



IEDA
iedadata.org

Supporting Data Stewardship Throughout the Data Life Cycle in the Solid Earth Sciences

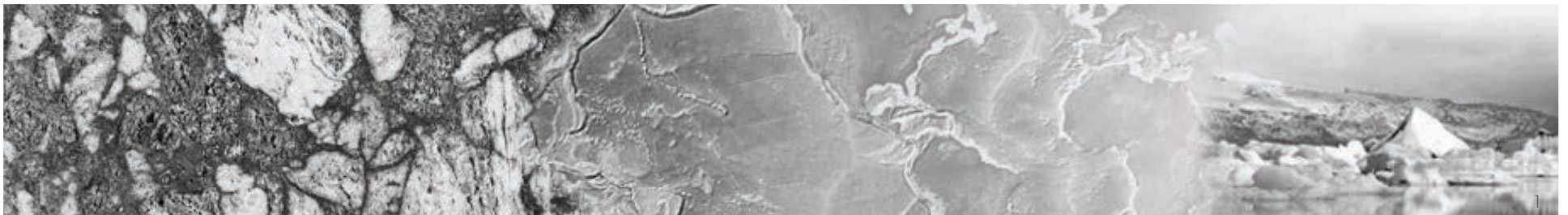
Vicki L. Ferrini, Kerstin A. Lehnert,
Suzanne M. Carbotte, and Leslie Hsu

Lamont-Doherty Earth Observatory

What is IEDA??

A community-based data facility funded by NSF to support, sustain, and advance the geosciences by providing data services for observational solid earth data from the Ocean, Earth, and Polar Sciences.

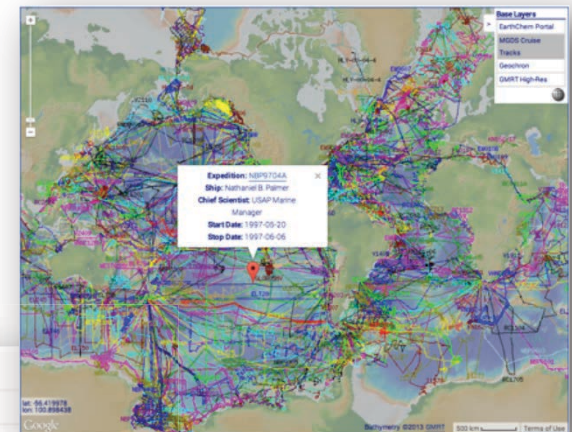
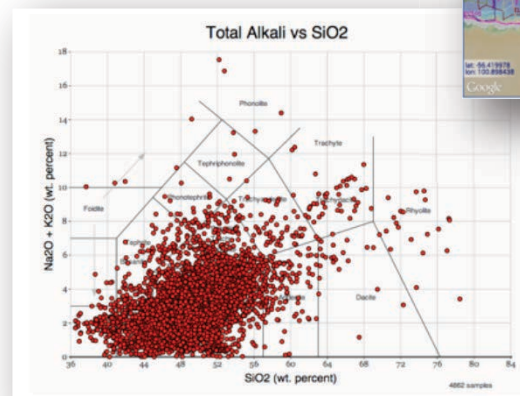
- Domain-specific Data Libraries
- Data Syntheses
- Visualization & Analysis Tools
- Data & Metadata Registries
- Data Publication
- Investigator Support Services



Data Curated in IEDA Systems

- Sensor-based Data
 - Marine geophysical data (e.g. bathymetry, sidescan, subbottom)
 - Complementary datasets
 - Navigation, bottom photos
 - Technical reports

- Sample-based Data
 - Sample metadata
 - Analytical results
 - Technical reports



Domain-specific Data Curation

- Familiarity with content, acquisition, processing, & use of data
- Ability to evolve services in response to both:
 - User-community needs
 - Technology
- Construct & curate high-quality syntheses
- Curate rich metadata
 - Enhanced discoverability & usability
 - “Fitness for re-use”

**Collect &
Assure**

Plan

**Document
& Preserve**

**Data Life
Cycle**

**Integrate
& Share**

Analyze

**Document
& Preserve**





Data Life Cycle: Plan

- Data Management Plan Tool
 - Facilitate assembly
 - Inform Investigators
 - Inform down-stream repositories
 - Promote dialogue
- Data Acquisition Plan
 - Metadata & data templates
 - Promote & facilitate *contemporaneous* documentation

