

**STORAGE AND
INFRASTRUCTURE
OPPORTUNITIES**

GREG MADEY

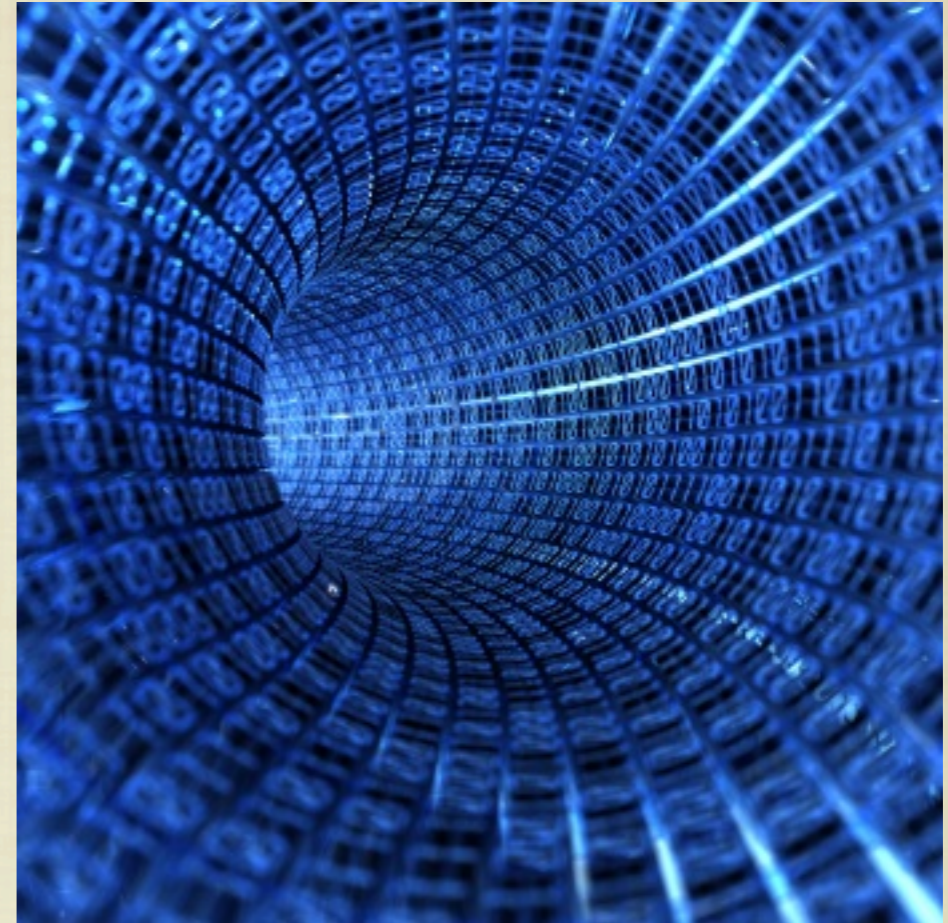
COMPUTER SCIENCE & ENGINEERING

2008 CRC WORKSHOP - SCIENTIFIC COMPUTING

MAY 1, 2008

IT'S ALL ABOUT THE DATA

- CHEAPER TO STORE
- CHEAPER TO OBTAIN
- DATA DRIVEN DISCOVERY
- DATA DRIVEN SCIENCE
- RESEARCH OPPORTUNITIES & CHALLENGES
- EXAMPLES / FUNDING OPPORTUNITIES

