



6-2017

2017 DWH Long-Term Data Management Coordination Workshop Report

Coastal Response Research Center (CRRC)

Follow this and additional works at: <https://scholars.unh.edu/crrc>

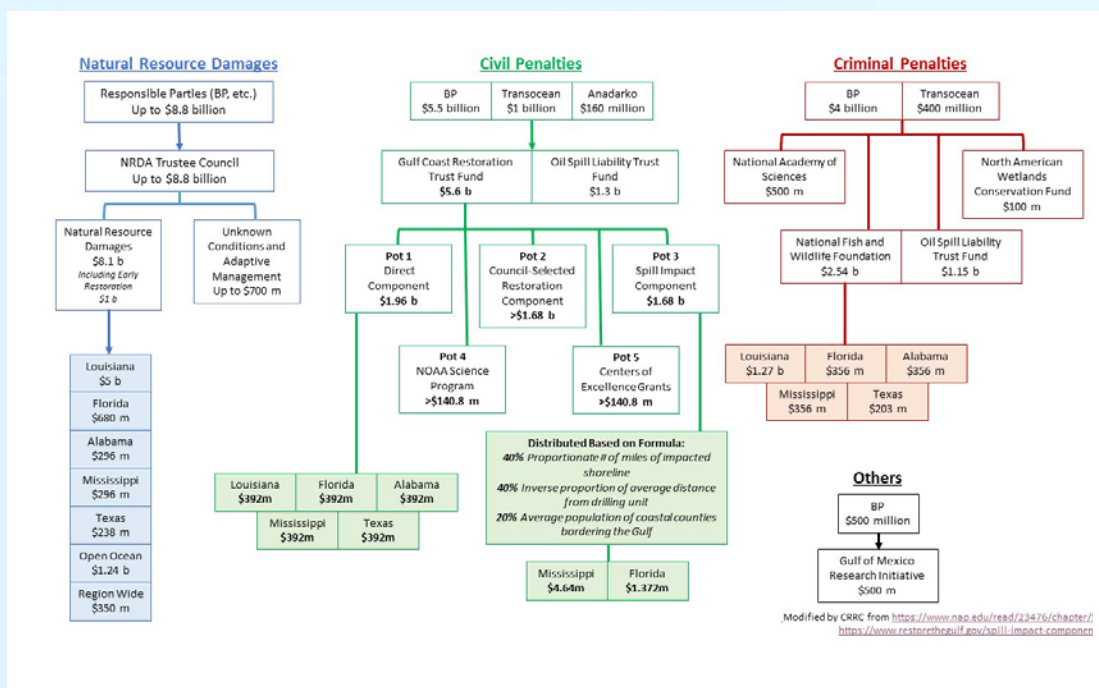
Recommended Citation

Coastal Response Research Center (CRRC), "2017 DWH Long-Term Data Management Coordination Workshop Report" (2017).
Coastal Response Research Center. 5.
<https://scholars.unh.edu/crrc/5>

This Workshop Report is brought to you for free and open access by the Research Institutes, Centers and Programs at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Coastal Response Research Center by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.

DWH LONG-TERM DATA MANAGEMENT COORDINATION WORKSHOP

JUNE 7 - 8, 2017



NOAA'S GULF OF MEXICO DISASTER RESPONSE CENTER
MOBILE, AL

WORKSHOP REPORT

COASTAL RESPONSE RESEARCH CENTER



Acronyms

BOEM	Bureau of Ocean Energy Management
CIMS	Coastal Information Management System
CMAWG	Council Monitoring and Assessment Workgroup
COC	Containment of Concern
COP	Common Operating Picture
CPRA	Coastal Protection and Restoration Authority
CRRC	Coastal Response Research Center
CWA	Clean Water Act
DIVER	Data Integration, Visualization, Exploration, and Reporting
DOI	Digital Object Identifier
DRC	Disaster Response Center
DWH	Deepwater Horizon Oil Spill
DWHPT	Deepwater Horizon Project Tracker
EDDM	Environmental Disaster Data Management
EHS	Environmental Health and Safety
ERMA	Environmental Response Management Application
ESRI	Environmental Systems Research Institute
FGDC	Federal Geographic Data Committee
FIO	Florida Institute of Oceanography
FTP	File Transfer Protocol
FWRI	Fish and Wildlife Research Institute
GCOOS	Gulf of Mexico Coastal Ocean Observing System
GEOSS	Global Earth Observing System of Systems
GIS	Geographic Information System
GOM	Gulf of Mexico
GOMA	Gulf of Mexico Alliance
GoMRI	Gulf of Mexico Research Initiative
GOOS	Global Ocean Observing System
GRIIDC	Gulf of Mexico Research Initiative Information and Data Cooperative
GRP	Gulf Research Program
ICP	Initial Comprehensive Plan
IEc	Industrial Economics, Inc.
IOOS	Integrated Ocean Observing System
IRB	Institutional Review Board
ISO	International Organization of Standardization
LTDM	Long Term Data Management
MAM	Monitoring and Adaptive Management
NARA	National Archives and Records Administration
NAS	National Academy of Sciences

NCEI	National Centers for Environmental Information
NFWF	National Fish and Wildlife Foundation
NGO	Non-Governmental Organization
NOAA	National Oceanic and Atmospheric Administration
NRDA	Natural Resource Damage Assessment
OAIS	Open Archival Information System
OPA 90	Oil Pollution Act of 1990
ORR	Office of Response and Restoration
PI	Principal Investigator
QA	Quality Assurance
QC	Quality Control
RESTORE	Resources and Ecosystems Sustainability, Tourist Opportunities, and Revived Economies
SCAT	Shoreline Cleanup Assessment Technique
SOP	Standard Operating Procedures
SRM	Standard Reference Material
TIG	Trustee Implementation Group
USGS	United States Geological Survey

I. Acknowledgements

The content for this workshop was funded by the National Oceanic and Atmospheric Administration's (NOAA's) Office of Response and Restoration and National Marine Fisheries Service (NMFS) Restoration Center. NOAA's National Centers for Environmental Information (NCEI) was also involved in the development of the workshop. The following individuals served on the Organizing Committee:

- Jonathan Blythe, BOEM
- Julie Bosch, NOAA NCEI
- Laura Bowie, GOMA
- Libby Featherston, FIO
- Jim Gibeaut, GoMRI
- Jessica Henkel, RESTORE Council
- Steve Jones, Geological Survey of AL
- Barb Kirkpatrick, IOOS
- Kirsten Larsen, NOAA NCEI
- Matt Love, Ocean Conservancy
- Laurie McGilvary, Dept. of Treasury
- Amy Merten, NOAA ORR
- Tamay Ozgokmen, University of Miami
- Mike Peccini, NOAA NMFS
- Jon Porthouse, NFWF
- Jamey Redding, NOAA RC
- Dave Reed, FL FWRI
- Lauren Showalter, NAS
- Greg Steyer, USGS

The workshop was facilitated by Dr. Nancy Kinner from the Coastal Response Research Center (CRRC; www.crrc.unh.edu). CRRC has extensive experience with issues related to oil spills. The Center is known for its independence and excellence in the areas of environmental engineering, marine science, and ocean engineering as they relate to spills. CRRC has conducted numerous workshops bringing together researchers, practitioners, and scientists of diverse backgrounds (including from government, academia, industry, and non-governmental organizations) to address issues in spill response, restoration and recovery.

We wish to thank all the presenters for their participation in the workshop:

- Julie Bosch, NOAA NCEI
- Laura Bowie, GOMA
- Craig Conzelmann, USGS
- Jim Gibeaut, GRIIDC
- Jessica Henkel, RESTORE Council
- Matt Howard, GCOOS
- Dan Hudgens, IEC
- Michele Jacobi, NOAA ORR
- Larry Langebrake, Connectsix
- Matt Love, Ocean Conservancy
- Laurie McGilvary, Dept. of Treasury
- Marti Goss McGuire, NOAA RC
- Amy Merten, NOAA ORR
- William Nichols, GRIIDC
- Rost Parsons, NOAA NCEI
- Mike Peccini, NOAA RC
- Jon Porthouse, NFWF
- Steve Ramsey, Social & Scientific Systems
- Jamey Redding, NOAA RC
- Ben Shorr, NOAA ORR
- Lauren Showalter, NAS
- Greg Steyer, USGS

We would like to thank Mark Miller, Dan Hudgens, Rost Parsons, Steve Ramsey, Larry Langebrake and Ann Jones for being breakout group leaders as well as Jay Coady, Nadia Martin, JB Huett, Megan Verfaillie, Kathy Mandsager and Melissa Gloekler for their note taking during the workshop.

We thank the ORR Gulf of Mexico (GOM) Disaster Response Center (DRC) for hosting the workshop.

II. Introduction

On June 7 and 8, 2017, the Coastal Response Research Center (CRRC)¹, NOAA Office of Response and Restoration (ORR) and NOAA National Marine Fisheries Service (NMFS) Restoration Center (RC), co-sponsored the Deepwater Horizon Oil Spill (DWH) Long Term Data Management (LTDM) workshop at the ORR Gulf of Mexico (GOM) Disaster Response Center (DRC) in Mobile, AL.

There has been a focus on restoration planning, implementation and monitoring of the on-going DWH-related research in the wake of the DWH Natural Resource Damage Assessment (NRDA) settlement.

This means that data management, accessibility, and distribution must be coordinated among various federal, state, local, non-governmental organizations (NGOs), academic, and private sector partners. The scope of DWH far exceeded any other spill in the U.S. with an immense amount of data (e.g., 100,000 environmental samples, 15 million publically available records) gathered during the response and damage assessment phases of the incident as well as data that continues to be produced from research and restoration efforts. The challenge with the influx in data is checking the quality, documenting data collection, storing data, integrating it into useful products, managing it and archiving it for long term use. In addition, data must be available to the public in an easily queried and accessible format.

Answering questions regarding the success of the restoration efforts will be based on data generated for years to come. The data sets must be readily comparable, representative and complete; be collected using cross-cutting field protocols; be as interoperable as possible; meet standards for quality assurance/quality control (QA/QC); and be unhindered by conflicting or ambiguous terminology.

During the data management process for the NOAA Natural Resource Damage Assessment (NRDA) for the DWH disaster, NOAA developed a data management warehouse and visualization system that will be used as a long term repository for accessing/archiving NRDA injury assessment data. This serves as a

¹ A list of acronyms is provided on Page 1 of this report.

foundation for the restoration project planning and monitoring data for the next 15 or more years. The main impetus for this workshop was to facilitate public access to the DWH data collected and managed by all entities by developing linkages to or data exchanges among applicable GOM data management systems.

There were 66 workshop participants (Appendix A) representing a variety of organizations who met at NOAA's GOM Disaster Response Center (DRC) in order to determine the characteristics of a successful common operating picture for DWH data, to understand the systems that are currently in place to manage DWH data, and make the DWH data interoperable between data generators, users and managers. The external partners for these efforts include, but are not limited to the: RESTORE Council, Gulf of Mexico Research Initiative (GoMRI), Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC), the National Academy of Sciences (NAS) Gulf Research Program, Gulf of Mexico Alliance (GOMA), and National Fish and Wildlife Foundation (NFWF).

