

Rolling Deck to Repository (R2R) Perspectives from a Decade of Ocean Data Management

Karen Stocks¹, Dru Clark¹, Vicki Ferrini², Rebecca Hudak³, Kevin McLain⁴, Suzanne O'Hara⁴, Christopher Olson¹, Shawn Smith⁵, Laura Stolp⁶, Rafael Uribe⁷, Emily Miller⁷, and Suzanne Carbotte⁸

¹Scripps Institution of Oceanography

²LDEO

³Woods Hole Oceanographic Institution

⁴Lamont-Doherty Earth Observatory of Columbia University

⁵Florida State University

⁶WHOI

⁷Lamont -Doherty Earth Observatory

⁸Lamont-Doherty Earth Obs

November 25, 2022

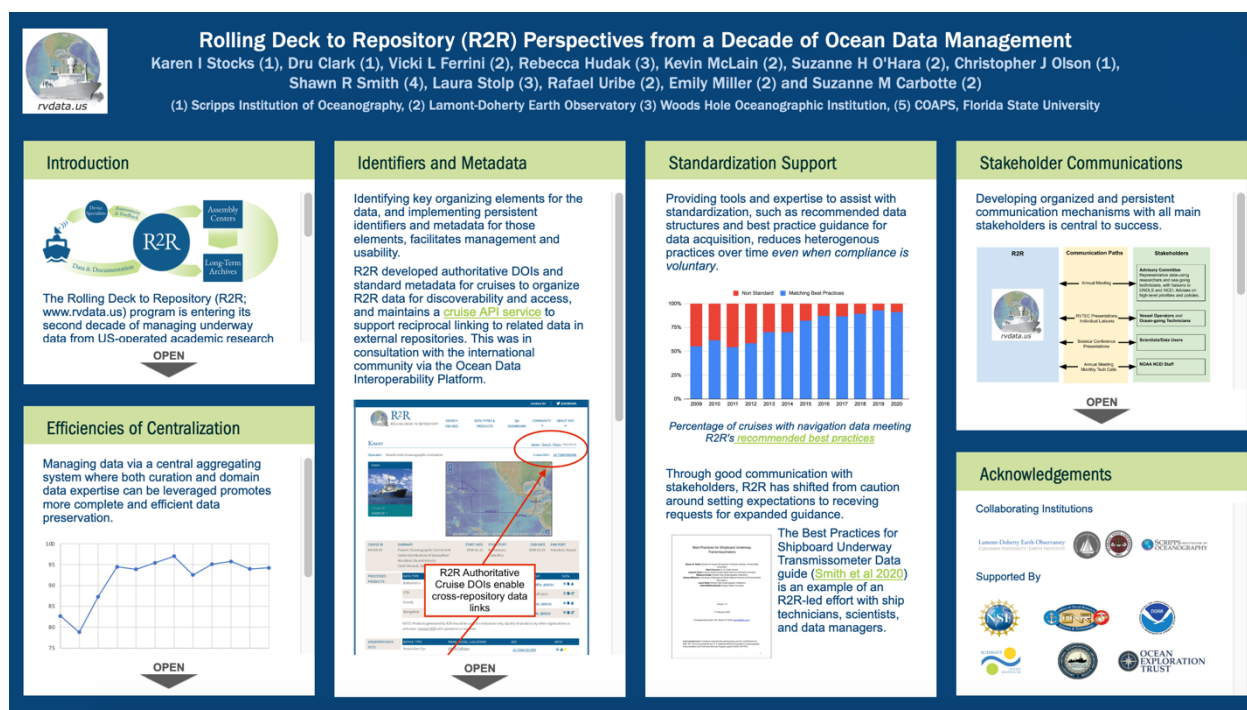
Abstract

The Rolling Deck to Repository (R2R; www.rvdata.us) program is entering its second decade of managing underway data from US-operated academic research vessels to ensure preservation of, and access to, these national oceanographic research assets. Reflecting on the move from decentralized data submission by chief scientists to an operational centralized facility has brought insights that may inform other communities with distributed networks of data acquisition providers with diverse practices and resources. 4,000 cruises and 100+TB of data later, here are lessons R2R has learned. - Managing data via a central aggregating system where both curation and domain data expertise can be optimally leveraged promotes more complete and efficient data preservation. - Identifying key organizing elements for the data, and implementing persistent identifiers and metadata for those elements, facilitates management and usability. R2R developed authoritative DOIs and standard metadata for cruises to organize R2R data for discoverability and access, and facilitate reciprocal linking to related data in external repositories. When data submissions from diverse providers are heterogeneous, standardizing data at ingest supports data aggregation and synthesis that promote broad data re-use. - Providing tools and expertise to assist with standardization, such as recommended data structures and best practice guidance for data acquisition, reduces heterogeneous practices over time even when compliance is voluntary. - Developing organized and persistent communication mechanisms with all main stakeholders is central to success. R2R has annual community-level meetings, as well as more frequent individual interactions, with vessel operators/technicians, the NOAA National Centers for Environmental Information staff, and oceanographic research scientists. These communications have been critical to informing high level priorities, overall approaches, and specific technical details and decisions.

