

ARL & CNI Forum on Re-inventing Science Librarianship: Models for the Future

E-Science: Trends, Transformations & Responses

October 16, 2008

Chris Greer

Director, US National Coordination Office

Networking and Information Technology Research and Development Program

Acronyms:

NITRD

Networking and Information Technology Research and
Development Program

NCO

National Coordination Office

**White House
Executive Office of the
President**

**Office of Science
and Technology Policy**

**National Science and
Technology Council**

**Committee on
Technology**

**NITRD
Subcommittee**

**National
Coordination
Office (NCO)**

**High End Computing
(HEC I&A - R&D)**

**Cyber Security and
Information Assurance**

**Human Computer
Interaction and
Information
Management**

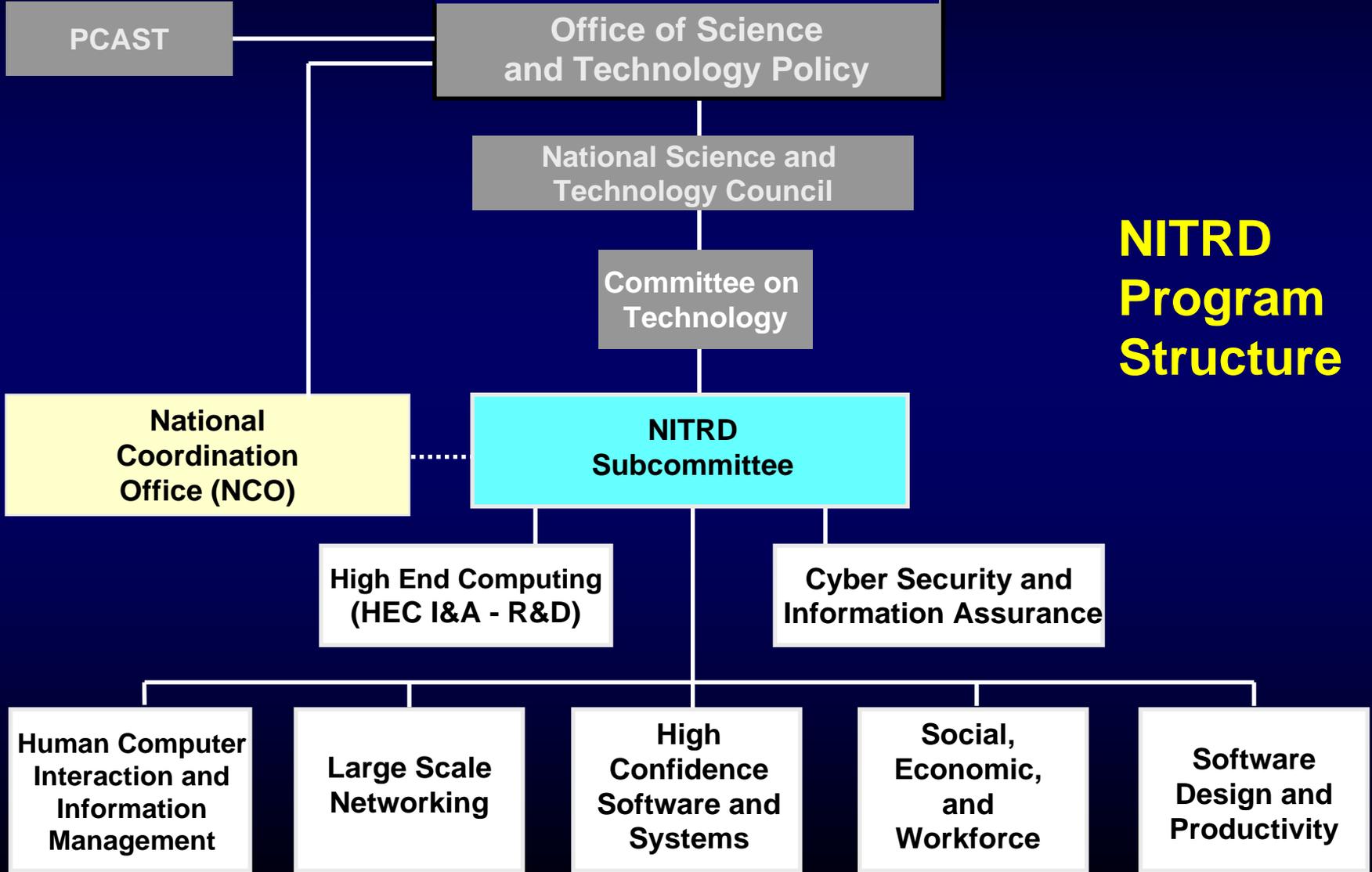
**Large Scale
Networking**

**High
Confidence
Software and
Systems**

**Social,
Economic,
and
Workforce**

**Software
Design and
Productivity**

**NITRD
Program
Structure**





