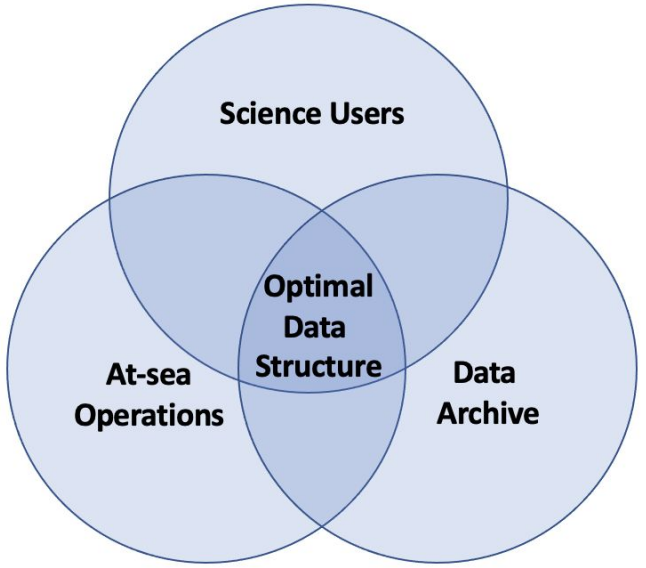
Developing a New Vision for NDSF Data Management

Vicki Ferrini & Tina Haskins

NDSF NATIONAL DEEP SUBMERGENCE FACILITY

Vision



- Develop a new strategy for the management and delivery of NDSF data that meets the needs of at-sea scientists and operations teams and down-stream data repositories that serve the broader science community and public
- Ensure that data management activities, and the maintenance and development of software that supports data management and access, are coordinated and are moving toward a common vision
- Standardize data structures and at-sea protocols to gain efficiencies and achieve cross-vehicle consistency
- Optimize the management, versioning, and maintenance of software tools

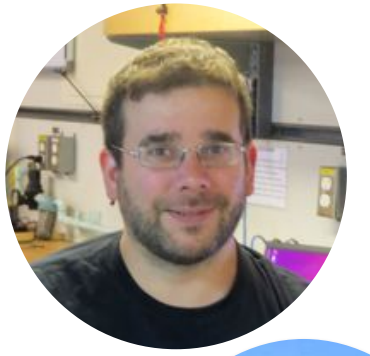
New NDSF Data Team

Software Team

Data Archiving and Access Team

Associate Data Director

Vehicle Data Staff



- Vicki Ferrini: NDSF Data Director
  - Vision & Strategy
  - Expertise: Shoreside Data
- Tina Haskins: NDSF Associate Data Director
  - **NDSF Data POC - ndsf_info@whoi.edu**
  - Expertise: At-Sea Data
- Stefano Suman: Software Lead
  - Expertise: Acquisition and Systems

Year 1 Goals

Document & review current at-sea practices

At-sea SOPs, metadata, data directory structures, naming conventions.

Identify vehicle-specific successes, bottlenecks and challenges

Leverage successes and strive toward facility-wide solutions.

Develop cross-vehicle guidelines

Where possible, develop cross-vehicle standardized directory structure, file naming conventions, formats, and generating machine-readable metadata.

