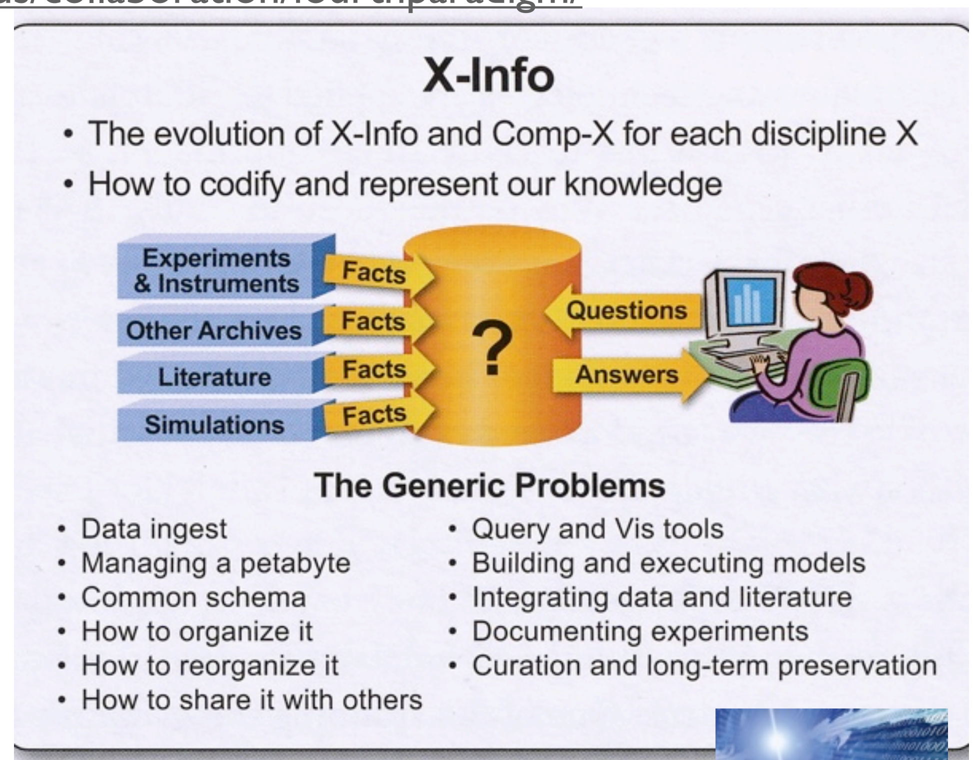
http://research.microsoft.com/en-us/collaboration/fourthparadigm/



Science Paradigms

- Thousand years ago: science was **empirical**
  describing natural phenomena
- Last few hundred years: **theoretical** branch
  using models, generalizations
- Last few decades: a **computational** branch
  simulating complex phenomena
- Today: **data exploration** (eScience)
  unify theory, experiment, and simulation
  – Data captured by instruments or generated by simulator
  – Processed by software
  – Information/knowledge stored in computer
  – Scientist analyzes database/files using data management and statistics

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi Gp}{3} - K\frac{c^2}{a^2}$$



X-Info

- The evolution of X-Info and Comp-X for each discipline X
- How to codify and represent our knowledge

The Generic Problems

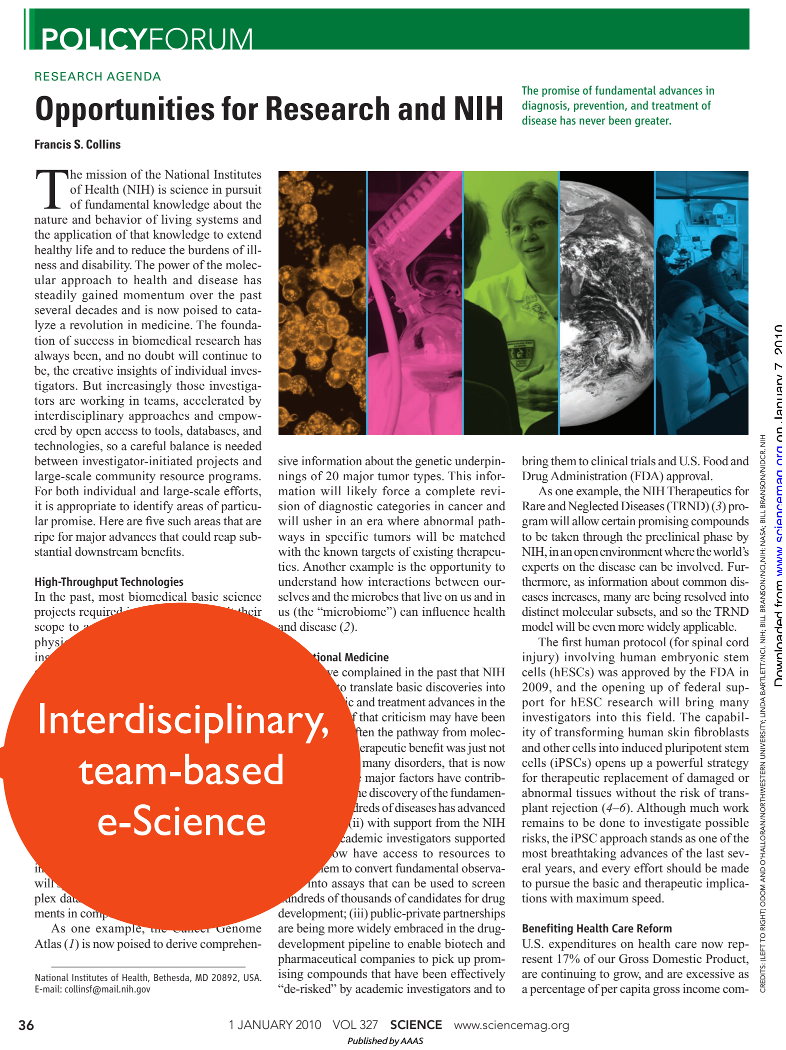
- Data ingest
- Managing a petabyte
- Common schema
- How to organize it
- How to reorganize it
- How to share it with others
- Query and Vis tools
- Building and executing models
- Integrating data and literature
- Documenting experiments
- Curation and long-term preservation

The Fourth Paradigm: Data-Intensive Scientific Discovery

Opportunities for Research and NIH by Francis S. Collins, Director

The power of the molecular approach to health and disease has steadily gained momentum over the past several decades and is now poised to catalyze a revolution in medicine. The foundation of success in biomedical research has always been, and no doubt will continue to be, the creative insights of individual investigators. <u>**But increasingly those investigators are working in teams, accelerated by interdisciplinary approaches and empowered by open access to tools, databases, and technologies**</u>, so a careful balance is needed between investigator-initiated projects and large-scale community resource programs. For both individual and large-scale efforts, it is appropriate to identify areas of particular promise. Here are five such areas that are ripe for major advances that could reap substantial downstream benefits.

1 Jan 2010, vol. 326 SCIENCE <u>www.sciencemag.org</u>



POLICYFORUM

RESEARCH AGENDA

Opportunities for Research and NIH

The promise of fundamental advances in diagnosis, prevention, and treatment of disease has never been greater.

Francis S. Collins

The mission of the National Institutes of Health (NIH) is science in pursuit of fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to extend healthy life and to reduce the burdens of illness and disability. The power of the molecular approach to health and disease has steadily gained momentum over the past several decades and is now poised to catalyze a revolution in medicine. The foundation of success in biomedical research has always been, and no doubt will continue to be, the creative insights of individual investigators. But increasingly those investigators are working in teams, accelerated by interdisciplinary approaches and empowered by open access to tools, databases, and technologies, so a careful balance is needed between investigator-initiated projects and large-scale community resource programs. For both individual and large-scale efforts, it is appropriate to identify areas of particular promise. Here are five such areas that are ripe for major advances that could reap substantial downstream benefits.

High-Throughput Technologies

In the past, most biomedical basic science projects required ... their scope to a ... physi... in... will ... plex data ... ments in comp... As one example, the Cancer Genome Atlas (*1*) is now poised to derive comprehensive information about the genetic underpinnings of 20 major tumor types. This information will likely force a complete revision of diagnostic categories in cancer and will usher in an era where abnormal pathways in specific tumors will be matched with the known targets of existing therapeutics. Another example is the opportunity to understand how interactions between ourselves and the microbes that live on us and in us (the "microbiome") can influence health and disease (*2*).

...tional Medicine

...ve complained in the past that NIH ...to translate basic discoveries into ...c and treatment advances in the ...f that criticism may have been ...ften the pathway from molec... ...erapeutic benefit was just not ... many disorders, that is now ... major factors have contrib... ...he discovery of the fundamen... ...dreds of diseases has advanced ...(ii) with support from the NIH ...cademic investigators supported ...ow have access to resources to ...em to convert fundamental observa... ...into assays that can be used to screen ...undreds of thousands of candidates for drug development; (iii) public-private partnerships are being more widely embraced in the drug-development pipeline to enable biotech and pharmaceutical companies to pick up promising compounds that have been effectively "de-risked" by academic investigators and to bring them to clinical trials and U.S. Food and Drug Administration (FDA) approval.

