

From Data to Knowledge for Cyber-Enabled Discovery and Innovation

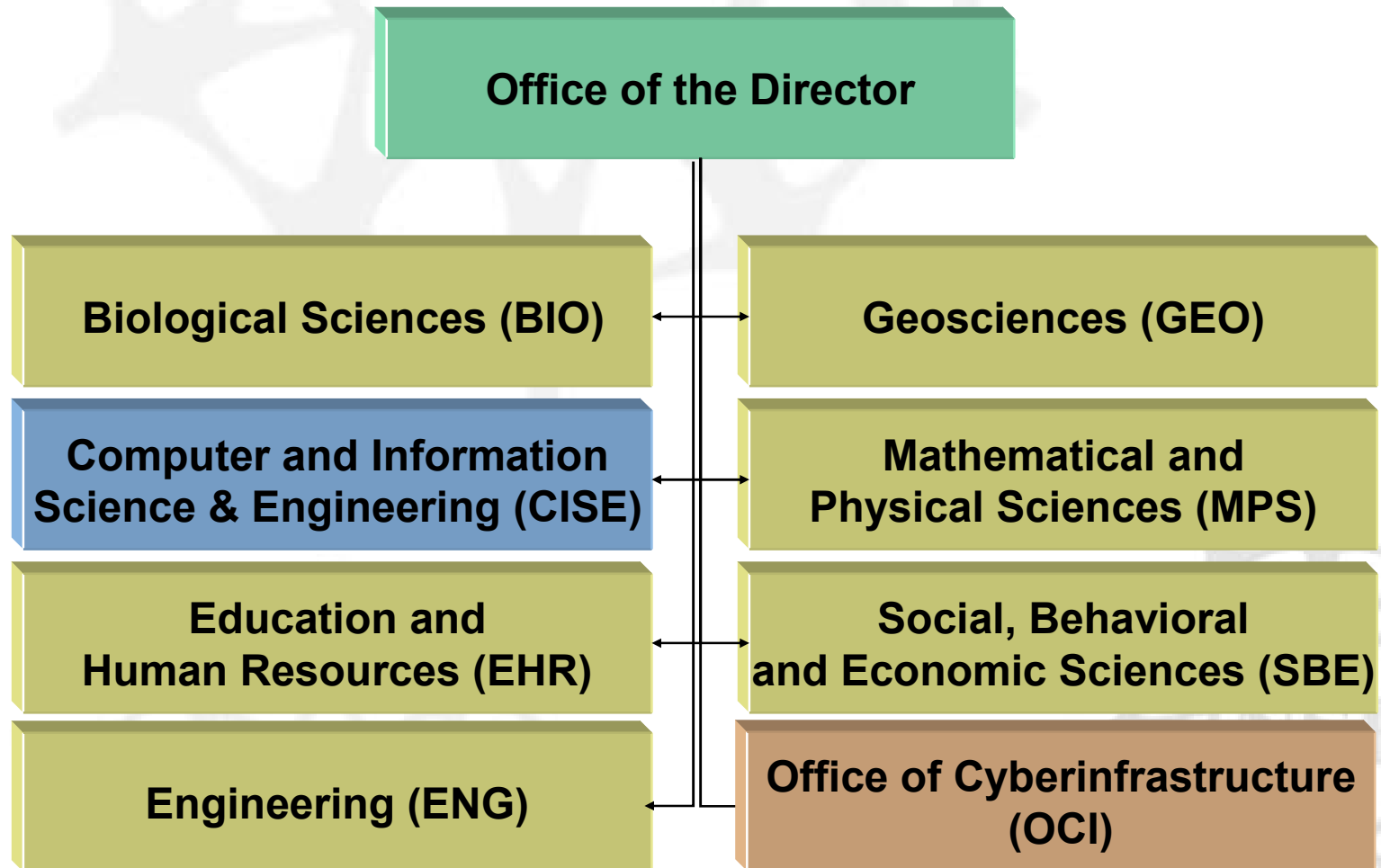
Haym Hirsh

National Science Foundation

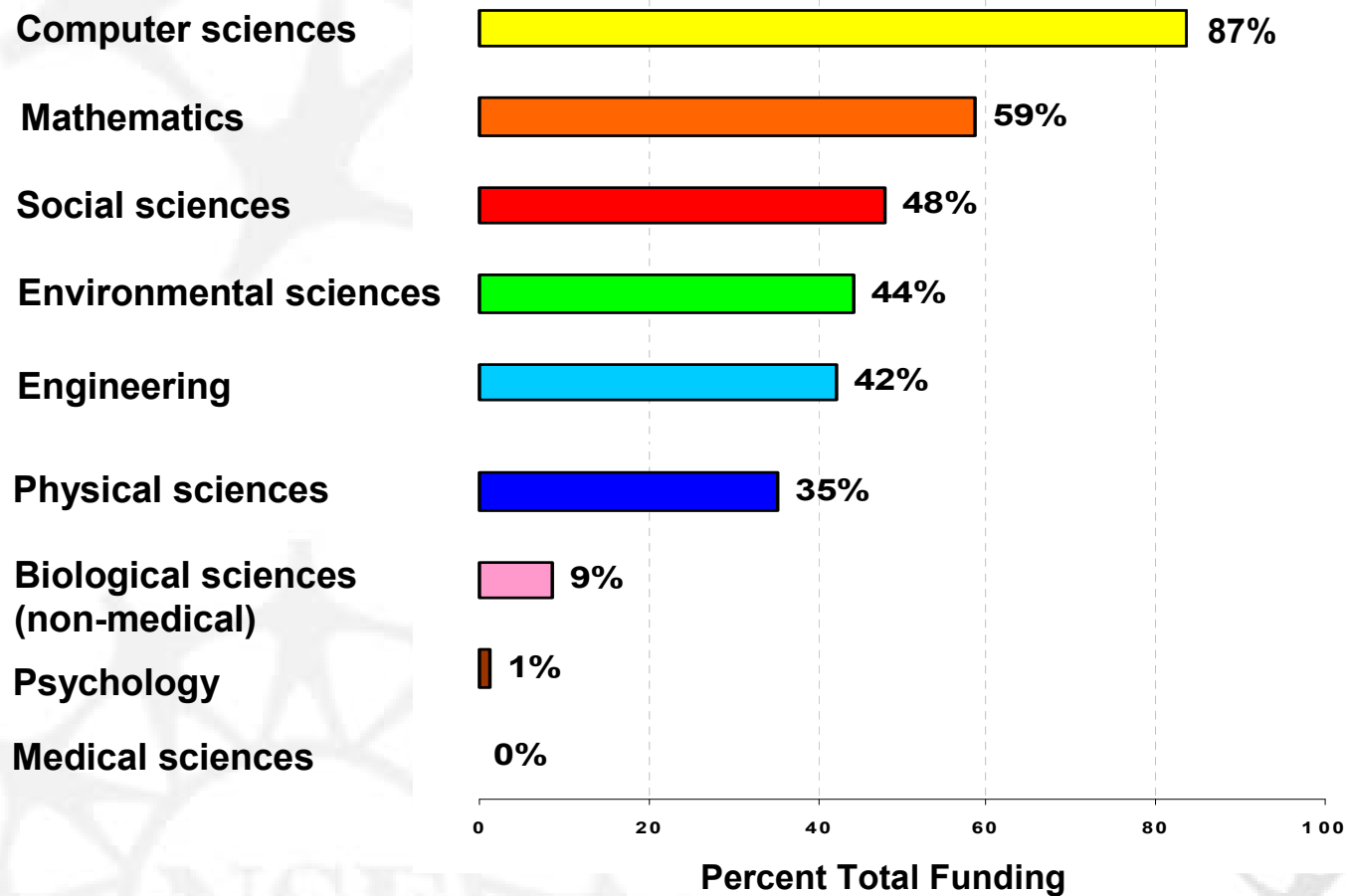
Rutgers University



National Science Foundation Overview



NSF's Share of Total Federal Support for Basic Research at Academic Institutions



Cyber-enabled Discovery and Innovation

- The advent of computing has changed how science and engineering is done
 - Computers as metaphors
 - Computers as “wet lab”
 - Computers as infrastructure