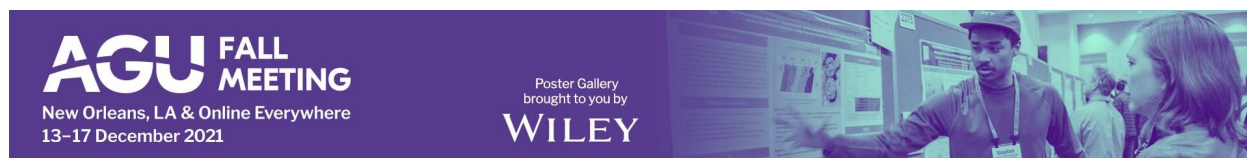
Rolling Deck to Repository (R2R) Perspectives from a Decade of Ocean Data Management

Karen I Stocks (1), Dru Clark (1), Vicki L Ferrini (2), Rebecca Hudak (3), Kevin McLain (2), Suzanne H O'Hara (2), Christopher J Olson (1), Shawn R Smith (4), Laura Stolp (3), Rafael Uribe (2), Emily Miller (2) and Suzanne M Carbotte (2)

(1) Scripps Institution of Oceanography, (2) Lamont-Doherty Earth Observatory (3) Woods Hole Oceanographic Institution, (5) COAPS, Florida State University



Presented at:



ABSTRACT

The Rolling Deck to Repository (R2R; www.rvdata.us) program is entering its second decade of managing underway data US-operated academic research vessels to ensure preservation of, and access to, these national oceanographic research assets. Reflecting on the move from decentralized data submission by chief scientists to an operational centralized facility has brought insights that may inform other communities with distributed networks of data acquisition providers with diverse practices and resources. 4,000 cruises and 100+TB of data later, here are lessons R2R has learned.

- Managing data via a central aggregating system where both curation and domain data expertise can be optimally leveraged promotes more complete and efficient data preservation.
- Identifying key organizing elements for the data, and implementing persistent identifiers and metadata for those elements, facilitates management and usability. R2R developed authoritative DOIs and standard metadata for cruises to organize R2R data for discoverability and access, and facilitate reciprocal linking to related data in external repositories.
- When data submissions from diverse providers are heterogeneous, standardizing data at ingest supports data aggregation and synthesis that promote broad data re-use.
- Providing tools and expertise to assist with standardization, such as recommended data structures and best practice guidance for data acquisition, reduces heterogeneous practices over time even when compliance is voluntary.
- Developing organized and persistent communication mechanisms with all main stakeholders is central to success. R2R has annual community-level meetings, as well as more frequent individual interactions, with vessel operators/technicians, the NOAA National Centers for Environmental Information staff, and

oceanographic research scientists. These communications have been critical to informing high level priorities, overall approaches, and specific technical details and decisions.

