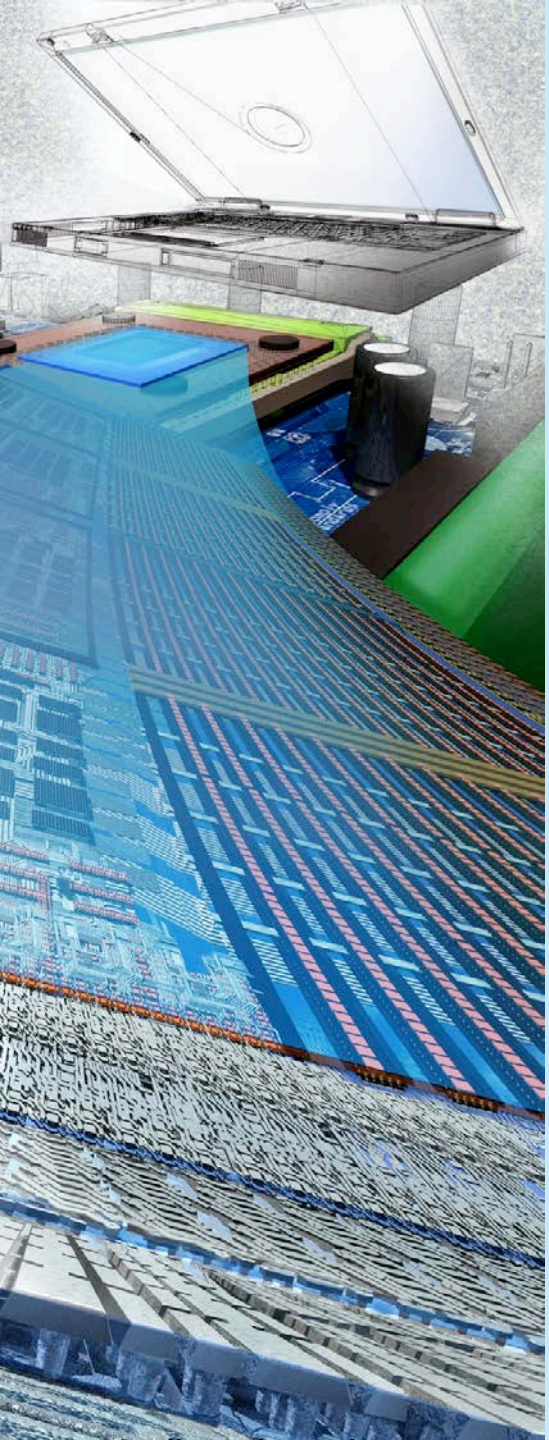


The Promise of Big Data

Farnam Jahanian
National Science Foundation

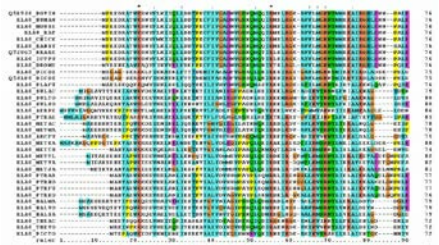
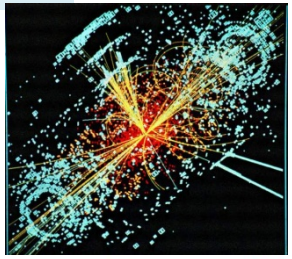
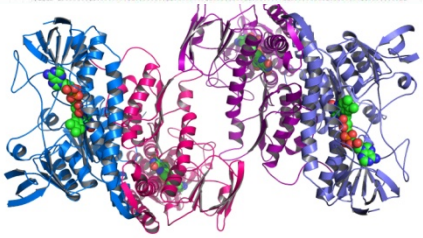
Big Data Partners Workshop
May 3, 2013

Advances in information technologies are transforming the fabric of our society, and data represents a transformative new currency for science, engineering, education and commerce.

