A Vision for a National Cyberinfrastructure for Biodiversity Research and what NSF can do Enable it

Digital Data In Biodiversity
Ann Arbor
June 5-6, 2017

Peter McCartney

Division of Biological Infrastructure

National Science Foundation

What is NSF Thinking?

2005 - Cyberinfrastructure

2006 - Computational Thinking

2009 - Cyber-infrastructure Framework for 21st Century Science

2010 - Data Enabled Science

2011 - Computational and Data Enabled Science and Engineering

2012 - Big Data

2013 - Data Science

What does the community hear?
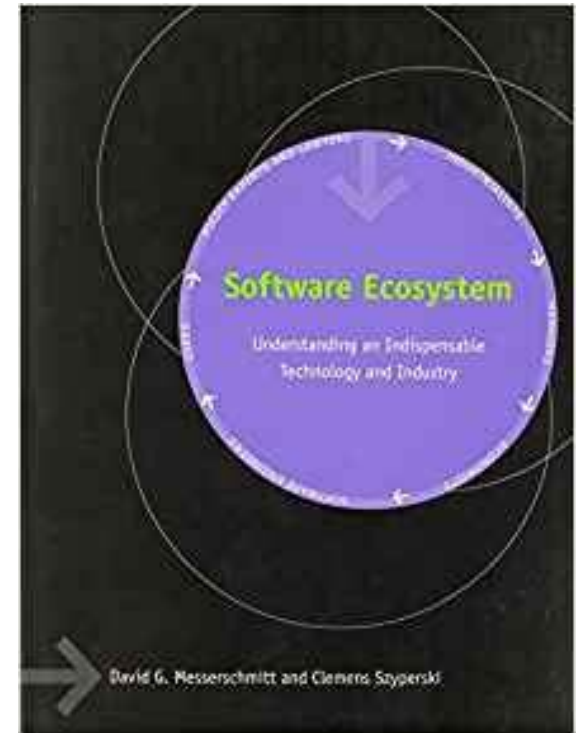
Program Title:

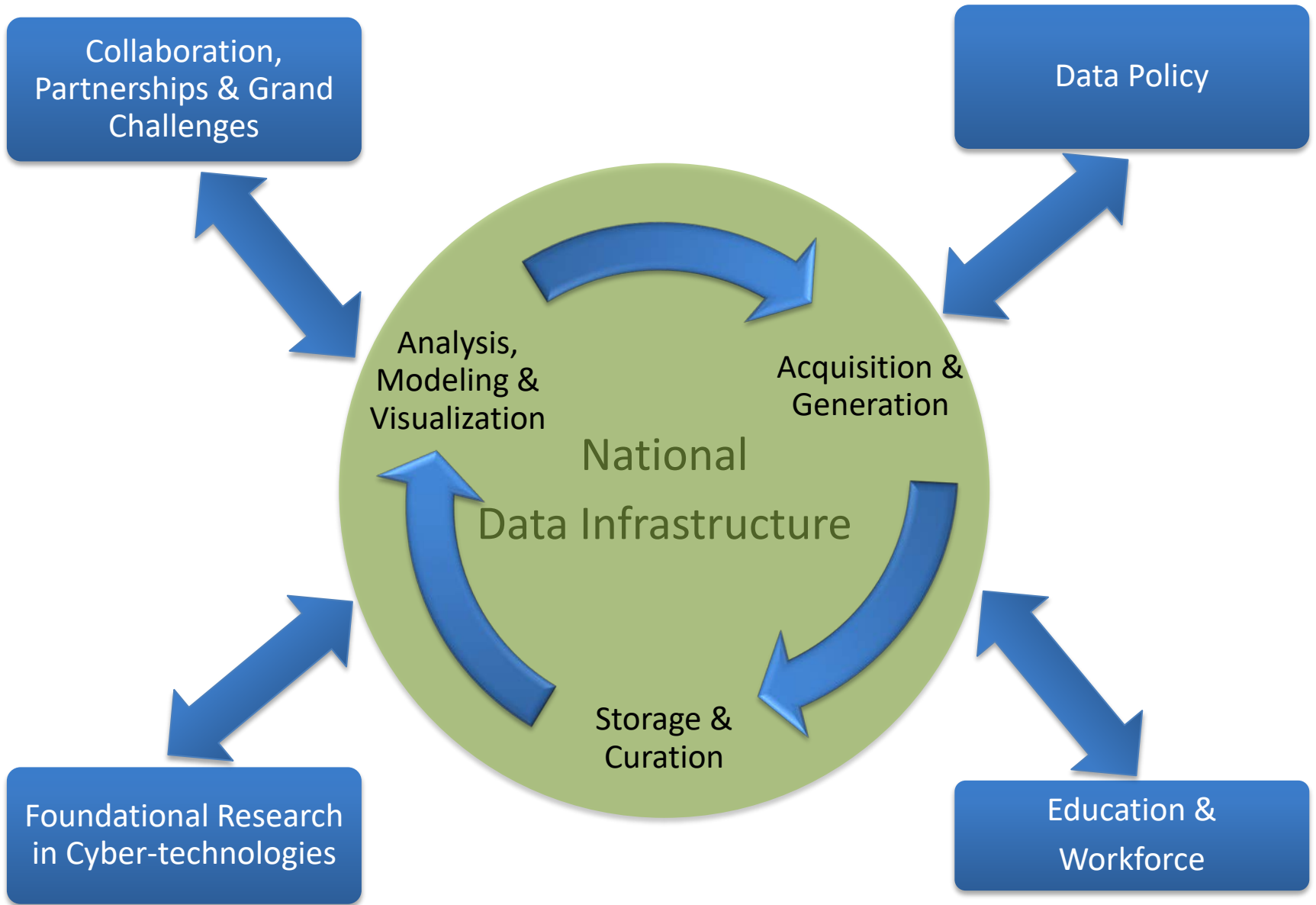Quisque rhoncus risus nec sapien dapibu (BIGMONEY)