The workshop objectives were to:

- Foster collaboration among the GOM partners with respect to data management and integration for restoration planning, implementation and monitoring;
- Identify standards, protocols and guidance for LTDM being used by these partners for DWH NRDA, restoration, and public health efforts;
- Obtain feedback and identify next steps for the work completed by the Environmental Disasters Data Management (EDDM) Working Groups; and
- Work towards best practices on public distribution and access of this data.

The workshop consisted of plenary presentations and breakout sessions. The workshop agenda (Appendix B) was developed by the organizing committee. The workshop presentations topics included: results of a pre-workshop survey, an overview of data generation, the uses of DWH long term data, an

overview of LTDM, an overview of existing LTDM systems, an overview of data management standards/protocols, results from the EDDM working groups, flow diagrams of existing data management systems, and a vision on managing big data.

The breakout sessions included discussions of: issues/concerns for data stakeholders (e.g., data users, generators, managers), interoperability, ease of discovery/searchability, data access, data synthesis, data usability, and metadata/data documentation.

III. Plenary Presentations

The workshop presentations topics included: results of a pre-workshop survey, an overview of data generation, the uses of DWH long term data, the overview of LTDM, the overview of existing LTDM systems, existing data management standards/ protocols, the products of EDDM working groups, flow diagrams of existing data management systems, and a vision on managing big data. Most of the speakers provided a summary of their presentations (below) and presentation slides are located in Appendix C.

Survey Results

Jessica Henkel (RESTORE Council) presented the results of a pre-workshop survey (Appendix D) which collected information on the perspective of the participants regarding data management and their goals and objectives of LTDM of DWH data. Of the 47 survey responses received, 55% of participants described themselves as data managers or administrators, 17% as data users, 15% as program managers or funders, 10% as data generators, and 2.1% as decision makers. The majority of respondents wanted GOM research/monitoring data over the next 15 years follow a common set of standards, be accessible and interoperable for all users, and be stored in a long term data repository. However, they were not optimistic about that being achieved. Many saw developing and adhering to a common set of data standards across GOM data generators as one of the biggest challenges for GOM LTDM.

Overview of Data Generation of the DWH Oil Spill

Michele Jacobi (NOAA ORR) described the DWH Oil Spill from the perspective of a data generator. The DWH incident falls far outside of the “normal” spill in terms of data generation with 20,000 trips to the field to collect data, 100,000 environmental samples collected and 15 million records publically available.

The DWH spill affected five states (TX, LA, AL, FL, MS) which became the focus areas for data collection. The DWH data aided the ecosystem overview and helped determine the actual impact of the oil at each location. There was a heavy reliance on technology to capture the impacts of the oil spill. Data sets came from the principal investigators (PIs), NGOs, state and federal agencies, academic institutions and independent parties. Much of the data was stored in the Environmental Response Management Application platform (ERMA®) which served as the common operational picture during the response. ERMA showed the results of NOAA's oil trajectory modeling and where clean up already took place; locations for Shoreline Cleanup Assessment Technique (SCAT) and scientists in the field; and aided in public transparency. Additionally, it supported the NRDA process and includes monitoring data generated during the restoration.

Response data includes public safety, response activities, SCAT, closures and advisories to inform recreation effects; provides evidence of exposure; and documents the extent of the spill. NRDA activities during the response included: (1) setting a baseline, collecting ephemeral data to document conditions before and after the spill, estimating fish kills, (2) fingerprinting of oil on shorelines, and in the water column, and (3) studying changes to recreation use and socio-economic impacts. Understanding and capturing these pieces of information aids in planning restoration activities. On-going monitoring is required to determine how the various restoration programs are progressing and meeting their objectives; and whether they are necessary vs. natural recovery.

The large influx of data throughout and after the spill helped determine best practices for data collection and management; data documentation became extremely important and scientists needed to work with data managers to make the data useful. Data remains accessible to future users in databases and may be accessed by the public through data repositories. There were many data generation lessons learned throughout the DWH spill, including having: a strong sampling design, multi-disciplinary questions, coordination across lab studies and field studies, clearly defined objectives of data collection

relative to action or decision making, a budget to perform data collection and management, and adherence to existing federal requirements and standards.

Uses of DWH Long Term Data

Matt Love (Ocean Conservancy) presented many uses for DWH data as well as overview of data products and end users. His presentation stressed the large scale marine ecosystem restoration process and the numerous data sets and types to be collected during GOM restoration. The only other scientific effort that compares in terms of data collection is the Census of Marine Life, a Decade of Discovery. The Census of Marine Life is a good representation of what can be done with DWH data, in the ways it facilitates free and open access to data, integrates existing data with new surveys to establish a complete picture, and compiles a data assimilation framework.

The goal of the DWH LTDM is to create an interoperable infrastructure that allows for data sharing and accessibility. Prior to DWH, the discussion of data management often stopped at the generators. Moving forward, generators need to collaborate with end users to expand planning through data use and synthesis. Data generators should envision a data system that enables an end goal that allows for development of data products to aid decision making and long term resource management.

A network of data users will rely on data and synthesized data products to make informed decisions (e.g., business or NGO research, response, restoration, management) based on their shared stake in the ultimate outcome of restoration. Spill responders will need access to real-time data and a common operation picture to help support functional decision making to control environmental damages. Using the generated data, the research community has many opportunities to collaborate on assessing the status and trends of GOM ecosystems to guide restoration and long term management strategies. The full-scale restoration process requires a unique set of data to inform what actions must be taken to repair the full suite of priority damages from the spill and long term degradation.

GOM restoration will be a long term process that will require enhanced analytical data applications such as ecosystem modeling. Models are data hungry and require vast amounts of information. Models can assist decision makers in determining if proposed restoration activities will help an area or ecosystem component of concern. They can serve as the foundation for assessing the value of implemented if restoration actions and to help determine if changes to implemented actions should be made.

There is a vast array of monitoring programs to inform GOM management and monitoring targeted at DWH restoration, all initiated by different organizations and funders. The data from both sources will serve the needs of the broader management community, and provide added benefits for enabling collaborative science to address priority questions pertaining to recovery and management of the GOM. The long term vision for these diverse programs (i.e., 15+ years) is moving towards a collaborative restoration effort. A successful restoration outcome hinges on reasonable decisions based on open, accessible data that can be synthesized by a variety of users. Restoration programs and data generators, as defined in this workshop, must envision the potential uses for the data beyond their immediate application through insuring data accessibility for future applications and broader scientific inquiries.

Overview of Long Term Data Management (LTDM)

Lauren Showalter (NAS) gave an overview of steps taken to ensure the legacy of science that came from DWH. The NAS focuses on making data accessible to researchers studying future disasters, and ensuring the research products are well documented and in stable formats. In order to frame the discussion of LTDM as it relates to the DWH disaster, the NAS identified a number of key topics for initial discussion: metadata, standards, federal mandates, data sharing, and interoperability. A basic overview of these terms was presented to help frame discussions over the course of the workshop.

DWH data types vary from human health, restoration and monitoring, oil systems safety, environmental, social science to real time data. The data is from a variety of sources and will be managed by private, state, and federal archives. Due to the interdisciplinary research products coming from the generated data, it is important to make the data interoperable between archives and repositories. The presentations noted specific standards and recommendations that have been used by ongoing GOM data efforts to ensure ease of collaboration as programs develop. The terms data archive and data repository were clarified vis-a-vis federal requirements; although these terms are sometimes used interchangeably they have distinct meanings within the federal data structure.

Data documentation (i.e., metadata) must be done well and standardized in order to make data usable and accessible in future. For example, the International Organization of Standardization (ISO) 19115 standards should be adopted by future data generators. The challenge is getting the research community to actually follow standards and share data.

Data citation supports proper attribution and credit of the data generator which facilitates the future use and collaboration between researchers and their data. It enables reproducibility of findings and fosters faster and more efficient research progress. Ensuring data is properly attributed and documented is essential for ease of redistribution and reuse. The creation of standard, machine readable metadata, the use of digital object identifiers (DOI), and adherence to data collection standards are important aspects of the data management process. As more data is collected, the use of tools and distributed data frameworks can improve interoperability and facilitate data synthesis. The use of data visualization can also help display the value of complex datasets and increase their use for other purposes than they were originally collected.

Accessibility of data is important to determining baselines of data, and determining what conditions were known prior to a disaster. Availability of data also provides opportunity for users to retrieve data

for exploration, analysis and decision making. Additionally, data interoperability is needed to compare metrics and baselines to better understand monitoring data and allow human readable and machine-to-machine compatibility. Consistency with metadata helps data sharing, which is essential when using real time data. Ideally, there would be an existing framework allowing data users to search by “text” or “keywords”. These frameworks must portray data in an aesthetic and easy-to-use manner. Data visualization aids and good data management practices allow data to be easily reused and synthesized to develop useful products.

IV. Overview of Existing LTDM Systems: Speed Presentations

Data Integration, Visualization, Exploration, and Reporting (DIVER) & Environmental Response Management Application (ERMA)

Ben Shorr (NOAA ORR) presented two data synthesis/management systems: DIVER and ERMA®. These two programs require standardized data and contain thorough metadata to facilitate data sharing and exchange. DIVER is a data warehouse, query tool, and collaboration application. DIVER integrates standardized datasets so users can query across data holdings and download information and results. It is a warehouse that can accommodate various formats of raw data and integrates it into a common format that many users can query and download. DIVER is also a one-stop repository for those working on a spill or site to submit their data and quickly transform it into a usable format for tracking, reporting and analysis from response to restoration. It was designed for incorporation of data coming in through multiple pathways including, quick provision field data and laboratory results that have been through the QA/QC process. DIVER has key capabilities to file and load collections from field forms, Contaminates of Concern (COCs), photos and notes. It allows multiple users to examine raw data that was collected in the field, and data managers to transcribe or input it into the new DIVER platform for application. The program transcribes, processes, and parses the data appropriately based on the DIVER common data

model; after transcription the data is available for query, export and loading into ERMA for visualization purposes.

ERMA is an online geographic information system (GIS) and visualization tool that allows users to view response, assessment, and restoration mapping layers in context with other environmental information. ERMA is available for the Pacific Northwest, Pacific Southwest, Great Lakes, the Atlantic Coast, and the GOM. Early in the DWH response, a GOM ERMA was created. It has standard layers specialized for each region/state and is accessible to the public. Some data is privileged and requires log-in for the use of additional tools. It functions as a common operating picture (COP) that was used for the NRDA. It can also be used to plan and monitor restoration efforts. It is a system that allows others to load and exchange spatial data from state, federal, NGO, tribal and academic organizations.

DIVER Portal Restoration Tracking

Mike Peccini (RC) presented a brief overview of how the RC is using DIVER to manage DWH restoration project tracking and monitoring data. Restoration project tracking within DIVER was developed to meet data management and reporting requirements outlined the DWH Consent Decree, Trustee Council Standard Operating Procedures (SOPs), and the DWH Monitoring and Adaptive Management Manual.

The DIVER portal provides project managers from all state and federal trustee agencies with role-based access to common workspaces and data entry tools allowing for distributed management of project-level data. Project tracking data includes information describing: project status, location(s), budget, implementation activities, accomplishments, and monitoring results. Project data is being used to serve Trustee Council reporting needs and to inform the public via maps, dashboards, project information pages, and query tools.

Project monitoring field data will be managed within the DIVER data warehouse and made accessible through integrated DIVER search tools. To the extent possible, monitoring data will be standardized

within resource types and integrated into DIVER data models to maximize interoperability across Gulf monitoring and assessment data.

