



Rolling Deck to Repository (R2R)

Status and New Developments -2019

S. O'Hara, D. Clark, J. Elya, K. McLain, C. Olson, C. Sellers, S. Smith, K. Stocks, L. Stolp, S. Carbotte
LDEO, FSU, SIO, WHOI

R2R Services for the ARF

Mission: To support acquisition, documentation, preservation, and enhanced usability of the underway environmental sensor data from scientific cruises conducted with the U.S. Academic Research Fleet (ARF).

Core Services

1. Digital curation & preservation for all routinely operated sensors
2. R2R Cruise Catalog -resource for discovery of information and data
3. Data documentation/QA and Level 1/2 data products to broaden access



R2R Services for the ARF

- **Support acquisition**: Event Log and NRT MET/TSG QC via SAMOS
- **Document**: Cruise level metadata, Datasets extracted and documented from cruise distro, QA for most common data types
- **Preservation**: Submission to NCEI for LTA
- **Dissemination**: via R2R Cruise Catalog, NCEI, Data Publication (DOIs)
- **Re-Use**: Standard navigation, QA and Post-field data products

Renewal Proposal -opportunity for ten
year review

Thanks to NSF and the community for
their support

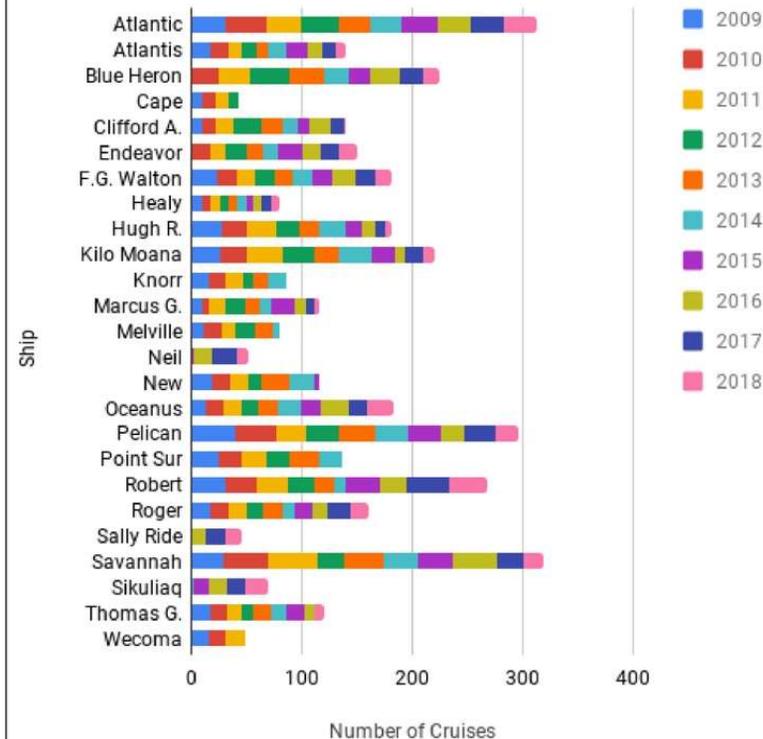
R2R Cruise Catalog

All ships of the ARF in operation from 2009 onward are cataloged (+ Falkor, Ron Brown, Nautilus with SOI+NOAA funding)