Collaboratively design and implement a new structure

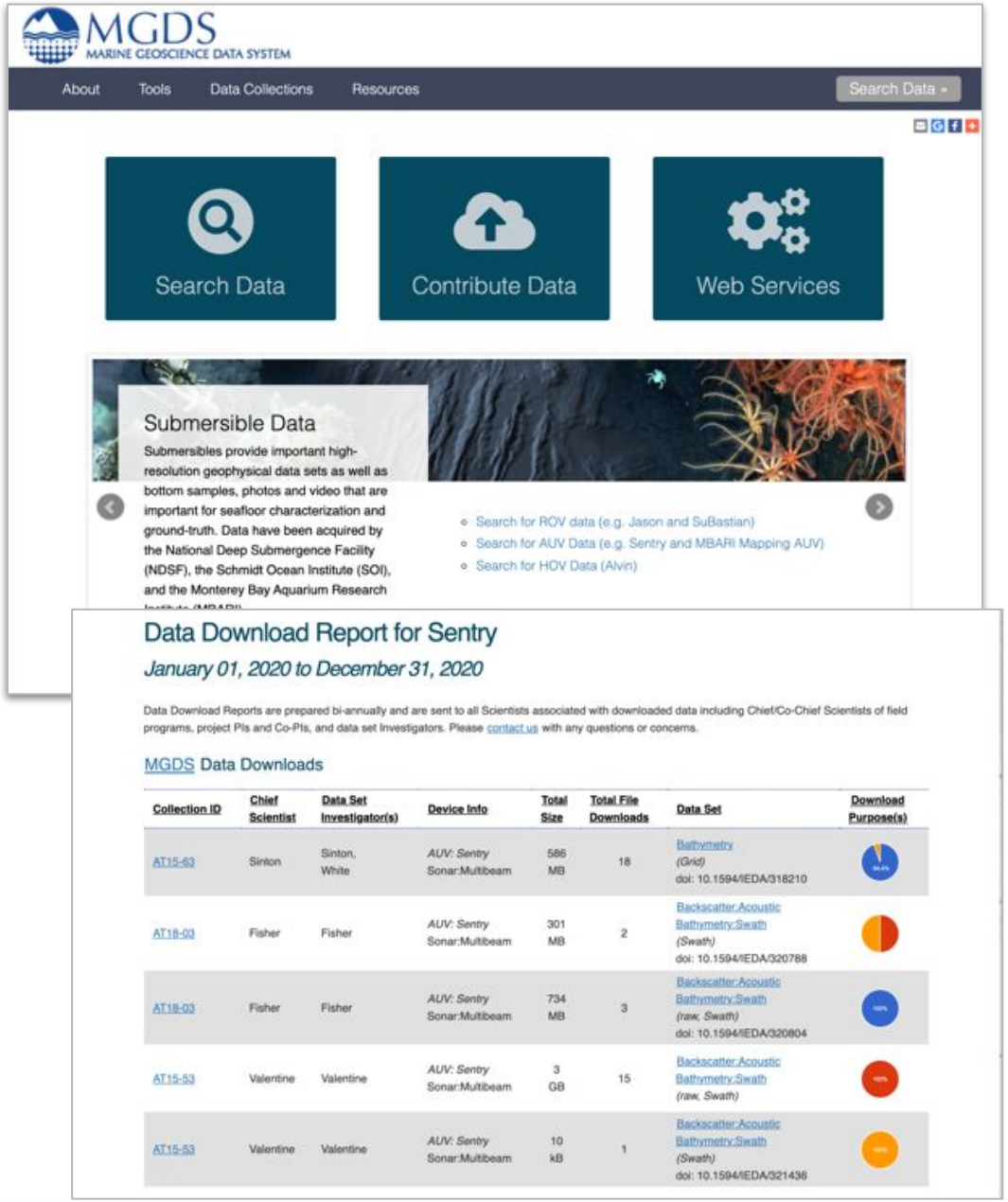
Improve efficiencies for at-sea operators and scientists, shoreside data managers, standardize data distributions and optimize data and metadata packages.

Long-Term Goals

- Improve shore-side data management to broaden access
- Work with WHOI Data Library to seek efficiencies and data redundancy
- Leverage efforts within MGDS for submersible data discovery and access
- Make data available for next-generation processing and integration (e.g. AI/ML)
- Curate NDSF data holistically as a coherent collection of data
- Address existing NDSF data that is not yet directly accessible to the community
- Make NDSF data resources available and interoperable with other similar data acquired by other groups both domestically and internationally
- Leverage evolving community solutions for video archiving and access