Image Credit: CCC and SIGACT CATCS

Era of Data and Information

Scientific Data


Digital Media


MOBILE VOIP

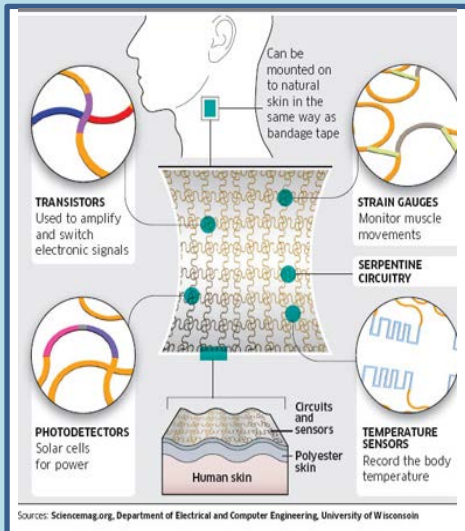
EMAIL

BLOGS

VIDEOS

MESSAGING

Human Sensors



Health Care

Credit: Image modified from University of Florida

Why is Big Data Important?

- Transformative implications for commerce and economy
- Critical to accelerating the pace of discovery in almost every science and engineering discipline
- Potential for addressing some of society's most pressing challenges

Paradigm Shift: from Hypothesis-driven to Data-driven Discovery

The Economist, The data deluge and how to handle it: A 14-page special report (Feb 25, 2010).

The Fourth Paradigm: Data-Intensive Scientific Discovery (2009, Microsoft Corporation).

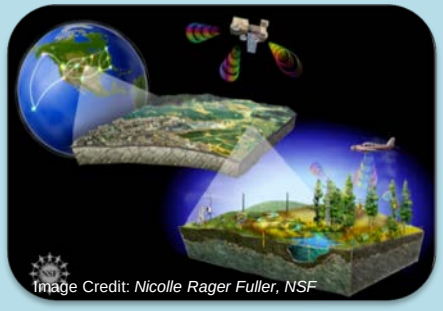
http://www.sciencemag.org/site/special/data/

http://www.economist.com/node/15579717

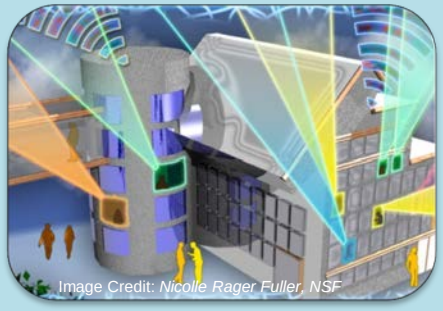
http://research.microsoft.com/en-us/collaboration/fourthparadigm/

Data are motivating a profound transformation in the culture and conduct of scientific research.