AHRQ Agency for Healthcare Research and Quality



DARPA Defense Advanced Research Projects Agency



DOE/NNSA Department of Energy - National Nuclear Security Agency



DOE/SC Department of Energy - Mathematical, Information, and Computational Science Division



EPA Environmental Protection Agency



NARA National Archives and Records Administration



NASA National Aeronautics and Space Administration



NIH National Institutes of Health



NIST National Institute of Standards and Technology



NOAA National Oceanic and Atmospheric Administration



NSA National Security Agency



NSF National Science Foundation



OSD and DoD Service research organizations, Office of the Deputy, Under Secretary of Defense (Science and Technology)

E-Science

- Digital data driven
- Distributed
- Collaborative
- Trans-disciplinary
- Fuses pillars of science:
 - Experiment
 - Theory
 - Model/Simulation
 - Observation/Correlation

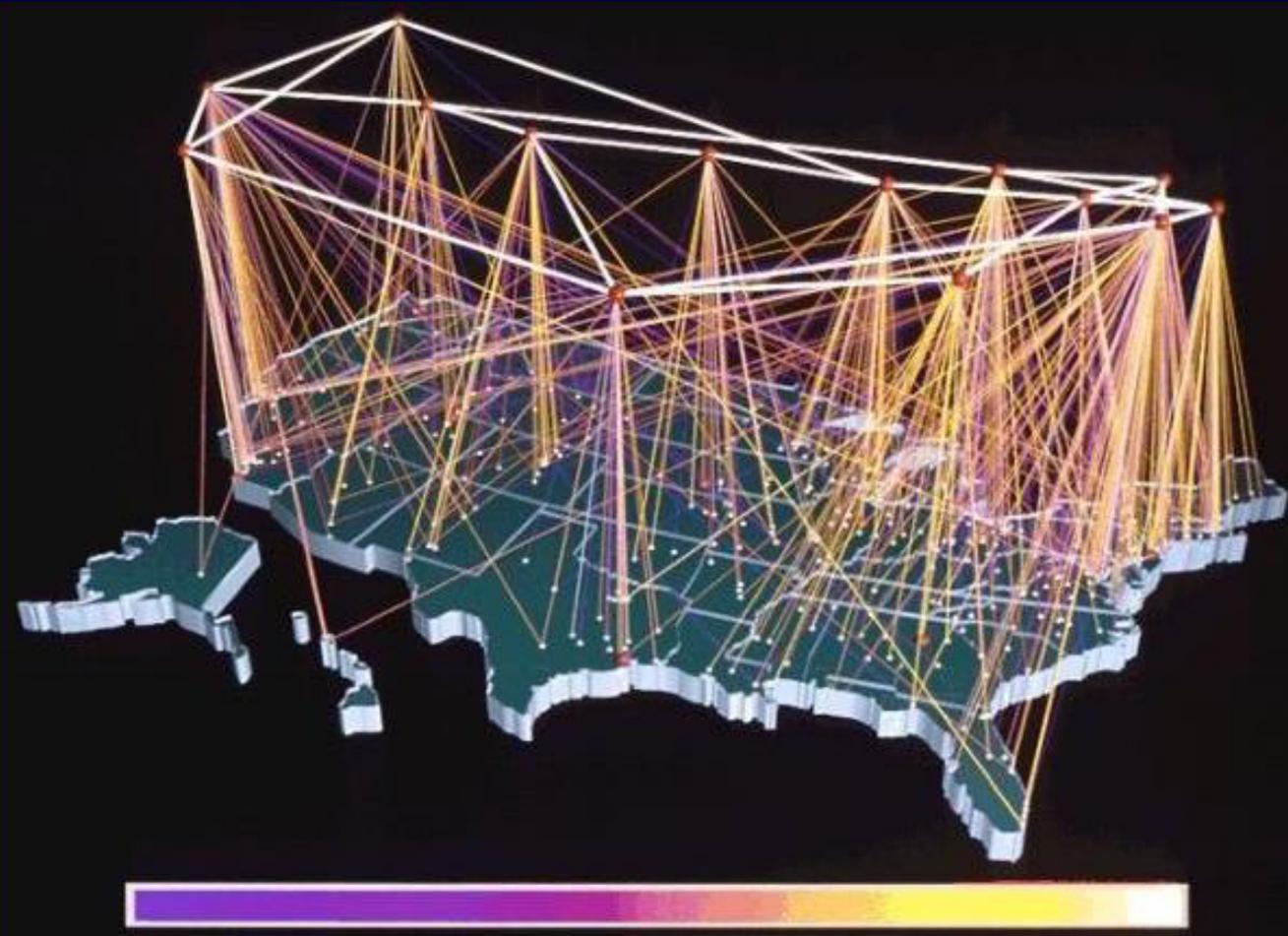
The End of Theory: The Data Deluge Makes the Scientific Method Obsolete