As one example, the NIH Therapeutics for Rare and Neglected Diseases (TRND) (*3*) program will allow certain promising compounds to be taken through the preclinical phase by NIH, in an open environment where the world's experts on the disease can be involved. Furthermore, as information about common diseases increases, many are being resolved into distinct molecular subsets, and so the TRND model will be even more widely applicable.

The first human protocol (for spinal cord injury) involving human embryonic stem cells (hESCs) was approved by the FDA in 2009, and the opening up of federal support for hESC research will bring many investigators into this field. The capability of transforming human skin fibroblasts and other cells into induced pluripotent stem cells (iPSCs) opens up a powerful strategy for therapeutic replacement of damaged or abnormal tissues without the risk of transplant rejection (*4*–*6*). Although much work remains to be done to investigate possible risks, the iPSC approach stands as one of the most breathtaking advances of the last several years, and every effort should be made to pursue the basic and therapeutic implications with maximum speed.

Benefiting Health Care Reform

U.S. expenditures on health care now represent 17% of our Gross Domestic Product, are continuing to grow, and are excessive as a percentage of per capita gross income com-

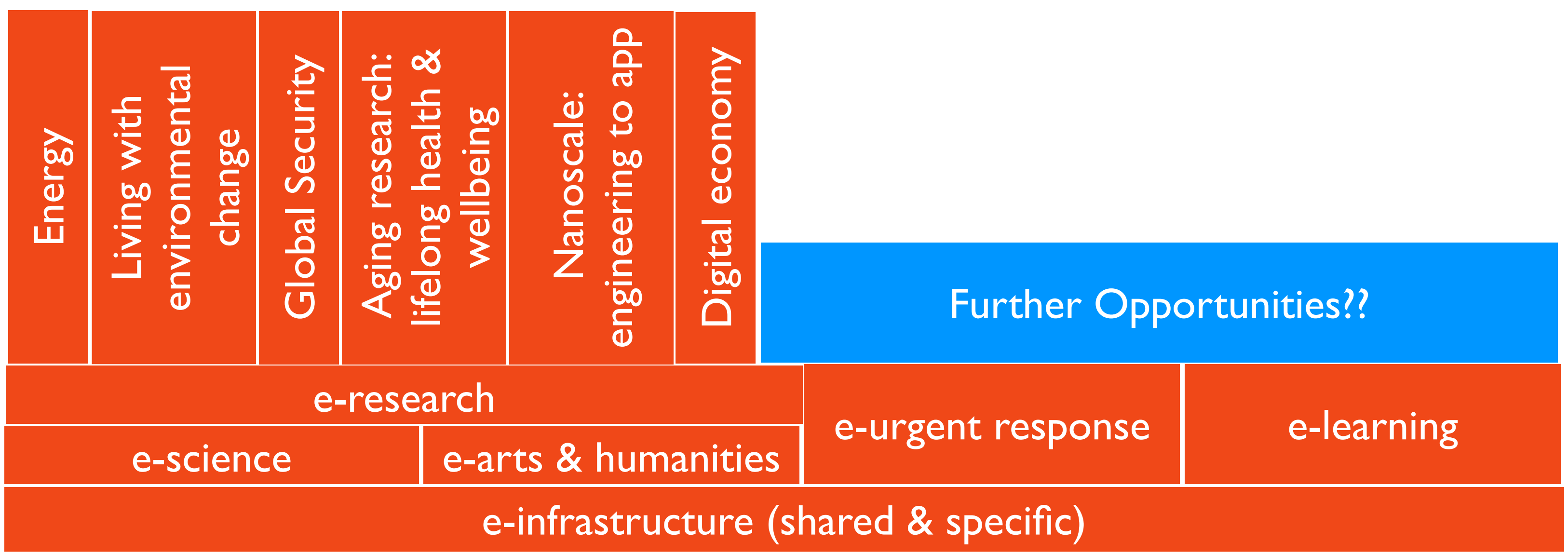
National Institutes of Health, Bethesda, MD 20892, USA. E-mail: collinsf@mail.nih.gov

Published by AAAS

Interdisciplinary, team-based e-Science

e-Infrastructure and e-Science as a Platform for Meeting Grand Challenges

Grand Challenge Research Initiatives



UK e-infrastructure for Science and Innovation

sets out the requirements for a national e-infrastructure to help ensure the UK maintains and indeed enhances its global standing in science and innovation in an increasingly competitive world.

Executive Summary

The growth of the UK's knowledge-based economy depends significantly upon the continued support of the research community and in particular its activities to engage with industry and to apply its world-leading innovations to commercial use. A national e-infrastructure for research provides a vital foundation for the UK's science base, supporting not only rapidly advancing technological developments, but also the increasing possibilities for knowledge transfer and the creation of wealth.

http://www.nesc.ac.uk/documents/OSI/index.html

Six Working Groups:
Data & information creation • Preservation & curation • Search & navigation • Virtual research communities • Networks, compute and data storage • Authentication, authorisation, accounting, middleware, and digital rights management

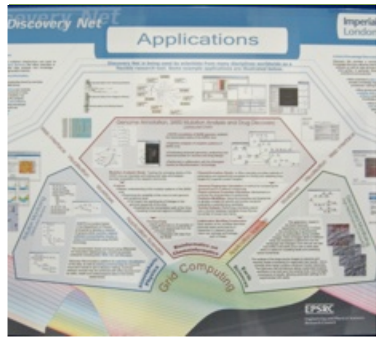
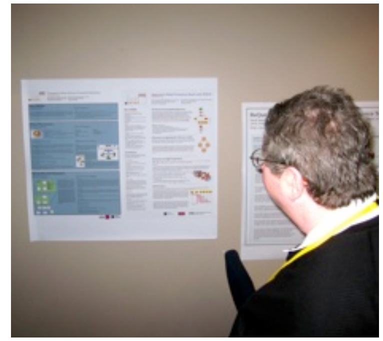
Revolutionizing Science and Engineering through Cyberinfrastructure:
e-Science
eGee
Enabling Grids for E-sciencE
Cyberinfrastructure Vision for 21st Century Discovery
e-Science
Our Cultural Commonwealth
HASTAC
Humanities, Arts, Science, and Technology Advanced Collaboratory
e-Humanities & Arts
The New Invisible College
Science for Development
e-development
Fostering Learning in the Networked World:
e-Learning
e-Development
The Research University in the Digital Age
PREPARING FOR THE REVOLUTION
Information Technology and the Future of the Research University
NATIONAL RESEARCH COUNCIL

(left to right)

Anders Ynnerman
Linköping University, Sweden
Paul Tackley
ETH Zürich, Switzerland
Albert Heck
Utrecht University, Netherlands
Dieter Heermann
U. of Heidelberg, Germany
Ian Foster
ANL & U. of Chicago, USA
Mark Ellisman
U. of California, San Diego, USA
Wolfgang von Rüden
CERN, Switzerland
Christine Borgman
U. of California, Los Angeles, USA
Daniel Atkins
University of Michigan, USA
Alexander Szalay
Johns Hopkins University, USA
Julia Lane
US National Science Foundation
Nathan Bindoff
University of Tasmania, Australia
Stuart Feldman
Google, USA
Han Wensink
ARGOSS, The Netherlands
Jayanth Paraki
Omega Associates, India
Luciano Milanesi
National Research Council, Italy

International Panel for 2009 Review of the RCUK e-Science Program




Discovery Net
Applications
Grid Computing
EPSRC







DATA_PLEX (ECONOMY)

