Rolling Deck to Repository (R2R): Envisioning the U.S. Academic Fleet as an Integrated Ocean Observing System

International Time-series Methods Workshop
Bermuda Institute of Ocean Sciences
28-30 November 2012

rvdata.us

http://www.rvdata.us/

Suzanne Carbotte[2], Stephen Miller[1], Andrew Maffei[3], Shawn Smith[4], Cynthia Chandler[3], Vicki Ferrini[2], Robert Arko[2], Karen Stocks[5], Suzanne O'Hara[2], Mark Bourassa[4], Dru Clark[1], Aaron Sweeney[1], John Morton[2]

[1] Scripps Institution of Oceanography
[2] Lamont-Doherty Earth Observatory (Lead Institution)
[3] Woods Hole Oceanographic Institution
[4] Florida State University
[5] San Diego Supercomputer Center

Rolling Deck to Repository (R2R) Program

Launched in 2009, R2R is a systematic effort to capture, catalog and archive U.S. underway shipboard data. The R2R program plan is to leverage and augment the existing centralized information resources of the UNOLS office, vessel operators, and National Data Centers to facilitate the documentation and delivery of data from "rolling deck" to "repository." As of November 2012, data from 2,886 cruises on 26 vessels had been submitted, totaling 12,749,775 files (>9 TB).

With its global capability and diverse array of sensors, the U.S. academic research fleet is an essential mobile observing platform for ocean science. Each vessel in the U.S. academic fleet is equipped with a multidisciplinary suite of sensors that are available for continuous operation during each expedition. The resulting "underway" geophysical, water column, and meteorological datasets describe basic environmental conditions for the oceans and are of high value for building global syntheses, climatologies, satellite validation and historical time series of air-sea fluxes and other ocean properties.

The R2R Portal (www.rvdata.us) is the central gateway through which underway data are routinely cataloged and securely transmitted to the appropriate national data center, ensuring long-term access and relieving chief scientists of their individual obligations under NSF policy to submit underway data.

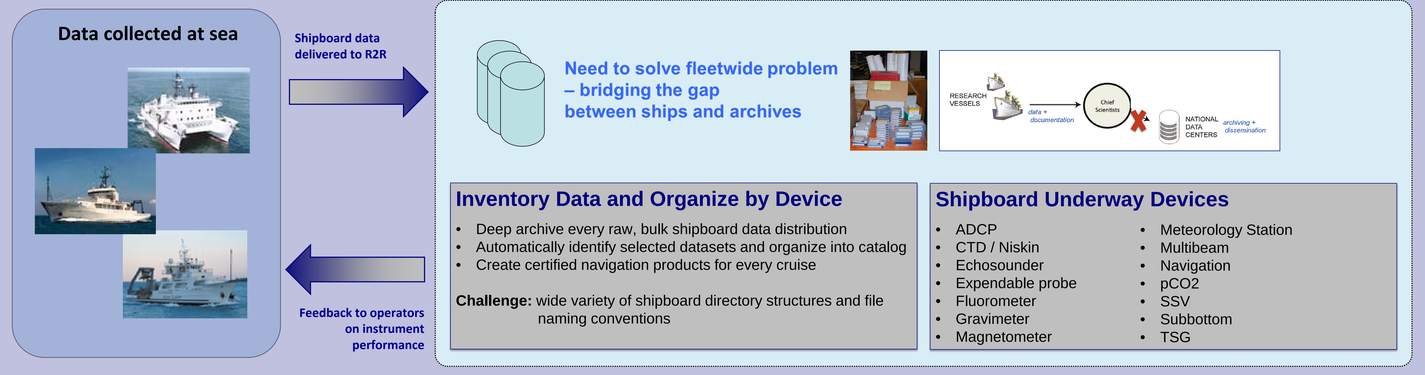
Protocols are being developed for quality assessing high priority underway data types, to provide feedback to shipboard instrument operators and to inform end users. Standard metadata will be supplied with each dataset, including provenance and quality information. Standard products, such as quality-controlled navigation, are being created.