Google's founding philosophy is that we don't know why this page is better than that one: If the statistics ... say it is, that's good enough. No semantic or causal analysis is required. That's why Google can translate languages without actually "knowing" them (given equal corpus data, Google can translate Klingon into Farsi as easily as it can translate French into German).

We've come a long way ...

The Department of Defense's ARPANET project, launched in 1966 to explore methods for "resource sharing among computers", initially connected 4 nodes. Today's Internet links more than 1.4 billion users over more than 200,000 networks worldwide; with 14 new users added every second.

NSFNet Traffic– September 1991



The first commercial mobile phone service, MTA, deployed in 1956 by Ericsson in Sweden, had a few hundred subscribers and a phone that weighed 88 lb. Today cell phones weighing as little as 2 ounces are being connected at the rate of 30 per second with more than 3.8 billion in use – equal to more than half the world's population.



ENIAC, dedicated in 1946, was one of the first fully-functional digital computers, using more than 17,000 vacuum tubes to accomplish up to 5,000 addition operations per second. Today's petascale machines, Roadrunner at Los Alamos National Lab and the planned Blue Waters system at the University of Illinois, are designed to sustain more than one quadrillion (1,000,000,000,000,000) operations per second. A calculation these machines could complete in a week would take a machine operating at ENIAC speeds several billion years.

Redefining “Computer”

- “All the calculations that would ever be needed in this country could be done on three digital computers.”
-- Douglas Hartree, Cambridge, 1954
- “There is no reason for an individual to have a computer in their home.”
-- Ken Olsen, DEC, 1977
- “For the full year [2007], IDC said 269 million PCs were shipped worldwide”
-- International Herald Tribune, January, 2008.
- “In a sense, there are only five computers on earth.”
-- Yahoo Research Chief Prabhakar Raghavan, December, 2007.
- “...some researchers at IBM believe that five computers may be four too many.”
-- Nick Carr, The Guardian, February 21, 2008.

We've come a long way ...

... to reach the beginning

Imagine a world where ...

... all of the text in all of the libraries worldwide is in a storage device - in your pocket

... the network responds at the speed of light to a plain language question with a perceptive answer