Synopsis of Program:

Nulla non massa vitae augue condimentum suscipit. Suspendisse Funding Opportunity ultrices sapien augue, a dictum velit accumsan et. Sed vulputate augue sit amet mi hendrerit, et pharetra felis malesuada. Aliquam convallis commodo lorem ac suscipit. Proin adipiscing dapibus tellus, vel aliquam risus eleifend at. Nullam imperdiet magna ut arcu congue, quis ultricies sem consectetur. Curabitur Funding Opportunity viverra bibendum risus.

A National Data Infrastructure

- Not a single entity.
- Not a new program, nor the product of any one existing program, or one agency's programs.
- Not solely a technological problem

Collaboration, Partnerships & Grand Challenges

Data Policy

Analysis, Modeling & Visualization

Acquisition & Generation

National Data Infrastructure

Foundational Research in Cyber-technologies

Storage & Curation

Education & Workforce

DBI Programs

15-582: Advances in Biological Informatics

13-557: Collections In support of Biological Research

13-569: Advancing Digitization of Biodiversity Collections

13-561: Instrument Development for Biological Research

15-527: Research Coordination Networks

DEB, IOS, MCB programs

- Core program clusters

- 17-523: Dimensions of Biodiversity

- 16-521: MacroSystems Biology and Early NEON Science

- 16-614: Plant Genome Research Program

Programs outside BIO

- 17-526:Software Infrastructure for Sustained Innovation

- 17-500:Data Infrastructure Building Blocks

- 16-514:EarthCube

- 17-534:Critical Techniques, Technologies and Methodologies for Advancing Foundations and Applications of Big Data Sciences and Engineering
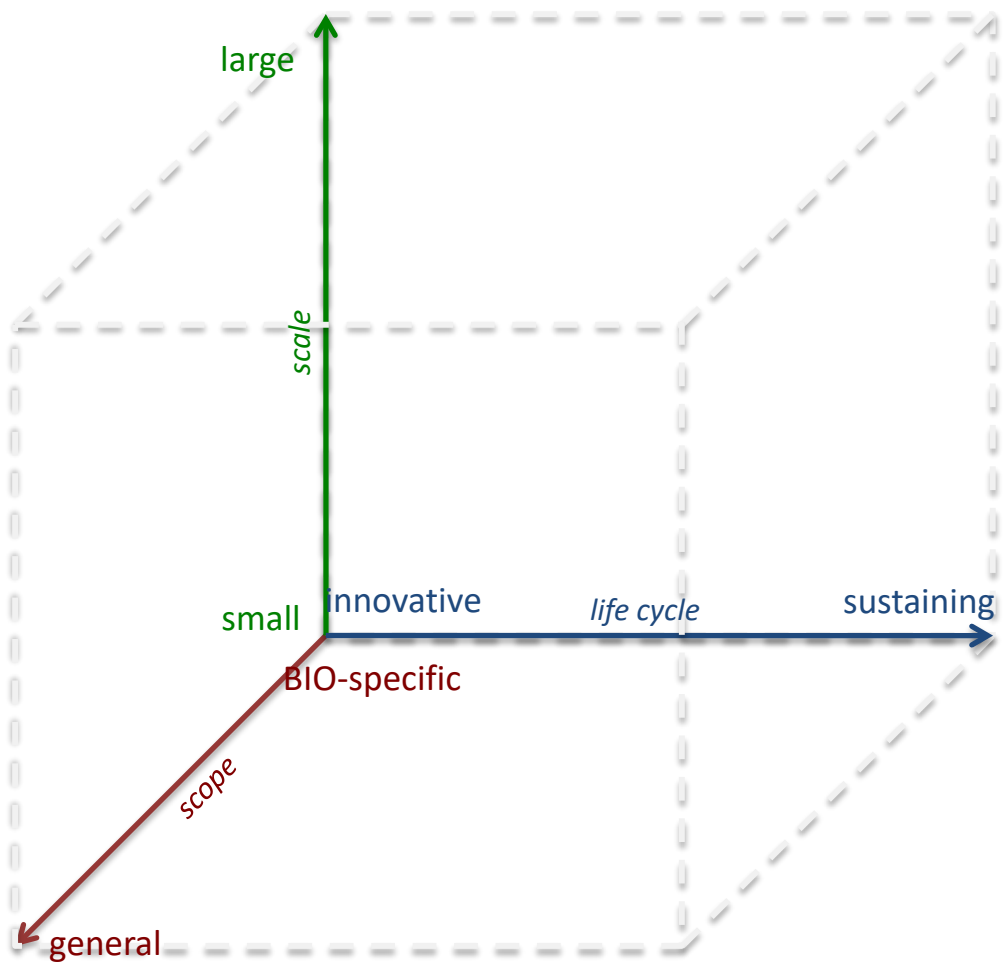
Building a National Infrastructure with a process designed for individual awards to institutions