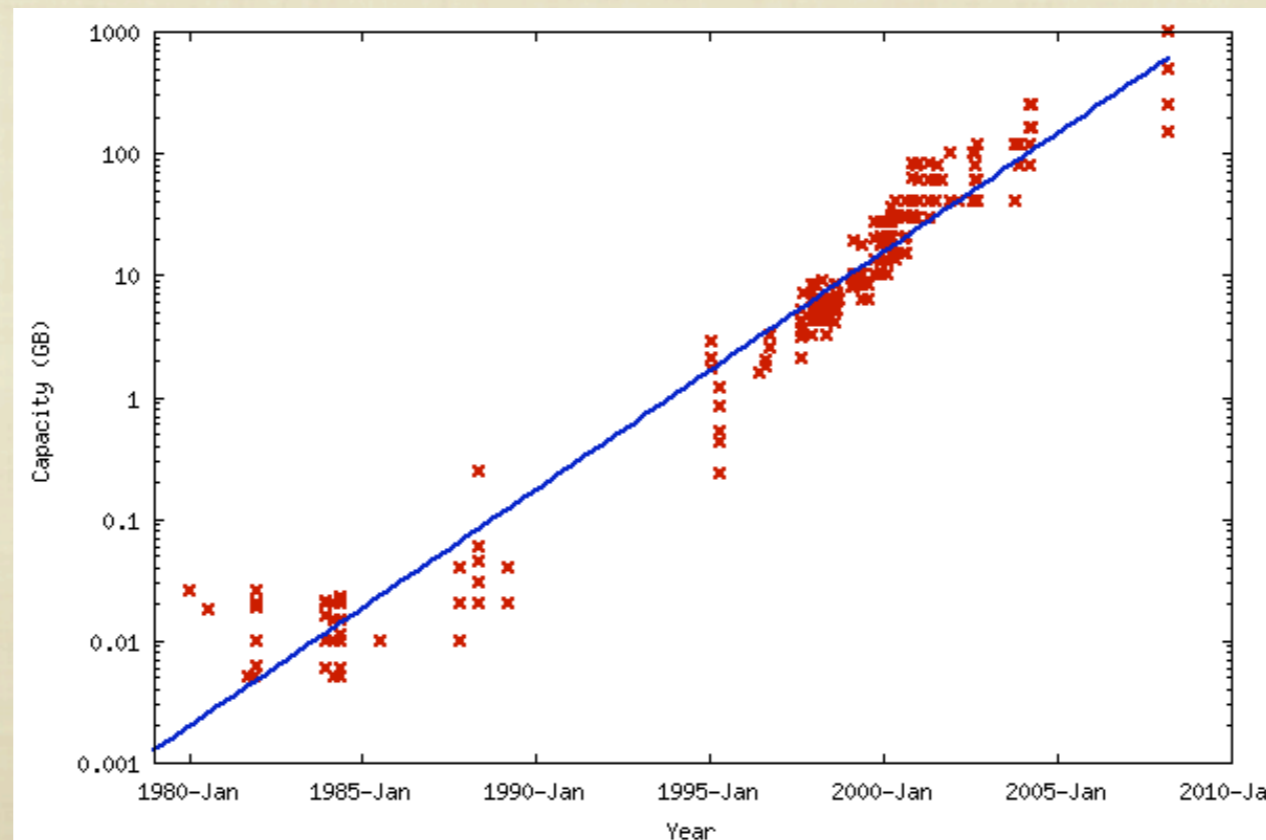


KRYDER'S LAW

- FASTER THAN MOORE'S LAW
- SEAGATE RESEARCH VP & CTO



CAPACITY
(GB)



YEAR

KRYDER'S LAW - \$COST

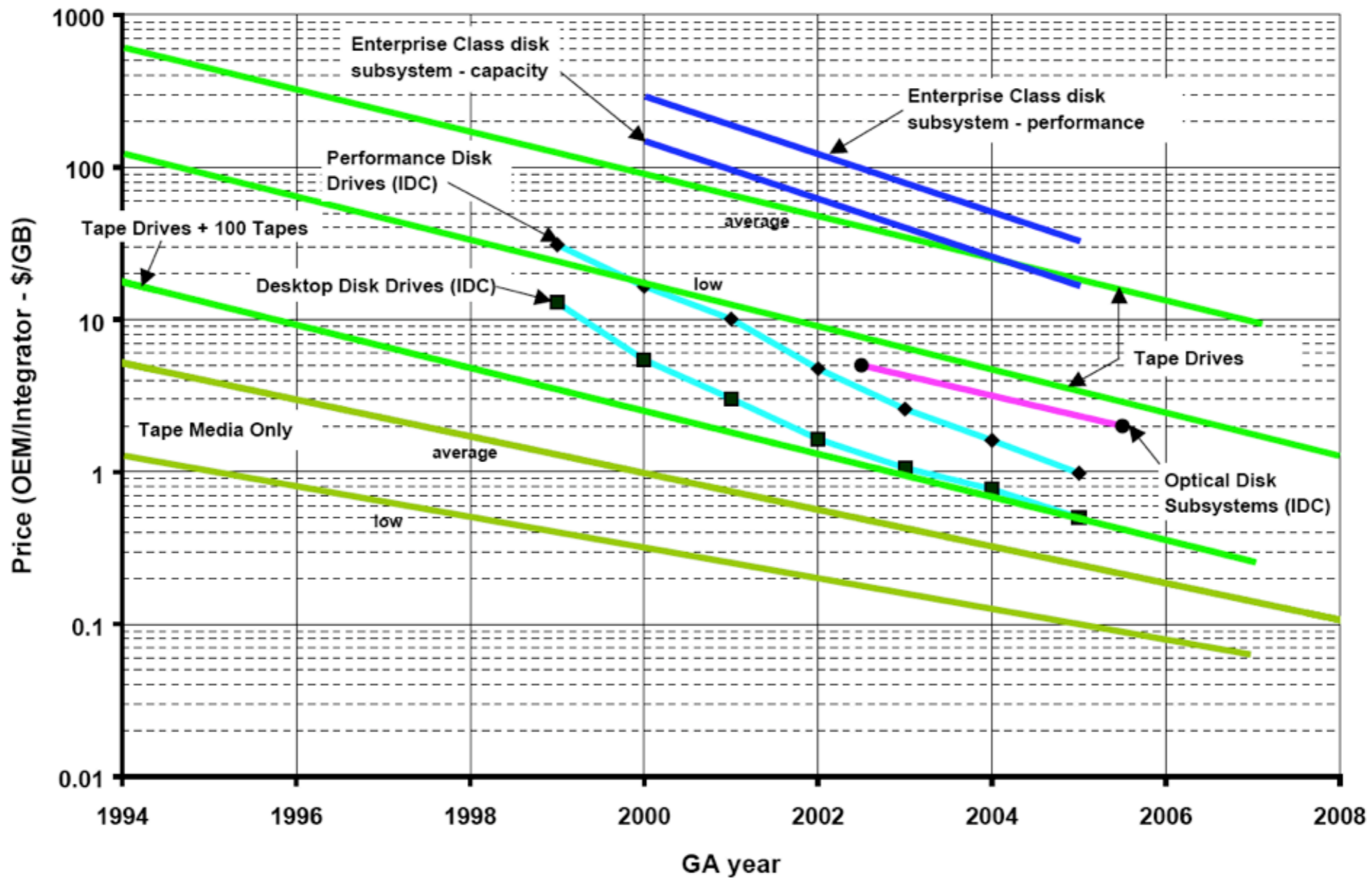


STORAGETEK

INFORMATION *made* POWERFUL

Storage Subsystem Price Trends

(OEM price/equiv. unless otherwise noted; no capacity compression or utilization factors)



\$/GB

YEAR

CHEAPER TO STORE

TODAY'S PRICES



1.0TB Seagate Barracuda 7200.11 SATA II 7200RPM 32MB w/Perpendicular Drive Technology, 5yr Warranty

1.0TB Seagate Barracuda 7200.11 SATA II 7200RPM 32MB W/ Perpendicular Drive Technology New w/5 Year Seagate Warranty! [more](#)

Same Day

\$247.99

Brand: Seagate OWC Item # SEAST31000340AS

\$0.25/GB



500GB Seagate Barracuda 7200.11 SATA II 7200RPM 32MB Cache - 5yr Seagate Warranty