Data Management Plan Tool

Use this form to generate a data management plan PDF for inclusion in your NSF proposals. Please note that data management plans are limited to two pages in length. For more information, see the [NSF Data Management & Sharing FAQ](#) (* indicates a required field)

Proposal Information

Lead PI:

Lead Institution:

Project Title:

Collaborators:

NSF Division submitted to:

NSF Program Info:

Proposal Submission Deadline:

Data Acquisition/Processing Summary

Overview:

Data Description:

Includes field work? No

Description of existing data/samples to be used:

Data Analysis Summary:

Proposed Data Products

—Data Product #1—

Type of Data Product: Observational Analytical

Responsible Investigator(s):

Data Product Description:

Intended Repository:

Timeline for data release:

[Add Another Expected Data Product](#)

Data Management Plan

Primary Investigator: John Morton
Institution: Lamont-Doherty Earth Observatory of Columbia University
Project: Inactivation of the Passive Margin of Eastern Laurentia
NSF Division: OCE - Solicitation Info: Marine Geology and Geophysics - **Submission Date:** 01/16/2013

Overview: Our project will use active source seismology on the Marcus G. Langbein to image the oceanic crust on the continental shelf of the Eastern U.S. after the Dec. 21, 2012 earthquake.

Data description: The proposed research will result in several new seismic transects along and across the new active margin.

Data analysis summary: CMP stacking and migration will be performed using the open source seismic utilities package Seismic Unix. Gravity data will be processed using the open source R2R_Gravity data processing tools. Multibeam bathymetry will be processed using MBSystems.

Includes field work? Yes
Description of field work: Active source seismology, multibeam bathymetry, and gravity (BGM 3) data will be collected.

Expected data product #1
Data type: Observational, Analytical
Responsible Investigator: John Morton
Product description: .egg files from seismic transects.
Intended repository: IRIS
Timeline for data release: Immediate Release

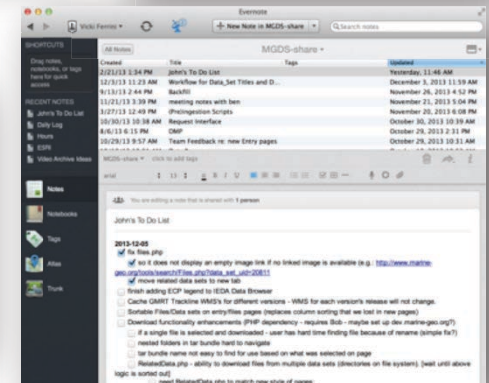
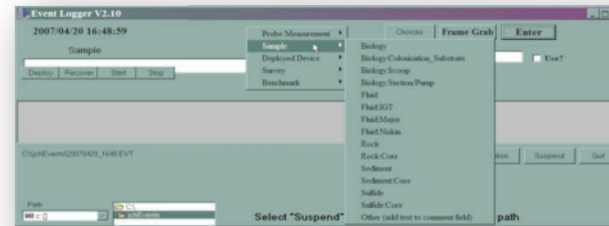
Expected data product #2
Data type: Observational
Responsible Investigator: John Morton
Product description: Processed free-air anomaly data in MCD07-T format
Intended repository: NGDC
Timeline for data release: Immediate Release

Expected data product #3
Data type: Observational
Responsible Investigator: Nick L. Ferrari
Product description: Multibeam bathymetry data
Intended repository: MGDS
Timeline for data release: Immediate Release

Page 1 Created: 2010-11-09 10:10:07 EST

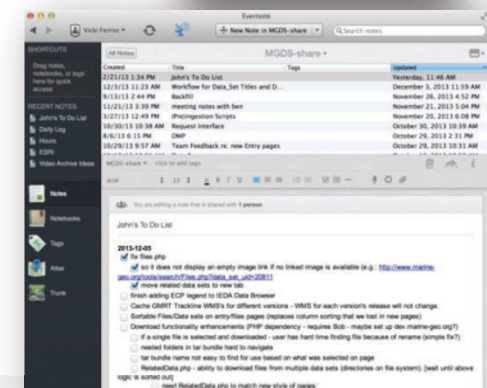
Data Life Cycle: Collect & Assure

- Promote Best Practices
 - What to document
 - How to document
- Tools and workflows to facilitate digital documentation
 - Metadata & Data Templates