+ older than 2009: Basic cruise metadata & some data for several retired ships

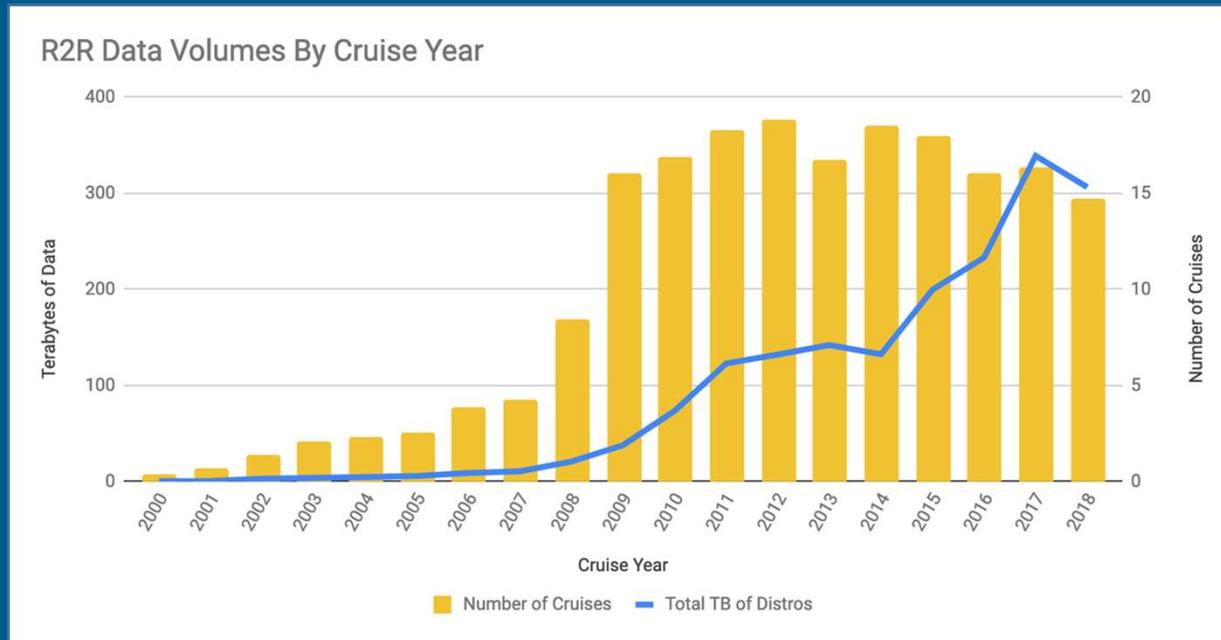
Discussions underway to bring Palmer and Gould into R2R

#Cruises/Year Curated by R2R by Ship



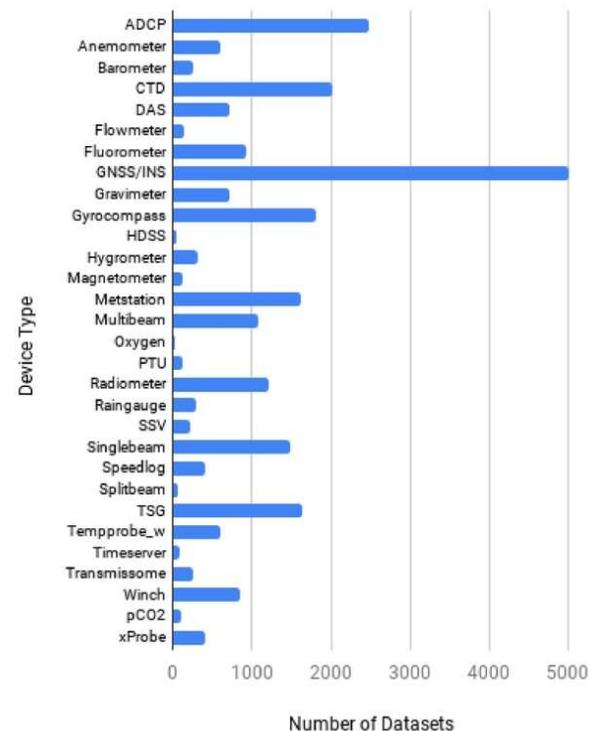
Cruises & Data Volumes

Number of cruises declining, but TB and # files increasing – new devices.



Curation totals for
ARF 2009-2018:
3775 cruises

Number and Type of Datasets Extracted and Curated



Device/Data Types Extracted and Accessible

Curation Totals for ARF 2009-2018

>26,000 datasets

>9 million files (broken out, documented and preserved)

Catalog Total: 29,173 datasets (12% more)

R2R ELOG to document underway metadata

R2R ELOG can be used by technicians and scientists to capture device metadata during a cruise. This information can be invaluable when data is re-used.

