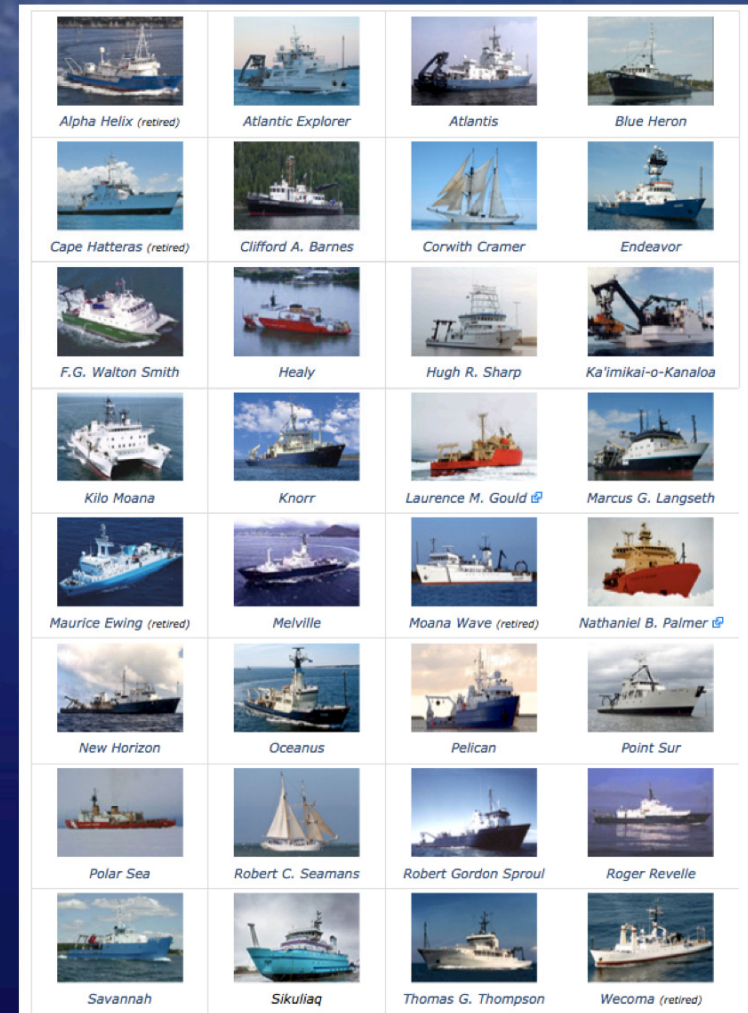# Rolling Deck to Repository (R2R) Program

Mission:

Stewardship of routinely-acquired environmental sensor data +documentation from U.S. academic research fleet

Funding:

NSF GEO Ocean+Polar Divisions
ONR

Services:

- Publish master cruise catalog
- Organize, archive, and disseminate data+documentation
- Assess data quality
- Create post-processed and/or quality-controlled data products
- Support at-sea event logging

Alpha Helix *(retired)*	Atlantic Explorer	Atlantis	Blue Heron
Cape Hatteras *(retired)*	Clifford A. Barnes	Corwith Cramer	Endeavor
F.G. Walton Smith	Healy	Hugh R. Sharp	Ka'imikai-o-Kanaloa
Kilo Moana	Knorr	Laurence M. Gould	Marcus G. Langseth
Maurice Ewing *(retired)*	Melville	Moana Wave *(retired)*	Nathaniel B. Palmer
New Horizon	Oceanus	Pelican	Point Sur
Polar Sea	Robert C. Seamans	Robert Gordon Sproul	Roger Revelle
Savannah	Sikuliaq	Thomas G. Thompson	Wecoma *(retired)*