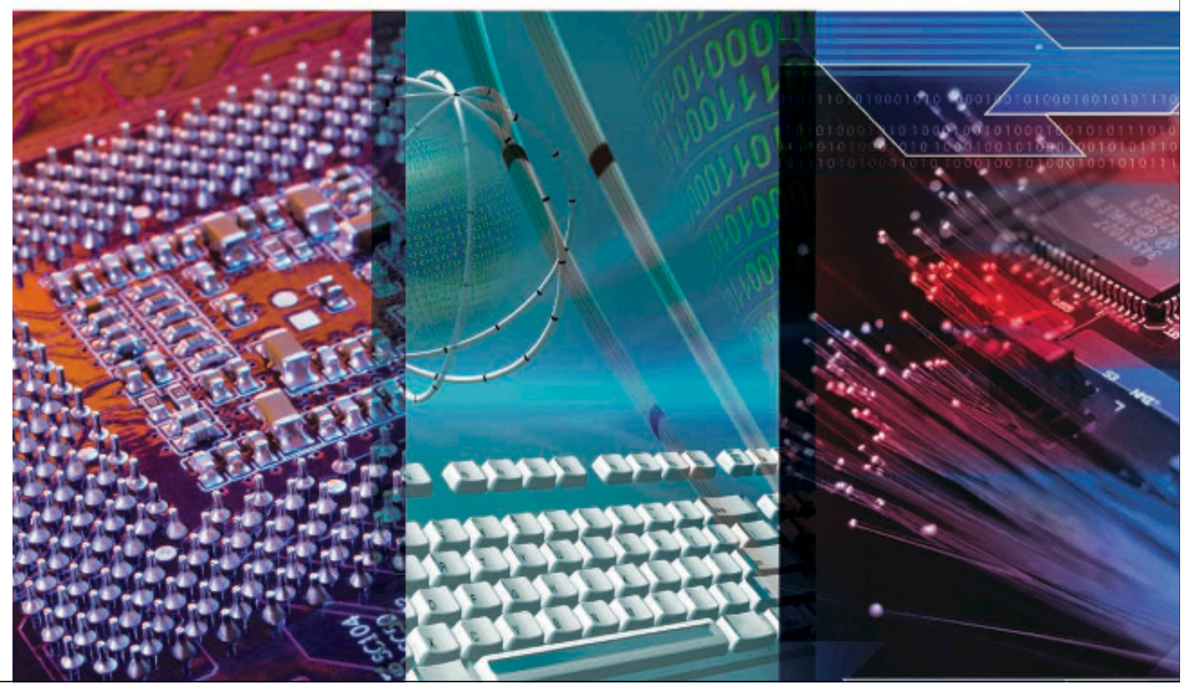


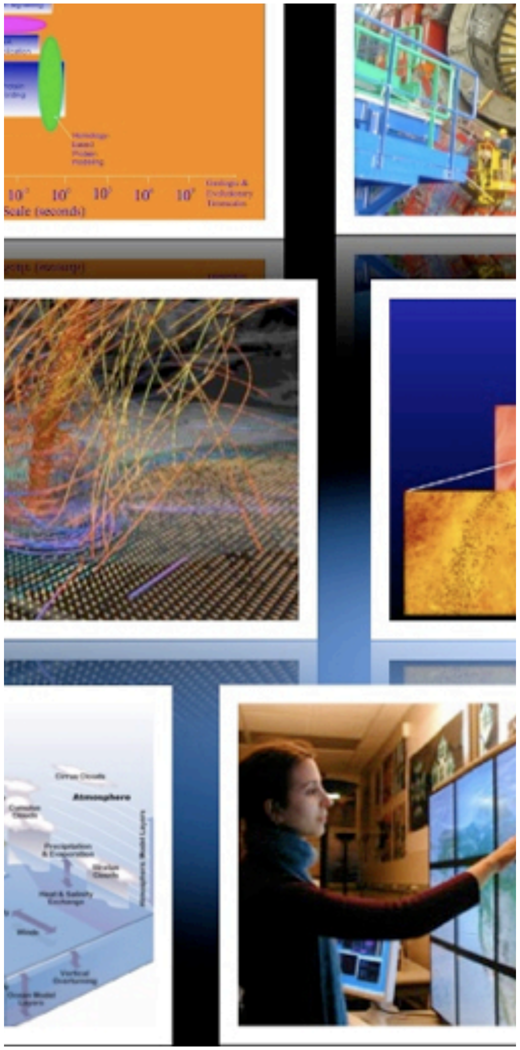

Available at

- http://www.epsrc.ac.uk/SiteCollectionDocuments/Publications/reports/RCUKe-ScienceReviewReport.pdf

RCUK Review of e-Science 2009

BUILDING A UK FOUNDATION FOR THE TRANSFORMATIVE ENHANCEMENT OF RESEARCH AND INNOVATION





Findings and Recommendations

The UK has created a "jewel" -- a pioneering, vital activity of enormous strategic importance to the pursuit of scientific knowledge and the support of allied learning.


Golden Jubilee Diamond

- The UK e-Science Programme is in a **world-leading** position.
- Investments are already empowering **significant contributions** in the UK and beyond.
- The UK **must decide** whether to create the necessary combination of financial, organisational, and policy commitments to capitalise on their prior investments,.
- e-Science is an organic, emergent process requiring ongoing, **coordinated investment** from multiple funders and **coordinated action** by multiple research and infrastructure communities. It is both an **enabler of** research and an **object of** research,
- The UK should continue to **nurture a robust infrastructure** that couples and balances research, application development, and training processes.
- None of this is easy, but done well could achieve **enormous reward**.

SIGNIFICANT FINDINGS: STRENGTHS (1)

- Produced competent people for both academia and industry.
- Created and enriched *interdisciplinary* collaborations (faculty & students) that are increasingly important for progress at the frontiers of scientific discovery.
- Created interdisciplinary e-science doctoral training programs.
- Stimulated industry up-take of e-science.
- Situated the UK reputation in e-science among the top nations.
- Accelerated the penetration of e-science capability in UK and Europe.
- Contributed significantly to establishing standards and tools for national and european consortial networked science projects.

SIGNIFICANT FINDINGS: STRENGTHS (2)

- Stimulated university leaders to establish e-science buildings and organizations on campus.
- Promoted recognition of digital data as an asset and the need for greater attention to its stewardship.
  - But still lacking adequate repository capacity and the trained human capacity that is needed.
- Supported directly and indirectly scientific international collaborations.
  - But very little (any?) participation by developing countries.
- Enabled science not otherwise possible.
- Initiated a national framework for sharing large-scale resources that will be increasingly important to both large and small scale research and education.

SIGNIFICANT FINDINGS: STRENGTHS (3)

- Supported important work in the humanities and social sciences.
- Promoted increasing standardization for data interoperability and aggregation.
- Promoted interdisciplinary work and provided mechanisms to link projects to projects on an on going way. (All-hands Meetings; community building activities)
- Leveraged investments from other sources (regional, industrial, international)
- Stimulated local/regional economies.
- Accelerated the productivity of researchers and science processes.
- Accomplished some knowledge/technology transfer but there could be more.

SIGNIFICANT FINDINGS: WEAKNESS (1)

- Too early and rapid reduction of core funding and high level leadership, resulting in
  - Sense of abandonment by some researchers and career shift away from e-science
  - Failure to reap the benefits of prior investments.
    - The time constants for real transformative impact and significant competitive advantage is decades.
  - Reduced momentum of an important movement in which UK established a global leadership role
  - Perception that RCs don't appreciate the strategic importance of e-science and the special handling that it still needs

SIGNIFICANT FINDINGS: WEAKNESS (2)

- Lack of models for sustaining/maintaining software, platform/infrastructure operations, and data resources
  - e-infrastructure providers are being forced to compete for research funding (rather than targeted infrastructure/facilities funding) every two years
- Missed opportunities to transfer successful software tools/systems and best practices to other fields. Lack of regularized processes and organizations to do so.

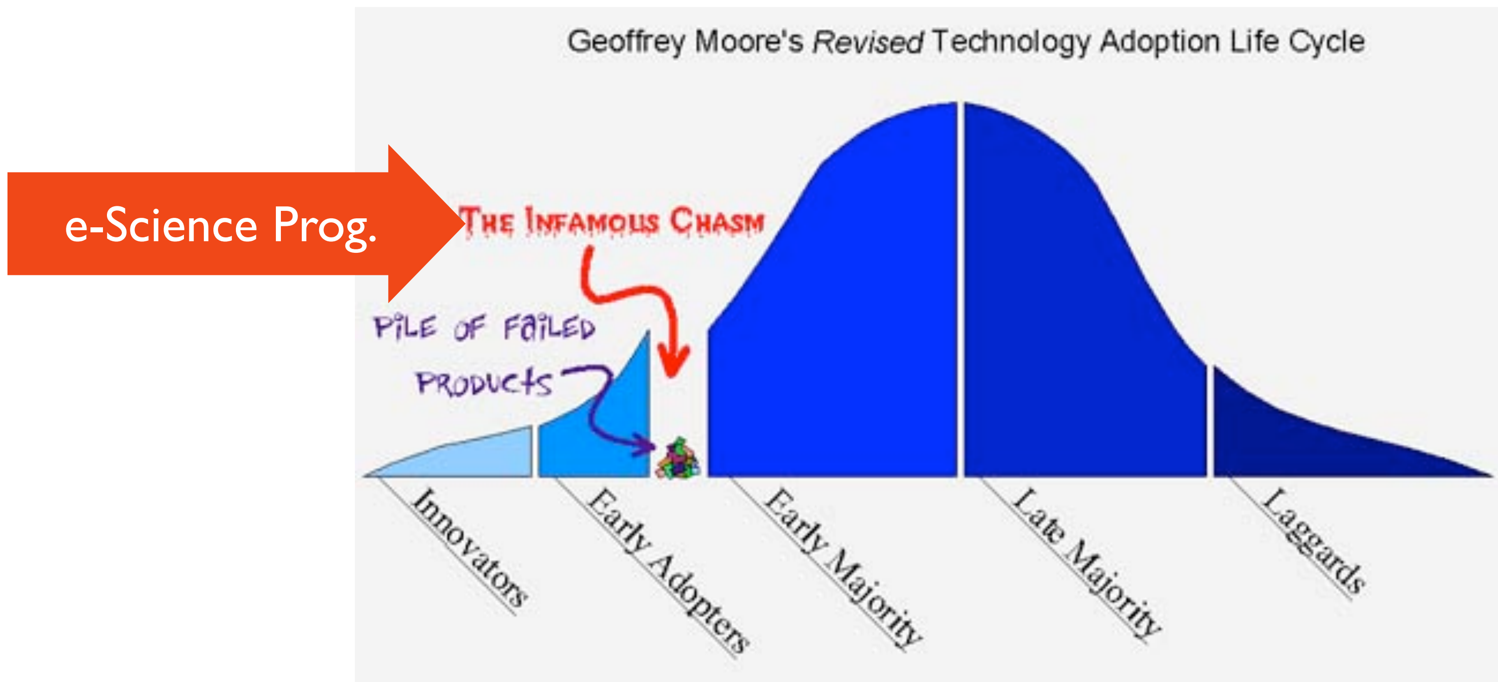
SIGNIFICANT FINDINGS: WEAKNESS (3)