Data-driven Discovery and Innovation Address Societal Challenges



Environment & Sustainability



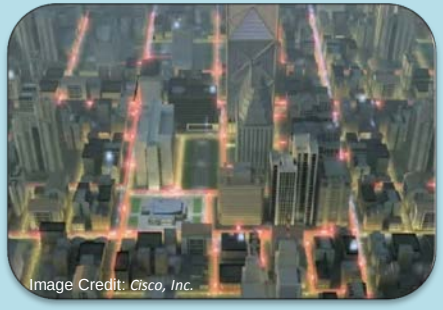
Broadband & Universal Connectivity

Manufacturing, Robotics, & Smart Systems

Emergency Response & Disaster Resiliency

Secure Cyberspace

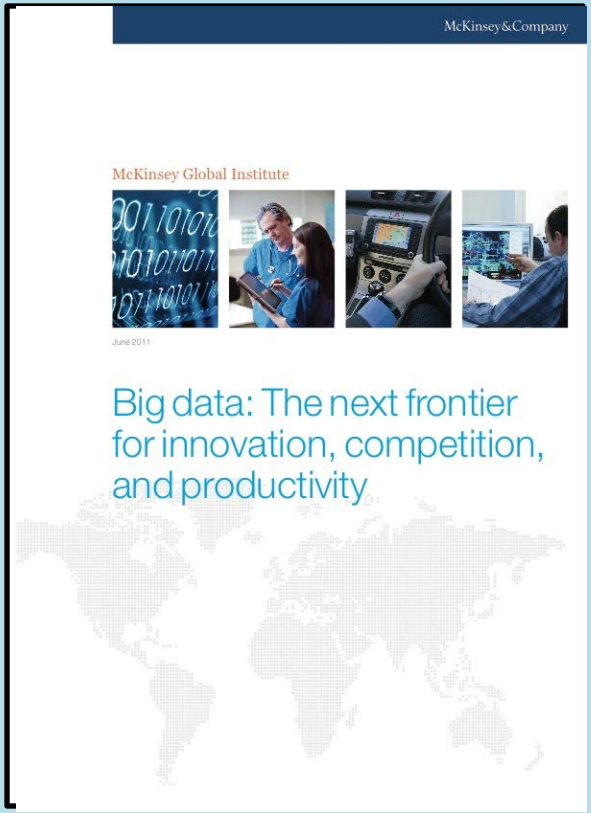
Health & Wellbeing



Transportation & Energy

Education and Workforce Development

Education, Learning, Workforce Development, Computational and Data-enabled Science



