DeSSC New-User Program: Data Management Overview Vicki Ferrini & Tina Haskins

NDSF DEEP SUBMERGENCE

Meet the Data Team

Vicki Ferrini, PhD NDSF Associate Director for Data Syster Strategy & Director of MGDS

NDSF submergence



Tina Haskins

NDSF Associate Director for Da and Science Operations



Hayley Drennon Data Manager Marine Geoscience Data System



Scott McCue Lead Jason Data Engineer NDSF



Catie Graver SSSG & Data Engineer, NDSF



Andrew Goodwillie, PhD, Senior Data Manager Marine Geoscience Data System



Joe Garcia Lead Alvin Data Engineer NDSF

Open Data Access

- Driven by:
 - Funding agency requirements
 - Journal requirements
 - Acquisition costs
- Enables New Opportunities:
 - Spatial & temporal change
 - Scientific reproducibility
 - Data synthesis

N) S E DEEP SUBMERGENCE

• New possibilities for analysis

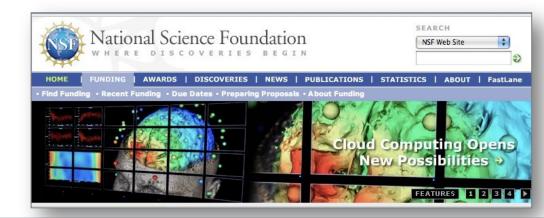
JANUARY 11, 2023

FACT SHEET: Biden-Harris Administration Announces New Actions to Advance Open and Equitable Research

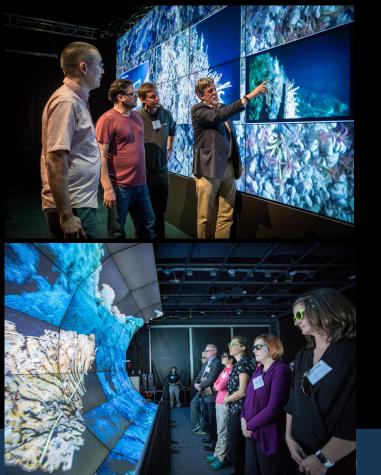
BRIEFING ROOM > PRESS RELEASES

OSTP launches Year of Open Science to advance national open science policies across the federal government in 2023





Beyond Analysis: *Experiencing* Data





Historic Scientific Workflow



Data Processing & Interpretation

Data Acquisition

NDSF DEEP SUBMERGENCE

DODD NUMBER



ABSTRACT

Using the near-bottom ARGO imaging system, we visually and acoustically surveyed the narrow (< 200 m wide) axial zone of the fast-spreading East Pacific Rise (EPR) along 83 km of its length (9°09'-54'N), discovered the Venture Hydrothermal Fields, and systematically mapped the distribution of hundreds of hydrothermal features relative to other fine-scale volcanic and tectonic features of the ridge crest. The survey encompasses most of a 2nd order ridge segment and includes at least ten 4th order (5-15 km) segments defined by bends or small lateral offsets of the ridge crest or axis (Devals). 4th order segmentation of the ridge crest is clearly expressed in the high-resolution ARGO data by the fine-scale behavior of the ridee axis and by changes in the characteristics of the axial zone (axial lava age, extent of fissuring, axial morphology and the roge and and by changes in the characteristics of the data bulk (Achi ava age, except of insuing, achi inorporting) and structure, etc.) across segment boundaries. The distribution and along-strike variability of hydrothermal features corresponds closely to the morphotectonic/structural segmentation of the ridge. On the 2nd order scale, we find that high T hydrothermal cervicy to the independent of the axial magnet chamber (AMC) reflector to depths <1.7 km beneath the ridge activity correlates with: (1) shallowing of the axial magnet chamber (AMC) reflector to depths <1.7 km beneath the ridge axis; and, (2) with the presence of a well-developed axial summit caldera (ASC). Previous work refers to this feature as in axial summit graben (ASG); however, the extent of volcanic collapse along the ASG revealed by the ARGO survey adds to evidence that on fast-spreading ridges it is an elongate volcanic caldera rather than a tectonic graben, and supports the citation of "axial summing addern" as a more accurate descriptor. All but 1 of the 45 active high T vent features identified with ARGO are located within 20 m of the margins of the ASC. Despite the significant extent of our track coverage outside the ASC, no important signs of venting were seen beyond the axial zone. On the 4th order scale, the abundance and distribution of hydrothermal features changes across 4th order segment boundaries. We find that high T vents are most abundant where: (1) the ASC is very narrow (40-70 m), (2) the AMC reflector is most shallow (< 1.55 km beneath the axia