- Lack of adequate structure, leadership, and resources to promote constructive interplay and achieving balance (1) between shared/generic and specific e-infrastructure and (2) between e-science core and disciplinary research.
  - This situation raises barriers to interdisciplinary work; to developing and exploiting common infrastructure, middleware, tools, systems; and to interoperability between disciplines, projects, platforms and shared use of resource
- Overly focused on computational "grid" platforms
  - Good fit to physics but not necessarily all other domains. But there are variations in what "grid" means. (computational, data, all types of resources)
- Knowledge transfer (KT) potential higher than was actually achieved. Need more systematic and staffed KT processes.

SIGNIFICANT FINDINGS: WEAKNESS (4)

- Shortage of women participants (as judged by data in evidence document and gender balance during review)
- Gaps in the e-science program conceptualization.
  - sensor networks
  - usability, human-computer-interaction
  - understanding of social/behaviors barriers to technology mediate science collaboration and ways to design systems and processes to reduce barriers

Crossing the Chasm – Modified



Recent novel examples of government, industry, academia partnerships on e-Science research

http://www.networkworld.com/community/node/27219

Google, IBM and NSF offer up $5M for large-scale computing research

By Layer 8 on Thu, 04/24/08 - 12:36pm.

Share Tweet This Email this page Comment Print Newsletter Sign-Up

The National Science Foundation in cooperation with Google and IBM today said it was seeking proposals for the group's Cluster Exploratory (CluE) initiative to explore innovative research ideas in data-intensive computing.

The NSF said it expects to award up to $5 million spread between

http://www.nsf.gov/news/news_summ.jsp?cntn_id=116336&org=NSF&from=news

Press Release 10-023

Microsoft and NSF Enable Research in the Cloud

Agreement will offer free access to new computational and collaborative services to accelerate scientific discovery for research communities.

http://www.nytimes.com/2010/02/05/science/05cloud.html

U.S. Scientists Given Access to Cloud Computing

By JOHN MARKOFF
Published: February 4, 2010

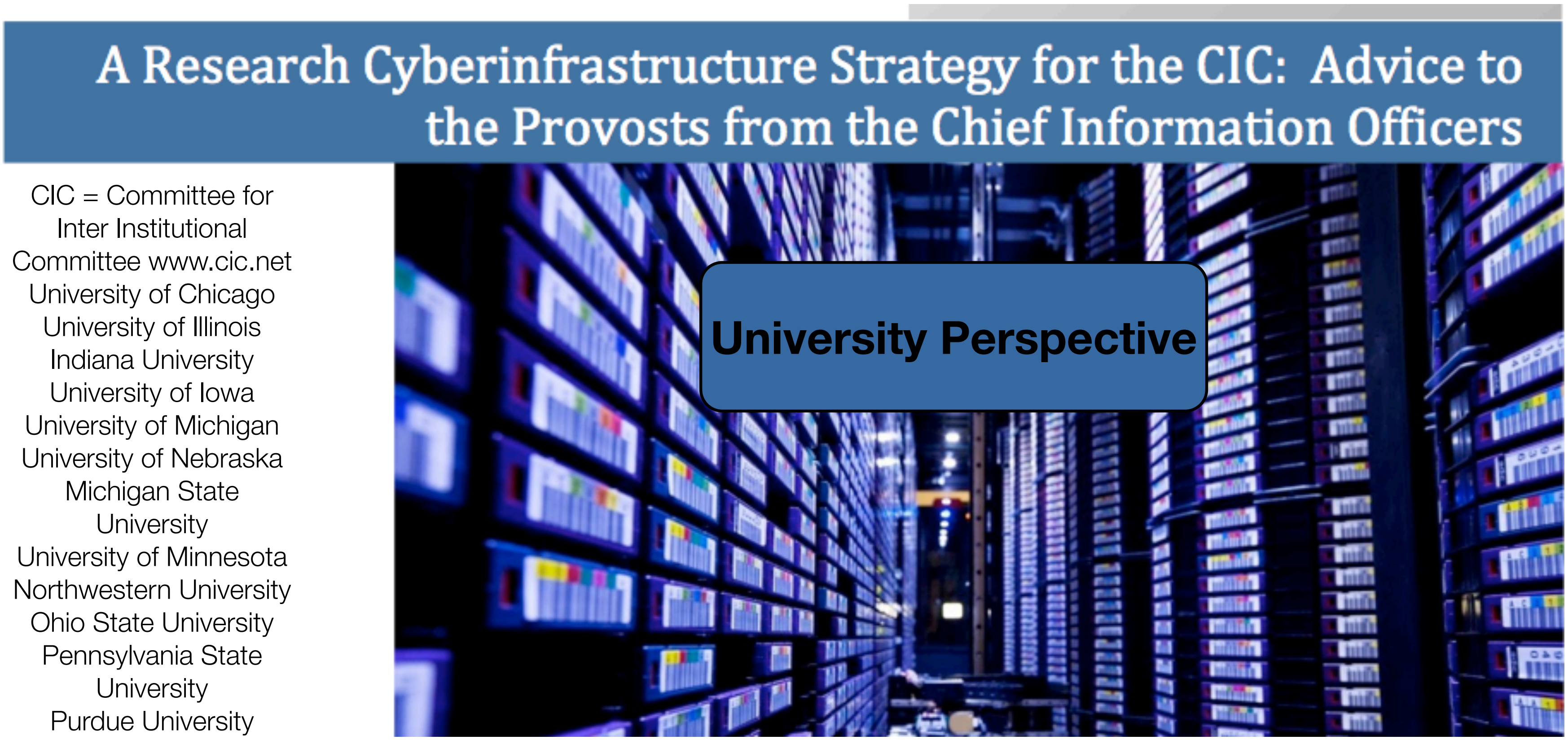
The National Science Foundation and the Microsoft Corporation have agreed to offer American scientific researchers free access to the company's new cloud computing service.

A goal of the three-year project is to give scientists the computing power to

RSS Feed

SIGN IN TO RECOMMEND
TWITTER
E-MAIL
SEND TO PHONE
PRINT

Creating a Transformative e-Science Programme is a Complex Balancing Act



CIC = Committee for Inter Institutional Committee www.cic.net
University of Chicago
University of Illinois
Indiana University
University of Iowa
University of Michigan
University of Nebraska
Michigan State University
University of Minnesota
Northwestern University
Ohio State University
Pennsylvania State University
Purdue University
University of Wisconsin-Madison