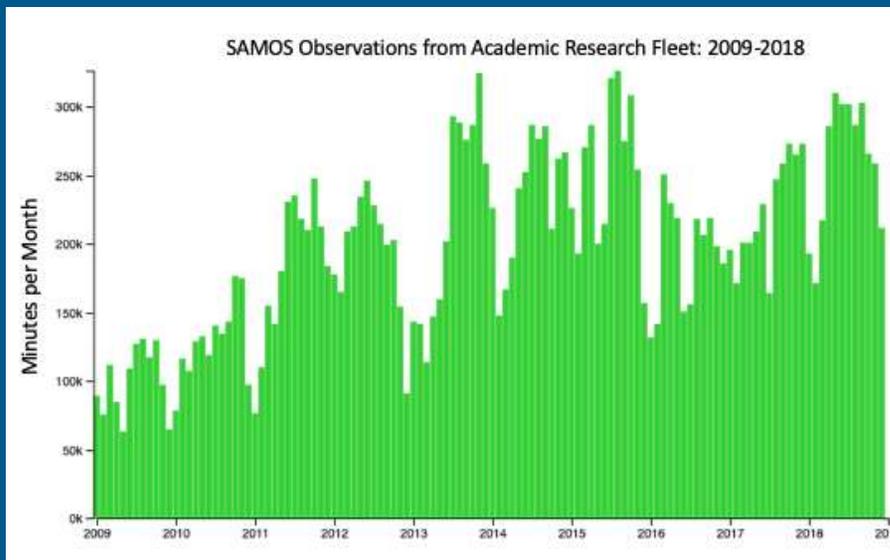


2019/07/07 12:35:02	Attune Flowcytometer	end	NaN	NaN	NaN	39.799803	-70.829629	tCrockford1	fixing leak in pump
2019/07/07	underway science seawater diaphragm	service	NaN	NaN	NaN	39.797496	-70.829631	tCrockford1	stop to fix leak at pump
	Attune Flowcytometer	start	NaN	NaN	NaN	39.750731	-70.829584	tCrockford1	
	underway science seawater diaphragm	start	NaN	NaN	NaN	39.750707	-70.829548	tCrockford1	

Comments can explain gaps in data, provide installation information, new calibration values and more....

Near RT QA/QC of MET and TSG: SAMOS

- One-minute data records sent each day as csv files
- Files are QA'd, converted to netCDF, QC'd to produce a science-ready data product
- Notifications to ship-board technicians of sensor and/or data flow problems
- Operators can make repairs during cruise - higher-quality dataset



10 vessels (*Roger Revelle*, *Atlantis*, *T.G. Thompson*, *Neil Armstrong*, *Kilo Moana*, *Atlantic Explorer*, *Pelican*, *Sikuliaq*, *Sally Ride*, *R.G. Sproul*), 3 ships retired.

R2R Quality Assessment

QA goals are to identify suspicious data not science quality

1. file/metadata integrity (exists, size, format, within temporal/spatial bounds)
2. device specific tests
3. Suitable to DP

R2R (Level 1/2) Data Products

Physical parameters rather than instrument values

Standard format across fleet

Suitable for/submitted to global syntheses

Broaden data access and re-use

Final Navigation for each cruise

The screenshot displays the R2R interface for cruise RR1210. It includes a ship image, a bathymetric map, and a summary table with columns for CRUISE ID, SUMMARY, START DATE, START PORT, END DATE, and END PORT. Below this is a table of 'R2R POST-FIELD PROCESSED PRODUCTS' with columns for DATA TYPE, DEVICE TYPE(S), FORMAT, and DATA. A 'Data Products' label is overlaid on the left side of this table. The 'UNDERWAY DATA SETS' table below it has columns for DEVICE TYPE, MAKE-MODEL (LOCATION), DOI, and DATA. Two black ovals highlight specific data rows in both tables.