... your contact lenses merge your digital and physical worlds

**Science is global and thrives in a world
that is not limited to 4-dimensions**

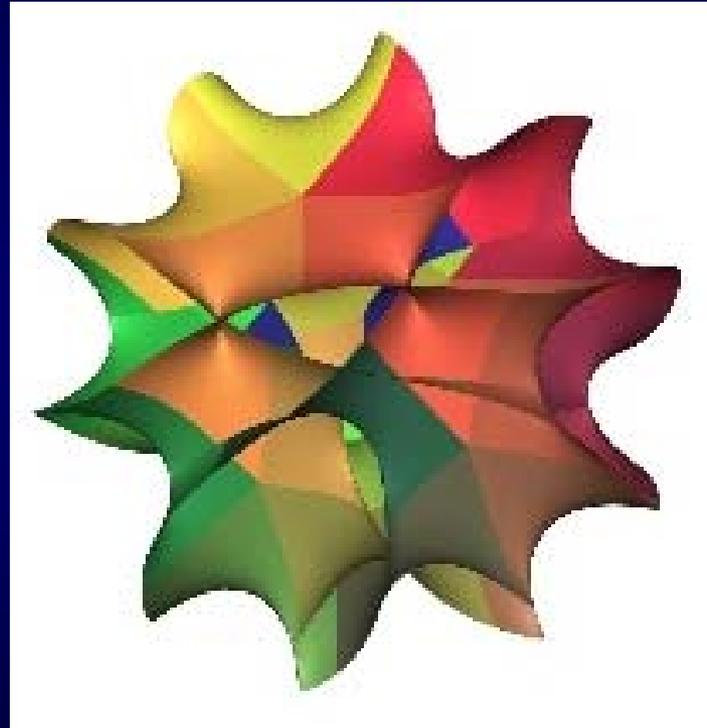
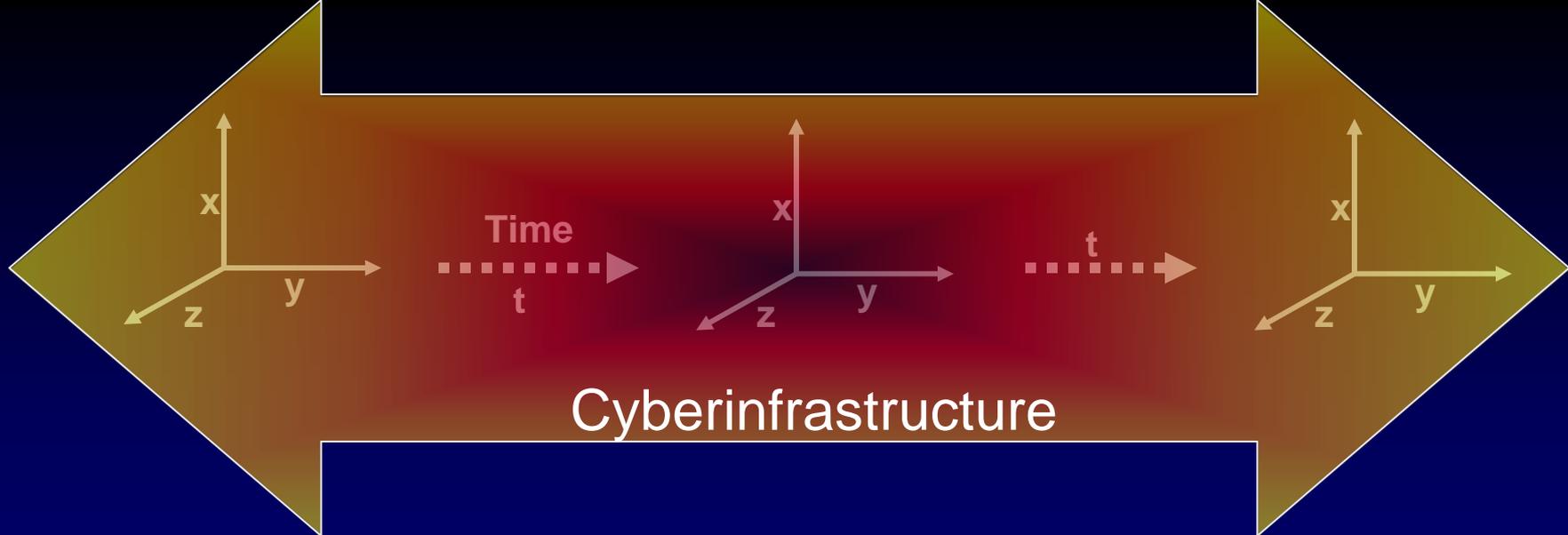