"By 2018 the United States alone faces a shortage of 140,000 to 190,000 people with analytical expertise and 1.5 million managers and analysts with the skills to understand and make decisions based on the analysis of big data."[1]

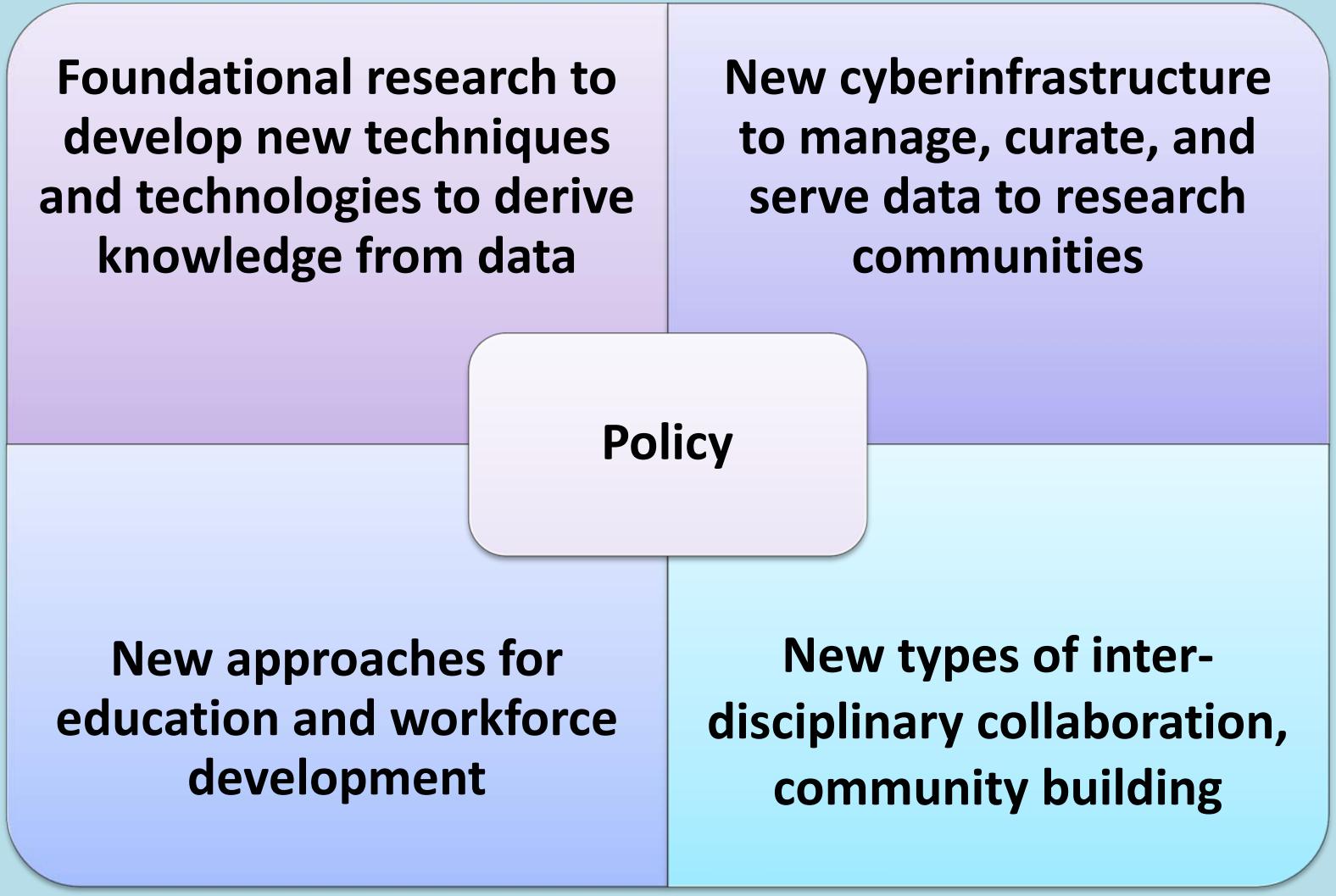
[1] McKinsey&Company (May 2011), "Big data: The next frontier for innovation, competition, and productivity." Available at: http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation

Classifying Breast Cancers via Image Analysis

Energy Savings in the Home

Reducing Traffic Congestion in Urban Areas

NSF Framework for Investments



Complex Policy Setting

- Practitioners and researchers want data.
- Public policy requires access to data.
- Public policy also requires protection of privacy, intellectual property, and other sensitive information.
- Policy and implementation plan for data sharing and open access are in progress. (WH OSTP Feb. 22nd memo on public access)

"Paradox of Innovation: no one knows how an invention will impact the world until it is widely used, leading to unintended consequences"

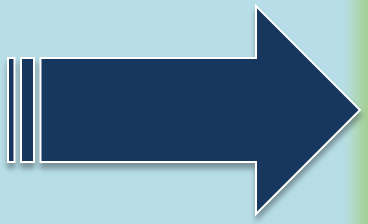
Why Now? Confluence of Social, Technical and Policy Interests

- Decades of advances in technology
- Data is no longer regarded as static:
  - now a raw material of business, potentially used to created economic value
- Scalability: collecting, organizing, storing and analyzing information
- Increasing transparency of democratic governance (open gov)
- Public access to high value datasets (data.gov)
- Democratization of data and tools

"Paradox of Innovation: no one knows how an invention will impact the world until it is widely used, leading to unintended consequences"

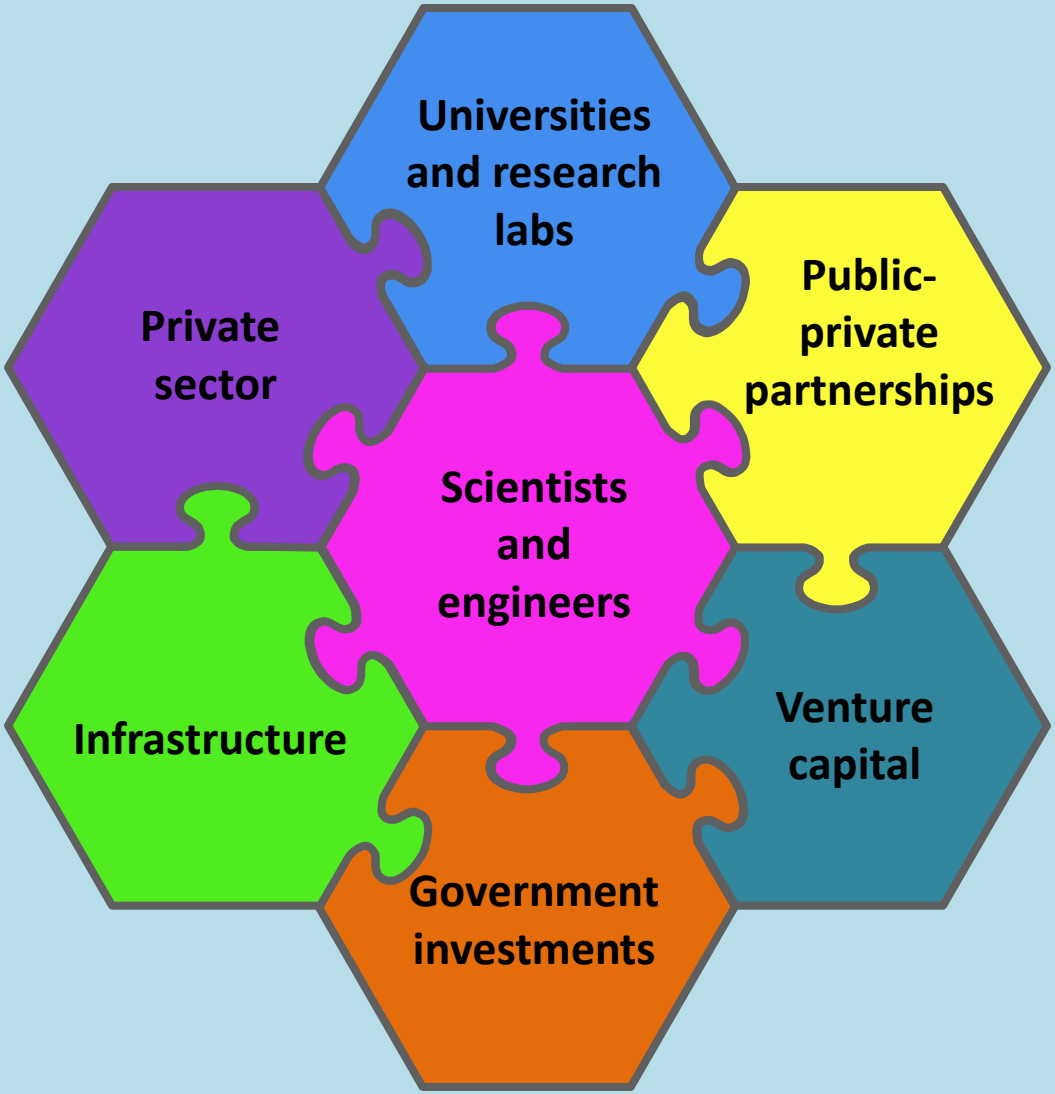
Why Now? Confluence of Social, Technical and

- Decades of advances in technology
- Data is no longer regarded as static:
  - now a raw material of business, potentially used to crea
- Scalability: collecting, organizing, storing and analyzin
- Increasing transparency of democratic governance (op
- Public access to high value datasets (data.gov)
- Democratization of data and tools



Moore's Law
Kryder's Law
Pervasive Sensors
Data Mining
Machine Learning
NL Understanding
Info Retrieval
Computer Vision
Video Analytics
Data Visualization
Crowd Sourcing
Social Networks
…

Discovery and Innovation Ecosystem



Thanks!

fjahania@nsf.gov

Credits

- Copyrighted material used under Fair Use. If you are the copyright holder and believe your material has been used unfairly, or if you have any suggestions, feedback, or support, please contact: ciseitsupport@nsf.gov.

- Except where otherwise indicated, permission is granted to copy, distribute, and/or modify all images in this document under the terms of the GNU Free Documentation license, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled “GNU Free Documentation license” at http://commons.wikimedia.org/wiki/Commons:GNU_Free_Documentation_License.

- The inclusion of a logo does not express or imply the endorsement by NSF of the entities' products, services, or enterprises.