CIC = Committee for Inter Institutional Committee
www.cic.net

RECENT REPORT

Recommendations

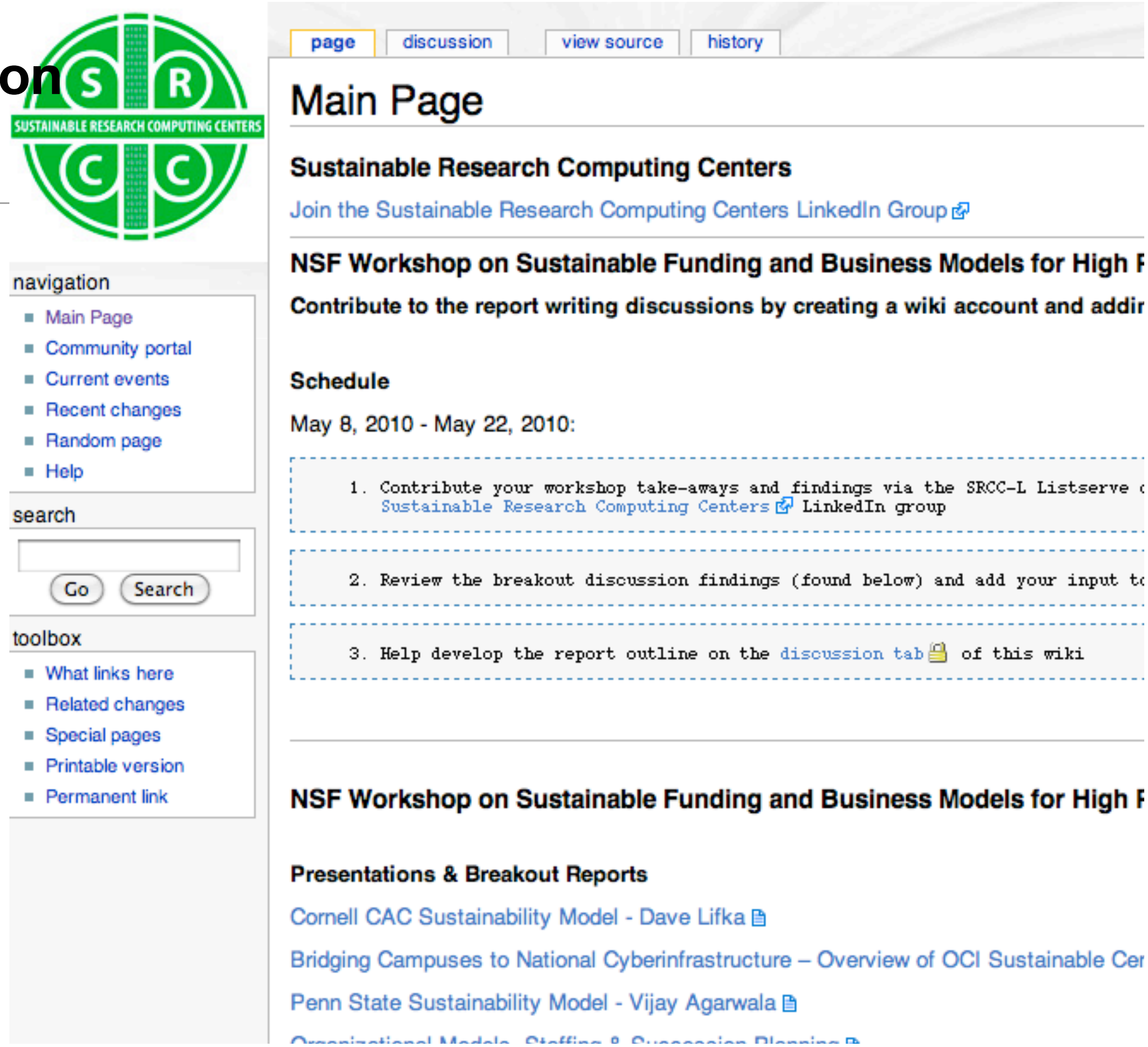
- Our goal should be to enable scholarship at the cutting edge of every discipline, while getting as much value as possible from every dollar spent on cyberinfrastructure.
- The CIC campuses are very richly endowed with cyberinfrastructure resources but can be even more effective by adopting good practices that support greater coordination at every level:
  - Plan.
  - Share (at the highest level possible).
  - Design funding models that promote scholarship and stewardship.
  - Rely on user governance.
  - Conduct cyberinfrastructure impact analysis.
- Over the long run, these practices should help our institutions produce **more scholarship (learning and research) per dollar invested.**

We should all currently be investing in:

- Preparing for **federated identity management** and other enabling technologies for virtual organizations.
- Maintaining **state-of-the-art** communication **networks**.
- Providing institutional **stewardship of research data**.
- **Consolidating** computing resources while **maintaining diverse** architectures.
- Expanding cyberinfrastructure **support to all** disciplines.
- Exploring **cloud computing** to manage longterm financial commitments.

Cornell NSF Workshop on Sustainable HPC

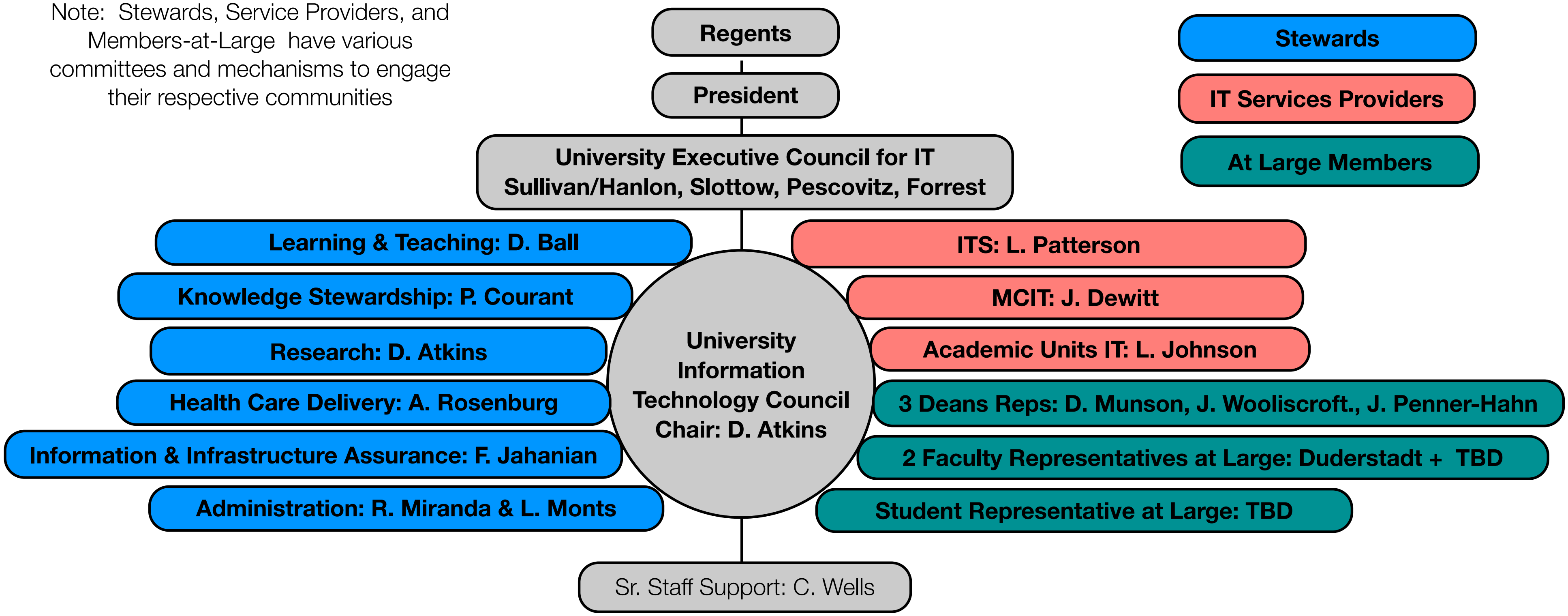
- https://mw1.osc.edu/srcc/index.php/Main_Page



D. E. Atkins • atkins@umich.edu

New University of Michigan IT Visioning, Planning, and Governance Model

Note: Stewards, Service Providers, and Members-at-Large have various committees and mechanisms to engage their respective communities



Mission Statement: UM Office of Research Cyberinfrastructure (ORCI)

To create and nurture a modern, fiscally responsible, and sustainable cyberinfrastructure (both resources and services), intellectual capacity, and academic environment in support of high-end computing- and data-intensive research activities that ensures UM leadership as a 21st century research institution.

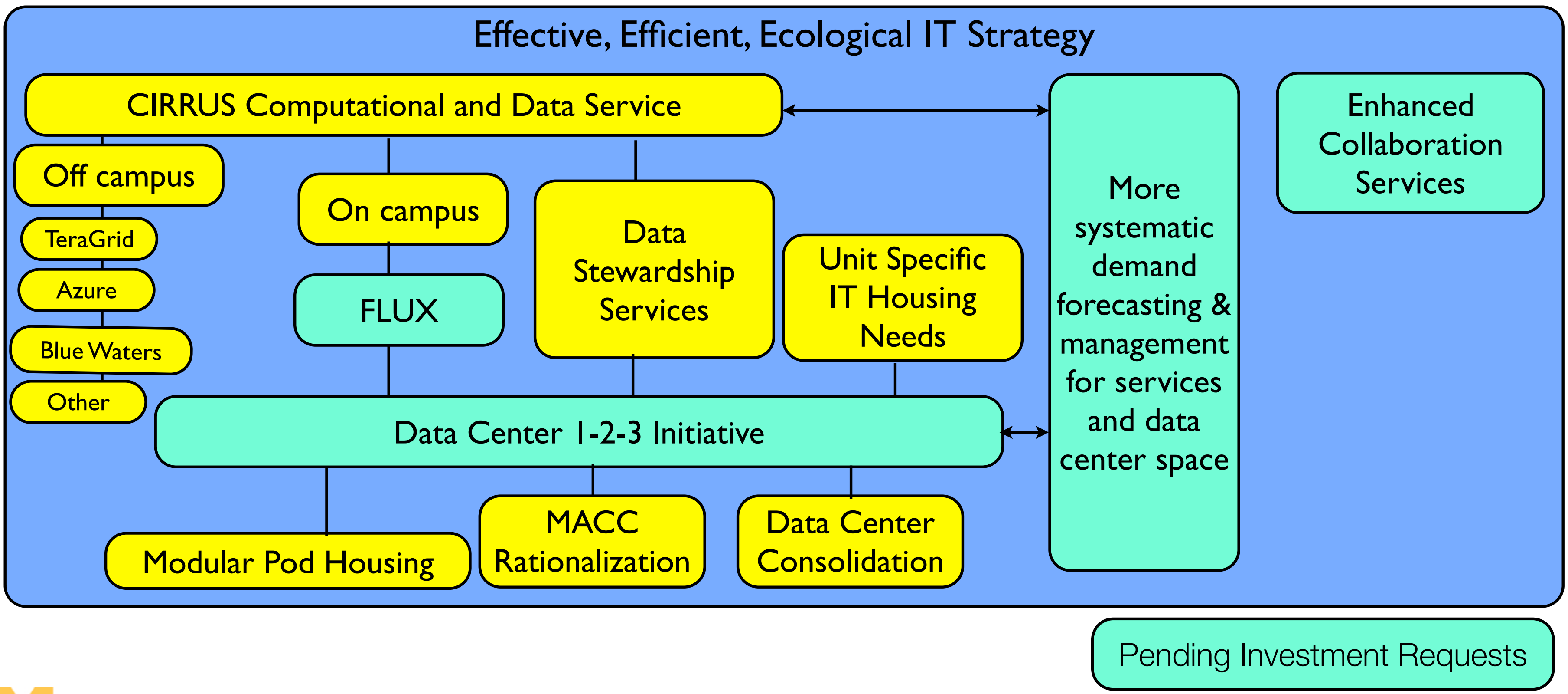
Adopted by ORCI Steering Committee, April xx