PRIVATE PROPERTY NO PUBLIC ACCESS





Modern Scientific Workflow



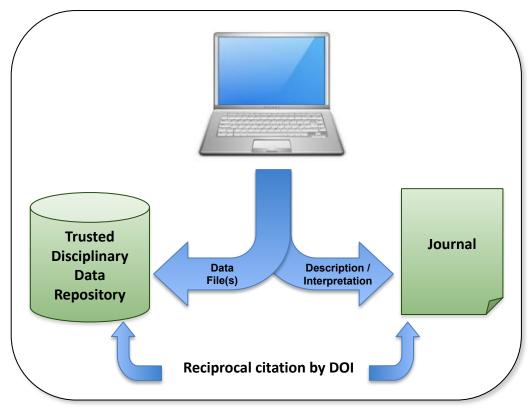


Data Acquisition

NDSF ANTIONAL DEEP SUBMERGENCE

DOCT NUMBER

"Best Practice"





What's in it for you?

- Scientific integrity & reproducibility
- Collaboration
- Increase research impact
- Data citation
- Compliance with data policies
- Outreach & Engagement
- Preserve data for your own future use

ELSEVIER	Marine Geology Volume 339, 1 May 2013, Pages 13-21		
Evidence of mass failure in the Hess Deep Rift			
from multi-resolutional bathymetry data		data Research data for this	article
Christopher J. MacLeod ^c 🖂 ,	onna J. Shillington ^a 쯔, Kathryn Gillis ^{, b} ठ, Damon A.H. Teagle ^d ठ, Antony Morris ^{, c} ठ, when Hurst ^f ठ, Masako Tominaga ^{, g} , the JC21 S are 카 Cite	Sonar Data from the	
https://doi.org/10.1016/j.margeo.	2013.03.006 🤊	Get rights and content a	
[



"The coolest thing to do with your data will be thought of by someone else." Rufus Pollock

al LORD MARCH 199

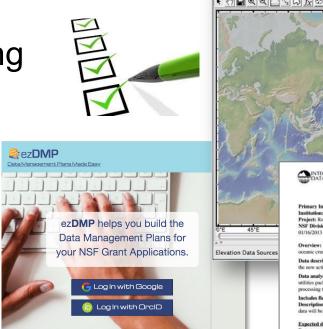
Cambridge University and Open Knowledge Foundation

Data Management Planning

- Concept/Proposal Development
 - Are existing data available?
- Data Acquisition Plan
 - Sensor Calibration
 - Survey Plans

NDSF EACHLITY

- Data Analysis + Reduction
- Data Documentation
- Data Management Plan (DMP)
 - Required in NSF Proposals
 - How will you preserver & document you data?





Primary Investigator: John Motton
Institution: Lanot Deherty Earth Observatory of Columbia University
Project: Reactivation of the Passive Margin of Eastern Laurentia
NSF Division: OCE Solicitation Info: Marine Geology and Geophysics Submission Date:
01/162013

Overview: Our project will use active source seismology on the Marcus G. Langseth 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 MBSystem.