NOAA National Centers for Environmental Information (NCEI)

Rost Parsons (NCEI) provided an introduction to NCEI which is the data management and archive entity for oceanographic, geophysical, and climatological information within the U.S. NCEI is a science-based organization that produces environmental information to enable individuals, businesses, and governments to make informed decisions. It provides the foundation for more tailored decision-support services to be developed and delivered by the public and private sector.

The Open Archival Information System (OAIS) is a reference model which mandates an important set of responsibilities and functions for the archive to perform. It is not an implementation architecture, but a system that can ingest data and allow access to it. NCEI may expand the application of OAIS to be a more integrated and robust ingest service. This would enable NCEI to add additional stewardship services (e.g., automated QA, granule metadata generation). Currently, metadata standards being implemented are ISO 19115. Having standardized metadata ensures easier archiving and aids in data interoperability. The data archive must follow the National Archive and Records Administration (NARA) requirements, uses Library of Congress guidance for data formats, and applies common or managed vocabularies as a NCEI standard practice. As standards evolve over time, and the structure for metadata and its content should not be static but be adaptable as well. NOAA Administrative Order 212-15 (Management of Environmental Data and Information) cites overarching mandates from the Federal Records Management Act to the National Archives and Records Administration (NARA) and Agency Directives.

NCEI outlined Tiers of Stewardship to help organize data management. The Tier 1 Long Term Preservation and Basic Access of Data; this tier is concerned with preservation of the original data with

robust metadata for data discovery/access, and safeguarding the data over its entire life-cycle. Tier 2, is Enhanced Access and Basic Quality Assurance, entailing the creation of complete metadata to enable automated quality assurance and statistic collection, and improved overall data access through specialized software services for users and applications. Tier 3 is Scientific Improvements to allow data to be reprocessed in new and improved versions to distribute to users. Additionally, it improves the data quality or accuracy with scientific quality assessments, controls, warning flags and corrections. Tier 4, derived products, builds upon archived data to create new products that are more broadly useful and analyzes/combines products and data to create new or blended scientific data products. Tier 5 is Authoritative Records that focus on combining multiple time series into a single, inter-calibrated product and establish authoritative quality and uncertainties, and ensure full documented and reproducible products. Tier 6, National Services and International Leadership, would lead coordination or implementation of scientific stewardship activities for a community across disciplines and establish highly specialized levels of data services and product assessments.

GRIIDC

Jim Gibeaut (GRIIDC) spoke on behalf of the GoMRI and GRIIDC; their objective is ensuring data access and an information legacy that promotes continual scientific discovery and public awareness for the GOM. GRIIDC serves the entire life cycle of data, beginning with planning for collection, tracking the process, providing proper documentation, archiving the information, and disseminating of the data. All data that is collected under GoMRI must be publically available and usable within one year or at the time of publication (whichever comes first). It is a repository for citable data packages for future users to access and validate results of their scientific research. The datasets are interdisciplinary (e.g., field and laboratory data), and the researchers must meet data management plans previously set GoMRI/GRIIDC. The data is tracked in order to ensure it meets data sharing standards, such as proper data

documentation. GRIIDC will support GoMRI data until at least 2030 and will hopefully expand to services beyond GoMRI-funded research to integrate with other repositories.

GOMA Portal

William Nichols (GRIIDC/Harte Research Institute) presented the GOMA Portal which is a data catalogue and repository that provides data discovery and access to GOM geospatial datasets. Users have the ability to search for data using a metadata catalog, topic, keyword and spatial search; view information; and download it via File Transfer Protocol (FTP). Most of the records reside on the server which allows for direct downloads. The GOMA Portal houses upwards of 800 data sets that do not have a stable repository. The organization wrote metadata for these 800 sets; the data came from a variety of sources with different documentation methods. The metadata follows Federal Geographic Data Committee (FGDC) standards and supports ISO 19115-2 and ISO 19119 (web services) as well as the Dublin Core for non-geospatial data types. The portal is based on an open-sourced Environmental Systems Research Institute (ESRI) geoportal platform, which allows interoperability between platforms.

DWH Project Tracker (DWHPT)

Laura Bowie (GOMA) presented the DWH Project Tracker (DWHPT) which is a system designed to track projects that are funded by programs resulting from the DWH – voluntary and negotiated settlement programs. The concept for a project tracker was developed by the GOM states as a way to try to understand what is being funded through the myriad of programs. The DWHPT only contains awarded projects (not proposed or unfunded) and it is currently 99% complete with 597 projects. The DWHPT categorizes projects in four primary types: environmental, human and social/planning, recreational use, and science/research. Each project “dot” on the map links to a “popup” box that provides basic information about the project and a link to the funding program’s database for more information. The system is queriable using a wide variety of metadata including geographic location. It also has some

standard “canned” reports. Primary users of the system tend to be the public, media, and restoration planning programs.

Louisiana Coastal Information Management System (CIMS)

Craig Conzelmann (U.S. Geological Survey (USGS)) spoke on behalf of the LA Coastal Information Management System (CIMS) Portal. Topics covered were: (1) observational data inputs, quality control, and downloads, (2) the CIMS spatial framework, and (3) the Coastal Protection and Restoration Authority’s (CPRA) use of a digital library.

CIMS is a suite of data driven applications and tools used to manage, visualize, share and analyze coastal data. A variety of data types (e.g., tabular, spatial, unstructured) can be used. Standardization is done through data documentation; metadata is currently required to follow FGDC formatting. Moving forward, it will be ISO standards. CIMS has a clean interface, mapping, data and library to ensure public usability. The library identifies documents by various type, project, name and location. There are options to add layers and control visibility which enhances the shareability of data. Contractors have the ability to enter data from the field on a Smartphone. The CIMS Portal only houses Louisiana data.

Gulf of Mexico Coastal Ocean Observing System (GCOOS)

Matt Howard (GCOOS) explained that the Gulf of Mexico Coastal Ocean Observing System (GCOOS) is one of 11 Regional Associations organized under the U.S. Integrated Ocean Observing System (IOOS) which is the U.S. contribution to the Global Ocean Observing System (GOOS), the oceanic component of the Global Earth Observing System of Systems (GEOSS). GCOOS’ goal is to deliver high-quality data from sensors to the desktop through networked systems without loss of information. GCOOS adheres to community standards and best practices in data stewardship and specializes in physical oceanographic, marine meteorological and biogeochemical data; and numerical model outputs. It has recently begun to work with marine biological data (e.g., plankton, fisheries). GCOOS works with near real-time data (i.e., 1600+ sensors), and delayed-mode data, and has extensive historical data collections including

climatology and quasi-static datasets (e.g., bathymetry, coastline). GCOOS serves data through standard interfaces in preferred formats and produces data-based products. GCOOS aggregates products from outside data collectors and combines data into a usable format. GCOOS a full-time education and outreach coordinator, holds stakeholder workshops, and hosts and serves data and products for Citizen Scientist groups. GCOOS is funded thorough 2021 and has 3+ full-time equivalents devoted to data management issues.

V. Data Management Standards and Protocols: Speed Presentations

Data Management Frameworks

RESTORE Council

Jessica Henkel (RESTORE) discussed the structure of the Gulf Coast Ecosystem Restoration Council (RESTORE Council), and the many funding recipients that will be generating data from RESTORE Council-funded activities. Her presentation discussed current data requirements for grant recipients. All data is to be digital and machine-readable, have the ability to be made publically available, and must comply with all federal laws and policies. In 2017, the Council staff will be exploring metadata development tools for funding recipients, and working with the Council Monitoring and Assessment Workgroup (CMAWG) to develop data management framework options for Council consideration.

NRDA Restoration Project Monitoring

Jamey Redding's (NOAA RC) presentation on NRDA restoration monitoring included a description of what was outlined in the Trustee Council SOP, what may be further developed within the Monitoring and Adaptive Management (MAM) Manual with the Cross-TIG Monitoring and Adaptive Management Workgroup (Cross-TIG MAM), and what next steps will be taken.

The DWH NRDA Restoration effort focuses on environmental data specific to monitoring and adaptive management. This data may be generated during any phase/component of restoration implementation,

as part of any project-specific monitoring, or non-project specific data collection. Within the SOP, the general standards for monitoring, data format, and data management are outlined. The Cross-TIG MAM Workgroup will develop these topics further and include this information in the MAM Manual.

Standardization of monitoring data with respect to parameters and metrics, precision, units, performance criteria, and collection protocols, will increase consistency, allow further analysis across TIGs and restoration types, and enhance compatibility with existing datasets.

The MAM Manual will include protocols for data review and clearance, storage and accessibility, sharing, and analysis and synthesis. Standardizing the aforementioned protocols will assist with QA/QC, validation of data, interoperability, and public accessibility. Trustees will follow standards and protocols set in the Federal Open Data Policy. Data can be accessed through warehouses such as DIVER and through the Trustee Council website. The data management section, outlined in the MAM Manual Version 1.0, outlines specific standards and management procedures to build within DIVER the capability and functionality for MAM data.

Direct Component & Centers of Excellence- U.S. Department of the Treasury

Laurie McGilvray (Treasury) presented the RESTORE Act and the data management framework, on behalf of the U.S. Treasury and the Office of Gulf Coast Restoration. The Clean Water Act (CWA) penalty funds for the DWH went to the Gulf Coast Restoration Trust Fund. It is the Treasury's responsibility allocate funds for the Direct Component portion of these funds and for the Centers of Excellence Research Grants. The Direct Component section allocates 35% of the Gulf Restoration Trust Fund among five states (i.e., AL, FL, LA, MS, TX) to help with ecosystem restoration, economic development and tourism promotion.

Additionally, the Treasury administers 2.5% plus interest earned from the Trust Fund's investments for research on the Gulf Coast Region. The funds are allocated to the same five states. Within these states,

Centers of Excellence are awarded funding for research grants. Research topics include, but are not limited to: coastal sustainability, restoration and protection, offshore energy development, coastal fisheries and wildlife monitoring and research, sustainable and resilient growth, and economic and commercial development in the Gulf Coast Region.

The RESTORE Act grant requires performance reporting which includes: summarizing any significant findings or events, including compiled, collected or created data; description of activities to disseminate or publicize results of the activity; and designation of the project or program responsible for the generation of that data. In many instances, data being collected use a common repository and are being put into an existing data framework.

National Fish and Wildlife Foundation (NFWF)

Jon Porthouse (NFWF) provided an overview of NFWF which is a non-federal entity focusing on ecosystem restoration and monitoring. There is no regulatory framework in place and the organization works with multiple states. NFWF works with a variety of stakeholders (i.e., federal, state, county, municipal, NGO, academia) to generate and manage data. The organization funds data generation, but does not house it; in some instances, the organization collects metadata as it sees fit. NFWF has no overarching data management requirements for its grantees.

NAS Gulf Research Program (GRP)

Lauren Showalter explained that GRP will be requiring all grantees to make data or information products that result from its funded research publically available within one year of the end of the grant. The GRP will provide grantees with a list of acceptable data repositories that have been identified with assistance from the GRP Advisory Board. The GRP will also create a catalog of data and information resources so interested parties can access all of these products.

The GRP has a public data policy and requires that all submitted proposals include a data management plan that is reviewed by the Program Officer prior to panel review of the proposal. Once a project is funded, the GRP works with researchers to ensure they are familiar with GRP data procedures and policy and continues to assist them as data is created so they develop sufficient metadata and identify the most appropriate repository.