- Emphasis on shared infrastructure

- Culture of competition rather than collaboration.

- Inherently a bottom-up process

- Tension between innovation vs stability

- NIH: Not Invented Here

NSF responses

- ADBC, Ideas Labs: Structured programs to stimulate collaboration among individual awards

- CI re-use: Rewarding developers with supplements & plus-ups for incorporating, re-purposing, or linking to CI developed independently.

- ABI: Partitioned into Innovation, Development, and Sustaining tracks to balance portfolio between novelty and usability.

- RCNs, PI meetings: Encouraging dialogs on synergy, standards, interoperability.

Supply side infrastructure

Noun

1. The theory that investment in infrastructure will have intrinsic value that will accelerate the pace of discovery in science.

Demand side considerations

- All "infrastructure" spending outside of MREFC and MRI comes out of Research and Related Activities account.
- Funding decisions are as much about priority as they are merit.
- Risk (or at least perception) of growing a specialized infrastructure PI community with its own agenda.

You are not the customer

- Characterizing and engaging user communities.

- Encourage users to identify use of resources.

- Aligning assessment metrics with science-driven value proposition.

- Documenting downstream impacts.

- DBI: Engaging POs from other divisions in our decision process.

Sustainability

Transition from investment revenue model to a transactional one.

Diverse revenue portfolios.

Revenue scales with use.

Community ownership of the problem.

Thank you.

Peter McCartney
pmccartn@nsf.gov
Division of Biological
Infrastructure