500GB Seagate Barracuda 7200.11 SATA II 7200RPM 32MB Buffer with Perpendicular Drive Technology. New w/5 Year Seagate Warranty! [more](#)

Same Day

\$108.97

Brand: Seagate OWC Item # SEAST3500320AS

\$0.22/GB



NEW PRICE **750GB Seagate Barracuda 7200.11 SATA II 7200RPM 32MB W/ Perpendicular Drive Technology**

750GB Seagate Barracuda 7200.11 SATA II 7200RPM 32MB Buffer W/ Perpendicular Drive Technology. New w/5 Year Seagate Warranty! [more](#)

Same Day

\$159.99

Brand: Seagate OWC Item # SEAST3750330AS

\$0.22/GB



250GB Seagate Barracuda 7200.10 SATA II 7200RPM 16MB W/ Perpendicular Drive Technology

250GB Seagate Barracuda 7200.10 SATA II 7200RPM 16MB Buffer W/ Perpendicular Drive Technology. New w/5 Year Seagate Warranty! [more](#)

Same Day

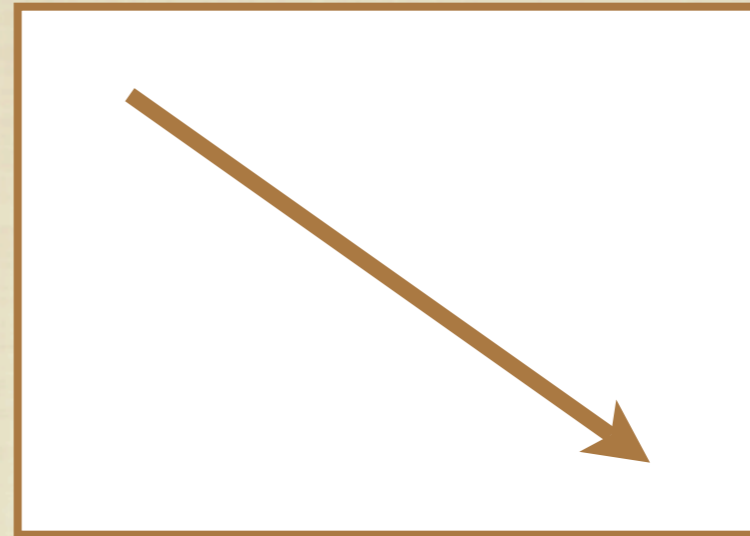
\$68.97

Brand: Seagate OWC Item # SEAST3250410AS

\$0.28/GB

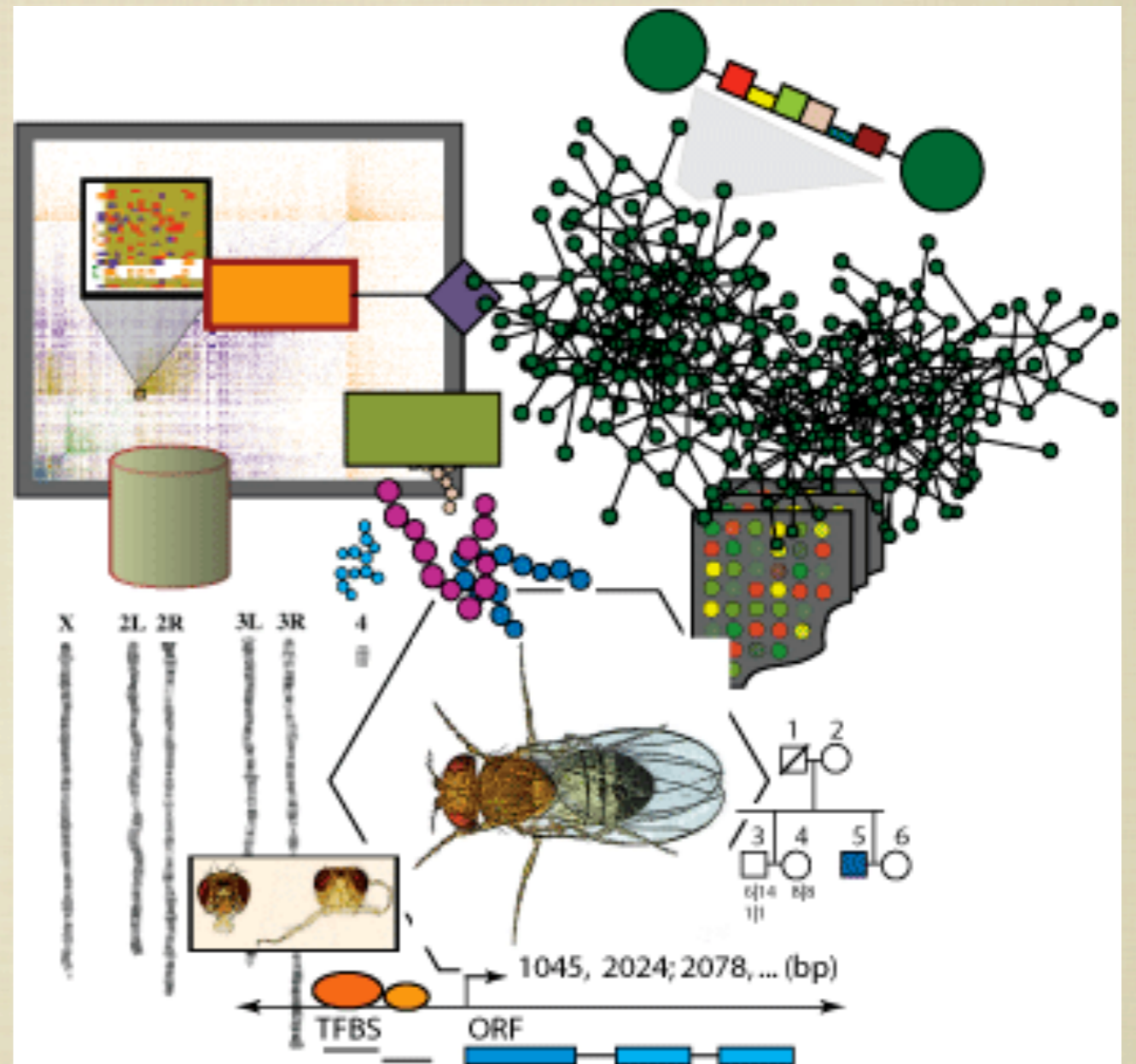
CHEAPER TO COLLECT

- **RFIDS**
- **SENSOR NETS**
- **THE WWW, SCREEN SCRAPING, GOOGLE SEARCHES**
- **LIFE IN CYBERSPACE - LOG FILES, DIGITAL TRACES, METADATA**
- **FASTER COMPUTERS ==> MORE DATA TO STUDY**



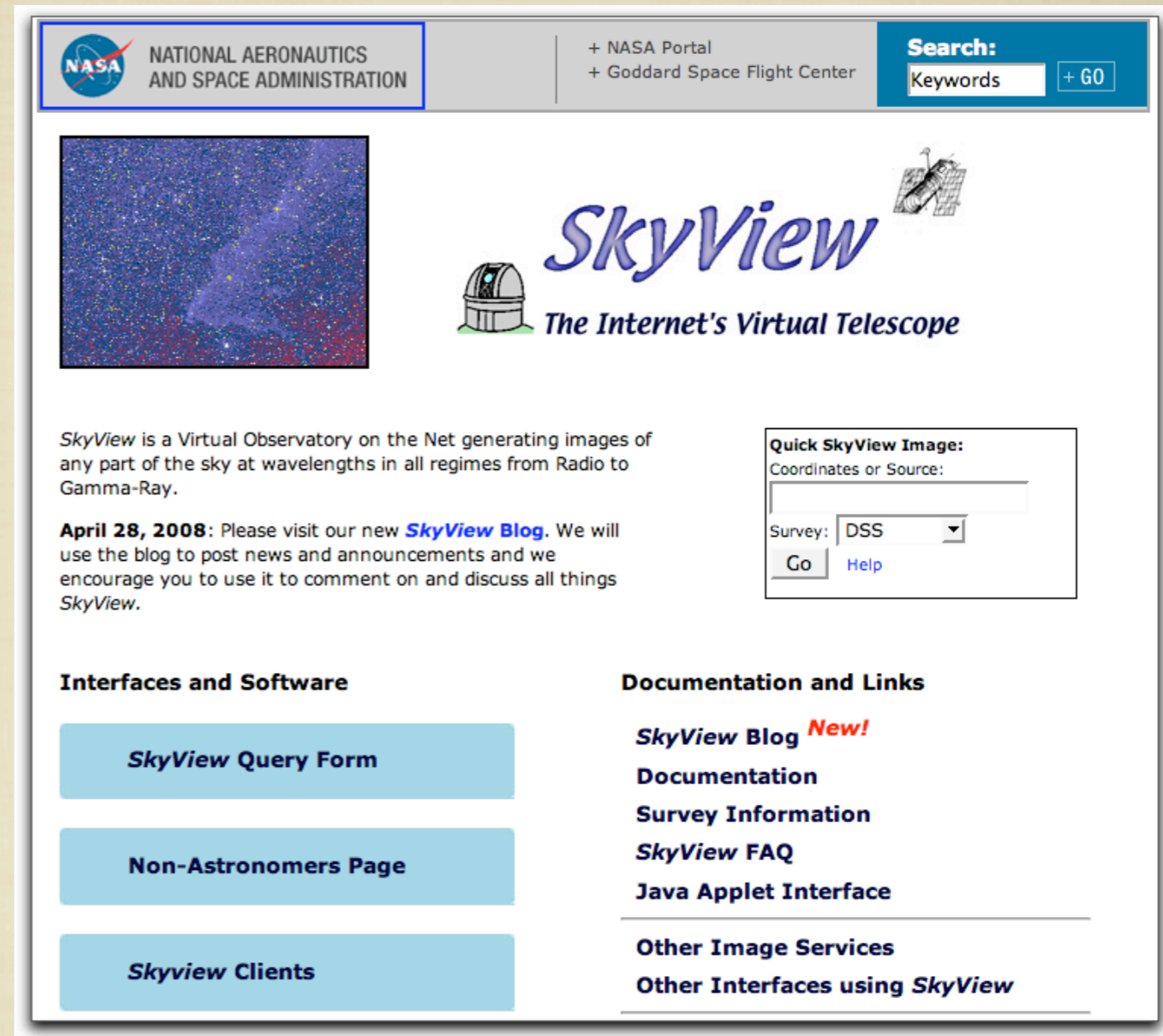
DATA DRIVEN DISCOVERY

- STATISTICS
- DATA MINING
- MACHINE LEARNING
- DATABASE QUERIES
- DATA VISUALIZATION



DATA DRIVEN SCIENCE

- ENGINEERING
- ASTROPHYSICS
- HIGH ENERGY PHYSICS
- BIOINFORMATICS
- CHEMINFORMATICS
- SOCIAL SCIENCES
- GEOGRAPHICAL INFORMATION SYSTEMS
- COLLABORATORIES, RESEARCH DATA ARCHIVES, VIRTUAL ORGANIZATIONS, CYBERINFRASTRUCTURE