Data Volumes

Currently:
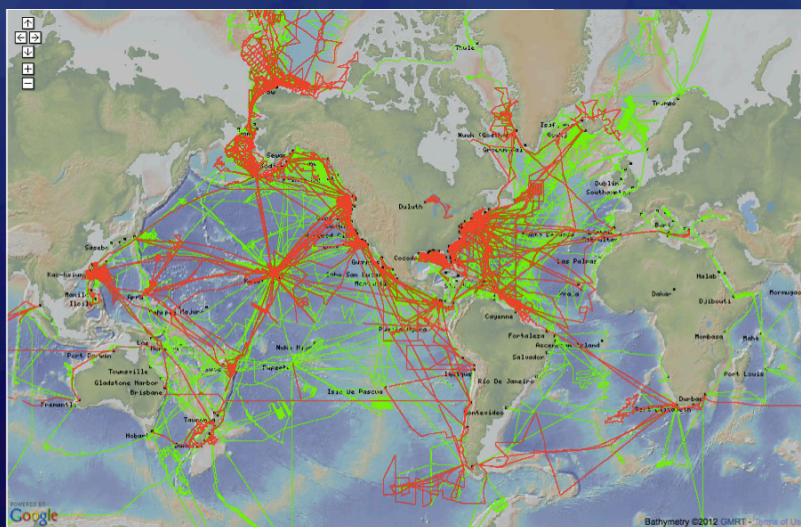   26 In-Service Vessels
    ~400 Cruises/Year
    ~400 Instrument Systems

Total as of November 2014:
    3,703 Cruises
    17.5M Files Archived

—— New cruises since R2R launch in 2009

—— Additional/ legacy cruises

Instrument Types:
    ADCP
    Anemometer
    CTD
    Expendable Probe
    Flowmeter
    Fluorometer
    GNSS /INS
    Gravimeter
    Gyrocompass
    Magnetometer
    Met Station
    Multibeam
    pCO2
    Radiometer
    Rain Gauge
    Singlebeam
    Speed Log
    SSV
    Thermometer
    Timeserver
    Transmissometer
    TSG
    Winch

Context

OOI
- ~Stationary platforms
- Long-term observations
- Regional/Local range

Fleet
- Mobile platforms
- Short-term observations
- Global range
- Ocean exploration

Academic Ocean Science Infrastructure

NSF
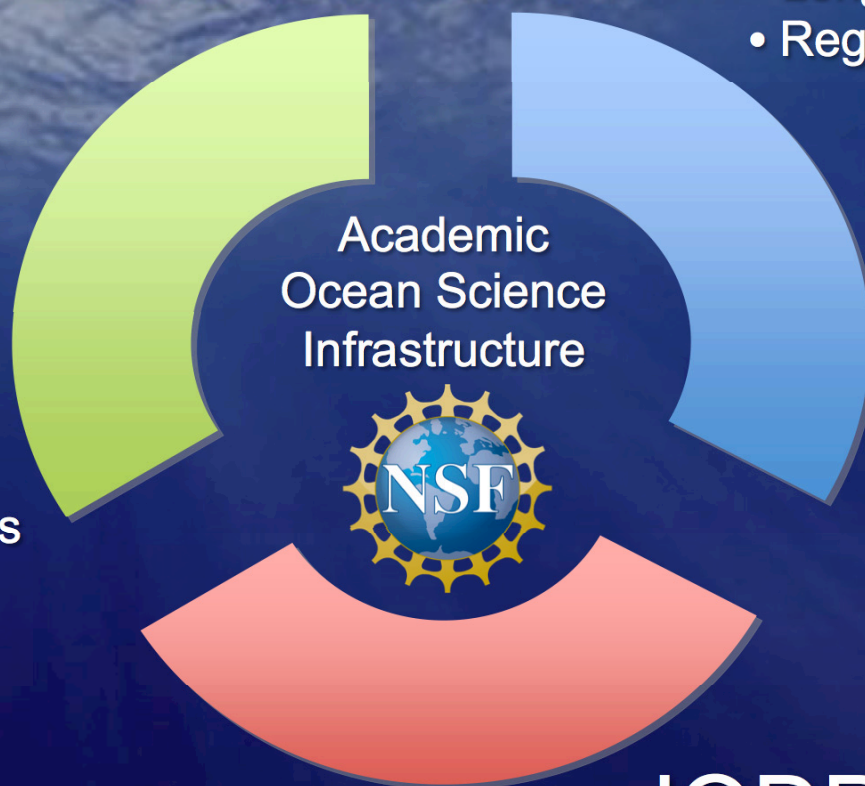
IODP
- Sampling/Logging
- Deep time
- Global range

