Lessons Learned: A Data Repository's Experience Requiring Data Management Training

Rosalie Rossi* and James Gibeaut

Harte Research Institute for Gulf of Mexico Studies, Texas A&M University-Corpus Christi, Corpus Christi, Texas

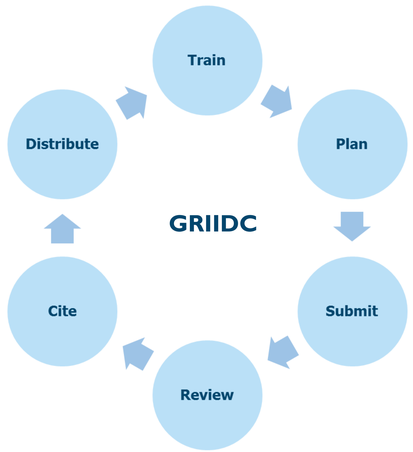
OUR MISSION

To ensure a data and information legacy that promotes continual scientific discovery & public awareness of the Gulf of Mexico ecosystem.

BACKGROUND

- BP oil spill in the Gulf of Mexico (2010)
- BP committed $500 million over 10 year period to fund the Gulf of Mexico Research Initiative (GoMRI)
- GoMRI researchers study Gulf of Mexico and oil spill science
- All data collected must be made available
- Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC) is the data repository for GoMRI data

GRIIDC DATA MANAGEMENT WORKFLOW



Train
- User guides
- Webinars

Plan
- Data Management Plan
- Dataset Information Form

Submit
- Dataset submission interface

Review
- Data package curation

Cite
- DOIs assigned to each dataset

Distribute
- Data available on website and ERDDAP server

GRIIDC DATA DIVERSITY BY THE NUMBERS

2249 DATASETS | 282 RESEARCH GROUPS | 2679 PEOPLE | 71 TERABYTES

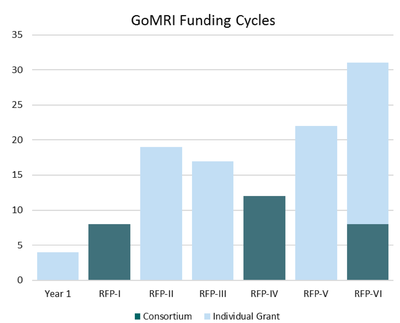


Figure 1. Number of GoMRI research groups by funding cycle. Funding cycle indicated by first year of funding (Year 1) or RFP (Request for proposal) number.

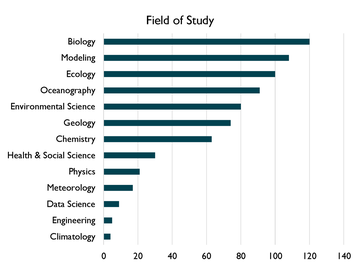


Figure 2. Number of datasets identified in a specific field of study from theme keywords provided by dataset submitters in metadata as of 2017 (n=722).

WHY REQUIRE TRAINING?

For RFP I-V, training and resources were available to help researchers with data submission and organization but were not required.

Common Issues:
- Poorly organized data
- Data in proprietary formats
- Issues opening data files
- Data not well described
- Data file names not descriptive
- Reluctance to share data

GRIIDC Training & Resources
- Tutorial videos
- User guides
- Data management planning resources
- Webinars
- Workshops
- One-on-one training

GRIIDC REQUIRED TRAINING

In 2017, in order to facilitate dataset submission and improve organization of data, RFP-VI PIs, co-PIs, and data managers were required to complete three trainings by webinar or workshop. Webinars were offered to all RFP-VI researchers (General), to research consortia (Group), or to individual researchers (Individual).

- **Introduction to the GoMRI Data Management Program** - introduction to the GoMRI data management program and GRIIDC data management planning requirements
- **Organizing Data - Best Practices and GRIIDC Submission -** overview of general best practices for organizing and managing data including file and folder organization, documentation, and backup
- **How to Submit Data to GRIIDC** - demonstration of how to use the GRIIDC Dataset Submission interface to submit data to GRIIDC and obtain a DOI

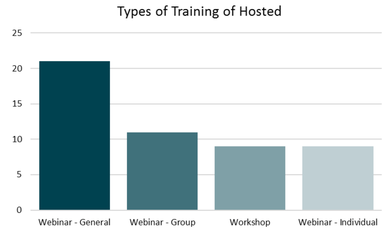


Figure 3. Number of trainings hosted by type.

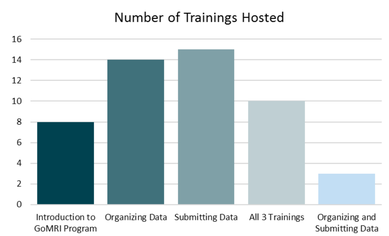


Figure 4. Number of trainings hosted by subject.

WEBINARS AND WORKSHOPS

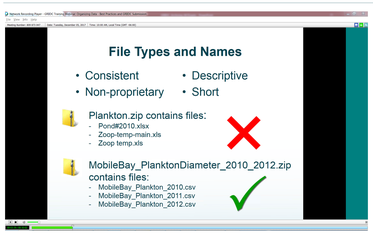


Figure 5. Webinar slide from the Organizing Data Webinar.

Figure 6. Workshop at the 2018 Gulf of Mexico Oil Spill & Ecosystem Science (GOMOSES) Conference.

TRAINING PARTICIPATION

WEBINARS
- Cisco Webex feature tracks webinar attention
- Attention: in-focus on attendee's computer (not clicking on another application)
- Attention to Attendance ratio: attentiveness based on time in session
- Attention to Duration ratio: attentiveness based on session duration

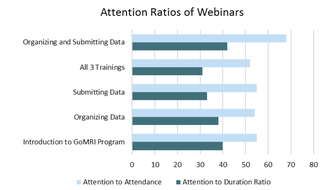


Figure 7. Cisco Webex attention ratios (%).

WORKSHOPS
- Workshops held during the 2018 GOMOSES Conference
- One 3 hour workshop for all 3 trainings
- Six 30 minute individual training workshops

Figure 8. Number of workshop participants by subject.

ISSUES ENCOUNTERED

- Individual webinars had to be scheduled to accommodate researcher's schedules
- Researcher's attendance had to be tracked for compliance
- Low attention rates
- Diverse attitudes and experiences sharing data
- Webinars could not be updated in order to keep consistent between researchers

FUTURE IMPROVEMENTS

- Use software to track presentation views
- Make webinars more interactive
- Host additional in-person workshops
- Require a quiz at the end of the training
- Allow researchers to test out of training

CONTACT

Rosalie Rossi
Program Manager
(361) 825-3719
Rosalie.Rossi@tamucc.edu
General Inquiries: griidc@gomri.org

ACKNOWLEDGEMENTS

This work was made possible by a grant from the Gulf of Mexico Research Initiative and support from the Harte Research Institute for Gulf of Mexico Studies at Texas A&M University-Corpus Christi.

Figure 9. The GRIIDC team at the 2018 staff retreat.