Image: Andrew J. Hanson
www.cs.indiana.edu/~hanson/



Computational
capacity and
capability



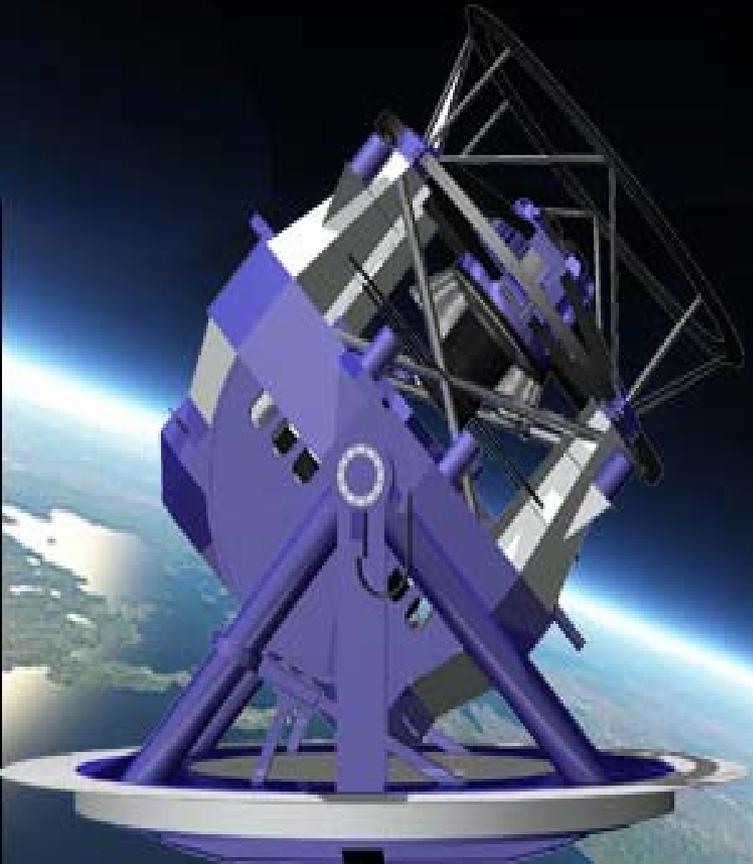
Connectivity
for access and
interaction



Information for
innovation and
discovery

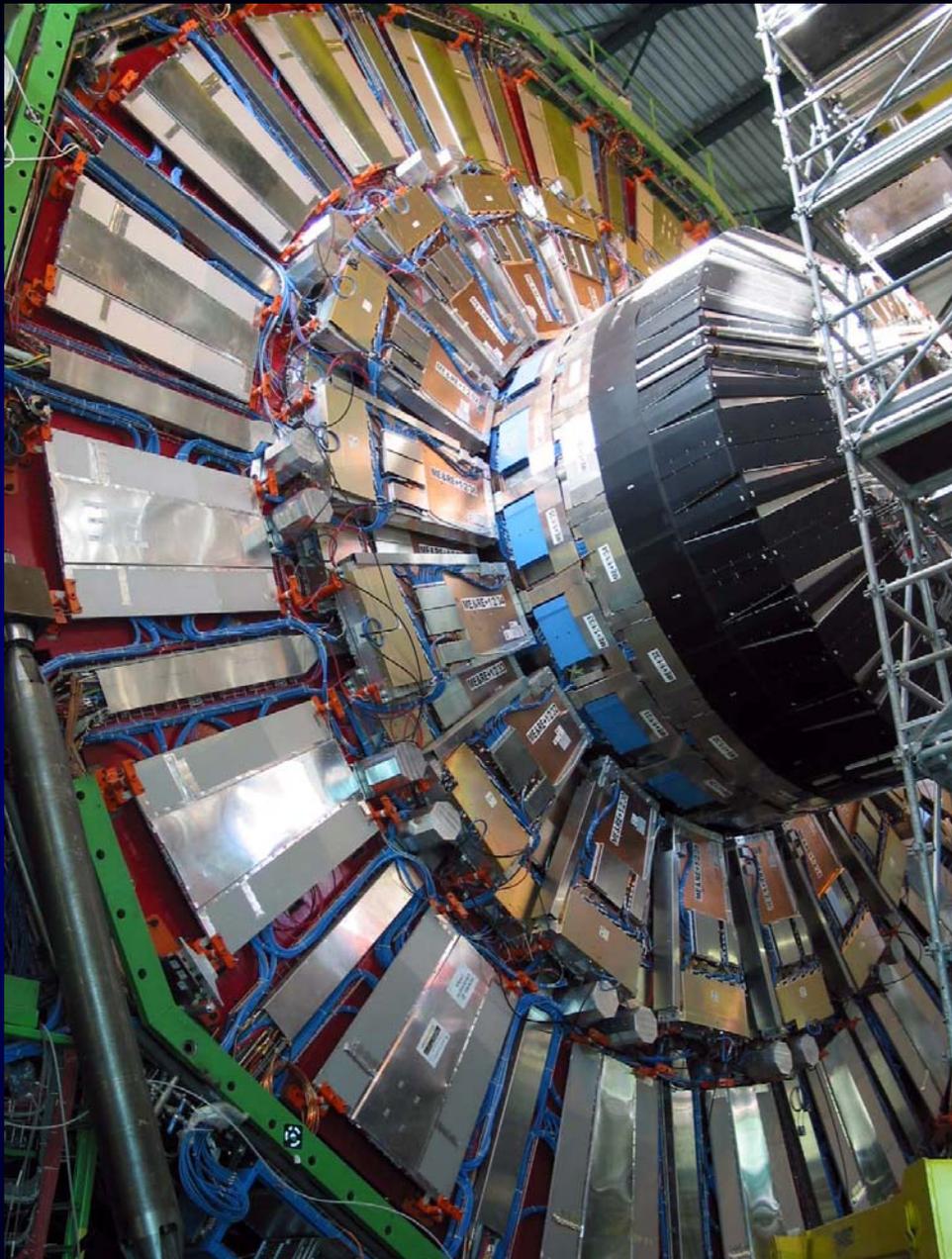
LSST

Large Synoptic Survey Telescope



“Sometime in the 2010s, if all goes well, the Large Synoptic Survey Telescope (LSST) will start to bring a vision of the heavens to Earth. Suspended between its vast mirrors will be a three billion-pixel sensor array, which on a clear winter night will produce 30 terabytes of data. In less than a week this remarkable telescope will map the whole night sky And then the next week it will do the same again ... building up a database of billions of objects and millions of billions of bytes.”

Nature 440:383



Large Hadron Collider

Physicists will use the LHC to recreate the conditions just after the Big Bang, by colliding two beams [of hadrons] head-on at very high energy.

When LHC begins operations, it will produce roughly 15 Petabytes of data annually, which thousands of scientists around the world will access and analyse ... The mission of the LHC Computing Project (LCG) is to build and maintain a data storage and analysis infrastructure for the entire high energy physics community that will use the LHC.

An IDC White Paper - sponsored by EMC

The Expanding Digital Universe

A Forecast of Worldwide
Information Growth Through 2010

March 2007

John F. Gantz, Project Director

David Reinsel

Christopher Chute

Wolfgang Schlichting

John McArthur

Stephen Minton

Irida Xheneti

Anna Toncheva

Alex Manfrediz

 **IDC**
Analyze the Future