Context *(cont.)*

History of data management in the U.S. academic fleet

1960's+   digital acquisition, institutional archiving (LDEO, SIO, WHOI, etc)

1980's+   data management at global program scale eg. JGOFS DMO

2000's+   integrated Web-based data systems (BCO-DMO, MGDS, SAMOS, etc)

2007   UNOLS Committee– fleetwide best practices

2008   R2R Pilot
2009+   R2R Program

Lamont-Doherty Earth Observatory
COLUMBIA UNIVERSITY | EARTH INSTITUTE

Program Team
(LDEO, FSU, SIO, WHOI)

International Partners

ODIP

Academic Fleet Office

R2R Partners

Fleet Operators

Disciplinary Data Systems

DataONE

SAMOS

BCO-DMO

IEDA

Geoinformatics Community

ESIP

National Data Centers

NOAA

R2R Benefits

- Facilitate data preservation +dissemination (+free operators from archiving)

- Standardize +simplify documentation

- Assess +improve data quality

- Enable scientists to integrate data sets +build visualization/analysis applications

- Facilitate cruise planning, permits, clearances

R2R Benefits *(cont.)*

Example: Contribute Multibeam data to Global Multi-Resolution Topography (GMRT) Synthesis
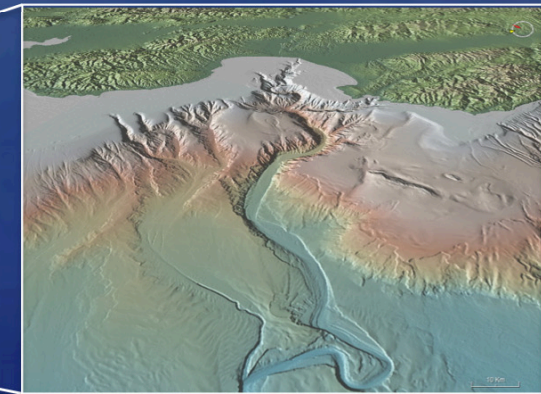
Applications:
• Tsunami modeling
• Ocean mixing/ Bottom currents
• Maritime navigation/ Cables
• Benthic habitats

Database Model

• Operator
• Vessel
• Ports/Dates
• Science Party
• Project
• Funding
(etc.)

CRUISE

DATA SET
(Instrument System +Interface)
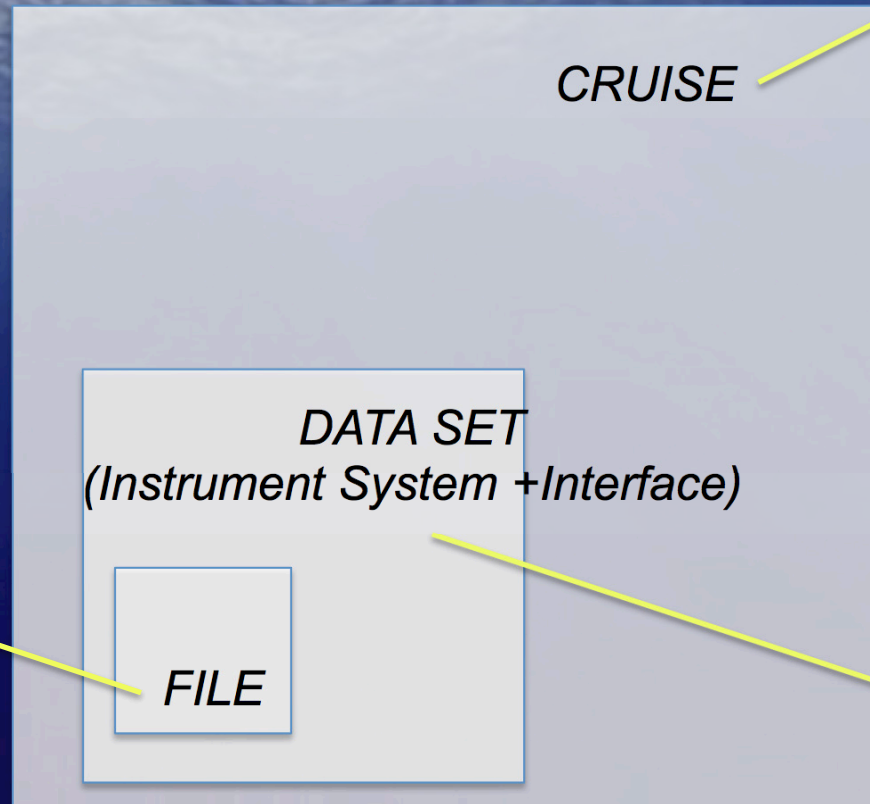
• Name
• Date
• Size
• Checksum
• Format
(etc.)

