Rolling Deck to Repository (R2R)

Current Status and New Developments

*E. Miller, S. O’Hara, D. Clark, R. Hudak, S. Smith, K. Stocks, L. Stolp, R. Uribe, V. Ferrini, S. Carbotte
LDEO, FSU, SIO, WHOI
*presenter
Our Mission

Raw Data & Metadata

Feedback, Best Practices

R²R

Raw Data & Metadata, QA Documentation, Data Products

Best Practices

Science, education users

NCEI Long-term Archives

Data Synthesis/Assembly Centers

www.rvdata.us
Data You Contribute Circles the Globe

R2R makes your data available to all
Cruises & Data Volumes at R2R

values generated 2022/10/26 for all funded cruises
Arrival of Data to R2R

Number of days between end of cruise and ingestion of metadata into R2R

- Blue bars: number of cruises
- Red line: number of days

Number of days between end of cruise and arrival of distro at R2R

- Blue bars: number of cruises
- Red line: number of days

Statistics generated 2022/10
Arrival of Data to R2R Varies by Vessel

Some operators have more difficulty getting data to R2R quickly.

Remember that the data isn’t archived properly until it reaches R2R.

We offer multiple methods for transfer of data. What can we do to make it faster and easier for you?
Current Data Transmission Methods

R2R accepts data and metadata using many methods

- Globus
- SFTP
- FTP
- Cloud downloads
- USB drives via mail
- Email
- Dropbox
- ...

RVTEC 2022
<table>
<thead>
<tr>
<th>Transfer Method</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Globus</strong></td>
<td>Faster than SFTP, no need for a paid subscription to push data to us</td>
<td>Requires system configuration by IT Operator will require a paid option for R2R to pull data. R2R hopes to setup a push option that would be free for users</td>
</tr>
<tr>
<td></td>
<td>Interrupted transfer are resumed automatically</td>
<td></td>
</tr>
<tr>
<td><strong>SFTP</strong></td>
<td>Secure, relatively quick transfer times, reliable, we pull from you</td>
<td>Requires system configuration on your end (we do not have a space you can push to that is secure)</td>
</tr>
<tr>
<td><strong>Cloud downloads</strong></td>
<td>Easily adaptable depending on what the operator is already utilizing, use of Rclone has made these downloads more traceable</td>
<td>Have encountered some issues with interrupted data uploads/losing files &amp; directories, as well as corrupting data on download, some require paid subscriptions</td>
</tr>
<tr>
<td>(dropbox, google, box, …)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Physical drive</strong></td>
<td>Can receive large swath of cruises at once (especially with unreliable network connections)</td>
<td>Less convenient for frequent data transfer, if there’s a mistake/data is missing it takes more time to fix the issue, delivery delays, must be mounted in-person</td>
</tr>
</tbody>
</table>
Users can now request any file or folder from a cruise data distribution!

New Distro File Request Tool

Distro File Request for Cruise AT42-17

CRUISE ID
AT42-17

SUMMARY
Project: Collaborative Research Pythias Oasis - Access to Deep Subduction Zone Fluids
Cruise Type: ep_science
DOI: 10.1234/56789
Vessel: Atlantis
Chief: Kelley, Deborah

This is a list of all files in the data distribution for the selected cruise that are not available and downloadable on the website. Please fill out the contact information below and then select the data you wish to request. We will contact you about when we need to clear release of the data with the chief scientist.

Name: 
Email: 

Main purpose:
- Commercial
- Research
- Personal
- Other

Requested data:
- Click on the submit button after selecting files below.

Submit

- Request entire cruise distro (1175 files, total size: 1.25GB)
- or select from below (click on folder icon to expand/contract, click on file or directory name to select):
  - AT42-17_cruise_letter.pdf (Size: 1223KB, Checksum: 31256d59a7e5e004457a)
  - AT42-17_datebook.pdf (Size: 256KB, Checksum: 31256d59a7e5e004457a)
  - CRUICE_ID (Size: 256KB, Checksum: 31256d59a7e5e004457a)
  - cruise.metadata (Size: 468KB, Checksum: 31256d59a7e5e004457a)
  - ctd (Size: 1.52GB, Checksum: 31256d59a7e5e004457a)
  - big
    - adig_files.txt (Size: 819KB, Checksum: 31256d59a7e5e004457a)
    - cruise.metadata (Size: 1.94GB, Checksum: 31256d59a7e5e004457a)
    - blog_data_descriptions.txt (Size: 154MB, Checksum: 31256d59a7e5e004457a)
    - gravimeter.txt (Size: 1.73GB, Checksum: 31256d59a7e5e004457a)
    - magnetometer.txt (Size: 8188KB, Checksum: 31256d59a7e5e004457a)
    - nanobeam
    - ctf
    - shiptrack
    - underway
API Services

Get information and data about:
- Cruises
- Devices
- File Manifests
- Filesets and Products
- Documentation
- Vessel Operator Summaries
- Vocabularies
- ... and more

Explore services at R2R API Services to see everything that is available.

https://www.rvdata.us/about/technical-details/services/api
Acknowledgements

Providing access to and ensuring the preservation of national oceanographic research data.
We’re hiring a Systems Analyst/Programmer!

Responsibilities:
● Work on the development of software tools for front-end and back-end applications (using Javascript, Python, and/or Java)
● Troubleshoot and resolve moderate to complex logic problems
● Support all phases of testing leading to deployment
● Address UI/UX needs of applications/systems
● Maintain and develop tools for:
  ○ Data curation
  ○ API-based search and discovery of geospatial data
  ○ (Meta)data exchange and data system interoperability
  ○ Exploring and visualizing earth science data, including web map interfaces and Desktop tools

Scan code to be brought to job posting!
Questions, Comments, Suggestions?

Contact us:

- Email info@rvdata.us
- Visit www.rvdata.us
- @R2RData