Cyber-enabled Discovery and Innovation

- Multi-disciplinary research advancing multiple areas of science or engineering via innovation in and/or innovative use of **computational thinking**
- Computational thinking refers to computational
 - Concepts
 - Methods
 - Models
 - Algorithms
 - Tools

Cyber-enabled Discovery and Innovation

- Planned five-year program seeking *{revolutionary, bold, radical, paradigm-changing, transformative}* science and engineering research outcomes made possible by innovations and advances in computational thinking
- Research that contributes to more than one area of science or engineering
- Projects that make straightforward use of existing computational concepts, methods, models, algorithms and tools to significantly advance only one discipline should be submitted to an appropriate program in that field instead of to CDI

Long-term Funding Plan for CDI

- All NSF directorates are participating in this activity (*subject to budget approval*)

FY 2008	FY 2009	FY 2010	FY2011	FY 2012
\$52M 50% pooled	\$100M	\$150M	\$200M	\$250M

Three CDI Themes

CDI seeks *transformative research* via innovations in and/or innovative use of computational thinking:

- **From Data to Knowledge:**
Derive new science and engineering knowledge from the growing abundance of heterogeneous digital data
- **Understanding Complexity in Natural, Built, and Social Systems:**
Insights on systems of many interacting elements throughout science and engineering
- **Virtual Organizations:**
Advance our ability to build and leverage cyberinfrastructure that links people and resources across institutional, geographic, cultural, and temporal boundaries, to change both the practice and the outcomes of science and engineering

Inter-related and overlapping!

Three CDI Themes

CDI seeks *transformative research* via innovations in and/or innovative use of computational thinking:

- **From Data to Knowledge:**

Derive new science and engineering knowledge from the growing abundance of heterogeneous digital data

- Can explore new approaches to data mining, data federation, knowledge extraction, knowledge representation, visualization, and other data technologies in demanding scientific and engineering applications

From Data to Knowledge

S&E Technological Drivers:

- Large quantities of data
- High rates of data
- High dimensionality of data
- Heterogeneity of data:
 - No central coordinator
 - Different generator and user
 - Multiple modalities, resolutions, scales
 - (AKA information integration, data mediation, data federation, etc.)
- Complex data objects:
 - Data from S&E applications often don't cleanly match existing data technologies

From Data to Knowledge

S&E Technological Drivers:

- Socio-political constraints on the data:
 - Privacy
 - Regulatory
 - Provenance of derived works
- Scientific interpretability
- Platform scalability

Types of Projects

- ◆ Three types of projects are defined:
Types I, II, and III
- ◆ Type III, center-scale efforts, will not be supported in the first year of CDI

Type I Projects

- ◆ Research and education efforts roughly comparable to that of up to two investigators with summer support, two graduate students, and their research needs (e.g., materials, supplies, travel), for a duration of three years



Type II projects

- Several intellectual leaders, multiple multidisciplinary objectives
- Significant education component
- Likely to be distributed collaborative projects with more extensive project coordination needs
- Greater effort than in Type I, and, for example, roughly comparable to that of up to three investigators with summer support, three graduate students, one or two other senior personnel (post-doctoral researchers, staff), and their research needs (e.g., materials, supplies, travel), for a duration of four years

Timeline

- Nov 30, 2007: Letters of Intent (required) due
- Jan 8, 2008: Preliminary Proposals due
- **Full proposals by invitation only!**
- Apr 29, 2008: Full proposals due
- July 2008: Award recommendations:

More Information on CDI:

➔ CDI URL:

- <http://www.nsf.gov/crssprgm/cdi>

➔ CDI Contacts:

- Sirin Tekinay (CISE)
- Tom Russell (MPS)
- Eduardo Misawa (ENG)

➔ cdi@nsf.gov

➔ 1-703-292-8080

Bottom Line

Have *significant* ideas that matter *significantly*
to collaborating scientists and/or engineers