VI. Standards, Parameters, and Challenges

Greg Steyer (USGS), presented on standards, guidance and challenges that affect the monitoring community. The community of monitoring and data acquisition is complex and encompasses state and federal agencies, the RESTORE Council, the academic community and others. A tremendous amount of data is being collected and the various DWH programs are working together to determine common monitoring standards and protocols. Standardizing the protocols would simplify the aggregation and synthesis of data following data collection.

The approach to standardizing data management is a 3-year, Phase 1 program which lays the foundation for a structure and implementation strategy related to monitoring. This strategy would enable the Restore Council to achieve the goals, objectives and commitments in its Initial Comprehensive Plan (ICP). The approach is to use coordination and collaboration to build upon the numerous existing monitoring activities and programs in the GOM. It is necessary to engage expertise within groups such as GCOOS, the state and federal resource agencies, state Centers of Excellence, academia, NAS, NGOs, industry and other interested stakeholders to move towards a coordinated GOM-wide monitoring and assessment program. Successful data acquisition would include: a catalog of existing data, an understanding of how existing data can be fully used, the quality of data being generated, and by following a minimum set of standards to ensure proper metadata and QA/QC.

DWH data should be aggregated, stored, and the quality assured such that it can be disseminated. Using existing capabilities (e.g., web portals, catalogues, archives) is advantageous to avoid creating new systems. Standardizing data description, formats and services would promote interoperability between existing systems.

Challenges with data management include bringing together existing data monitoring and management agencies from inception to develop an integrated process; and communicating and coordinating across both DWH and non-DWH programs. When developing data management systems, user needs should be considered, and data generators should clearly articulate measurable objectives from project to programmatic scales. The generators should delineate common sets of questions that need to be addressed so researchers are not asking the same questions. Other major challenges are adopting common data standards, following minimum monitoring standards and data requirements, and governing across programs.

Environmental Disasters Data Management Workshop (EDDM) Working Groups

The Environmental Disasters Data Management (EDDM) Working Groups are an outcome of a CRRC workshop held in September 2014, the objectives were to promote the use of the protocols and practices during data collection, as well as recommend data management limitations to be later discussed in workshops. The overall goal of the working groups is to provide information and data services that improve the quality and speed of decision-making in response to environmental disasters. The EDDM Working Groups were coordinated by CRRC, NOAA's National Coastal Data Development Center and NOAA ORR. EDDM Working Groups focused on Field Protocols, Common Data Models, and "Gold" Standards. Additional information regarding EDDM can be found at <https://crrc.unh.edu/EDDM>.

1. Field Protocols

Steve Ramsey (Social & Scientific Systems) presented a summary of the work done by the Field Protocols working group. The objectives set for this group were to: (1) Inventory existing resources for field data collection; (2) inventory existing equipment, devices, and monitors for field data collection; and (3) apprise academics and NGOs of sampling protocols they should use to get data included in existing systems. The group sought to bring existing tools together in one location which have been developed by agencies to better empower the environmental health and safety (EHS) community to gather useful data. The working group compiled existing protocols and surveys, but found that institutional review board (IRB)/ethics guidance need to be further developed to support researchers, particularly for public health data. Compiling this information will allow researchers to quickly assemble survey instruments, protocol templates, and search for existing standardized collection methods for EHS topics. This is especially important because researchers often want to be involved in environmental disasters on short notice and need a readily available source of accepted protocols for collecting data. Creating a network of information/metadata will help researchers to communicate with one another and understand what protocols “work” as well as allow continuous improvement of tools and information to better aid the research community.

2. Common Data Model

Dan Hudgens (Industrial Economics, Inc. (IEc)) presented the output of Common Data Model working group. The group’s objectives were to: (1) document what specific data models, portals (data sets), and web services are being used across different disciplines and compile details regarding each one (i.e., portal name, description, type of data accessible, data base compatibility, URL, key contacts); (2) crosswalk existing data models to find similar elements; and (3) at all levels (field collection, synthesis, analysis), inventory/identify existing ways to be interoperable. The outcome from the first objective was

a spreadsheet of data systems pertinent to environmental disasters; the group identified 24 data systems and explored seven of them in more detail.

The second objective resulted in identifying redundancy and compatibility across data models, or cross-walking between systems. The group recognized the importance of “federated” data, and the importance of connecting systems. Cross-walking facilitates information sharing between agencies. In order to do this, a common vocabulary must be developed. Currently, cross-walking data is challenging because systems refer to certain types of data by different titles. This creates challenges when importing data from one system to another (i.e., different nomenclature amongst systems limits data compatibility). As an example, the group cross-walked two systems - NOAA’s DIVER and USEPA’s SCRIBE. The third objective is an ongoing task to make recommendations where researchers and data managers can leverage approaches to interoperability and data security.

3. Gold Standard

Julie Bosch (NOAA NCEI) presented the objectives and outcomes for the “Gold” Standard Working Group. Its objectives were to: (1) Identify the functionality needed for information management and decision support tools for different disaster types and where these functionalities are located, (2) identify criteria to evaluate data and procedures (i.e., QA/QC, data transport, security, data use analytics) that can be considered a Gold Standard, (3) identify critical data types for baseline data for different environments and types of disasters, and (4) define terms (data dictionaries).

Objective 1 was addressed by a table including a series of matrices of tools for different disaster scenarios. Objective 2 developed a list of criteria based on approximately 25 different data types, subdivided depending on types of data, methodology, and disaster. An evaluation worksheet of criteria and ranking was also developed. The working group noted that suggestions for improving QA/QC would

help to create a consistent Standard Reference Material (SRM) and released source material within a program would allow for accurate assessment of inter-and intra-laboratory variability.

Objective 3 required listing the critical data types and recommending authoritative sources. The table developed included: greater than 170 parameters/media and their critical data types for baseline data, parameters for the data, media and category, and recommended sources. Objective 4 contained a list of different data dictionaries as a function of environmental disaster type and provided access to them.

The list included 56 vocabularies, data dictionary names, links, and critical data types.

Vision of Managing Big Data

Larry Langebrake (ConnectSix) gave a keynote presentation on vision of managing big data from an outsider's (?? DWH/GOM) perspective. When considering the infrastructure needed to handle big data and its transformation into actionable information. It can be daunting to decide what is valuable, especially when the end users are from disparate groups. Best practices, such as the process of value-creation, can help clarify vision and bring focus to actions that should be pursued. Common language, champions, and "alignment" are crucial elements but the most important is a deep understanding of customer needs. He presented was an overview of the process of value creation, industry examples of its application, and a specific example of how value can be created for researchers (i.e., especially those relevant to the GOM and DWH).

Flow Diagram of Existing Systems

An end-to-end use flow diagram (Appendix E) of five existing data systems (i.e., GRIIDC, the USGS National Water Quality Portal, GCOOS, CIMS, DIVER) were presented to better understand each system that is used for data management and synthesis, as well as the challenges of each system. Presentations touched upon the major topics of: interoperability, data access, data usability, metadata, ease of discovery and data synthesis. Understanding current systems is the first step in creating an

interoperable DWH data community. Data interoperability facilitates the use of DWH data to produce interdisciplinary products which can better assist decision making, setting baselines, and ultimately improving restoration efforts.

VII. Breakout Sessions

Each breakout group had a leader to help facilitate discussion among all participants and a note taker equipped with a laptop computer and projector to capture the discussion. Each group completed a workshop template (Appendix E).

For Breakout Session I, participants were divided into three groups corresponding to their roles in data management: data users, data generators and data managers/governors. Breakout Session II, participants were divided into six breakout groups: (1) interoperability, (2) ease of discovery/searchability, (3) data access, (4) data synthesis, (5) data usability, and (6) metadata/data documentation. Specific questions for each topic were developed by the organizing committee based on the highest priority challenges reported in Breakout Session I. An effort was made by the organizing committee to distribute participants into their respective groups. A list of the breakout groups is located in Appendix G.

The summary and distillation of key points from the breakout sessions are presented below. Breakout session notes can be found in Appendix H.

Breakout Session I

The three groups (data users, data generators, data managers) documented their challenges on interoperability, ease of discovery/searchability, etc. and ranked them. This section includes a summary from each breakout group and a table that identifies the high priority challenges used for Breakout Session II.

1. Data Users

The data users discussed the bounds of the workshop specific to DWH, but noted the findings can be used in a broader arena. Additionally, they agreed that the workshop's findings will help to set a baseline for data management practices in future environmental disasters. The data users discussed how to engage a wider community of researchers that were not involved (e.g., county data managers), and how their data could be helpful during restoration activities in the Gulf. The discussion was largely based around standards for data generation, and the challenges that arise regarding the quality when there is no documentation of best practices.

A variety of disciplines and agencies were represented within the working group, and therefore the term "user" was defined/discussed. Users have varying levels of domain knowledge. When creating a data repository, the level of domain knowledge was very important. For example, "high" domain data users would be researchers, regulators, government agencies, and funders. "Low" domain users would be people who use the data for litigation, media purposes, or are resource users (e.g., fishermen). The data repository should provide the highest quality data or decide what user type they want to target. The user can then decide the level (high vs. low) of data quality necessary for their purpose.

The group filled out a table explaining the needs/features that exist with respect to interoperability, data discovery, and other challenges. The highest priority set by data users was the standardization of data, such that it is interoperable and sharable between organizations. The challenge is having data generators provide sufficient data documentation enabling users to compare data and properly understand it during synthesis. The group suggested that the outcomes of this workshop should be applied to a broader arena and include sharing and engaging with other agencies.

2. Data Generators

The data generators breakout group (1) gave reactions about the plenary sessions, (2) discussed the challenges/incentives/costs for each of the topics (interoperability, ease of discovery, data access, data

synthesis, data usability, metadata), and (3) identified which challenge was of highest priority.

Additionally, the group discussed the following topics:

- Common set of metadata standards across the community,
- Maintenance of datasets in perpetuity,
- Timely availability to the user community, and
- Accessibility (use/reuse) of data in future.

The group thought that major data systems are very different and that the expectation of easily finding integrated data was unreasonable. It is more realistic to expect people to package data in a standard way rather than have integrated data available for immediate use. Additionally, the group wondered if a federated data system was a goal; they believed it would be more valuable to build systems on a common framework. Overall, it would be mutually beneficial to get stakeholders to “do the right thing” (e.g., provide quality data documentation) so that their data is “re-usable”. Quality data documentation would provide better recognition/acknowledgement of the generators regarding their data. This would create a feedback loop for data generators to produce quality data because their work would be better cited, and this would help advance their careers.

The highest priority challenge is for data generators to make the concept of interoperability a part of everyday collection efforts. The incentive proposed was that the more researchers provide and thoroughly document their data, the more funding opportunities arise for them and the more visibility they receive. Another challenge is that data generators may not have the understanding or funding to meet a minimum standard.

Another challenge regarding data usability was clearly displaying/documenting the quality of the data. This entails that the generator is providing sufficient information about the data quality. Quality data can be used on multiple occasions rather than for a single use. It helps the data generator to be more

widely known. One challenge that stems from this is the concept of citing the data generator. This could be an incentive to promote data sharing, but the citation is done by the data user. A recommendation was made to create tracking DOI's for data generators to make citation more uniform and easier for data users.

