NDSF Data Management Update

Vicki Ferrini - NDSF Associate Director for Data System Strategy
Tina Haskins - NDSF Associate Director for Data and Science Ops
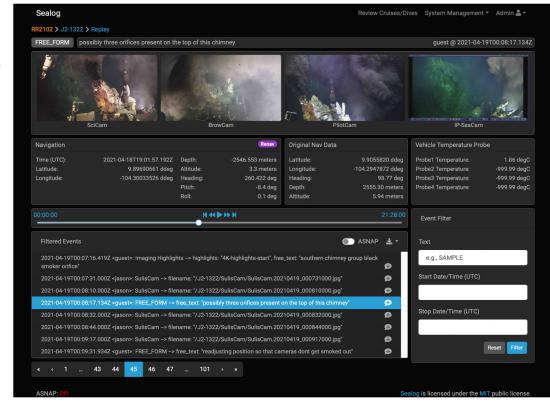
NDSF NATIONAL DEEP SUBMERGENCE FACILITY

Sealog Improvements - Phase 1

- **UI Updates** - devoting significant effort to modifying Sealog to improve usability and to streamline back-end components to help make future changes easier to achieve.
  - Specific focus on addressing user requests with Sealog interface improvements.
- **Code consolidation** - normalized versions used across vehicles
- **Containerization** - to facilitate and simplify deployment
  - Smaller update packages allowing for at sea updates/bug fixes if necessary.
- **Continuous integration** - more efficient code validation and testing
- Work is on-going and overseen by NDSF Data Leadership
- Deployed on recent Beinart cruise

Data QA/QC

- Goals
  - To provide timely feedback to at-sea operations team to ensure that
    - Intended data were acquired and sensors are working properly
    - Data high quality standard are met
- Strategy
  - Developing a framework for tests and code that will enable cross-vehicle implementation
  - Leverage R2R QA approach
  - Seek input from domain specialists where possible
- Status & Timeline
  - Currently laying the framework to identify tests, criteria and thresholds
  - Development to begin in coming months

Directory Structure Standardization

- Goals
  - Standardize data structures and at-sea protocols to achieve cross-vehicle consistency
  - Improve user experience for scientists and down-stream data users
  - Gain efficiency with code and operations
- Strategy
  - Review current structure
  - Develop new/consistent structure & implementation plan
  - Update documentation
- Status & Timeline
  - Initial plan developed
  - Expecting to make progress on vehicle directory structure unification in the coming months.

Review/Update NDSF Data Documentation

- Goals
  - Review and update NDSF Data Policy to ensure it is clear, current
- Strategy
  - Refined language and improved flow
  - Clarify the important idea that the facility is NOT responsible for archiving of PI-provided equipment/sensors
- Status & Timeline
  - DeSSC review underway
  - Once updated policy is posted, data distribution documentation will be reviewed and normalized to be consistent with revised facility-wide directory structure

Data Access via MGDS

- Goals
  - Enable broad community/public access to NDSF data via MGDS
  - Integrate NDSF data with other submersible data from US research community
- Strategy
  - Leverage tools/workflows already built by MGDS
  - Prioritize data for migration/ingestion starting with geophysical data
- Status & Timeline
  - Start with Sentry data post-2018, then move on to other vehicles
  - In parallel, review new data and begin to build workflows that leverage new directory structure.
  - Develop strategy for prioritizing ingestion of older data

Ongoing Endeavors

- Documentation and reports under development
  - Documentation of modifications and plans for Sealog
  - Report describing the new NDSF directory structure
- NDSF Website
  - Ongoing site improvements
  - Planned training section for videos and documentation
- Questions/Suggestions
  - Feedback on Data Policy Update?
  - Reach us at ndsf_info@whoi.edu