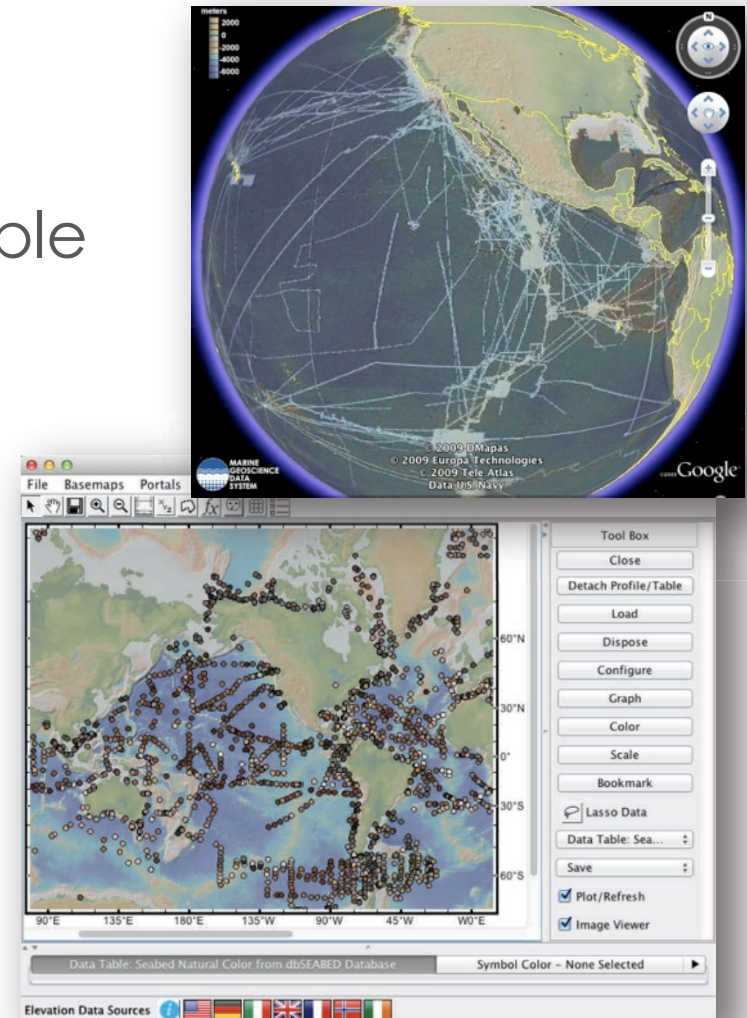
Data Life Cycle: Document & Preserve

- Document & capture data & metadata as soon as it is available
 - Simple interfaces & guidelines
- Sample metadata registry
- Link to complementary data & metadata