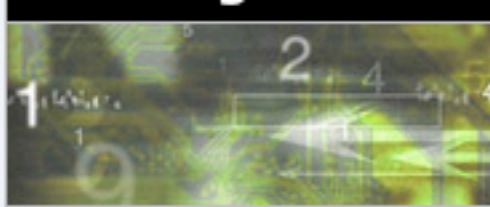


The screenshot shows the SkyView website interface. At the top left is the NASA logo and the text "NATIONAL AERONAUTICS AND SPACE ADMINISTRATION". To the right of this is a search bar with the text "Search: Keywords" and a "+ GO" button. Below the NASA logo is a large image of a star field. To the right of the star field is the "SkyView" logo, which includes a small telescope icon and the tagline "The Internet's Virtual Telescope". Below the star field and logo is a paragraph of text: "SkyView is a Virtual Observatory on the Net generating images of any part of the sky at wavelengths in all regimes from Radio to Gamma-Ray." Below this text is a date announcement: "April 28, 2008: Please visit our new SkyView Blog. We will use the blog to post news and announcements and we encourage you to use it to comment on and discuss all things SkyView." To the right of this text is a "Quick SkyView Image:" form with a text input field for "Coordinates or Source:", a "Survey:" dropdown menu set to "DSS", and "Go" and "Help" buttons. Below the main text are two columns of links. The left column is titled "Interfaces and Software" and contains three blue buttons: "SkyView Query Form", "Non-Astronomers Page", and "Skyview Clients". The right column is titled "Documentation and Links" and contains several links: "SkyView Blog **New!**", "Documentation", "Survey Information", "SkyView FAQ", "Java Applet Interface", "Other Image Services", and "Other Interfaces using SkyView".

RESEARCH OPPORTUNITIES & CHALLENGES

- **SENSORS, SENSOR NETWORKS**
- **STORAGE TECHNOLOGIES**
- **COLLABORATORY DESIGN**
- **DATABASE DESIGN**
- **DATAMINING ALGORITHMS**
- **ONTOLOGIES, CONTROLLED VOCABULARIES**





Sustainable Digital Data Preservation and Access Network Partners (DataNet)

[Funding](#)

[About Funding](#)

[A-Z Index of Funding Opportunities](#)

[Advanced Funding Search](#)

[Guide to Programs/Browse Funding Opportunities](#)

[Recent Funding Opportunities](#)

[Upcoming Due Dates](#)

Proposals and Awards

[Proposal and Award Policies and Procedures Guide](#)

[Introduction](#)

[Proposal Preparation and Submission](#)

- [Grant Proposal Guide](#)
- [Grants.gov Application Guide](#)

[Award Administration](#)

- [Award Administration Guide](#)

[Award Conditions](#)

CONTACTS

Name	Email	Phone	Room
Lucy Nowell	lnowell@nsf.gov	(703) 292-8970	
Sylvia Spengler	sspenle@nsf.gov	(703) 292-8930	

PROGRAM GUIDELINES

[07-601](#) Solicitation

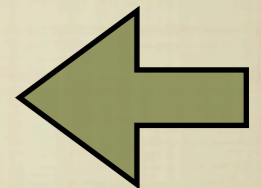
DUE DATES

Preliminary Proposal Deadline Date: October 6, 2008

Full Proposal Target Date: February 16, 2009

SYNOPSIS

Science and engineering research and education are increasingly digital and increasingly data-intensive. Digital data are not only the output of research but provide input to new hypotheses, enabling new scientific insights and driving innovation. Therein lies one of the major challenges of this scientific generation: how to develop the new methods, management structures and technologies to manage the diversity, size, and complexity of current and future data sets and data streams. This solicitation addresses that challenge by creating a set of exemplar national and global data research infrastructure organizations (dubbed DataNet Partners) that provide unique opportunities to communities of researchers to advance science and/or engineering research and learning.





Foundations of Data and Visual Analytics (FODAVA)

Funding

[About Funding](#)

[A-Z Index of Funding Opportunities](#)

[Advanced Funding Search](#)

[Guide to Programs/Browse Funding Opportunities](#)

[Recent Funding Opportunities](#)

[Upcoming Due Dates](#)

Proposals and Awards

[Proposal and Award Policies and Procedures Guide](#)

Introduction

[Proposal Preparation and Submission](#)

- [Grant Proposal Guide](#)
- [Grants.gov Application Guide](#)

[Award Administration](#)

- [Award Administration Guide](#)

[Award Conditions](#)

[Other Types of Proposals](#)

[Merit Review](#)

[NSF Outreach](#)

[Policy Office Website](#)

CONTACTS

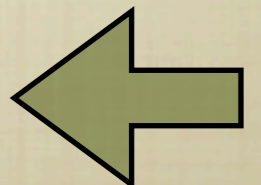
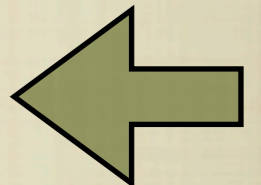
Name	Email	Phone	Room
Lawrence Rosenblum (CS Contact)	lrosenbl@nsf.gov	(703) 292-8910	1115 N
Tie Luo (Math Contact)	tluo@nsf.gov	(703) 292-8448	1025 N
Sankar Basu	sbasu@nsf.gov	(703) 292-7843	1115N
Ephraim Glinert	eglinert@nsf.gov	(703) 292-8930	1125 N
Leland Jameson	lameson@nsf.gov	(703) 292-4883	1025N
Maria Zemankova	mzemanko@nsf.gov	(703) 292-8930	1125 N

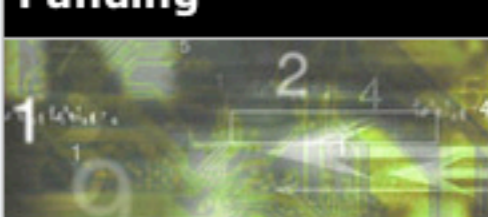
PROGRAM GUIDELINES

[07-583 Solicitation](#)

SYNOPSIS

Those involved with science, engineering, commerce, health, and national security all increasingly face the challenge of synthesizing information and deriving insight from massive, dynamic, ambiguous and possibly conflicting digital data. The goal of collecting and examining these data is not to merely acquire information, but to derive increased understanding from it and to facilitate effective decision-making. To capitalize on the opportunities provided by these data sets, a new, interdisciplinary field of science is emerging called Data and Visual Analytics, which is defined as the science of analytical reasoning facilitated by interactive visual interfaces. Data and Visual Analytics requires interdisciplinary science, going beyond traditional scientific and information visualization to include statistics, mathematics, knowledge representation, management and discovery technologies, cognitive and perceptual sciences, decision sciences, and more. This solicitation is concerned only with a subset of the overall problem, namely the creation of the mathematical and computational sciences foundations required to transform data in ways that permit visual-based understanding.