Introduction

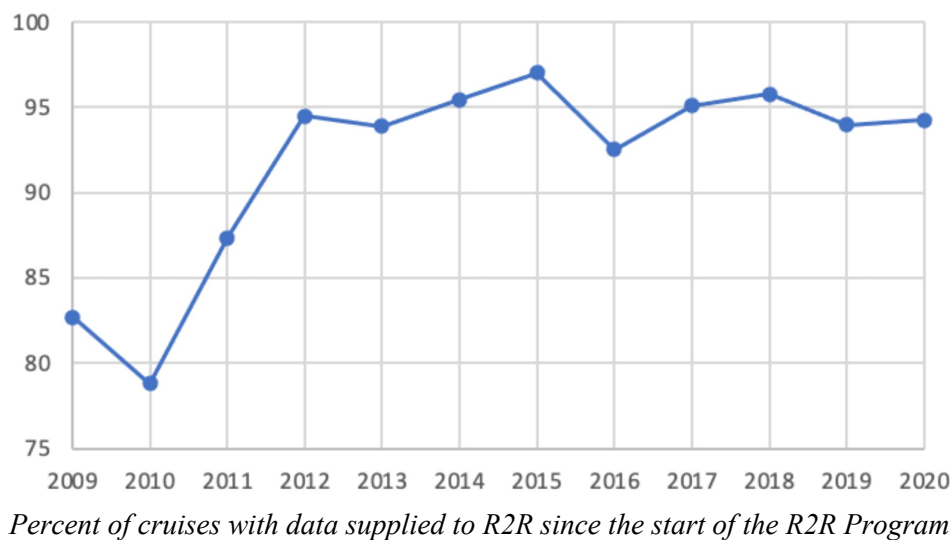


The Rolling Deck to Repository (R2R; www.rvdata.us) program is entering its second decade of managing underway data from US-operated academic research vessels to ensure preservation of, and access to, these national oceanographic research assets. Reflecting on the move from decentralized data submission by chief scientists to an operational centralized facility has brought insights that may inform other communities with distributed networks of data acquisition providers that include diverse practices and resources. **4,000 cruises and 200+TB of data later, here are lessons R2R has learned.**



Efficiencies of Centralization

Managing data via a central aggregating system where both curation and domain data expertise can be leveraged promotes more complete and efficient data preservation.




Identifiers and Metadata

Identifying key organizing elements for the data, and implementing persistent identifiers and metadata for those elements, facilitates management and usability.

R2R developed authoritative DOIs and standard metadata for cruises to organize R2R data for discoverability and access, and maintains a [cruise API service](#) to support reciprocal linking to related data in external repositories. This was in consultation with the international community via the Ocean Data Interoperability Platform.

[Contact Us](#)
[@R2RDATA](#)


R2R
 ROLLING DECK TO REPOSITORY

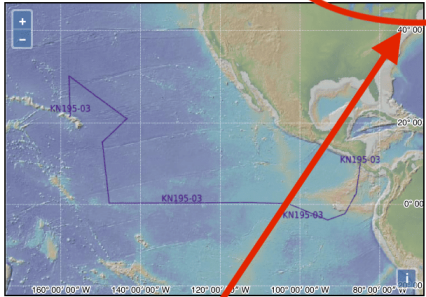
[SEARCH CRUISES](#)
[DATA TYPES & PRODUCTS](#)
[QA DASHBOARD](#)
[COMMUNITY](#)
[ABOUT R2R](#)

Knorr

Operator: Woods Hole Oceanographic Institution

Cruise ID: KN195-03

Cruise DOI: [10.7284/900496](#)



CRUISE ID	SUMMARY	START DATE	START PORT	END DATE	END PORT
KN195-03	Project: Oceanographic Control and Global Distributions of Subseafloor Microbial Life and Activity Chief: DHondt, St...	2009-01-12	Puntarenas, Costa Rica	2009-02-23	Honolulu, Hawaii

PROCESSED PRODUCTS	DATA TYPE	DATA
	Bathymetry	bathy_geocsv
	CTD	soft-proc
	Gravity	grav_geocsv
	Navigation	nav_geocsv

NOTE: Products generated by R2R should be used for evaluation only. Quality of products by other organizations is unknown. [Contact R2R](#) with questions or requests.


UNDERWAY DATA SETS	DEVICE TYPE	MAKE / MODEL [LOCATION]	DOI	DATA
	Acquisition Sys	WHOI Calliope	10.7284/101393	Data
	ADCP	Hawaii UHDAS	10.7284/101392	Data
	CTD	Sea-Bird SBE-911+	10.7284/101389	Data
	Gravimeter	Bell BGM-3	10.7284/101390	Data
	Multibeam Sonar	SeaBeam 3012-P1	10.7284/104733	Data
	Singl Beam Sonar	Knudsen 320B/R	10.7284/101391	Data

NOTE: This data set inventory may not be complete. Please [Contact R2R](#) with questions or requests.