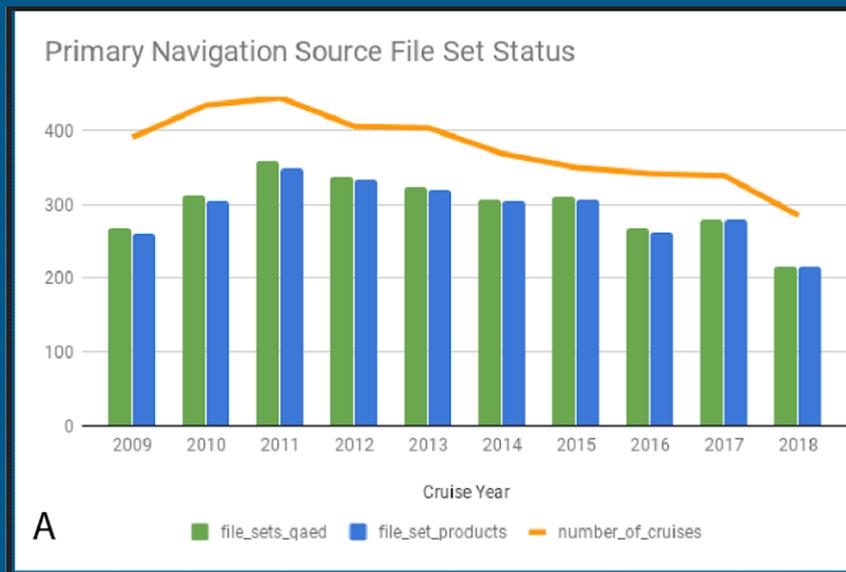
CRUISE ID	SUMMARY	START DATE	START PORT	END DATE	END PORT
RR1210	Project: Tonga Trench Microbiology Chief: Leon-Zayas, Rosa	2012-08-31	Apia, Samoa	2012-09-06	Suva, Fiji

R2R POST-FIELD PROCESSED PRODUCTS	DATA TYPE	DEVICE TYPE(S)	FORMAT	DATA
Bathymetry	singlebeam (Knudsen 3260), singlebeam (Knudsen 3208/R)		r2rbathy_geocsv	
CTD	ctd (Sea-Bird SBE-911+)		seasoft-proc	
Gravity	gravimeter (Bell BGM-3)		r2rgrav_geocsv	
Magnetics	magnetometer (Marine Magnetics SeaSPY)		r2rmag_geocsv	
Navigation	gnss (Furuno GP-150)		r2rnav_geocsv	

UNDERWAY DATA SETS	DEVICE TYPE	MAKE-MODEL (LOCATION)	DOI	DATA
ADCP		Hawaii UHDAS	10.7284/110409	
CTD		Sea-Bird SBE-911+	10.7284/110419	
Expendable Probe		Turo Quoli	10.7284/110562	
GNSS		Furuno GP-150	10.7284/110411	