FILE

• Type/Class
• Manufacturer
• Model
• Installation
• Calibration
• History
(etc.)

Instrument Model

Acquisition System

An instrument may have subcomponents
eg. a CTD or Met Station

↳ Instrument Type (eg. "echosounder")

↳ Make

↳ Model

↳ Location

↳ Interface (talker)

↳ Format (output files)
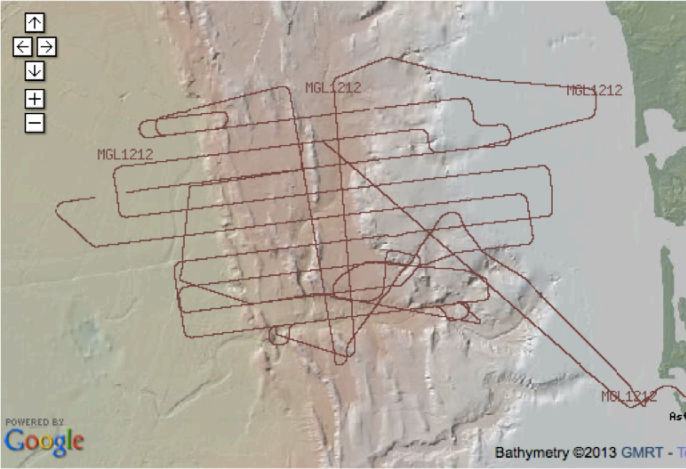
Cruise Catalog

Example: R/V Langseth cruise MGL1212

Cruise Summary + Underway Data in R2R Catalog

Seismic Data in partner system

Cruise Catalog *(cont.)*

Develop a fleet-standard cruise data directory structure to

1. Preserve original full-resolution data – especially navigation

2. Enable R2R to easily "break out" a data package by matching filename patterns against the vessel's instrument profile
   - Segregate resident sensor data from everything else
   - Segregate data from documentation
   - Segregate original data from sub-sampled /post-processed

3. Preserve "exact copy" of what the science party took home

Publish best practices for file compression, checksums, etc

http://www.rvdata.us/operators

Shiptrack Navigation
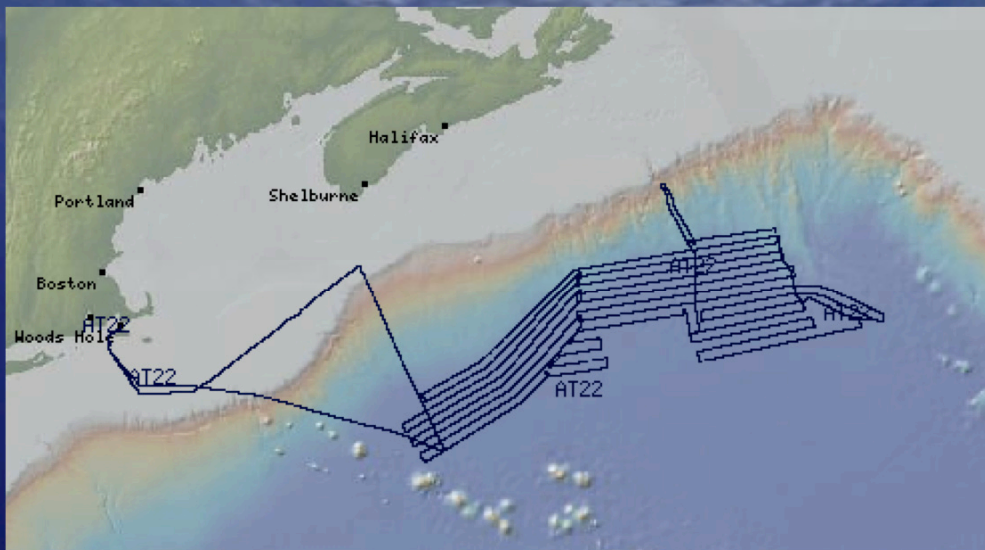
R/V Atlantis cruise AT22 (Scotian Shelf Survey, August 2012)