Funding

[About Funding](#)

[A-Z Index of Funding Opportunities](#)

[Advanced Funding Search](#)

[Guide to Programs/Browse Funding Opportunities](#)

[Recent Funding Opportunities](#)

[Upcoming Due Dates](#)

Proposals and Awards

[Proposal and Award Policies and Procedures Guide](#)

[Introduction](#)

[Proposal Preparation and Submission](#)

- [Grant Proposal Guide](#)
- [Grants.gov Application Guide](#)

[Award Administration](#)

- [Award Administration Guide](#)

[Award Conditions](#)

[Other Types of Proposals](#)

[Merit Review](#)

[NSF Outreach](#)

[Policy Office Website](#)

[Crosscutting/NSF-wide](#)

Cyber-Enabled Discovery and Innovation (CDI)

[Crosscutting Programs](#) [NSF Wide Programs](#)

CONTACTS

Name	Dir/Div	Name	Dir/Div
Sirin Tekinay	CISE/CCF	Thomas Russell	MPS/DMS
Eduardo Misawa	ENG/CMMI		

Drs. Tekinay, Russell, and Misawa are being assisted by a multidisciplinary team of Program Officers drawn from throughout NSF. CDI team members include: Kile Baker (GEO/ATM), Charles Bouldin (MPS/DMR), Maria Burka (ENG/CBET), Arlene de Strulle (EHR/DRL), Cheryl Eavey (SBE/SES), Anne Emig (OD/OISE), Anne-Francoise Lamblin (BIO/DBI), D. Terence Langendoen (SBE/BCS), Mary Lou Maher (CISE/IIS), Peter McCartney (BIO/DBI), Barbara Olds (EHR/OAD), Abani Patra (OD/OCI), Wayne Patterson (OD/OISE), Diana Rhoten (OD/OCI), William Wiseman (OD/OPP) and Eva Zanzerkia (GEO/EAR).

PROGRAM GUIDELINES

[07-603 Solicitation](#)

DUE DATES

Letter of Intent Window: August 30, 2008 - September 30, 2008

Preliminary Proposal Window: October 4, 2008 - November 4, 2008

Full Proposal Window: January 27, 2009 - February 27, 2009

SYNOPSIS

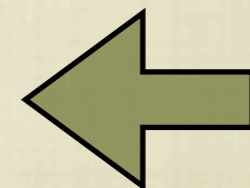
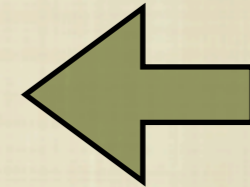
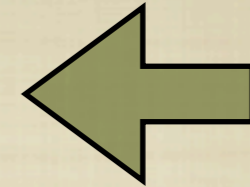
Cyber-Enabled Discovery and Innovation (CDI) is NSF's bold five-year initiative to create *revolutionary* science and engineering research outcomes made possible by innovations and advances in computational thinking. Computational thinking is defined comprehensively to encompass computational concepts, methods, models, algorithms, and tools. Applied in challenging science and engineering research and education contexts, computational thinking promises a profound impact on the Nation's ability to generate and apply new knowledge. Collectively, CDI research outcomes are expected to produce paradigm shifts in our understanding of a wide range of science and engineering phenomena and socio-technical innovations that create new wealth and enhance the national quality of life.

CDI seeks ambitious, transformative, multidisciplinary research proposals within or across the following three thematic areas:

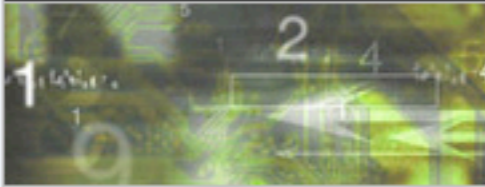
- **From Data to Knowledge:** *enhancing human cognition and generating new knowledge from a wealth of heterogeneous digital data;*
- **Understanding Complexity in Natural, Built, and Social Systems:** *deriving fundamental insights on systems comprising multiple interacting elements;* and
- **Building Virtual Organizations:** *enhancing discovery and innovation by bringing people and resources together across institutional, geographical and cultural boundaries.*

With an emphasis on bold multidisciplinary activities that, through computational thinking, promise radical, paradigm-changing research findings, CDI is unique within NSF. Accordingly, investigators are encouraged to come together in the development of far-reaching, high-risk science and engineering research and education agendas that capitalize on innovations in, and/or innovative use of, computational thinking. CDI projects are expected to build upon productive intellectual partnerships involving investigators from academe, industry and/or other types of organizations, including international entities.

Congruent with the three thematic areas, CDI projects will enable transformative discovery to identify patterns and structures in massive datasets; exploit computation as a means of achieving deeper understanding in the natural and social sciences and engineering; simulate and predict complex stochastic or chaotic systems; explore and model nature's interactions, connections, complex relations, and interdependencies, scaling from sub-particles to galactic, from subcellular to biosphere, and from the individual to the societal; train future generations of scientists and engineers to enhance and use cyber resources; and facilitate creative, cyber-enabled boundary-crossing collaborations, including those with industry and international dimensions, to advance



Funding



Funding

About Funding

A-Z Index of Funding Opportunities

Advanced Funding Search

Guide to Programs/Browse Funding Opportunities

Recent Funding Opportunities

Upcoming Due Dates

Proposals and Awards

Proposal and Award Policies and Procedures Guide

Introduction

Proposal Preparation and Submission

- [Grant Proposal Guide](#)
- [Grants.gov Application Guide](#)

Award Administration

- [Award Administration Guide](#)

Award Conditions

Other Types of Proposals

Merit Review

NSF Outreach

Policy Office Website

[Directorate for Biological Sciences](#)

Protein Data Bank Management (PDB Mgmt)

CONTACTS

Name	Email	Phone	Room
Peter H. McCartney	biopdb@nsf.gov	(703) 292-8470	

Investigators are encouraged to contact the program with questions about appropriateness for this solicitation before submitting a proposal.