QA Information

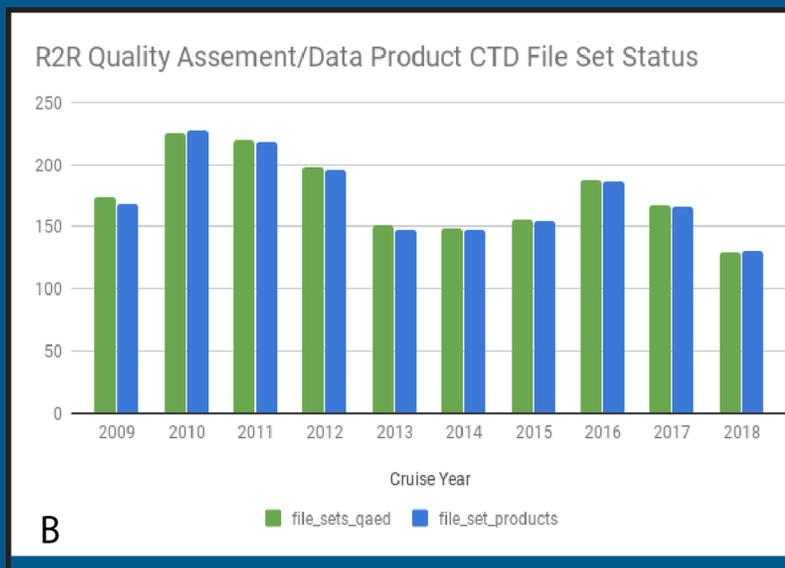
Navigation



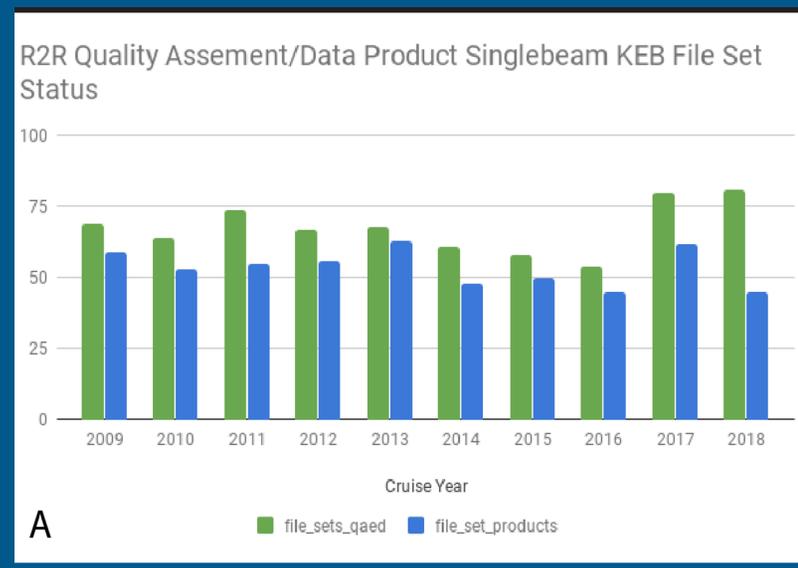
Cruises without navigation QA/DP (17%):

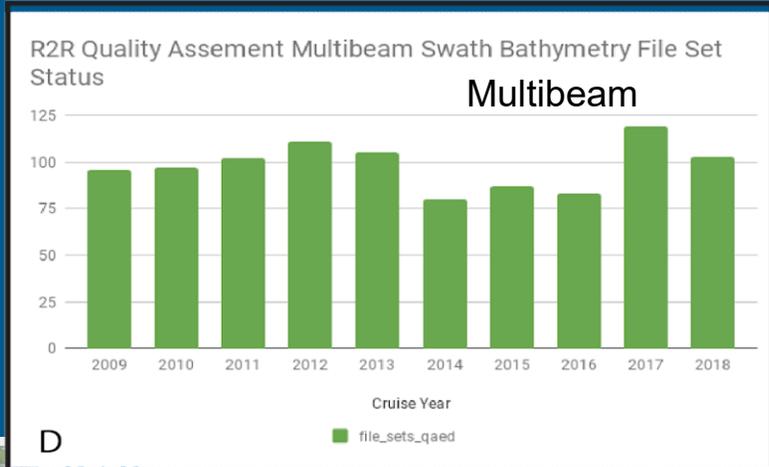
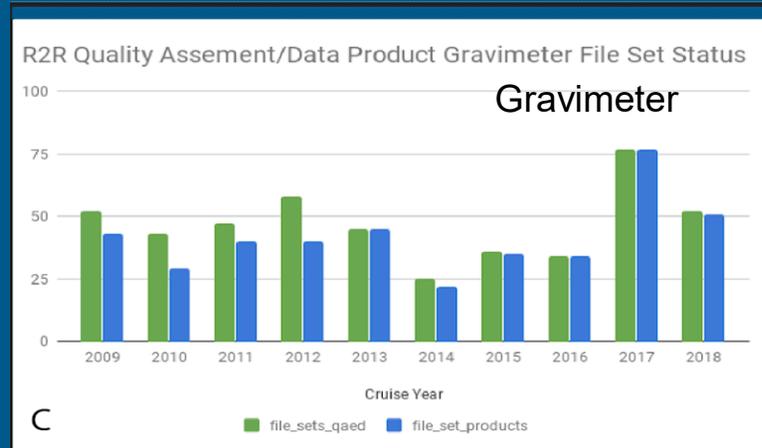
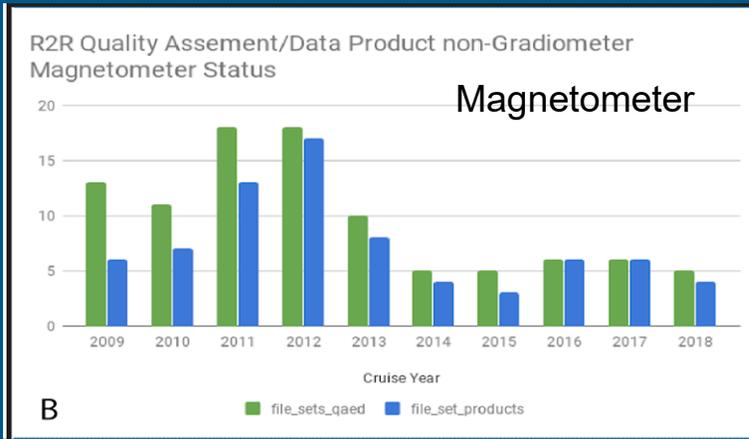
- 1) No identifiable GPS stream in cruise distro or nav info in multiplexed file (8%)
- 2) Nav not identified due to changes in file naming or directory structure (3% of cruises),
- 3) “one off” issues that prohibit the identification and breakout of navigation for a single cruise (6%).