3. Data Managers

The data managers (1) expressed their reactions to the plenary sessions, (2) delineated the challenges they thought were of concern with respect to each topic (e.g., data access) and the associated workload effort, (3) determined the challenges for creating a long term data repository, (4) documented metrics of success for data management systems and data repositories, (5) listed the barriers that generators face when submitting, finding and using data, (6) provided suggestions for enforcing data policy and data security.

The data managers noted the need to define certain terms (e.g., interoperability) to help facilitate discussion of issues regarding data management. Additionally, they noted creating a DOI assignment for data sets, and establishing quality control for different systems is important so that transparency is controlled. Quality metadata would help interoperability, as well as match different systems with their respective requirements.

The data managers believed that having common vocabularies would assist with better data documentation, ease of discovery, data usability, interoperability and data synthesis for interdisciplinary projects. The use of "themes" when searching for data would be helpful to sort between different types of data (e.g., biological vs. chemical). The ability to search by "keywords" would also help with ease of discovery. One major challenge came about when discussing data access; the ability to pay for data management infrastructure and the cost are directly related to the data volume being archived in the warehouse. The data managers discussed the challenges in defining who the data users will be in the

future, and therefore predicting the level of quality to ensure the data are preserved over time. The group assumed that data synthesis will become automated and therefore noted a common framework is essential to facilitate interoperability between various data archives/repositories.

The challenge is maintaining resources and funding to create, establish and preserve the long term repositories. Additionally, the technology must be maintained, new standards and sources pose additional financial and technological burdens. Because technology, funding, and standards change over time, consistent leadership is necessary to insure the usefulness of repositories is maintained. The leaders must understand the importance of quality data needed by future users. Data security is a major concern as it impacts the ability to protect and maintain the data quality.

The data managers were tasked to define successful characteristics of data management systems (i.e., transactional system: access, discovery portals, queries) and data repositories (i.e., warehouse, central system, stewardship management, preservation). The consensus was that “happy” data users, funders and data generators is a sign of a high quality data management system. The reuse or publication of data sets would indicate that the management system was providing useful and accessible data for users. Both of these metrics also apply to data repositories. Additionally, the use of analytics (i.e., assessing who is going to use data), open services and third party reuse would be indicators for a useful repository.

The data managers identified categories and barriers that are likely to be problematic for proper data management. Barriers to entice generators to submit data include, but are not limited to: time, money, ease, willingness to share, training, difficulty to submit data, guidance and mandates. If generators are not submitting data then all of the other categories become insignificant because there is no data.

Barriers to accessing data are the lack of knowledge regarding the proper language/keywords, a useful interface, and the flexibility in the search engine. Using data is challenging when data documentation is

lacking or there are no tools to synthesize/evaluate the data. Enforcing data policy is challenging because there are so many funders and each agency has their own policies. Data security is ever evolving and changes between funders and agencies.

4. Collated Priorities from Breakout Session I:

The major challenges identified by each of the three breakout groups are compiled Table 1.

Table 1. Challenges identified by the workshop data managers, generators, and users

	Data Managers	Data Generators	Data User
1. Interoperability	<ul style="list-style-type: none"> • Scale of interoperability down to the metadata record • Selective technical data standard 	<ul style="list-style-type: none"> • Little understanding or funding that allows meeting a minimum standard • Interoperability should be part of everyday collection efforts 	N/A
2. Ease of Discovery & Searchability	<ul style="list-style-type: none"> • Lack of common vocabulary • Multiple portals using a common internet search engine • Web page with different “themes” to help drill down for needed data • Search by “keywords” • Design for user experience is difficult 	<ul style="list-style-type: none"> • Level of funding available to hire people to monitor the input and sharing of data by generators • Data users should work with generators to determine the needs/search preferences • Data entered in a way that it may be used for visualization by users 	<ul style="list-style-type: none"> • Characteristics of a good repository
3. Data Access	<ul style="list-style-type: none"> • Paying for infrastructure • Data volume • Create a common interfaces for standards • Number/amount of people accessing data (infrastructure behind access) • Restrictions & sensitivity & patents & security; level 2 product can be accessed but not the raw data 	N/A	<ul style="list-style-type: none"> • Funding • IT security • Confidentiality of data
4. Data Synthesis	<ul style="list-style-type: none"> • Better interoperability feeds into better synthesis 	N/A	<ul style="list-style-type: none"> • Anticipation of user needs

	<ul style="list-style-type: none"> • Multiple data synthesis (human vs environmental; timescale and granularity) • Data ambiguity/ biased • Automated/computer synthesis • Funding • Definition of synthesis 		<ul style="list-style-type: none"> • Anticipation of diversity of user needs
5. Data Usability	<ul style="list-style-type: none"> • Accuracy, resolution, level of confidence, fitness of uses • Sufficient record level data • Known quality • No control of how data is used • Versioning 	<ul style="list-style-type: none"> • All data should be of known quality (must be enough information to judge quality) 	<ul style="list-style-type: none"> • Need for sufficient information to allow users to compare data
6. Metadata/Data Documentation	<ul style="list-style-type: none"> • Lack of common metadata vocabulary (units, time range, scale) • Multiple portals with easy find for internet searches • Consistency of implementing the standards • Training of what metadata is (dataset description) • Maintaining • Versioning 	<ul style="list-style-type: none"> • Templates/minimum standards must be created from the beginning of data collection (e.g., data management plans) 	<ul style="list-style-type: none"> • Very robust metadata takes a lot of work
7. Other: Communication to the user base	<ul style="list-style-type: none"> • Some data managers do not communicate well • Common data management strategies are needed (priority of the program) • Time required to do this may not be allotted 	N/A	N/A
8. Other: Longevity	N/A	N/A	<ul style="list-style-type: none"> • Long term maintenance of repositories • Magnitude of data

Breakout Session II

The participants were divided into six groups: interoperability, ease of discovery, data access, data synthesis, data usability and metadata/data documentation.

1. Interoperability

Interoperability occurs when data sets can be translated from system to system without extensive transformation. Interoperable data would be a set that blends across archives, repositories, domains, and sectors and not impacted by formatting, vocabulary, and metadata. The questions posed for the group included:

1. To whom is interoperability important?
2. Why is interoperability important?
3. How does interoperability happen and who is responsible?

Interoperability is important when crossing data from sector to sector (e.g., using climate data to cross over to the health sector to answer public health questions). It is important to those synthesizing the data and using products to answer broader societal/scientific questions. The conclusion to the second question was that interoperability is important because it supports activities such as synthesis, data discovery, access, and dissemination. Interoperability promotes interdisciplinary use of data and helps answer complex questions. The response to the third question was that establishing standards or a framework for the entire life of the data stream would enable system to system communication. The system standards could be established within the initial data management plans and possibly by government agencies.

The breakout group discussed the data management requirements to achieve successful interoperability and those included: clear plans that follow standards, proper resources and training, a catalogue of existing frameworks to better understand and establish a common vision across organizations that helps

translate data in the future. The group also discussed best practices and guidance for interoperability concluding that the development of homogenous standards would help promote interoperability between systems. This would work as long as sectors adhered to standards, and standards were documented, so that future generators would use the established standards.

2. Ease of Discovery / Searchability

The breakout group discussed three questions specific to ease of discovery:

1. What are the characteristics of a good repository- in terms of ease of discovery/searchability?
2. How is metadata quality ensured?
3. How are user needs met?

Characteristics of a good repository include abundant keywords with a common vocabulary, semantic search, and searchable data all within a “findable” repository. Ensuring the quality of metadata comes from investment in human resources, an early focus on complete/accurate metadata, and generators training on proper data documentation. The only way to ensure that user needs are met is to know the user and their level of knowledge.

Requirements to fulfill successful ease of discovery includes early involvement by the data management team, definition of user needs, and the ability to edit metadata once it has been collected to make it easily discoverable. Best practices for ease of discovery includes establishing a federated database, funding mandates, and identifying end users at the start of data collection.

3. Data Access

This breakout group discussed the meaning of data access, and established that there are subtleties within that topic that are more complex than just having the data be available. Data access is the successful endpoint to data discovery; the user can get what they need without a lot of extra work. The group answered three questions:

1. How is restricted/sensitive data addressed?
2. How is data security addressed?
3. What are the options for the data volume challenge (i.e., high volume; large data sets or many hits on one data set)?

The first question required an established definition; referring to the actual sensitivity of the data (e.g., human, archaeological, embargoed data, dark data). The conclusion was: limit the amount of restricted data, summarize forms/frameworks used by the data generator, obfuscate the GIS, build security into the system using summarized forms of the data, and roll up the data such that individuals cannot be identified. There is also the option to create credentials or differential access through log-in requirements.

The second question required a specific definition to aid the discussion of a broad interpretation of data integrity and system security. Again, the group agreed that granularity/credentials/differential access, and login access roles would be helpful along with meeting common IT security requirements. An alternative or additional option is certified data warehouses/centers to promote the security and integrity of data sets.

The third question touched upon the challenges of high volume/large data sets. The outcomes of the discussion were a scalable cloud. Challenges include the associated cost and procurement. Another option would be to create subsets or previews of the data prior to download. All of which could combine with the option to have multiple methods of access (e.g., FTP, direct cloud download, cold storage). It may also be possible to leverage a private industry to store the data (e.g., Google Earth engine) so that the data is accessible to the public.

The data management requirements for successful data access include: a common summarization approach, complete documentation, robust metadata, bolstered public accessibility, and effective user

interfaces (human) and services (machine). Best practices and guidance for data access must ensure data integrity within a system. The system is designed for public access, as well as credentialed logins. The user interface should be easy to use and share data in an enticing way (e.g., communicate with story maps). It is also important to give data owners/generators credit for their work including identification of the generator (as appropriate).

4. Data Synthesis

Data synthesis was defined as bringing together different data sets to do comparisons and analyses. It requires a multi-disciplinary approach which is guided by human activity to answer questions beyond the original purpose of the data. This group altered their original questions to better suit the challenges of data synthesis. The questions were:

1. How can data discovery and accessibility occur for unknown future users?
2. Why is data synthesis important?
3. How does data management facilitate data synthesis?
4. How can data be preserved for future use?

In order to make data discoverable and usable for unknown future purposes, the data must be interoperable. Interoperability is reliant upon metadata, key wording, standards, crosswalks, machine-to-machine-discovery, and standard archiving formats. Additionally, optimization of web searchability for data (e.g., Google-like search) must help all users who do not know about the vast amount of data available, and market the data so that people can learn of the variety of data sets. Data synthesis is important because it is used to answer questions (e.g., for the DWH NRDA restoration effort). One question is whether the resource recovered and the extent to which restoration efforts helped recovery? Data must be properly managed such that synthesis is easy for data users (e.g., through transformation and analytical tools). The establishment of an effective archive can enable proper dissemination of data and aid in the synthesis process. Coordinating synthesis centers and building the

capacity for researchers to do the work and use accepted tools would help standardize the synthesis process.

Data management requirements for successful data synthesis include: direct access to the data, searchability, consistent metadata, and knowledge/training on repositories and their tools. Best practices and guidance for data synthesis include: talking to experts, defining potential questions and knowing the audience.

5. Data Usability

This group defined their topic as making as little effort as possible to execute a task with an end goal of a usable product (e.g., opening the file, transforming it). Key factors include: attributes and characteristics of the data, knowledge of biases, confidence in the data quality (QA/QC), and awareness of the data format and collection methods. The goal of data is synthesis and product creation. Users need the data sets to be interoperable and ensure that they are extracted in the same or high quality than they were entered. This group discussed:

1. How the generators can be “encouraged” to define quality, resolution and accuracy of the data?
2. How the quality, accuracy, resolution of the data can be conveyed users?
3. How data quality, accuracy and resolution can be assessed and reported?