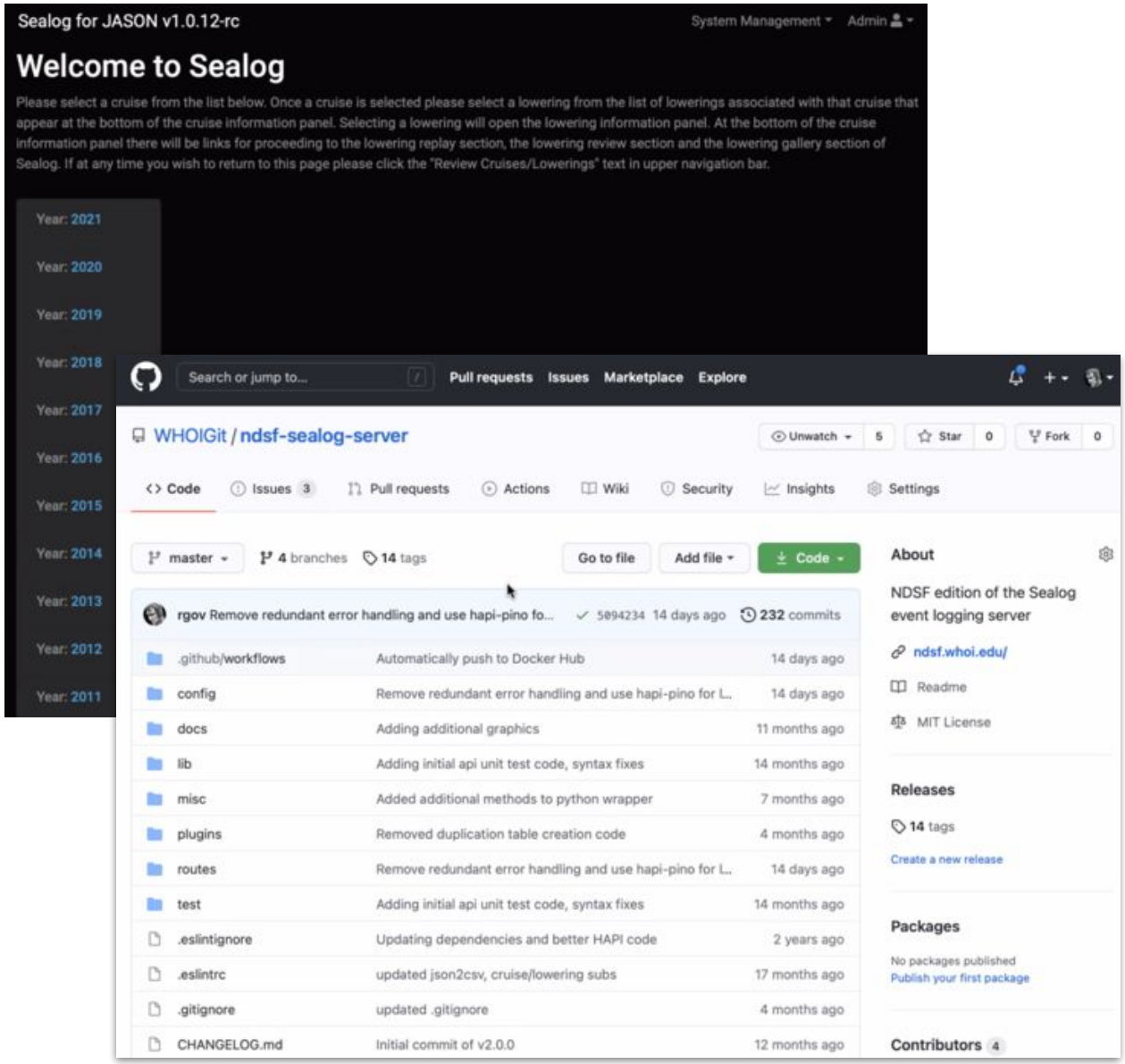
Submersible Data at MGDS

- Submersible data from 21 vehicles
  - Total: 6 TB, > 1 million files, 1,400 dives, 137 cruises
  - NDSF: 1.6TB, >160k files, 830 dives, 85 cruises
- Data Types:
  - Dive metadata
  - Dive logs and reports
  - Bathymetry, Magnetics, Sidescan, Navigation, Subbottom, Interpretations, Heatflow, Turbidity, Eh, Photomosaic, Photos, FrameGrabs, KMLs, (including Sealog), and more…
  - Virtual Van and FrameGrabber APIs/UIs
- User Interfaces for discovery and self-service
- Web services for interoperability



Sealog Updates

- **Code consolidation** - normalized versions used across vehicles
- **Containerization** to facilitate and simplify deployment
- **Continuous integration** - more efficient code validation and testing
- COVID-response: modifications for shoreside watchstander logging
    - Optimize bandwidth use via shore-side caching
- **Back-end modifications** position code base for more rapid development, diagnostics and testing
- New ideas for improving software architecture to gain further efficiencies
- Work is on-going and will be overseen by NDSF Data Leadership



NDSF Website

- Currently under renovation
- Feature updates
  - Training videos
  - Documentation
  - Newsletter Archive
- NDSF Newsletter Sign Up

https://ndsf.whoi.edu/sign-up-for-the-nsdf-newsletter/

- Questions/Suggestions? Please contact us via ndsf_info@whoi.edu