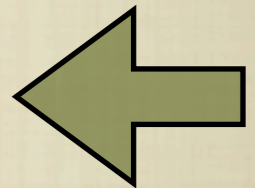
PROGRAM GUIDELINES

[08-507](#) Solicitation

SYNOPSIS

The goal of this solicitation is to ensure the continuing, highest possible quality of management for the central, U.S. component of the Protein Data Bank (PDB). The Protein Data Bank has achieved global recognition as the sole, authoritative repository for macromolecular structure information. To secure the continued success of this international resource, the U.S. supporting agencies are committed to ensuring that the management of the U.S. component is of the highest quality, providing a reliable, capable, and effective partner for its counterparts overseas. This competition for PDB management is designed to fulfill that commitment.

Proposals submitted in response to this solicitation must describe a plan for PDB management that is forward-looking and provides an effective framework for anticipating and responding to rapidly changing technologies and to the constantly changing needs, expectations, and composition of the user community. Proposers should: present a compelling rationale and vision for the management organization including its vision for the future of PDB; describe a strategy for collaborating with other national and international partners, resource providers, and community groups; address the full data life cycle management needs and requirements; describe mechanisms to ensure that PDB will be a community resource for research, education, and training; and describe a plan for leadership, administration, and community input and oversight that ensures that the management organization will be responsive to the community and successful in achieving its vision for the Protein Data Bank.



Funding



Funding

About Funding

A-Z Index of Funding Opportunities

Advanced Funding Search

Guide to Programs/Browse Funding Opportunities

Recent Funding Opportunities

Upcoming Due Dates

Proposals and Awards

Proposal and Award Policies and Procedures Guide

Introduction

Proposal Preparation and Submission

- [Grant Proposal Guide](#)
- [Grants.gov Application Guide](#)

Award Administration

Award

[Division of Biological Infrastructure](#)

Biological Databases and Informatics (BD&I)

CONTACTS

Name	Email	Phone	Room
Peter McCartney	dbidba@nsf.gov	(703) 292-8470	

PROGRAM GUIDELINES

[05-577](#) Solicitation

DUE DATES

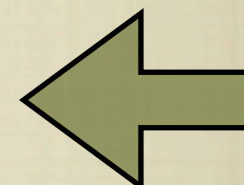
Full Proposal Target Date: July 14, 2008
annually thereafter

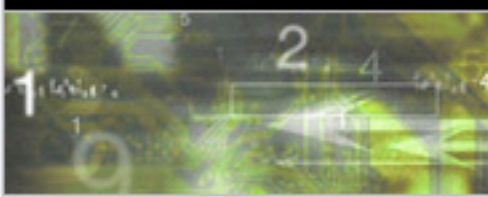
SYNOPSIS

The Biological Databases and Informatics (BD&I) program seeks to encourage new approaches to the management, analysis, and dissemination of biological knowledge for the benefit of both the scientific community and the broader public. The BD&I program is especially interested in the development of informatics tools and resources that have the potential to advance all fields of biology under the purview of the Directorate for Biological Sciences at the National Science Foundation.

THIS PROGRAM IS PART OF

[Research Resources Cluster](#)





Funding

[About Funding](#)

[A-Z Index of Funding Opportunities](#)

[Advanced Funding Search](#)

[Guide to Programs/Browse Funding Opportunities](#)

[Recent Funding Opportunities](#)

[Upcoming Due Dates](#)

Proposals and Awards

[Proposal and Award Policies and Procedures Guide](#)

Introduction

[Proposal Preparation and Submission](#)

- [Grant Proposal Guide](#)
- [Grants.gov Application Guide](#)

[Award Administration](#)

- [Award Administration Guide](#)

[Award Conditions](#)

[Other Types of Proposals](#)

[Merit Review](#)

[NSF Outreach](#)

[Policy Office Website](#)

Division of Computer and Network Systems

Computer Systems Research (CSR)

CONTACTS

Name	Email	Phone	Room
David Du	ddu@nsf.gov	(703) 292-8950	
Helen Gill	hgill@nsf.gov	(703) 292-8950	
Anita J. LaSalle	alasalles@nsf.gov	(703) 292-8950	

PROGRAM GUIDELINES

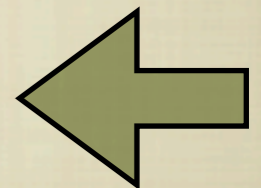
[08-538 Solicitation](#)

SYNOPSIS

Computer systems are being applied to increasingly demanding applications. The environments in which they function and the resources they manage are increasingly diverse, distributed and dynamic. While the time scales for control decisions are shrinking, the scale and complexity of the systems are increasing. Further, many of the assumptions behind today's most common computer systems no longer hold. As a result, these systems often fail in unpredictable ways, become compromised or perform poorly. Accordingly, the frontiers of computer systems research must be moved forward, and new bold research directions must be established to draw upon interdisciplinary research capabilities across science and engineering.