The first major conclusion was that data usability must be approached from the perspective of the funder/repository as well as the data generator. Funders could provide a template/framework that generators use while collecting data. This system would be established within the data management plan/contract at the beginning of the process. The generator needs to understand that data documentation is crucial for synthesis, usability and longevity. The group determined that conveying the quality, resolution and accuracy of data to users should be done through intensive data documentation. Information should be conveyed through the repository. This would include a disclaimer regarding the

purpose of the data, and why it was created and further background on the data set. Assessing the quality, accuracy and resolution is challenging because it encompasses the data generator, user and repositories. The generator must document collection methods with sufficient metadata and the data must be checked for quality assurance. The user must be able to determine the use of the data (e.g., resolution, accuracy) to ensure that it meets the need. Overall, standardization of quality assurance should be done through a peer review process using the communities of practice. The category of “other” brought to light challenges that data usability faces. One example of this is defining key policy questions that need to be answered, and using data to answer these bigger picture questions.

Coordination within the DWH community to enable common metrics that can be aggregated to answer key policy questions would be helpful when determining which data sets to use. Restoration monitoring requires a baseline to compare the data pre- and post-incident. Often, baseline data is unavailable which results in gaps in monitoring efforts.

Data management requirements for successful data usability include: common data descriptors; shared knowledge and understanding of the data (e.g., robust metadata), and engaging the data users early so that collection of data is not random and purposeless. Machine readable metadata allows for easier usability of the data because it enables systems to pull in data without extensive transformation.

Allowing data users to preview data prior to download would reduce strain on the repository and speed the search. The best practices and guidance recommended for data usability are to: develop communities of practice, determine if there is a consensus approach across many communities, and bridge gaps between diverse communities. Creation of a user report that outlines the quality of data and provides a summary of which standards it meets, and trains users about access via warehouses/repositories. The repository should consider the user interface and how existing tools can reduce the struggle between user and data.

6. Metadata / Data Documentation

The metadata breakout group created a scope for their discussion, this included NRDA restoration/RESTORE Act as their primary data of discussion. However, there will be other end users of this data and therefore the need for quality metadata is even more important to maintain the longevity of data sets. Overall, it is important that greater value is placed upon archived data sets related to the research that will be funded in the future. This group answered the following questions:

1. How is the consistency of geospatial metadata standards established and ensured?
2. How is a controlled vocabulary implemented in metadata?
3. How is completeness of the metadata ensured for users in the future?
4. How is the burden of metadata reduced on generators?

The group concluded that if data documentation is done properly then the rest of the data management process becomes much easier. Establishing consistent metadata standards can be accomplished through mandate from the funding agency, particularly for tags, supplemental information, and definitions for data fields. Establishing training for the data generators at the start of collection would help data set description. The funding agency or repository can assist in data documentation by providing a template for data generators to complete. Engaging communities who work together to agree on protocols/procedures and establish a written contract would benefit metadata quality. The contract would hold data generators accountable for high quality metadata, and help data managers to store it. Implementation of a controlled vocabulary could be made easier with the use of templates, but first the community of practice must agree upon the vocabulary.

Creating value for metadata documentation at the start of data collection helps the entire data management process. For example, generators who have quality data documentation have more discoverable data, and therefore it can be synthesized on a more frequent basis. Complete metadata is ensured through a review process (human and automated) with established rubrics. Templates and

guidance encourages comprehensive metadata. Establishing a minimum required set of standards within the data management plan provides generators with clear expectations.

Data management system requirements for successful metadata included complete metadata standards for all stakeholders. Additionally, when metadata is machine readable or automated it reduces the time data managers spend on maintenance. Complete metadata allows for more discoverable, transferable and adaptively managed data. High quality metadata assists with connecting data to the source, and directing any questions to the data generator. Best practices and guidance for metadata include: generating metadata continuously, collecting metadata as soon as data collection begins, and implementing existing workflows (e.g., rolling deck to repository), and listing best practices accumulated by communities of practice and stakeholders.

VIII. Conclusion

Workshop Outcomes

The final day of the workshop included a plenary session where the participants discussed the outcomes. Participants identified the need to determine the barriers that data generators perceive to following directives and standards. Any enhancements or improvements to data management systems should consider the needs of data users. The first step in making more robust data management systems is a compilation and review of existing data management directives including federal requirements with respect to data delivery, annotation and documentation for grants. [N.B., This is already in progress for biological and monitoring standards.]

Moving Forward

During the final plenary session, the participants identified steps to improve LTDM centering on data management standards, interoperability and data discovery/searchability.

1. Data Management Standards

1. Form a small, short-term working group to define data management components.
2. Identify categories of standards needed (e.g., data acquisition including sampling protocols and quality control, data management).
3. Create a DWH LTDM standards working group (DWH LT DMSWG) to determine what gaps need to be filled for data management standards. The gap analysis will inform the list of standards (e.g., metadata) that need to be established. This must be done in concert with the RESTORE Council's monitoring and assessment work group (CMAWG) and the Cross-TIG MAM (NRDA) analysis for data acquisition and monitoring.
4. Provide feedback to funding entities on standards needed to manage data long term that are recommended by the DWH LT DMSWG, CMAWG and Cross-TIG MAM working group.

2. Interoperability

1. Create a working group to determine what could optimize interoperability efficiency between DWH LTDM systems and drive collaboration among them.
2. Compile strategic goals and key features for data warehouses and repositories.
3. Determine the intended, current and future use of DWH LTDM systems.

3. Discovery/Searchability

1. Develop and share technology used by DWH data management services for keyword, semantic, geospatial, and temporal searches.
2. (In parallel with 1.) Create a community driven, vocabulary working group to identify definitions for specific data types as well as incorporating keywords into data and metadata.
 - a. Populate with individuals from the NRDA TIGs and RESTORE Act communities of practice as well as DWH LTDM workshop participants.
3. Leverage architecture of existing systems where possible (e.g., USGS Sciencebase, NOAA OneStop).
 - a. Compile approaches regarding data, links, and metadata (e.g., embedded ESRI maps that delineate study areas).
 - b. Note whether data is service-enabled (e.g., machine readable, consumable by other programs) because serviceability enables interoperability.

Appendix

Appendix A: Workshop participant list

Appendix B: Agenda

Appendix C: Presentation slides

Appendix D: Pre-workshop survey

Appendix E: Flow diagrams

Appendix F: Breakout session templates

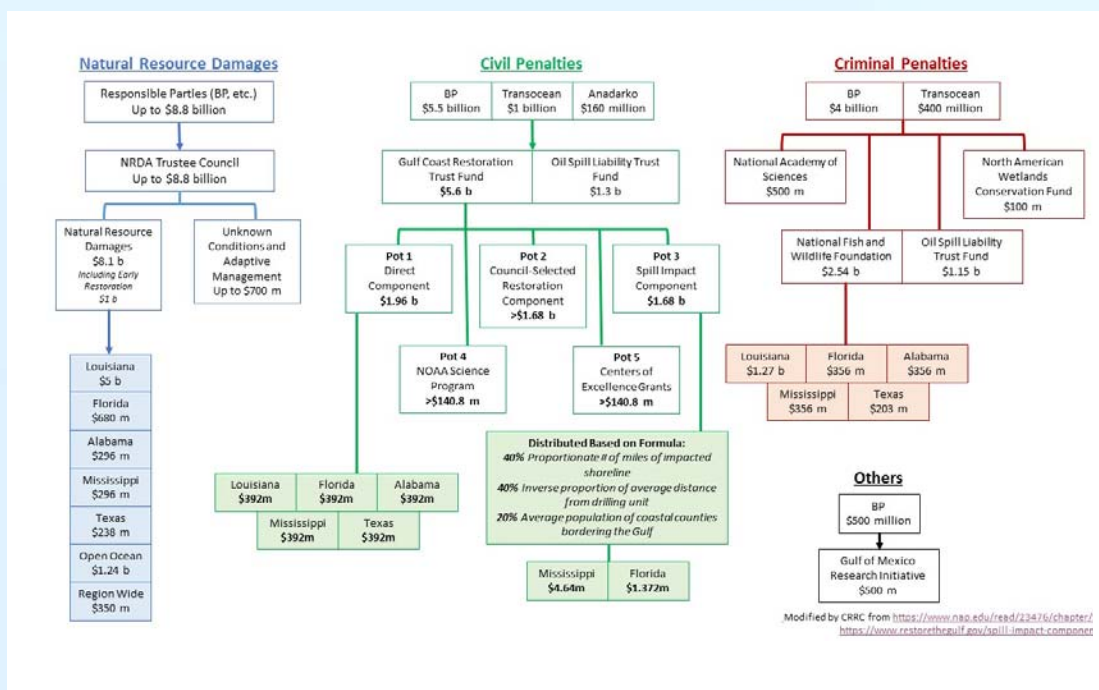
Appendix G: Breakout group participant list

Appendix H: Breakout group notes

Appendix I: Funding Diagram

DWH LONG-TERM DATA MANAGEMENT COORDINATION WORKSHOP

JUNE 7 - 8, 2017



NOAA'S GULF OF MEXICO DISASTER RESPONSE CENTER
MOBILE, AL

APPENDIX

COASTAL RESPONSE RESEARCH CENTER



Appendix

Appendix A: Workshop participant list

Appendix B: Agenda

Appendix C: Presentation slides

Appendix D: Pre-workshop survey

Appendix E: Flow diagrams

Appendix F: Breakout session templates

Appendix G: Breakout group participant list

Appendix H: Breakout group notes

Appendix I: Funding Diagram

Appendix A: Workshop participant list

DWH LONG-TERM DATA MANAGEMENT COORDINATION WORKSHOP

JUNE 7 - 8, 2017

PARTICIPANTS

Courtney Arthur
Industrial Economics, Inc (IEC)
carthur@indecon.com

Jennifer Bauer
National Energy Technology Laboratory (NETL)
jennifer.bauer@netl.doe.gov

Holly Binns
The Pew Charitable Trusts
hbinns@pewtrusts.org

*Jonathan Blythe
Bureau of Ocean Energy Management (BOEM)
jonathan.blythe@boem.gov

*Julie Bosch
NOAA National Centers for Environmental Information (NCEI)
julie.bosch@noaa.gov

*Laura Bowie
Gulf of Mexico Alliance (GOMA)
laura.bowie@gomxa.org

Melody Chimahusky
MS Department of Environmental Quality
mchimahusky@mdeq.ms.gov

Jay Coady
NOAA ORR ARD
jay.coady@noaa.gov

Craig Conzelmann
U.S. Geological Survey (USGS)
Wetland and Aquatic Research Center
conzelmannc@usgs.gov

Alyssa Dausman
Gulf Coast Ecosystem Restoration Council
alyssa.dausman@restorethegulf.gov

Steve Delgreco
NOAA National Centers for Environmental Information (NCEI)
blackswaninnovations@gmail.com

Ryan Druyor
Florida Fish and Wildlife Conservation Commission Fish and Wildlife Research Institute
ryan.druyor@myfwc.com

Nicolas Eckhardt
NOAA ORR Spatial Data Branch
nicolas.eckhardt@noaa.gov

Sandra Ellis
Gulf of Mexico Research Initiative Information & Data Cooperative (GRIIDC)
sandra.ellis@tamucc.edu