As part of this work, R2R has collaborated with NOAA to create an XML-based, ISO 19115-compliant cruise metadata template. This describes the basic elements of a seagoing expedition: cruise identifier, vessel name, operating institution, dates/ports, navigation track, survey targets, science party, funding sources, scientific instruments, daughter platforms, and data sets. Controlled vocabulary terms are directly embedded as Uniform Resource Identifier (URI) references. We envision a hierarchical framework where a single "cruise-level" record is linked to multiple "dataset-level" records that may be published independently.

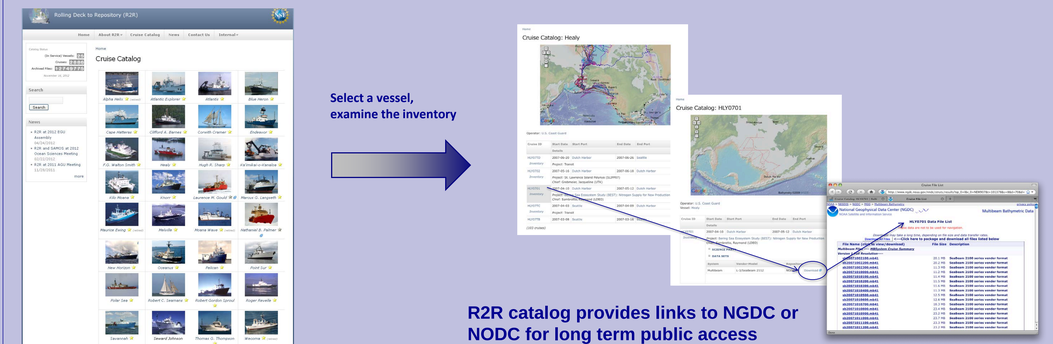
One of the subprojects within R2R is the development of a shipboard scientific event logging system that incorporates best practice guidelines, controlled vocabularies, a cruise metadata schema, and a scientific event log. The ELOG-based cruise event logging system, currently being tested, enables researchers to record digitally all scientific events and assign a unique event identifier to each entry, to assist in the ingestion of these data into oceanographic data repositories and subsequent reuse of the datasets.

Rolling Deck to Repository is a collaboration between Lamont-Doherty Earth Observatory (lead institution), Scripps Institution of Oceanography, San Diego Supercomputer Center, Woods Hole Oceanographic Institution, and Florida State University; and works with the vessel operating institutions, UNOLS Office, NOAA National Data Centers, and disciplinary data assembly centers (DACs). Funding is provided by the National Science Foundation (NSF), Oceanographic Instrumentation and Technical Services (OITS) Program, NSF OCE-0947828.

Systematic Capture and Archive of Routine Underway Data from U.S. Academic Fleet Vessels



Explore R2R cruise catalog and discover links to download files from national data centers



Event Logger – a flexible, extensible approach to capturing field sampling events based on open-source software.

CH0511-SE
R/V Cape Hatteras Science Event Log; LatMix; June 1, 2011 - June 21, 2011; Chief Scientist: Jim Ledwell (WHOI), Page 1 of 13

List | Find | Help

Full | Summary | Threaded — 242 Entries

Goto page 1, 2, 3 ... 11, 12, 13 Next All

Event	dateTimeUTC	Instrument	Action	Station	Latitude	Longitude	Author	Comment	Revisions
20110601.1316.001	20110601.1317	Ship	startCruise	NaN	34.694000	-76.670400	jLedwell		
20110601.1742.001	20110601.1840	ADCP150	other	NaN	34.114217	-76.150767	jLedwell	turning 90 degrees	
20110601.1820.001	20110601.1821	ADCP150	other	NaN	34.132433	-76.170750	jLedwell	Pole deployed; setting up software; stopped here about 30 min ago;	
20110601.1825.001	20110601.1826	ADCP150	startLine	NaN	34.129867	-76.170650	jLedwell	Calibration run; Depth ~ 100 m.	& 01 Jun 2011 18:28
20110601.1852.001	20110601.1853	ADCP150	endLine	NaN	34.096667	-76.164300	jLedwell	End of calibration run. Looks ok.	
20110601.2302.001	20110601.2305	calibrationFluorometer	start	NaN	33.678667	-75.631067	jLedwell	Started setting up around 2200 Z	
20110602.1355.001	20110602.1356	ADCP150	startLine	NaN	32.006933	-73.628517	sPierce	ok at 6 kts	& 02 Jun 2011 14:05