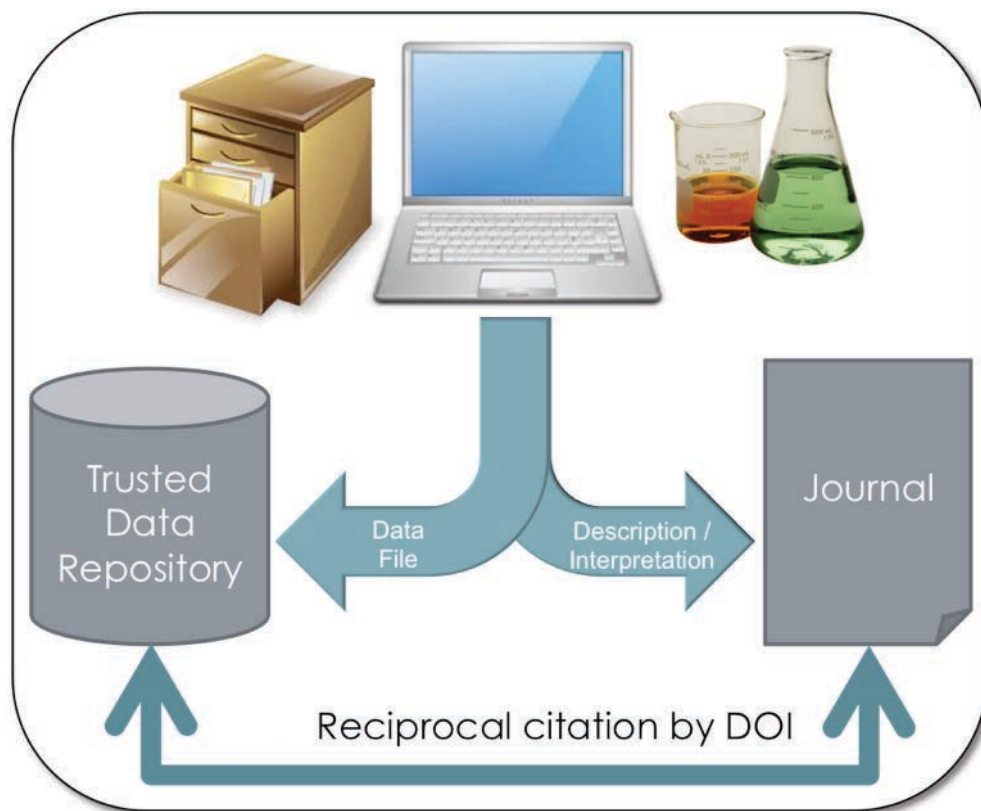
Data Life Cycle: Analyze

- Tools to:
 - Support domain specialists
 - Make specialist data accessible to non-specialist users
 - Integrate & visualize
- Access to Data Syntheses
- Access to complementary data & resources



Data Life Cycle: Integrate & Share


- Advise on what to preserve & how
 - Data supporting pubs
 - Data of value
- Facilitate data prep.
 - Metadata requirements
 - Templates
 - Format guidelines



Data Life Cycle: Document & Preserve

Develop simple workflows, interfaces & templates to capture sufficient information for:

- Long-term curation & access
 - Inclusion in syntheses
- Links to scientific publications
- Data Publication
- Data use, discovery & re-use
- Attribution & collaboration
- Data Download Stats
- Data Compliance Reporting



The screenshot shows the IEDA website interface. The top navigation bar includes links for HOME, ABOUT, CONTRIBUTE DATA, SERVICES, DATA COLLECTIONS, COMMUNITY, NEWS, HELP, and CONTACT US. The main content area is titled 'Contribute Data' and includes a list of data types such as 'Sample-based Data' (Analytical geochemistry datasets, Geochemical or petrological syntheses, etc.) and 'Sensor-based Data' (Derived Geophysical Data, Photos and images, etc.).

Overlaid on the bottom right is a search result from ScienceDirect for the article 'Evidence of mass failure in the Hess Deep: resolutional bathymetry data' from the journal 'Marine Geology'. The article is by Vicki Lynn Ferrini, Donna J. Shillington, Kathryn G. Damon, A.H. Teagle, Antony Morris, Pierre W. Czerny, and Tomlinagal, the JC21 Scientific Party. The IEDA MGDS logo is visible in the search results, indicating that data for this article is available in the IEDA data system.



Data Compliance Reporting Tool

- Tool for demonstrating compliance:
 - Award-based
 - Informed by DMP
- Report includes:
 - Data Inventory
 - Release Status
 - Links to data
 - Save as PDF

Data Compliance Reporting Tool (v1.0)

Instructions: The IEDA Data Compliance Report Tool enables the easy preparation of reports to demonstrate compliance with NSF Data Policies. Enter a NSF award and this service will provide a list of related data sets and their release status. Note that data will only be returned for awards that are currently cataloged within IEDA, and data sets returned are based on data and metadata received to date. Please [contact us](#) with comments or questions, or to [submit additional data or metadata](#).

Enter NSF Award:

[Locate a NSF Award through the Fastlane Award Search](#)

URL to this dynamic report: <http://www.iedadata.org/compliance/report>

NSF Award Info
Award Title: [Collaborative Research: Crustal Accretion and Mantle Center](#)
Investigator(s): Spahr Webb, William Menke
Award ID: 0426369

Sensor Data Linked to Award

Data System	Expedition/Compilation	Data Type(s)	# of Data Sets	Release Status
MGDS	MGL0903	Hydrographic	1	Released
MGDS	MGL0903	Geophysical	1	Released
MGDS	MGL0903	Geophysical	2	Released
MGDS	MGL0903	Navigation:Primary	3	Released
MGDS	MGL0903	Hydrographic	2	Released
MGDS	RR0915	Navigation:Primary	2	Released
MGDS	MGL0903	Seismic:Active-Subbottom	1	Released
MGDS	MGL0903	Bathymetry:Swath	1	Released
NGDC	RR0915	Bathymetry:Swath	1	Released
NGDC	RR0915	Current:Measurement	1	Released
NGDC	RR0915	Geophysical	1	Released
NGDC	RR0915	Hydrographic	1	Released
R2R	MGL0903	Underway Data and Final Navigation	1	Released
R2R	RR0915	Underway Data and Final Navigation	1	Released