Neal Etre
Industrial Economics, Inc (IEC)
netre@indecon.com

*Elizabeth Fetherston-Resch (unable to attend workshop)
Florida RESTORE Act Centers of Excellence Program
ehfetherston@usf.edu

Shawn Fisher
U.S. Geological Survey (USGS)
scfisher@usgs.gov

*Jim Gibeaut
Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC)
Harte Research Institute for Gulf of Mexico Studies
Texas A&M University - Corpus Christi
james.gibeaut@tamucc.edu

Melissa Gloekler
Coastal Response Research Center (CRRG)
University of New Hampshire (UNH)
mdx52@wildcats.unh.edu

George Graettinger
NOAA ORR ARD
george.graettinger@noaa.gov

*Rebecca Green (unable to attend workshop)
Bureau of Ocean Energy Management
rebecca.green@boem.gov

*Jessica Henkel
Gulf Coast Ecosystem Restoration Council
jessica.henkel@restorethegulf.gov

*Denotes workshop organizing committee member

Matt Howard
Texas A&M University
GCOOS/GOMRI/GRIIDC
mhoward@tamu.edu

Lei Hu
Northern Gulf of Mexico Sentinel Site Cooperative Center of
Excellence (AL) / Dauphin Island Sea Lab
rcollini@disl.org

Dan Hudgens
Industrial Economics, Inc (IEC)
dhudgens@indecon.com

Christina Hunnicutt
U.S. Geological Survey (USGS)
Wetlands and Aquatic Research Center
hunnicuttc@usgs.gov

Helga Huntley
CARTE
University of Delaware
helgah@udel.edu

JB Huyett
NOAA ORR ERD
jb.huyett@noaa.gov

Michele Jacobi
NOAA ORR
michele.jacobi@noaa.gov

Ann Jones
Industrial Economics, Inc (IEC)
ajones@indecon.com

*Steve Jones
Geological Survey of Alabama
Groundwater Assessment Program/Coastal Resources
sjones@gsa.state.al.us

*Nancy Kinner
Coastal Response Research Center (CRRC)
University of New Hampshire (UNH)
nancy.kinner@unh.edu

*Barb Kirkpatrick
IOOS
barb.kirkpatrick@gcoos.org

Larry Langebrake
Connectsix, LLC
larry@sayweather.com

*Kirsten Larsen
NOAA National Centers for Environmental Information
(NCEI)
kirsten.larsen@noaa.gov

Julien Lartigue
NOAA RESTORE Act Science Program
julien.lartigue@noaa.gov

*Matt Love
Ocean Conservancy
mlove@oceanconservancy.org

*Kathy Mandsager
Coastal Response Research Center (CRRC)
University of New Hampshire (UNH)
kathy.mandsager@unh.edu

Nadia Martin
Industrial Economics, Inc (IEC)
nmartin@indecon.com

Kate McClure
SeaGrant Fellow
Department of the Treasury
kate.mcclure@treasury.gov

*Laurie McGilvray
Department of the Treasury
laurie.mcgilvray@treasury.gov

Marti Goss McGuire
NOAA Restoration Center
marti.mcguire@noaa.gov

*Amy Merten
NOAA ORR ARD Spatial Data Branch
amy.merten@noaa.gov

Mark Miller
NOAA ORR ERD
mark.w.miller@noaa.gov

William Nichols
Gulf of Mexico Research Initiative Information and Data
Cooperative (GRIIDC)
Harte Research Institute for Gulf of Mexico Studies
Texas A&M University - Corpus Christi
william.nichols@tamucc.edu

Rost Parsons
NOAA National Centers for Environmental Information
(NCEI)
rost.parsons@noaa.gov

*Mike Peccini
NOAA Restoration Center
mike.peccini@noaa.gov

*Jon Porthouse
National Fish and Wildlife Foundation (NFWF)
jonathan.porthouse@nfwf.org

Steve Ramsey
Social & Scientific Systems/NIH GuLF Study
sramsey@s-3.com

Rick Raynie
Coastal Protection and Restoration Authority of Louisiana
richard.raynie@la.gov

*Jamey Redding
NOAA Office of Habitat Conservation
Restoration Center
jamey.redding@noaa.gov

*Dave Reed
Florida Fish & Wildlife Conservation Commission
Fish and Wildlife Research Institute
dave.reed@myfwc.com

Denise Reed
The Water Institute of the Gulf
dreed@thewaterinstitute.org

Angela Schrift
Texas Parks and Wildlife Department
(TX) Cross TIG MAM Data Management
angela.schrift@tpwd.texas.gov

Ben Shorr
NOAA ORR ARD Spatial Data Branch
benjamin.shorr@noaa.gov

*Lauren Showalter
National Academy of Science (NAS)
Gulf Research Program
lshowalter@nas.edu

*Greg Steyer
U.S. Geological Survey Gulf Science
gsteyer@usgs.gov

Kevin Suir
U.S. Geological Survey
suirk@usgs.gov

Hugh Sullivan
U.S. Environmental Protection Agency (EPA)
sullivan.hugh@epa.gov

Megan Verfaillie
Coastal Response Research Center (CRRC)
University of New Hampshire (UNH)
mpv1000@wildcats.unh.edu

Carrie Wall
University of Colorado at Boulder
NOAA NCEI
carrie.wall@noaa.gov

Jason Weick
Coastal Waters Consortium/LUMCON
jweick@lumcon.edu

Eric Weissberger
NOAA National Marine Fisheries Service
Office of Habitat Conservation - Restoration Center
eric.weissberger@noaa.gov

Danny Wiegand
U.S. Environmental Protection Agency (EPA)
wiegand.danny@epa.gov

Caitlin Young
NOAA RESTORE Act Science Program
NAS Gulf Research Program Science Policy Fellow
caitlin.young@noaa.gov

Dwane Young
U.S. Environmental Protection Agency (EPA)
young.dwane@epa.gov

Appendix B: Agenda

DWH LONG-TERM DATA MANAGEMENT COORDINATION WORKSHOP

JUNE 7 - 8, 2017

AGENDA

Workshop Partners: CRRC, NOAA ORR, NOAA NMFS RC, NOAA NCEI

Workshop Objectives:

- Foster collaboration among the Gulf of Mexico partners with respect to data management and integration for restoration planning, implementation and monitoring.
- Identify standards, protocols and guidance for long term data management being used by these partners for DWH NRDA, restoration, and public health efforts.
- Obtain feedback and identify next steps for the work completed by the Environmental Disasters Data Management (EDDM) Working Groups.
- Work towards best practices on public distribution and access of this data.

DAY 1: June 7, 2017

- 8:00 Registration
- 8:30 Welcome [*Amy Merten, Rost Parsons, Mike Peccini*]
- 8:45 Workshop Objectives [*Amy Merten*]
- 9:00 Participant Introductions
- 9:30 PLENARY: Participant Survey: Vision of Long Term Data Management in the Gulf [*Jessica Henkel*]
- 9:45 **Break**
- 10:00 PLENARY: Overview of Data Generation [*Michele Jacobi*]
- 10:30 PLENARY: Uses of DWH Long Term Data [*Matt Love*]
- 11:00 PLENARY: Overview of Long Term Data Management (LTDM) [*Lauren Showalter*]
- 11:30 PLENARY: Overview of Existing Long Term Data Management Systems
- NOAA ORR (DIVER, ERMA) [*Ben Shorr*]
 - NOAA Restoration Center [*Mike Peccini*]
 - NOAA NCEI [*Rost Parsons*]
 - GRIIDC [*Jim Gibeaut*]
 - GOMA Portal [*William Nichols*]
 - DWH Project Tracker [*Laura Bowie*]
 - LA Coastal Information Management System (CIMS) [*Craig Conzelmann*]
 - GCOOS [*Matt Howard*]
- 12:30 **Lunch** (please plan to contribute \$10 towards this lunch delivery)
- 1:15 PLENARY: Data Management Standards / Protocols

- Data Management Frameworks
 - Restore Council [*Jessica Henkel*]
 - NRDA Restoration [*Jamey Redding*]
 - Direct Component & Centers of Excellence – Treasury [*Laurie McGilvray*]
 - NFWF [*Jon Porthouse*]
 - NAS [*Lauren Showalter*]
- Standards identified / Parameters / Guidance / Challenges [*Greg Steyer*]
- EDDM Working Groups:
 - Field Protocols –*Steve Ramsey*
 - Common Data Model – *Dan Hudgens*
 - “Gold” Standard – *Julie Bosch*

2:15 **Breakout Group Session I:** Issues / Concerns for Data Stakeholders (identify top priorities for next day discussion)

Session I Breakout Groups:

- Data User
- Data Generator
- Data Manager/Governor

Questions to address in Breakout Group Session I:

- 1) Data user: List of requirements from the user community
 - a) Reactions to earlier plenary sessions
 - b) What are the challenges faced with each topic (i.e., interoperability, ease of discovery/searchability, data synthesis, data usability,) as a data user
- 2) Data generator: Ability to have people participate in data sharing and data collaboration
 - a) Reactions to earlier plenary sessions
 - b) What are the incentives (and impediments) to participating in a long term collaborative?
 - i) From Individual Agency Requirements
 - ii) From common set of metadata standards across community (and maintain the data set in perpetuity)
 - iii) What is the workload or level of effort required in order to be interoperable, searchable, etc.?
 - c) Readily available to user community, in a timely manner, with appropriate standards to allow for interoperability?
 - d) What are the costs?
 - e) Data generated from each program, in different locations, usable and searchable so that data is used or re-used in the future?
- 3) Data managers and data governors
 - a) Reactions to earlier plenary sessions
 - b) What is the workload or level of effort required in order to be interoperable, searchable, etc.?
 - c) Challenges of creating “long term” repositories
 - d) Funding a repository in perpetuity
 - e) How do you define success? What makes a useful data management system and repository
 - i) What are the program evaluation questions to determine a successful data management program?
 - f) What are barriers for getting people to submit data?
 - g) Barrier to finding and using data?
 - h) Cross cutting issues
 - i) Enforcing data policy
 - ii) Challenges of data security

3:45 **Break**

4:00 Group Reports from Breakout Session I

5:00 Adjourn

DAY 2: June 8, 2017

8:30 Recap & Recalibrate

8:45 **Keynote:** Big Picture Vision – An Outsider Perspective – Managing Big Data [Larry Langebrake]

9:15 **Breakout Group Session II:** Solutions / Actions to Address Issues / Concerns from Breakout Session I

Session II Breakout Group (6 mixed groups):

- Interoperability
- Ease of discovery/searchability
- Data access
- Data synthesis
- Data usability
- Metadata/data documentation

All groups should consider: examples of solutions from other long term data management disciplines.

Questions to address in Breakout Group Session II:

1. Assignment will be to fill in the 'block' regarding each topic.
2. List the requirements of a successful end-to-end data management process with respect to your group's topic
3. List the necessary guidance/best practices for funders/data generators with respect to data management for your group's topic.

10:30 **Break**

10:45 Group Reports from Breakout Session II

12:00 **Lunch** (please plan to contribute \$10 towards this lunch delivery)

12:45 PLENARY: End-to-end process/ flow diagram

2:15 **Break**

2:30 PLENARY: Moving Forward

- Is there agreement on an end-to-end process/flow diagram?
- Prioritize actions to move forward.
- Address ways to encourage participation of researchers and programs in long term management programs for post-DWH data (e.g., restoration monitoring data).