CRUISE DATA AT OTHER SYSTEMS	REMOTE SYSTEM	DATA
	Biological and Chemical Oceanography Data Management Office (Info)	Bio+Chem Data
	Index to Marine and Lacustrine Geological Samples (Info)	Samples
	EarthChem Library (Info)	GeoChem Dataset#890
	EarthChem Library (Info)	GeoChem Dataset#899
	EarthChem Library (Info)	GeoChem Dataset#900
	EarthChem Library (Info)	GeoChem Dataset#902
	EarthChem Library (Info)	GeoChem Dataset#905
	EarthChem Library (Info)	GeoChem Dataset#906
	EarthChem Library (Info)	GeoChem Dataset#907
	Shipboard Automated Meteorological and Oceanographic System (Info)	RT METOC Data
	System for Earth Sample Registration (Info)	Samples
	Marine Geoscience Data System (Info)	Marine Geo Data
	Marine Biological Laboratory Woods Hole Oceanographic Institution Library (Info)	WHOI Data Library Archives

An [example cruise landing page](#) at [rvdata.us](#)

R2R also mints DOIs for filesets, and leverages extensive external [controlled vocabularies](#).

 **R2R**
ROLLING DECK TO REPOSITORY

SEARCH
CRUISES

DATA TYPES &
PRODUCTS

QA
DASHBOARD

COMMUNITY
▼

ABOUT R2R
▼

WET Labs WETStar Fluorometer

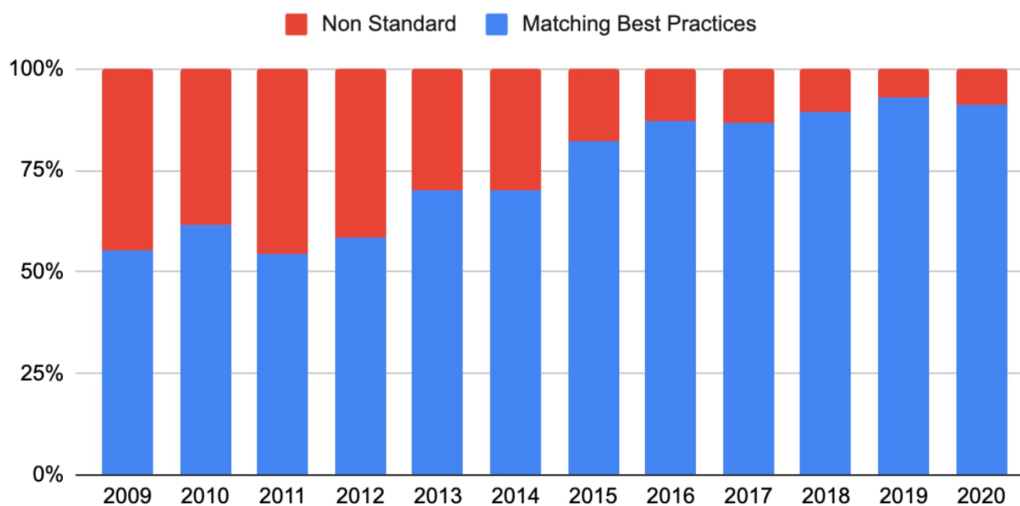
Home / Search / [model_vocabulary/100039](#)

URI	http://vocab.nerc.ac.uk/collection/L22/current/TOOL0075/
Identifier	L22::TOOL0075
Preferred Label	WET Labs WETStar Fluorometer
Alternative Label	WET Labs WETStar
Definition	Submersible fluorometer designed for through-flow or pumped CTD applications manufactured by WetLabs and which can be configured for various types of fluorescence. The probe has a temperature range of 0-30°C and a depth rating of 600m.

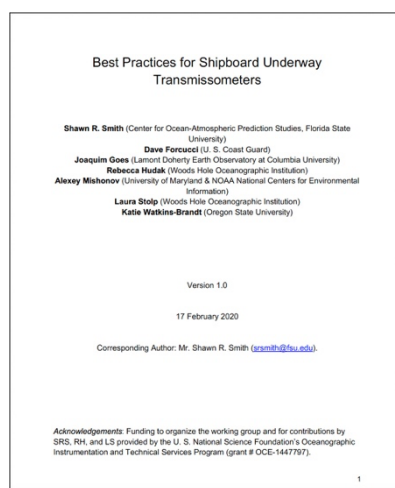
Example Device Model vocabulary term, leveraging the NERC vocabulary server

Standardization Support

Providing tools and expertise to assist with standardization, such as recommended data structures and best practice guidance for data acquisition, reduces heterogenous practices over time *even when compliance is voluntary*.