The UM CIRRUS Research Computing Project

- CIRRUS - Computation and Information Resources (for) Research as a Utility Service
- The centerpiece of our innovation in the provisioning of research cyberinfrastructure.
- Goal is to provide a *portfolio* of integrated research computing services that are mission effective, financially efficient, and environmentally friendly.

D. E. Atkins • atkins@umich.edu

Part of NextGen Michigan Roadmap and Pending Investment Requests for IT Council Consideration



Exploring Use of Pod Optimized Performance Datacenters

UNIVERSITY OF MICHIGAN

D. E. Atkins • atkins@umich.edu

https://research.umich.edu/blogs/ci/


Computational Discovery & Cyberinfrastructure at U-M

News and information from the Office of Research Cyberinfrastructure

Want to explore a very large Microsoft cloud for research computing? Let us know!

Posted February 17th, 2010 by OVPR-CI

Categories: Funding Opportunities, Resources

In case you missed our post about Microsoft teaming with NSF to offer free access to research computing, there is also supplemental and EAGER funding available. From the NSF

Search this blog

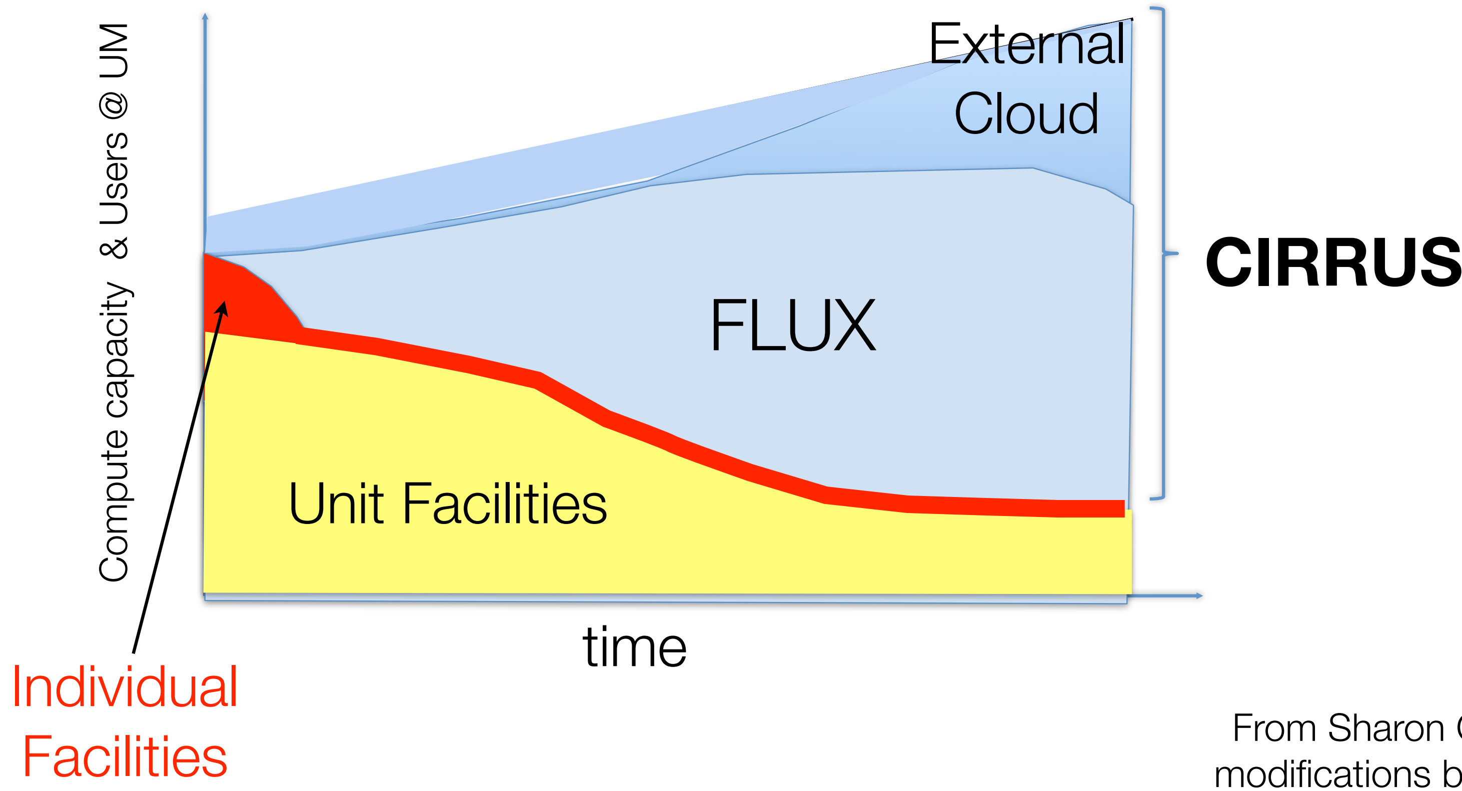
Search

Archives

- February 2010 (7)
- January 2010 (5)
- December 2009 (2)
- October 2009 (1)
- September 2009 (2)

ORCI Blog Site to Help Share Information and Build Community

A Vision of HPC @ UM



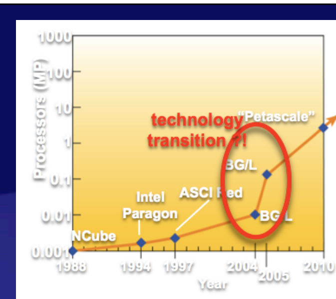
From Sharon Glotzer with modifications by Dan Atkins

Cyberinfrastructure Framework for 21st Century Science & Engineering (CF21)

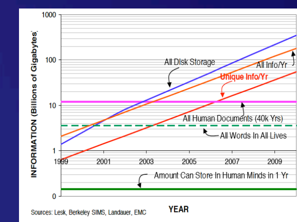
NSF-wide Cyberinfrastructure Vision

People, Sustainability, Innovation, Integration

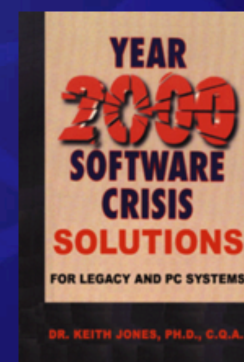
Five Challenges



- Computing Technology
  - Multicore: processor is new transistor
  - Programming model, fault tolerance, etc
  - New models: clouds, grids, GPUs,… where appropriate



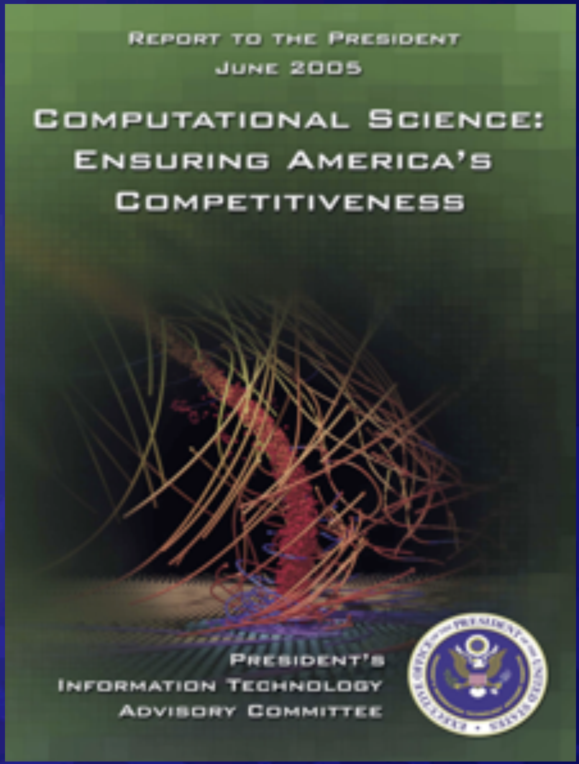
- Data, provenance, and visualization
  - Generating more data than in all of human history: preserve, mine, share?
  - How do we create "data scientists"?