“In 2006, the amount of digital information created, captured, and replicated was 1,288 x 10¹⁸ bits (or 161 exabytes) ... This is about 3 million times the information in all the books ever written”

A Challenge for Society

“If we are effectively to preserve for future generations the corpus of information in digital form that represents our cultural record, we need ... to commit ourselves technically, legally, economically, and organizationally to the full dimensions of the task.”

Report of the Task Force on Archiving of Digital Information, 1996
Commission on Preservation and Access and the Research Libraries Group

Role of Universities and Academic Libraries

The Universities

“Ever since their inception, universities have been occupied with the fundamental elements of what we now call 'knowledge management', i.e. the creation, collection, preservation and dissemination of knowledge.”

Andre Oesterlinck, Knowledge Management in
Post-Secondary Education: Universities

E-Education

- Information-Driven
- Accessible
- Distributed
- Interactive
- Context-Aware
- Experience & Discovery-Driven

The Academic Libraries

“It is to the research library community that others will look for the preservation of ... digital assets, as they have looked to us in the past for reliable, long-term access to the ‘traditional’ resources and products of research and scholarship.”

Association of Research Libraries (ARL)
Strategic Plan 2005-2009

Reinventing the Library

- Institutional commitment
- Sustainable funding model
- Defining the library user community
- Legal and policy frameworks
- Library workforce skills

Library as a computational center

- Sustainable technology framework



Contact:

greer@nitrd.gov