CTD



Knudsen KEB





QA-DP # difference due to

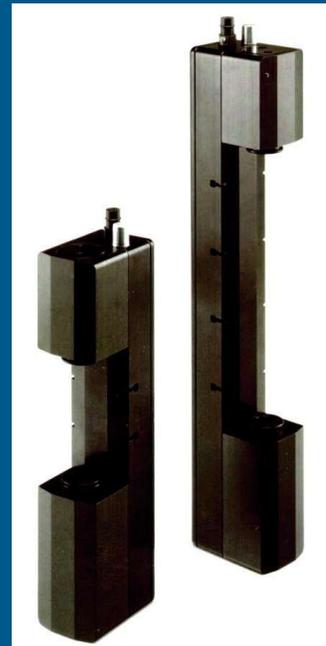
- lack of nav
- missing files
- files with no data
- missing calibration info

New Developments for 2018-2019

- Best Practices for Transmissometer
- XBT QA and DP
- Standard documentation of all QA & DP processes
- R2R Cybersecurity Plan- working with SCGI/Trusted CI 
- Improved curation efficiency and automation—
Internal Status Dashboard

Underway Flow-Water System Best Practices

- FSU and WHOI coordinated a transmissometer working group in 2018
- Drafted best practices (BP) that include
 - Data to record and provide to users
 - Basic C-Star Calculations
 - Metadata to document devices
 - Cleaning and in-situ calibration
 - Installation
 - Storage between cruise
- [https://www.rvdata.us/files/Transmissometer Best Practices v2 final.pdf](https://www.rvdata.us/files/Transmissometer%20Best%20Practices%20v2%20final.pdf)



SeabirdC-Star Transmissometers (source: datasheet_cstar.pdf)

- 2019 RVTEC Recommendations
 - Develop overall standard operating procedures (SOP)/ BP for common flow-water devices
 - Develop and conduct hands-on technician training for optical flow water devices
- Way Forward
 - Publish Transmissometer BP
 - Expand collaboration with RCRV team via electronic or in person meetings to leverage their SOP work for flow-water devices
 - Develop unified SOP/BP and encourage adoption within the academic research fleet