- Software
  - Complex applications on coupled compute-data-networked environments, tools needed
  - Modern apps: 10^6+ lines, many groups contribute, take decades

Five Crises con't

- Organization for Multidisciplinary Computational Science
  - "Universities must significantly change organizational structures: multidisciplinary & collaborative research are needed [for US] to remain competitive in global science"
  - "Itself a discipline, computational science advances all science… inadequate/outmoded structures within Federal government and the academy do not effectively support this critical multidisciplinary field"
- Education
  - The CI environment is becoming more **complex** and more **fundamental** for research
  - How do we develop a workforce to provide leadership, expertise and support ?



Some observations

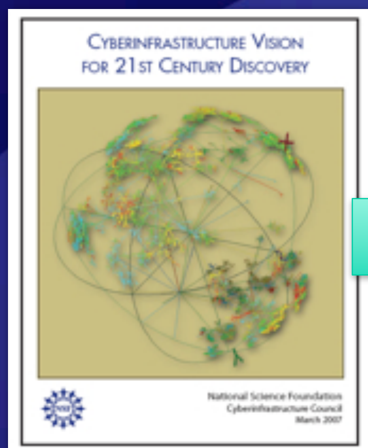
- Science and Scholarship are team sports
  - Competitiveness and success will come to those who can put together the best team, and can marshal the best resources and capabilities
- Collaboration/partnerships will change significantly
  - Growth of dynamic coalitions and virtual organizations
  - International collaboration will become even more important
- Ownership of data plus low cost distribution fuels growth and number of data systems
  - Growth in both distributed systems and local systems
  - More people want to access more data
  - Federation and interoperability become more important

More observations

- More discoveries will arise from search approaches
  - Mining vast amounts of new and disparate data
  - Collaboration and sharing of information
- Mobility and personal control will continue to drive innovation and business
- Gaming, virtual worlds, social networks will continue to transform the way we do science, research, education and business
- The Internet has collapsed six degrees of separation and is creating a world with two or three degrees.

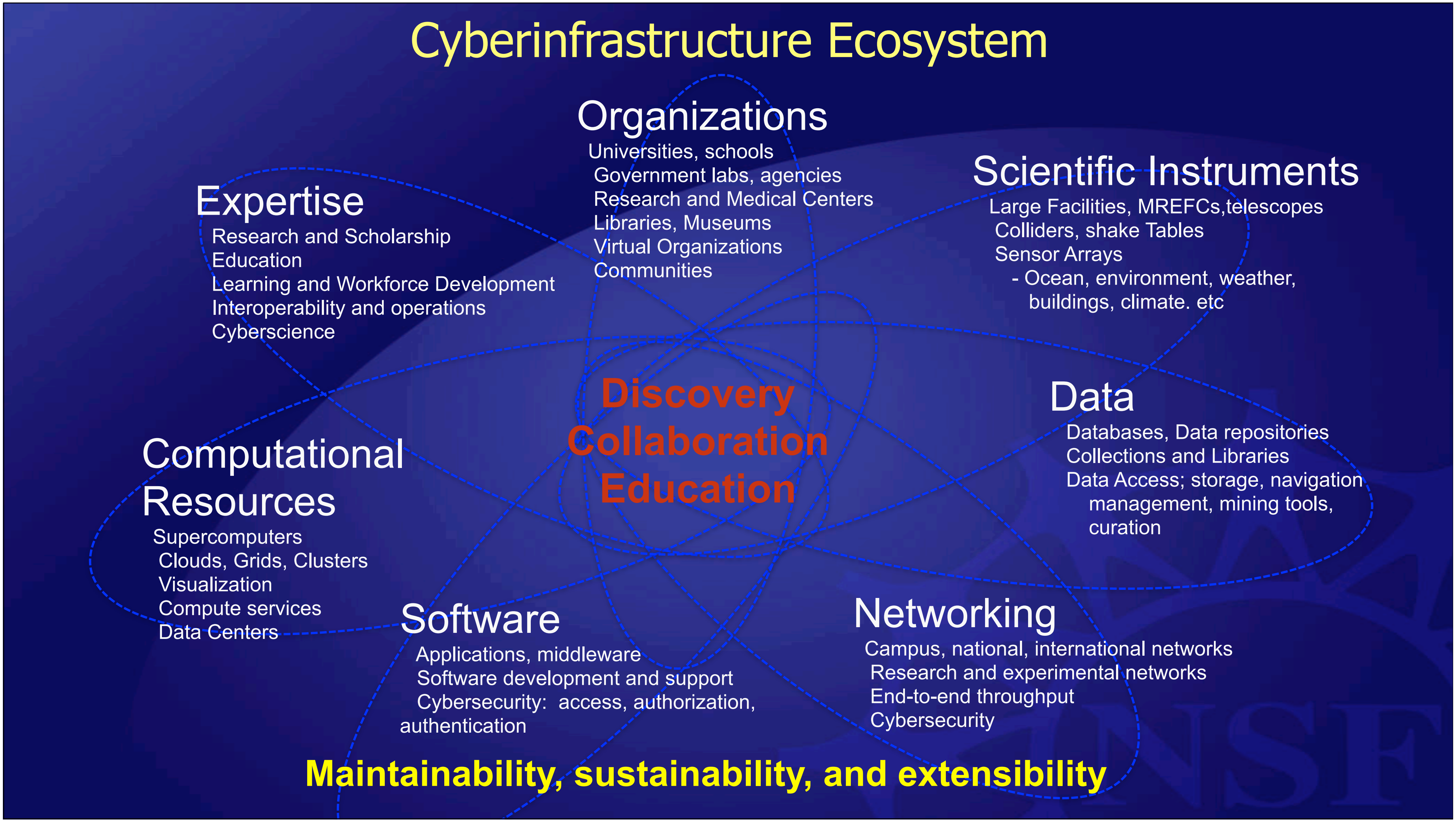
What is Needed?

An **ecosystem**, not just components…



NSF-wide CI *Framework* for 21st Century Science & Engineering

People, Sustainability, Innovation, Integration


Cyberinfrastructure Ecosystem
Organizations
Universities, schools
Government labs, agencies
Research and Medical Centers
Libraries, Museums
Virtual Organizations
Communities
Scientific Instruments
Large Facilities, MREFCs,telescopes
Colliders, shake Tables
Sensor Arrays
- Ocean, environment, weather, buildings, climate. etc
Expertise
Research and Scholarship
Education
Learning and Workforce Development
Interoperability and operations
Cyberscience
Discovery
Collaboration
Education
Data
Databases, Data repositories
Collections and Libraries
Data Access; storage, navigation management, mining tools, curation
Computational Resources
Supercomputers
Clouds, Grids, Clusters
Visualization
Compute services
Data Centers
Software
Applications, middleware
Software development and support
Cybersecurity: access, authorization, authentication
Networking
Campus, national, international networks
Research and experimental networks
End-to-end throughput
Cybersecurity
Maintainability, sustainability, and extensibility

ACCI Task Forces

Campus Bridging: Craig Stewart, IU (BIO)

Data & Viz: Tony Hey, , Dan Atkins

- Completion by end of year
- Advising NSF
- Conducting Workshop(s)
- Recommendations
- Input to NSF informs CF21 programs, 2011-12 CI Vision

Software: David Keyes, Columbia/ KAUST (MPS)

Computing: Thomas Zacharia, ORNL/UTK (DOE)

Education & Workforce: Alex Ramirez, CEOSE

Grand Challenge Communities/VOs: Tinsley Oden, Austin (ENG)

CF21 Impact on Data

- All science is becoming data-dominated
  - Experiment, computation, theory
- Totally new methodologies
  - Algorithms, mathematics
  - All disciplines from science and engineering to arts and humanities
- End-to-end networking becomes critical part of CI ecosystem
- How do we train "data-intensive" scientists?
- Data policy and sustainability becomes critical!

US NSF Data and Visualization Strategy Task Force

Co-chairs Tony Hey and Dan Atkins; Member: Liz Lyon

- Charge: Examine the increasing importance of data and their visualization in driving grand challenge science, engineering, education.
- Emphasis: Value of the data, complexity, and organic aspects. The role of data in research, the value of of metadata and ontologies for integration, etc.
- Goal: Catalyze a network of science and engineering data collections that is open, extensible and sustainable. Enable multiple fields of science and engineering research and education -- including new types of data-driven computational science, interdisciplinary research and cross-disciplinary education.