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

Expected data product #1 Data type: Observational, Analytical Responsible investigator: John Mortes

Responsible investigator: John Morton Product description: .segy files from seismic transits. 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 MGD77-T format Intender repository: NGDC Timeline for data release: Immicilate Release

Expected data product #3 Data type: Observational

Data type: Observational Responsible investigator: Vicki L. Ferrini Product description: Multibeam bathymetry data Intended repository: MGDS Timeline for data release: Immediate Release

Field Data

- Facility-Managed Data
 - Ship (R2R)
 - Raw sensor data
 - Vehicle Data (WHOI + MGDS)
 - Raw sensor data
 - First-order at-sea products

PI-Managed Data

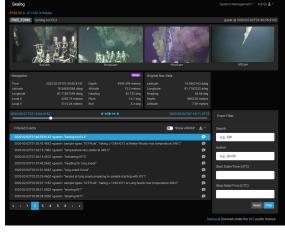
- Documentation
- Cruise report

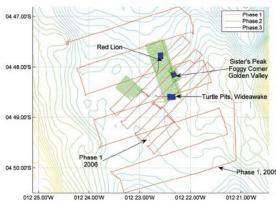
NDSF DEEP SUBMERGENCE

- Sample metadata
- Processing metadata
- Physical samples
- Science party instrumentation









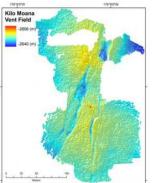
Processed/Derived Data

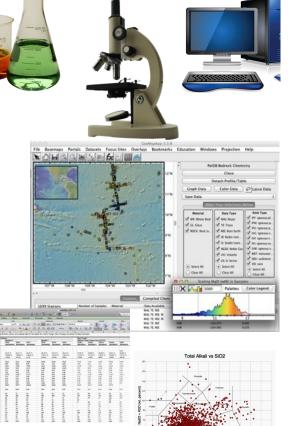
- Which data should be preserved?
 - Data supporting publications
 - Processed data of value
 - Results of lab analysis
- Where should it be curated?
 - Domain-specific repository?
 - What are requirements of repository?
- Documentation

NDSE NATIONAL DEEP SUBMERGENCE

- What does a new user need to know?
- How were products generated?
- What are caveats of data?





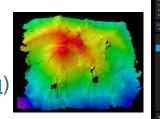


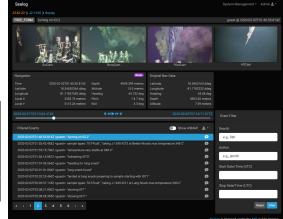
Deep-Submergence Data Resources

- WHOI/NDSF (<u>whoi.edu</u>)
 - Dive metadata & data at WHOI Data Library
 - NDSF vehicles
- Marine Geoscience Data System (<u>marine-geo.org</u>)
 - Dive metadata, field & derived data
 - Navigation, geophysical data, event logs, bottom photos
 - Data from many operators/vehicles
 - SOI: ROV SuBastian
 - NDSF: Alvin, Jason, Sentry
 - Nautilus: Hercules
 - MBARI Mapping AUV
 - REMUS
 - LEGACY: DSL120, IMI130, ABE
- NOAA/NCEI (<u>www.ncei.noaa.gov</u>)
 - OER dive metadata & data
 - Searchable dive video archive
- Scientific publications

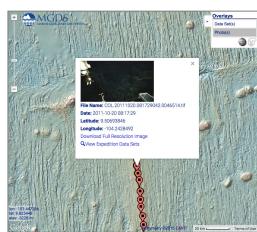
NDSF DEEP SUBMERGENCE

• Other...











Navigating Data Management

- Plan ahead
- Know what resources are available
 - Software Tools
 - Guidelines & Templates
- Communicate
 - Upstream (Operations Team)
 - Downstream (Data Managers)
- Organize consistently
- Document contemporaneously
- Treat data as a valuable community resource
- Participate!
 - Metadata & data format standards
 - System usability





Which Repository?

- Know data policies
- Seek domain-specific repositories
- System features to consider
 - Long-term Archiving
 - Data Usage Reports
 - Data Publication
 - User Support
 - Usability

NDSF EACH ITY

• Interoperability

Ready, set, share: Researchers brace for new data-sharing rules



science for a changing world

MARINE GEOSCIENCE DATA SYSTEM













rvdata.us

funded by NSP

unidata





CRITICAL ZONE OBSERVATORIES