Percentage of cruises with navigation data meeting R2R's [recommended best practices](#)

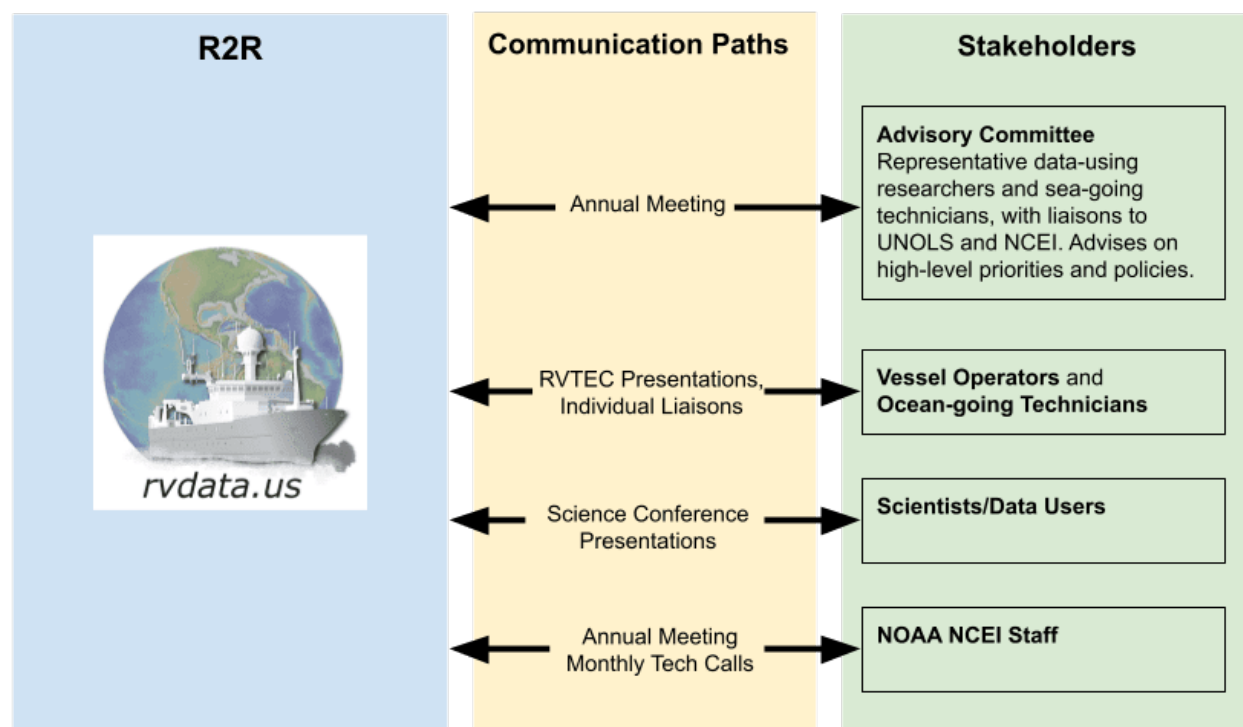


Through good communication with stakeholders, R2R has shifted from caution around setting expectations to receiving requests for expanded guidance.

The Best Practices for Shipboard Underway Transmissometer Data guide ([Smith et al 2020](#)) is an example of an R2R-led effort with ship technicians, scientists, and data managers.

Stakeholder Communications

Developing organized and persistent communication mechanisms with all main stakeholders is central to success.



R2R has annual community-level meetings, as well as more frequent individual interactions, with vessel operators / technicians, the NOAA National Centers for Environmental Information staff, and oceanographic research scientists. These communications have been critical to informing high level priorities, overall approaches, and specific technical details and decisions.

Acknowledgements

Collaborating Institutions

Lamont-Doherty Earth Observatory
COLUMBIA UNIVERSITY | EARTH INSTITUTE



SCRIPPS INSTITUTION OF
OCEANOGRAPHY

Supported By