12. Towards functionally complete, four-quadrant, research environments

- Pursue capacity for collaborative, international, interdisciplinary team science to occur routinely in "functionally complete, four-quadrant environments" built upon e-infrastructure.
- "A four-quadrant environment" refers to a blended virtual-physical environment in which the activities of a group can flow easily between all four quadrants in a 2-by-2 matrix with same versus different for the two dimensions of both time and place. It subsumes the concept of virtual research environment.
- "Functionally complete" means that the environment supports access to all the people, the information and data services, the observatories and facilities, the computational services, and the collaboration and communication services necessary for a scientific team (or more generally, a community of practice) to carry out its work.

Four Quadrant Organizations (virtual + physical organizations) offer additional modes of interaction between People, Information, and

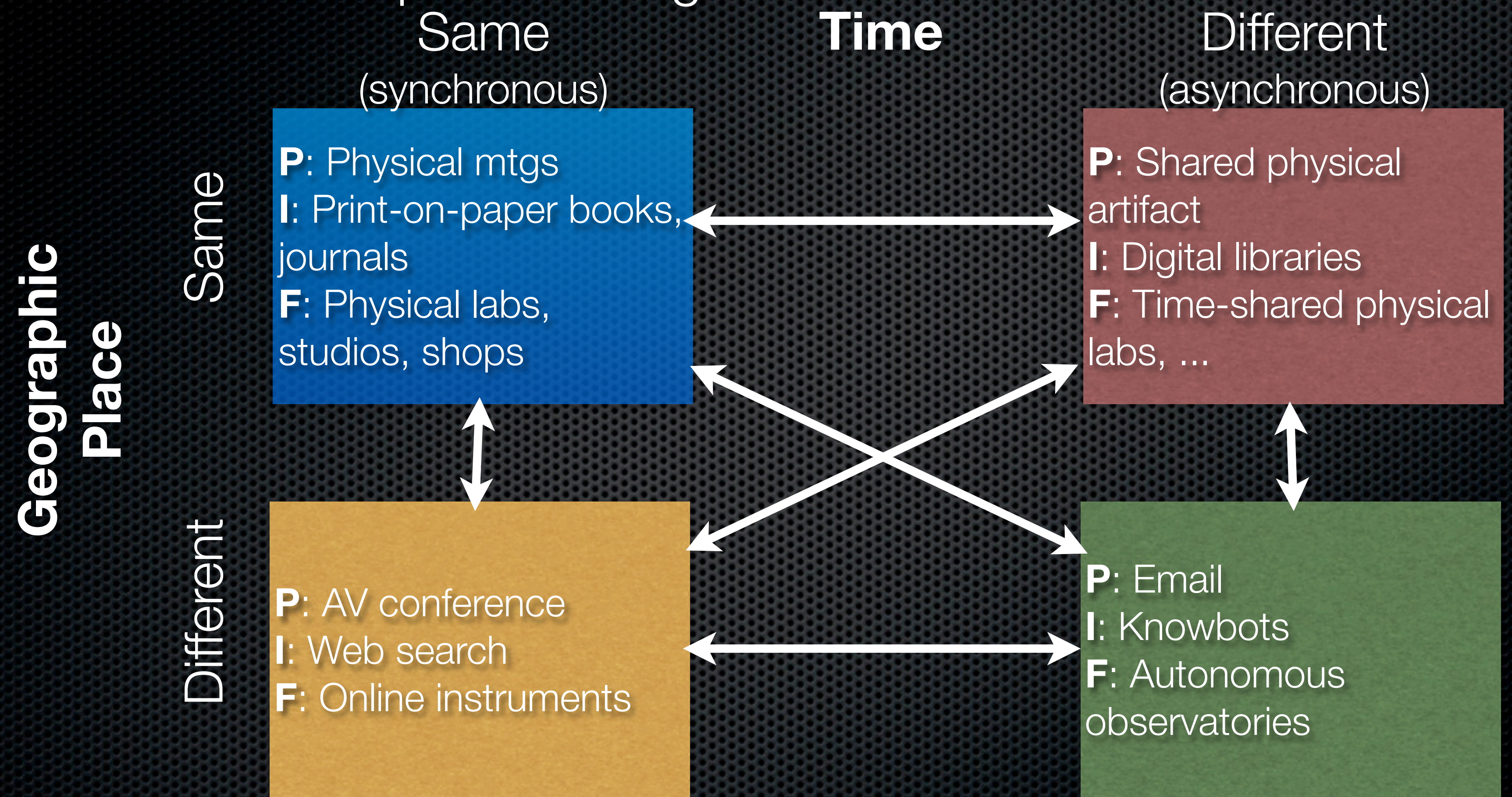
Geographic Place	Time: Same (synchronous)	Time: Different (asynchronous)
Same	**P**: Physical mtgs **I**: Print-on-paper books, journals **F**: Wet labs, studios, shops	**P**: Shared physical artifact **I**: Library reserve **F**: Time-shared physical labs, ...
Different	**P**: AV conference **I**: Web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

Physical + Virtual, Not Physical vs. Virtual

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

Daniel E. Atkins
atkins@umich.edu

Need to discover how to best map collaborative workflows onto various paths through these collaboration states.



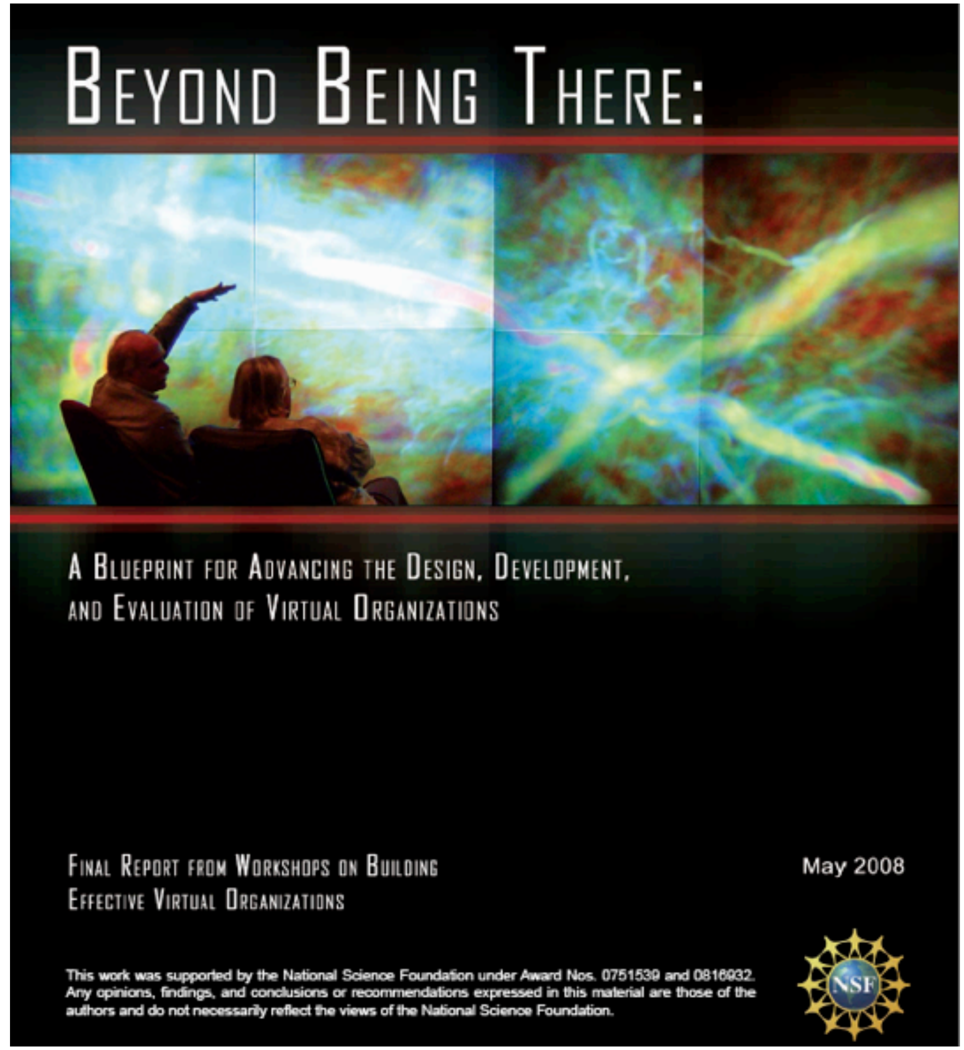
P: people, **I**: information, **F**: facilities, instruments

University of Michigan

Daniel E. Atkins
atkins@umich.edu

Beyond Being There: A Blueprint for Advancing the Design, Development, and Evaluation of Virtual Organizations

Available at: www.ci.uchicago.edu/events/VirtOrg2008/



CF21 Impact on Virtual Organizations

- What constitutes effective virtual organizations? How do they enhance research and education; what about production and innovation?
- Multi–disciplinary approach
  - Anthropology, complexity sciences, CS, decision and management sciences, economics, engineering, organization theory, organizational behavior, social and industrial psychology, public administration, sociology
- Broad variety of qualitative and quantitative methods
  - Ethnographies, surveys, simulation studies, experiments, comparative case studies, network analyses.
- Grounded in theory, rooted in empirical methods

Discussion & Questions