Additional Data Sets: Submission Status Unknown

Expedition/Compilation	Data Type(s)	Instrument Info	Investigator(s)	Release Status
RR0915	Gravity:Field	Roger Revelle Gravimeter	Dunn, Robert	Not Supplied
RR0915	Magnetic:Field	Roger Revelle Magnetometer	Dunn, Robert	Not Supplied
RR0915	Backscatter:Acoustic Bathymetry:Swath Sidescan	Roger Revelle Sonar:Multibeam	Dunn, Robert	Not Supplied
MGL0903	Seismic:WideAngle-OBS	Marcus G. Langbein Seismic:OBS	Dunn, Robert	Not Supplied
RR0915	Seismic:Passive-OBS	Roger Revelle Seismic:OBS	Wiens, Douglas	Not Supplied

INTEGRATED EARTH DATA APPLICATIONS
iedadata.org

Data Compliance Report

URL to this report: http://www.iedadata.org/compliance/report?award_id=0426369

NSF Award Info
Award Title: [Collaborative Research: Crustal Accretion and Mantle Processes Along the Subduction-Influenced Eastern Lau Spreading Center](#)
Investigator(s): Spahr Webb, William Menke
Award ID: 0426369

Sensor Data Linked to Award

Data System	Expedition/Compilation	Data Type(s)	Data Sets	Instrument Info	Investigators	Release Info	Citations
MGDS	MGL0903	Hydrographic	1	Marcus G. Langbein Thermosalinograph	Dunn, Robert	Released	Not Supplied
MGDS	MGL0903	Geophysical	1	Marcus G. Langbein Gravimeter	Dunn, Robert	Released	Not Supplied
MGDS	MGL0903	Geophysical	2	Marcus G. Langbein Magnetometer	Dunn, Robert	Released	Not Supplied
MGDS	MGL0903	Navigation:Primary	3	Marcus G. Langbein Navigation	Dunn, Robert	Released	Not Supplied
MGDS	MGL0903	Hydrographic	2	Marcus G. Langbein Probe: Expendable: XBT	Dunn, Robert	Released	Not Supplied
MGDS	RR0915	Navigation:Primary	2	Roger Revelle Navigation	Dunn, Robert	Released	Not Supplied
MGDS	MGL0903	Seismic:Active-Subbottom	1	Marcus G. Langbein Seismic:Subbottom	Dunn, Robert	Released	Not Supplied
MGDS	MGL0903	Bathymetry:Swath	1	Marcus G. Langbein Sonar: Multibeam	Dunn, Robert	Released	Not Supplied
NGDC	RR0915	Bathymetry:Swath	1	Roger Revelle Sonar: Multibeam	Wiens, Douglas	Released	Not Supplied
NGDC	RR0915	Current:Measurement	1	Roger Revelle Sonar: ADCP	Wiens, Douglas	Released	Not Supplied
NGDC	RR0915	Geophysical	1	Roger Revelle Gravimeter	Wiens, Douglas	Released	Not Supplied
NGDC	RR0915	Hydrographic	1	Roger Revelle Probe: Expendable: XBT	Wiens, Douglas	Released	Not Supplied
R2R	MGL0903	Underway Data and Final Navigation	1	Roger Revelle Underway Data and Final Navigation	Wiens, Douglas	Released	Not Supplied
R2R	RR0915	Underway Data and Final Navigation	1	Roger Revelle Underway Data and Final Navigation	Wiens, Douglas	Released	Not Supplied

Additional Data Sets: Submission Status Unknown

Expedition/Compilation	Data Type(s)	Instrument Info	Investigator(s)	Citations
RR0915	Gravity:Field	Roger Revelle Gravimeter	Dunn, Robert	Not Supplied
RR0915	Magnetic:Field	Roger Revelle Magnetometer	Dunn, Robert	Not Supplied
RR0915	Backscatter:Acoustic Bathymetry:Swath Sidescan	Roger Revelle Sonar:Multibeam	Dunn, Robert	Not Supplied

Page 1 Created: 2010-11-09 12:40:55 EST

<http://www.iedadata.org/compliance/>

Summary



- Domain-specific repositories have an important role to play throughout the data life cycle:
 - Increase awareness of resources & tools
 - Lessen “burden” of data management
 - Provide guidelines & tools for data curation, discovery & access
 - Respond to evolving needs of user community
- Greatest success when repository team members are integrated members of user community



Opportunities & Challenges

- Data management training
 - Minor modifications throughout life-cycle will yield great results
 - Balance between data management & science training
 - Barriers to adoption must be low
- Evolving technologies can be leveraged
 - Pace of technology vs. pace of adoption
- Evolving significance of data publication
 - Minimal metadata approaches are easy but less useful