- Published quality-controlled navigation products for 1,757 cruises to date
  1. full-resolution
  2. one-minute
  3. control points

- Upgraded file format includes embedded metadata and COG/SOG

- Annual report delivered to Tech Managers on primary navigation system

Quality Assessment

In production:
Nav
Met/TSG
Multibeam
Grav
Mag

On deck:
CTD

In development:
ADCP

In planning:
Subbottom
Fluoro

Real-time Met/TSG

Shipboard Automated Meteorological and Oceanographic System (SAMOS)

- 10 university vessels recruited to date

- Daily automated quality control

- Data archived monthly at US NODC

Supports study of global air-sea fluxes, bench-marking satellite products, modelling/reanalysis, etc.

Ship Key	
Atlantic Explorer	☑
Atlantis	☑
Kilo Moana	☑
Knorr	☑
Melville	☑
New Horizon	☑
Robert Gordon Sproul	☑
Roger Revelle	☑
T G Thompson	☑

Map Controls	
On / Off	☑

Orientation	
Landscape	◉
Portrait	○

Satellite   Hybrid

Imagery ©2013 NASA - Terms of Use
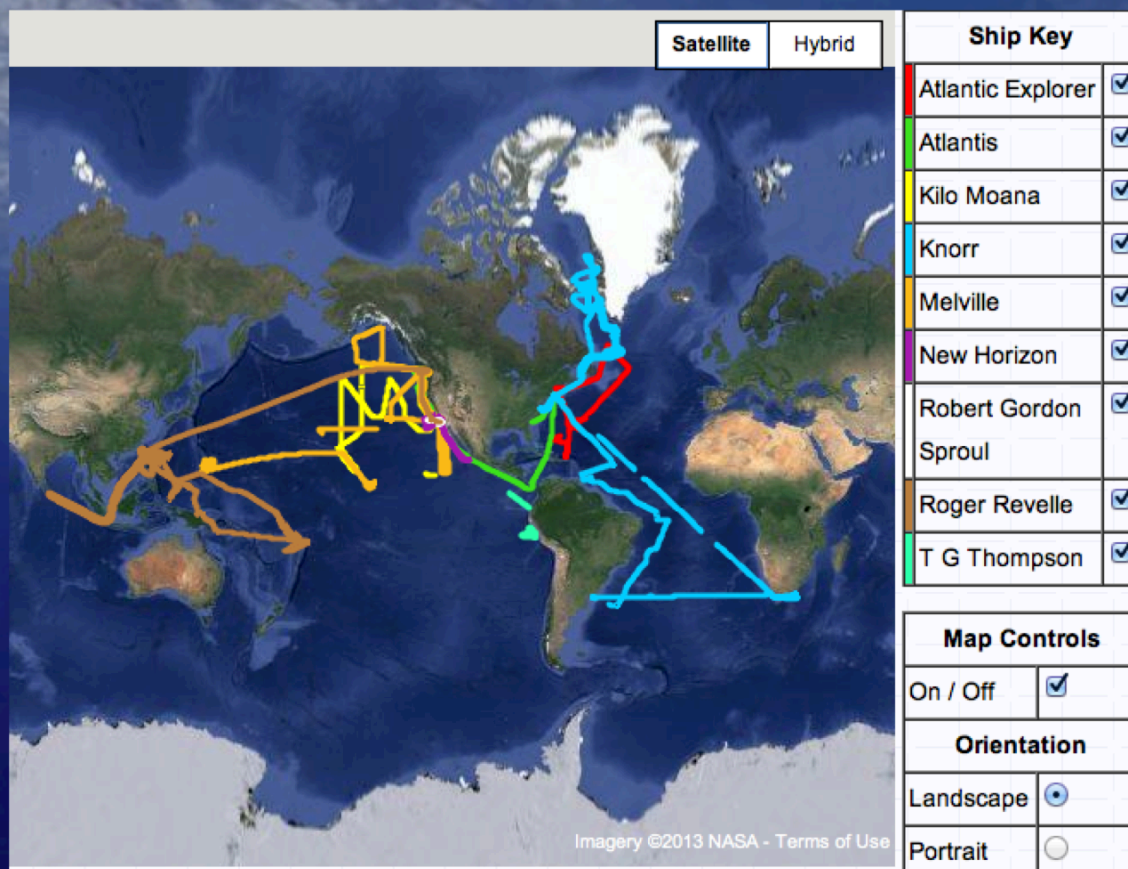
Data quality subscription service:

*https://samos.coaps.fsu.edu/html/subscription/*

Event Logger

- Goal: Create a digital record of all sampling events during a cruise

- Based on ELOG (open source Web log application)

- Deployed on 60+ cruises to date fleetwide

- Config file customized for each cruise, can be reused

- Deployment modes: Laptop ("science-owned") or DreamPlug ("ship-owned")

Collaboration

R2R participates in global/ community forums

- Int'l Oceanographic Data and Information Exchange (IODE) Workshops

- Research Data Alliance (RDA)

- Ocean Data Interoperability Platform (ODIP)
  - Publish R2R Cruise Summary Reports using EU/SeaDataNet template
  - Harmonize QC flags and parameter codes among eg. SAMOS, IODE-JCOMM, and SeaDataNet
  - Harmonize Event Logs between EUROFLEETS and UNOLS/R2R
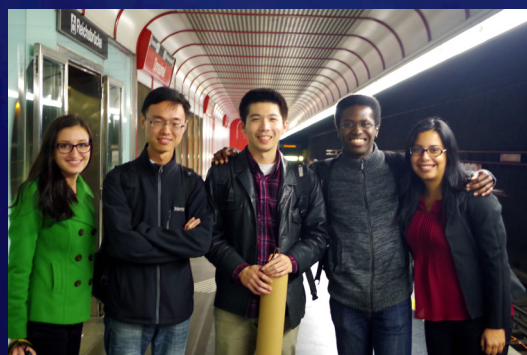
Workforce Development

R2R student fellows work with European +Australian developers to link data systems
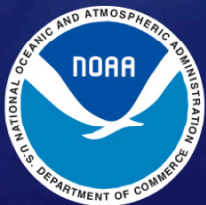
Collaboration *(cont.)*

- Earth Science Information Partners support best practices and prototype development

- NSF EarthCube Council of Data Facilities, Technical Governance, and "Building Block" projects


- Prototype development toward NOAA R2R Cruise Catalog

Lessons Learned

- Start with a clear mandate – original enviro data from resident sensor systems

- Keep the data model simple, built for high throughput

- Draw experience from existing data systems

- Reuse existing international/ community vocabularies eg. ICES Platform Codes

- Start pilot development work with small # of volunteer vessels

- Keep up in-person visits to operators /vessels

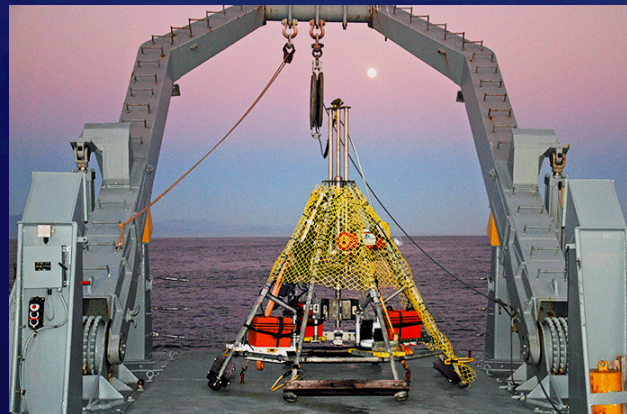- Stage the pipeline – upload, catalog, breakout, stage, release, archive, publish
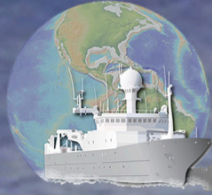
Questions and comments
   info@rvdata.us
   http://rvdata.us/contact

Team Members present this week:
   R.Arko, D.Clark, V.Ferrini, C.Nobre,
S.Smith, K.Stocks, L.Stolp

Thank you.

(Photo: Ellen Roosen, WHOI)