

**EAGER: Collaborative Research: Supporting  
Public Access to Supplemental Scholarly  
Products Generated from Grant Funded Research**

**Helen Berman, Kerstin Lehnert, Victoria Stodden**

**Research Data Alliance**

Denver, CO

September 17, 2016

# NSF requires..

- NSF DATA SHARING POLICY
- Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing.

# NSF DMP Requirements

- Proposals submitted or due on or after January 18, 2011, must include a supplementary document of no more than two pages labeled “Data Management Plan”. This supplementary document should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results.

## REQUIREMENTS BY DIRECTORATE, OFFICE, DIVISION, PROGRAM, OR OTHER NSF UNIT

---

Links to data management requirements and plans relevant to specific Directorates, Offices, Divisions, Programs, or other NSF units, are provided below. If guidance specific to the program is not provided, then the requirements established in [Grant Proposal Guide, Chapter II.C.2.j](#) apply.

Please note that if a specific program solicitation provides guidance on preparation of data management plans, such guidance must be followed.

- Biological Sciences Directorate (BIO)
  - [Directorate-wide Guidance](#)
  
- Computer & Information Sciences & Engineering (CISE)
  - [Directorate-wide Guidance](#)
  
- Education & Human Resources Directorate (EHR)
  - [Directorate-wide Guidance](#)
  
- Engineering Directorate (ENG)
  - [Directorate-wide Guidance](#)
  
- Geosciences Directorate (GEO)
  - [Directorate-wide Guidance](#)
  
- Mathematical and Physical Sciences Directorate (MPS)
  - [Division of Astronomical Sciences](#)
  - [Division of Chemistry](#)
  - [Division of Materials Research](#)
  - [Division of Mathematical Sciences](#)
  - [Division of Physics](#)
  
- Social, Behavioral and Economic Sciences Directorate (SBE)
  - [Directorate-wide Guidance](#)

# GEO Data Management Plan (DMP) Analysis

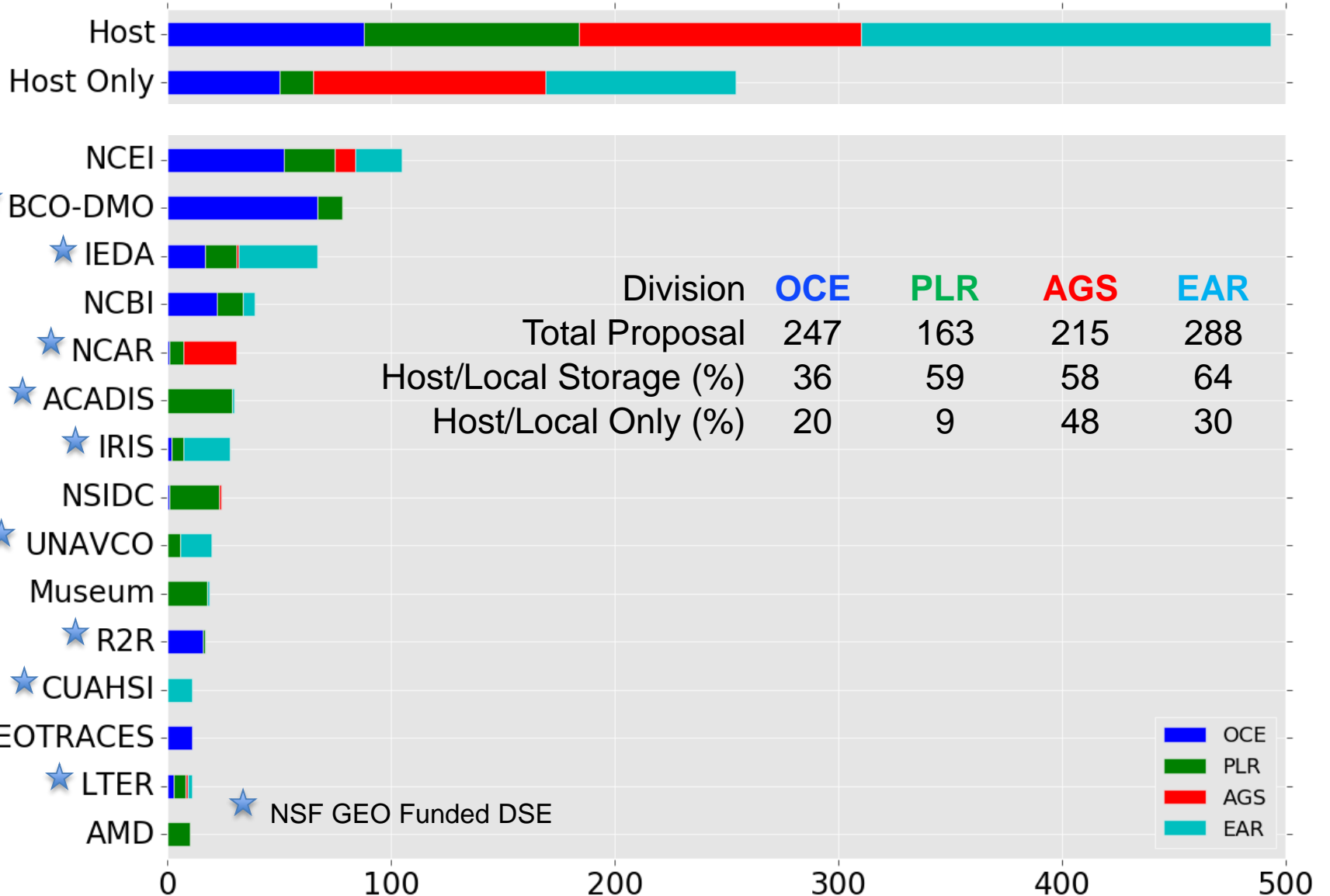
- In 2011 data management plans became a required part of an NSF proposal
- This study was conducted to take a **snapshot** of the current GEO data management landscape.
- Tasked with analyzing one year of funded projects in each division:

OCE	Aaron Rosenberg
PLR	Colene Haffke
AGS	Ed Bensman
EAR	Amy Chen

# GEO Data Management Plan (DMP) Analysis

- Identify Data Service Entities (DSE)
  - Data Routing and Metadata Services
  - Physical Sample Storage and Sample Unique Identification
  - Long-term Data Storage
  - Real-time Data Services
  
  - Host Institutions

# Top 15 Data Service Entities



Division	OCE	PLR	AGS	EAR
Total Proposal	247	163	215	288
Host/Local Storage (%)	36	59	58	64
Host/Local Only (%)	20	9	48	30

Number of Mentions in Analyzed DMPs

## Top 10 DSE for each Division

OCE (N=241)		PLR (N=163)		AGS (N=215)		EAR (N=283)	
DSE	#	DSE	#	DSE	#	DSE	#
BCO-DMO	67	ACADIS	29	NCAR	24	IEDA	39
NCEI	52	NCEI	23	NCEI	9	IRIS	21
NCBI	22	NSIDC	22	NASA	6	NCEI	20
R2R	16	Museum	18	GEM	3	UNAVCO	14
IEDA	16	IEDA	14	ICARTT	3	CUAHSI	11
GEOTRACES	11	NCBI	12	REU Website	3	CSDMS	6
DataONE	4	BCO-DMO	11	ORNL	2	SERC	5
PANGAEA	3	AMD	10	AMS DB	2	USGS	5
OBIS	3	GCMD	7	IRB	2	EarthCube	5
LTER	3	NCAR	6	NAW	2	LacCore/Github	4

GEO Funded DSE

Unique DSE:

OCE	PLR	AGS	EAR
50	68	52	74



# Characteristics of DSE

	Data Routing and Metadata	Physical Samples and Sample ID	Long Term Digital Storage	Real-time Data Services
BCO-DMO	X			
R2R	X			
IEDA	X	X	X	
GEOTRACES	X			
ACADIS	X		X	
NCAR	X		X	X
IRIS			X	X
UNAVCO			X	X
CUAHSI	X		X	
CSDMS	X			

- EarthCube is mentioned but it is unclear how investigators are using this service.
- Some DSE provide other services (visualization tools, templates, data collection) but in this study we looked specifically for use as a place for data storage.

# FINDINGS

- Host institution storage is a popular data management method
- Data management needs and practices are diverse
- 10 of the top 15 data storage entities are funded by GEO
- DMP templates provide advantages and disadvantages

# DIVISIONAL DATA POLICIES

- Revise GEO DMP Guidance linked to GPG
  - <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>
- Review and Revise GEO Data Policies
  - PLR – Revised Data Policy January 2016
  - AGS, EAR, OCE divisional reviews in 6 months, with recommendations to GEO Management
- Provide guidance for PIs, Reviewers, and GEO Staff

# NEED TO BALANCE MANY FACTORS

- The value to advancing geosciences in the future by the easy sharing, discovery and access of data and products;
- The balance between flexibility and diversity of research practice with creating GEO-wide principles;
- Guiding and evaluating effective Data Management Plans while ensuring that PI and community judgement is respected;
- Consideration for PI and PO workload and true cost to science in evaluating the burden of any policy.

# NSF GEO AC Study

- Had to be carried out by hand..

# DMP II

- Machine readable
- Flexible
- Captures descriptions of artifacts researchers use/produce
- Actionable



# Research

- Understand products in the submitted IEDA DMPs for three NSF Directorates.
- Provide comprehensive update to the IEDA DMP Tool.
- Pilot and review, develop Best Practices for researchers, repositories, and funders.