The R2R Event Logger uses the ELOG weblog software (https://midas.psi.ch/elog/). A custom configuration file is generated for each research cruise. This application is accessible to researchers using any Web browser that has access to the ship network.

A final event log that documents all instrument-oriented sampling events, with the use of controlled vocabularies, is included in and archived with the final cruise data distribution. As of November 2012 the event logger has been used on 35 cruises.

Initial accomplishments

Community collaboration

- Developing dataflow and submission agreements with NGDC and NODC
- Site visits and technical discussions with UNOLS, NOAA and 18 vessel operating institutions
- Defining ISO-standard cruise- and dataset-level metadata, and associated controlled vocabularies, in cooperation with NOAA. Develop templates for additional supporting documentation (see panel on right)
- Coordinating plans with international programs: SeaDataNet, Eurofleets, Geo-Seas

Developing the processing pipeline

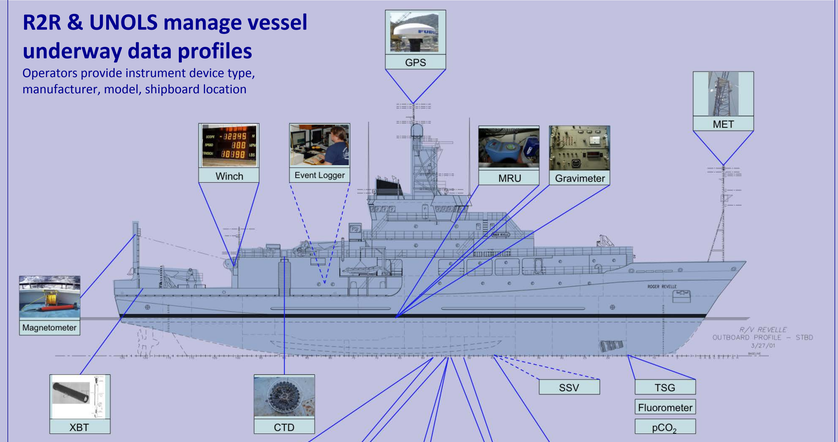
- Developing dataflow procedures for 26 vessels
- Capturing real-time MET and TSG data via the Shipboard Automated Meteorological and Oceanographic System (SAMOS) DAC
- Creating automatic cataloging scripts (sort through complex bulk shipboard distributions, identify data sets, populate standard directory structure, populate metadata database)
- Developing automated quality assessment for selected datasets

Routine provision of data to National Data Centers

- Saving all original, raw bulk shipboard distributions in deep archive (e.g. 9 TB at NGDC), a first step while data breakout and quality processes are in progress
- Routine creation and transfer of cruise-level metadata and standard navigation products
- Maintaining vessel instrument profile database, tracking changes, in coordination with UNOLS

R2R & UNOLS manage vessel underway data profiles

Operators provide instrument device type, manufacturer, model, shipboard location



GPS: Global Positioning Unit
SSV: Hull-mounted sound velocimeter
CTD: Conductivity, Temperature, Depth
XBT: Expendable Bathythermograph
ADCP: Acoustic Doppler Current Profiler

pCO_2: CO_2 partial pressure
MRU: Motion Reference Unit
MET: Meteorological Station
TSG: Thermosalinograph

A Community Effort

LDEO/SIO/WHOI/Florida State
UNOLS
NOAA