XBT QA/DP

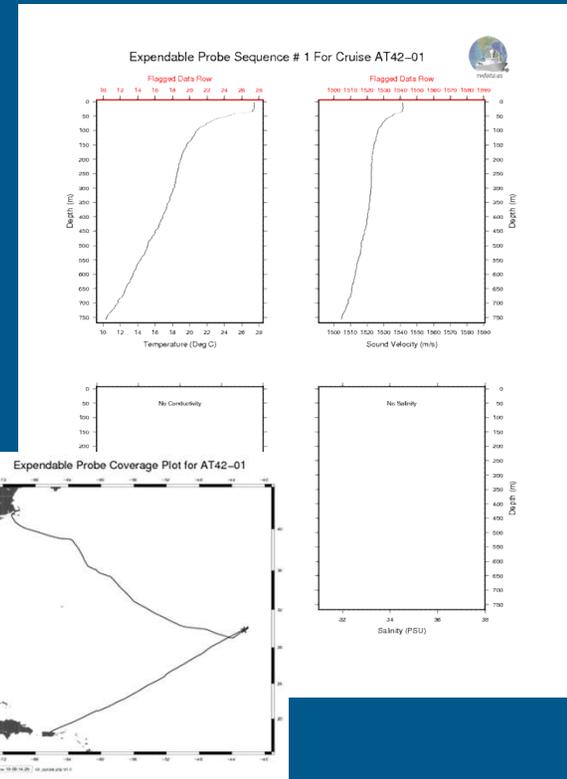
Expendableprobe Filesets for Atlantis

Total Filesets: 44

Click column headings to sort

View: Test Results

Rating	Device	Cruise	Fileset ID	Vessel	QA Summary	HasRequiredMetadata	PercentExpectedDataFiles	PercentFilesWithDefPairs	PercentFilesWithValidChecksum	PercentFilesWithValidConductivity	PercentFilesWithValidDepth	PercentFilesWithValidPosition	PercentFilesWithValidSalinity	PercentFilesWithValidSoundVelocity	PercentFilesWithValidSpatialRange	PercentFilesWithValidTemperature	PercentFilesWithValidTemporalRange	PercentNonZeroLengthFiles	PercentRecordsBelowTerminalDepth	PercentRecordsWithValidConductivity	PercentRecordsWithValidDepth	PercentRecordsWithValidSalinity	PercentRecordsWithValidSoundVelocity	PercentRecordsWithValidTemperature	PercentValidFormatFiles
🟢	Expendableprobe Sippican MK12	AT15-42	101380	Atlantis	QA Summary	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
🟢	Expendableprobe Sippican MK12	AT18-07	108569	Atlantis	QA Summary	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
🟢	Expendableprobe Sippican MK12	AT18-08	108572	Atlantis	QA Summary	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
🟢	Expendableprobe Sippican MK12	AT18-12	108583	Atlantis	QA Summary	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
🟢	Expendableprobe Sippican MK12	AT18-18	108596	Atlantis	QA Summary	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
🟢	Expendableprobe Sippican MK12	AT21-02	108600	Atlantis	QA Summary	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
🟢	Expendableprobe Sippican MK12	AT21-03	108602	Atlantis	QA Summary	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
🟢	Expendableprobe Sippican MK12	AT21-05	108476	Atlantis	QA Summary	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
🟢	Expendableprobe Sippican MK12	AT22	109675	Atlantis	QA Summary	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢



Standard format GeoCSV file of XBT metadata and data
 Flagging of suspect data values

R2R – Next 5 years

- Ongoing Data Curation and Cruise Catalog for the ARF
- Operator Dashboard of status of cruise ingestion
- New web services and R2R Catalog search functionality
- Extracting instrument documentation and additional data from distros
- Working with RVTEC to develop additional device best practices
- Explore (with RCRV group) near-realtime data streams- nav and archiving issues
- Tracking cruise publications via Crossref

Acknowledgements

Lamont-Doherty Earth Observatory
COLUMBIA UNIVERSITY | EARTH INSTITUTE



 SCRIPPS INSTITUTION OF
OCEANOGRAPHY

Providing access to and ensuring
the preservation of national
oceanographic research data.