The Computer Systems Research (CSR) Program supports innovative research and education projects that:

- increase our understanding of large-scale and increasingly data-intensive computer systems and applications, through the creation of new knowledge needed to improve their design, use, behavior, and stability;
- capitalize on research opportunities provided by new technologies and new classes of systems;
- expand the capabilities of computer systems by developing highly innovative new ways to exploit existing technologies;
- lead to systems software that is quantifiably more reliable, easier to use, and/or more efficient; and





Funding



Funding

[About Funding](#)

[A-Z Index of Funding Opportunities](#)

[Advanced Funding Search](#)

[Guide to Programs/Browse Funding Opportunities](#)

[Recent Funding Opportunities](#)

[Upcoming Due Dates](#)

Proposals and Awards

[Proposal and Award Policies and Procedures Guide](#)

[Introduction](#)

[Proposal Preparation and Submission](#)

- [Grant Proposal Guide](#)

- [Grants.gov Application](#)

[Division of Computer and Network Systems](#)

CISE Computing Research Infrastructure (CRI)

CONTACTS

Name	Email	Phone	Room
Tatiana Korelsky	tkorelsk@nsf.gov	(703) 292-8930	1125
Joseph Urban	jurban@nsf.gov	(703) 292-8910	1115
Anita J. La Salle	alabelle@nsf.gov	(703) 292-5006	1175

PROGRAM GUIDELINES

[06-597 Solicitation](#)

SYNOPSIS

The Computer and Information Science and Engineering (CISE) Computing Research Infrastructure (CRI) program supports the acquisition, development, enhancement, and operation of research infrastructure that enables discovery, learning, and innovation in all computing fields supported by CISE. Supported infrastructure includes instrumentation needed by research or research and education projects, major experimental facilities for an entire department or for multi-institutional projects, and testbeds or data archives for an entire subfield of CISE researchers.

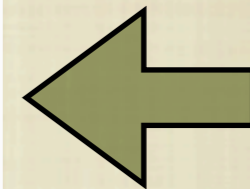
The CRI program will support a variety of infrastructure needs, such as general or specialized research equipment, technical support, and/or software. CRI will also support the development of infrastructure that can be used by others, such as data archives or libraries of software tools. The infrastructure must facilitate high-quality research and related education, and cannot be acquired or developed without funding resources beyond those available from individual research and education grants and the host institution.

The CRI program will make three kinds of awards.

- **Infrastructure Acquisition/Development.** These awards have budgets from \$50,000 and up to \$2,000,000.
- **Community Resource Development.** These awards have budgets from \$300,000 to \$2,000,000. Community Resource Development projects create a resource for an entire CISE research community, such as a testbed for evaluating research results or a large data resource for use by a research community (e.g., annotated speech data).
- **Planning.** These awards facilitate the preparation of a proposal for a medium or large Infrastructure Acquisition/Development or Community Resource Development grant. They have budgets up to \$50,000 for one institution or up to \$100,000 if more than one institution is involved.

The program supports projects in four size categories: large projects have budget requests from \$800,000 and up to \$2,000,000; medium projects have budget requests from \$300,000 and up to \$799,999; small projects have budgets from \$50,000 and up to \$299,999; Planning proposals may request budgets up to \$50,000 for one institution or \$100,000 for two or more institutions. Project sizes affect page limits, review processes, and eligibility.

The CRI program replaces and expands upon three previous CISE programs: Minority Institutional Infrastructure (MII), Research Infrastructure (RI), and Research Resources (RR). The most significant changes from the former programs are that CRI will support Community Resource Development grants in addition to Infrastructure Acquisition/Development grants.





Funding



Funding

[About Funding](#)

[A-Z Index of Funding Opportunities](#)

[Advanced Funding Search](#)

[Guide to Programs/Browse Funding Opportunities](#)

[Recent Funding Opportunities](#)

[Upcoming Due Dates](#)

Proposals and Awards

[Proposal and Award Policies and Procedures Guide](#)

Introduction

Proposal Preparation and Submission

- [Grant Proposal Guide](#)
- [Grants.gov Application Guide](#)

Award Administration

- [Award Administration Guide](#)

[Award Conditions](#)

[Other Types of Proposals](#)

Directorate for Biological Sciences

Planetary Biodiversity Inventories (PBI)

CONTACTS

Name	Email	Phone	Room
Patrick Herendeen	sbbi@nsf.gov	(703) 292-7184	635 N
Jerry Cook	jcook@nsf.gov	(703) 292-4821	635 N
W. Carl Taylor	sbbi@nsf.gov	(703) 292-7121	635 N

PROGRAM GUIDELINES

[06-500 Solicitation](#)

SYNOPSIS

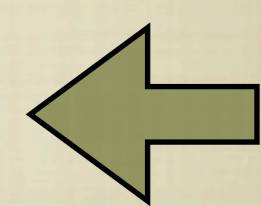
To accelerate the discovery and study of the world's biodiversity, proposals are invited from teams of investigators to conduct a worldwide, species-level systematic inventory of a major group of organisms. Each project should conduct fieldwork necessary to fill gaps in existing collections, produce descriptions, taxonomic revisions, web-searchable databases, and interactive keys (or other automated identification tools) for all new and known species in the targeted group, analyze their phylogenetic relationships, and establish predictive classifications for the group. Proposals may target any particular group of organisms, from terrestrial, fresh-water, or marine habitats, at any feasible level in the taxonomic hierarchy, but must be global in scope.

RELATED URLS

[Academy of Natural Sciences PBI website](#)

THIS PROGRAM IS PART OF

Additional Funding Opportunities for the DEB Community



EXAMPLES

- VECTORBASE - \$10M NIH/NIAID
- MALARIA TRANSMISSION CONSORTIUM - \$20M BMGF
- WIPER/DDDAS - \$500K, NSF
- SOURCEFORGE RESEARCH DATA ARCHIVE @ ND, \$360K, NSF
- VORTEXWINDS/ENGINEERING VIRTUAL ORGANIZATION, \$200K, NSF

HOW CAN CRC HELP?

- STORAGE
- BACKUP
- LARGE MEMORY MACHINES
- RELATIONAL DBMS / DBA SUPPORT
- STATISTICS / DATA MINING SUPPORT
- DATA VISUALIZATION
- COLLABORATION SUPPORT