4:30 Adjourn

Appendix C: Presentation slides

WELCOME

**NOAA's GOM Disaster
Response Center**

DWH LONG-TERM DATA MANAGEMENT

**Nancy Kinner, UNH Co-Director
Coastal Response Research
Center (CRRC)**

DWH LONG-TERM DATA MANAGEMENT

2

WORKSHOP LOGISTICS

- Emergency Exits
- Restrooms
- Cell phones / laptops
- Breaks (coffee, tea, snacks)
- Meals
 - \$10/day for special lunch delivery
 - Dinners on your own
 - See restaurant map in packet
- Logistical questions – see Kathy Mandsager or me

DWH LONG-TERM DATA MANAGEMENT

3

Coastal Response Research Center (CRRC)

- Partnership between NOAA's Office of Response and Restoration and the University of New Hampshire
 - Emergency Response Division (ERD)
 - Assessment and Restoration Division (ARD)
 - Marine Debris
- Since 2004
- Co-Directors:
 - UNH – Nancy Kinner
 - NOAA – Mark W. Miller

DWH LONG-TERM DATA MANAGEMENT

4

Overall CRRC Mission

- Conduct and oversee basic and applied research and outreach on spill & environmental hazard response and restoration
- Transform research results into practice
- Serve as hub for spill /environmental hazards R&D
- Facilitate workshops bringing together ALL STAKEHOLDERS to discuss spill/hazards issues and concerns

DWH LONG-TERM DATA MANAGEMENT

5

FACILITATION PLEDGE

- I will recognize and encourage everyone to speak
- I will discourage side conversations
- I commit to:
 - Being engaged in meeting
 - Keeping us on task and time
- Stop me if I am not doing this!

DWH LONG-TERM DATA MANAGEMENT

6

PARTICIPANT INTRODUCTIONS

- **Name**
- **Affiliation**
- **Work related to DWH LT Data Management**

DWH LONG-TERM DATA MANAGEMENT

7

Workshop Organizing Committee

NCEI	Julie	Bosch
GOMA/NGO	Laura	Bowie
RESTORE Council	Jessica	Henkel
Centers of Excellence	Libby	Fetherston
Academic/GOMRI	Jim	Gibeaut
BOEM	Jonathan	Blythe
State Rep (AL) - Geological Survey of Alabama	Steve	Jones
IOOS	Barb	Kirkpatrick
NCEI	Kirsten	Larsen
NGO: Ocean Conservancy	Matt	Love
Department of the Treasury	Laurie	McGilvray
NOAA ORR EDDM Lead	Amy	Merten
Data Generator	Tamay	Özgökmen
NOAA NMFS	Mike	Peccini
NFWF	Jon	Porthouse
NOAA NMFS	Jamey	Redding
State Rep (FL)	Dave	Reed
NAS	Lauren	Showalter
DOI and Restore Council (USGS)	Greg	Steyer

DWH LONG-TERM DATA MANAGEMENT

8

PARTICIPANT PLEDGE

- Be Engaged
 - Turn off cell phones & laptops(except at breaks)
- Listen to Others
- Contribute
- Speak Clearly; Use Microphones
- Learn from Others
- Avoid Side Conversations

DWH LONG-TERM DATA MANAGEMENT

9

AGENDA – DAY 1, June 7 Plenary Sessions

- Report on Participant Survey re: Vision of Long Term Data Management in the Gulf [*Jessica Henkel*]
- Overview of Data Generation [*Michele Jacobi*]
- Overview of Data Users [Matt Love]
- Overview of Long Term Data Management [*Lauren Showalter*]
- Overview of Existing Long Term Data Management Systems
 - NOAA ORR (DIVER, ERMA) [*Ben Shorr*]
 - NOAA Restoration Center [*Mike Peccini*]
 - NOAA NCEI [*Rost Parsons*]
 - GRIIDC [*Jim Gibeaut*]
 - GOMA Portal [*William Nichols*]
 - DWH Project Tracker [*Laura Bowie*]
 - LA Coastal Information Management System [*Craig Conzelmann*]
 - GCOOS [*Matt Howard*]

DWH LONG-TERM DATA MANAGEMENT

10

AGENDA – June 7 WEDNESDAY AFTERNOON

- Data Management Standards / Protocols
 - Data Management Frameworks
 - Restore Council [*Jessica Henkel*]
 - NRDA Restoration [*Jamey Redding*]
 - Direct Component & Centers of Excellence – Treasury [*Laurie McGilvray*]
 - NFWF [*Jon Porthouse*]
 - NAS [*Lauren Showalter*]
 - Standards Identified / Parameters / Guidance / Challenges [*Greg Steyer*]
- EDDM Working Groups
 - Field Protocols [*Steve Ramsey*]
 - Common Data Model [*Dan Hudgens*]
 - “Gold” Standard [*Julie Bosch*]

DWH LONG-TERM DATA MANAGEMENT

11

Breakout Session I: Issues / Concerns for Data Stakeholders

- Stakeholders:
 - Data User
 - Data Generator
 - Data Manager/Governor
- Discussion:
 - Reactions to plenary sessions
 - LTDM Requirements from Stakeholder Perspective:
 - Interoperability
 - Ease of Discovery/Searchability
 - Data Access
 - Data Synthesis
 - Data Usability
 - Metadata/Data Documentation
 - Other?

**Determine
Priority Issues
Used During
Day 2
Discussions**

DWH LONG-TERM DATA MANAGEMENT

12

Session I Breakout Groups:

Group A: Data User	Group B: Data Generator	Group C: Data Manager/Governor
Lead: Mark Miller	Lead: Dan Hudgens	Lead: Rost Parsons
Recorder: Melissa Gloekler	Recorder: Megan Verfaillie	Recorder: Kathy Mandsager
Holly Binns	Courtney Arthur	Jonathan Blythe
Melody Chimahusky	Jennifer Bauer	Julie Bosch
Jay Coady	Ryan Druyor	Laura Bowie
Linda Cook	Neal Etre	Craig Conzelmann
Alyssa Dausman	Carl Ferraro	Steve Delgreco
Nic Eckhardt	Shawn Fisher	Sandra Ellis
Jim Gibeaut	George Graettinger	Lei Hu
Jessica Henkel	Mark Howard	Christina Hunnicutt
Amy Hunter	Dan Hudgens	Michele Jacobi
Helga Huntley	Ann Jones	Steve Jones
JB Huyett	Kirsten Larsen	Laurie McGilvray
Syed Khalil	Matt Love	Marti McGuire
Barb Kirkpatrick	Kate McClure	Mike Peccini
Julien Lartigue	Amy Merten	Dave Reed
Gareth Leonard	Tamay Ozgokmen	Denise Reed
Nadia Martin	Steve Ramsey	Angela Schrift
William Nichols	Rick Raynie	Greg Steyer
Jon Porthouse	Jamey Redding	Tom Strange
Lauren Showalter	Ben Shorr	Kevin Suir
Danny Wiegand	Hugh Sullivan	Jason Weick
Eric Weissberger	Carrie Wall	Caitlin Young
Dwane Young		

DWH LONG-TERM DATA MANAGEMENT

13

Data Users:

2. List of requirements from the user community:

Requirements	Features/Needs	Challenges	Priority (high, med, low)
1. Interoperability			
2. Ease of Discovery/Searchability			
3. Data Access			
4. Data Synthesis			
5. Data Usability			
6. Metadata/Data Documentation			
7. Other?			

DWH LONG-TERM DATA MANAGEMENT

14

Data Generators Topic 1:

Requirements	Incentives:	Challenges:	Cost (high, med, low)	Priority (high, med, low)
1. Interoperability				
2. Ease of Discovery/Searchability				
3. Data Access				
4. Data Synthesis				
5. Data Usability				
6. Metadata/Data Documentation				
7. Other?				

DWH LONG-TERM DATA MANAGEMENT

15

Data Generators: Topic 2

	Incentives:	Challenges:	\$ Cost (optional) (high,med,low):
Individual agency requirements			
1. ??			
2.			
3.			
Common set of metadata standards across community			
Maintenance of datasets in perpetuity			
Timely availability to user community			
Accessibility (use/reuse) of data in future			

DWH LONG-TERM DATA MANAGEMENT

16

Data Managers/Governors: Topic 1

1. Workload Effort

Requirements	Challenges:	Priority (high, med, low)
1. Interoperability		
2. Ease of Discovery/Searchability		
3. Data Access		
4. Data Synthesis		
5. Data Usability		
6. Metadata/Data Documentation		
7. Other?		

DWH LONG-TERM DATA MANAGEMENT

17

Data Managers/Governors: Topics 2 and 3

2. Long term repository

Challenges:	Costs:

3. Definition of success

	Characteristics/ Metrics
Data management system	
Data repository	

DWH LONG-TERM DATA MANAGEMENT

18

Data Managers/Governors: Topic 4

Barriers

	Barriers
Getting generators to submit data	
Finding data	
Using data	
Enforcing data policy	
Data security	

DWH LONG-TERM DATA MANAGEMENT

19

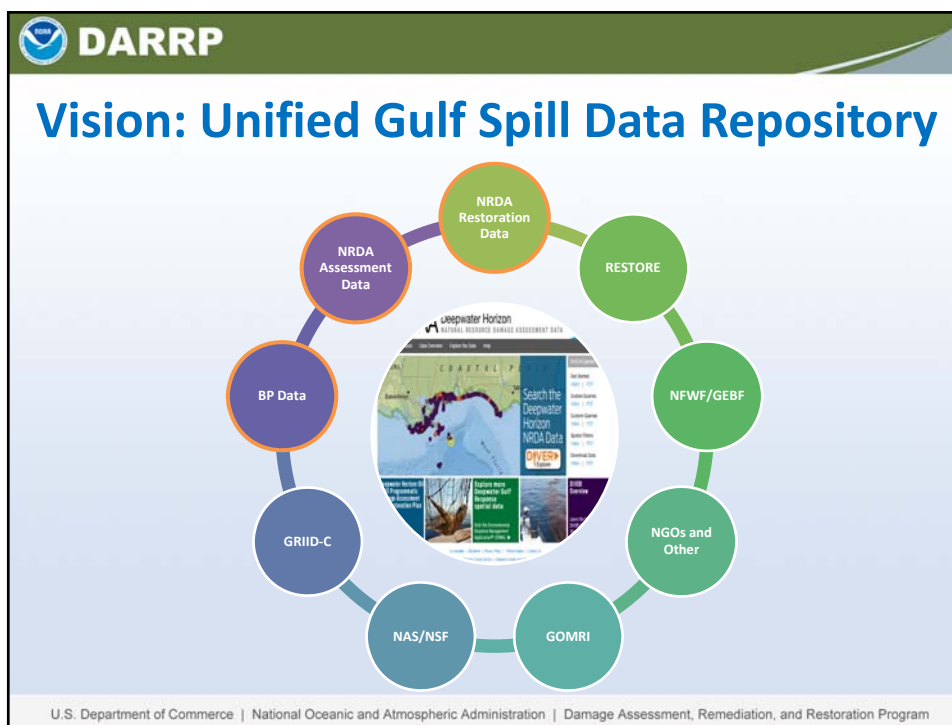
General Counsel for Natural Resources • Office of Habitat Conservation • Office of Response and Restoration

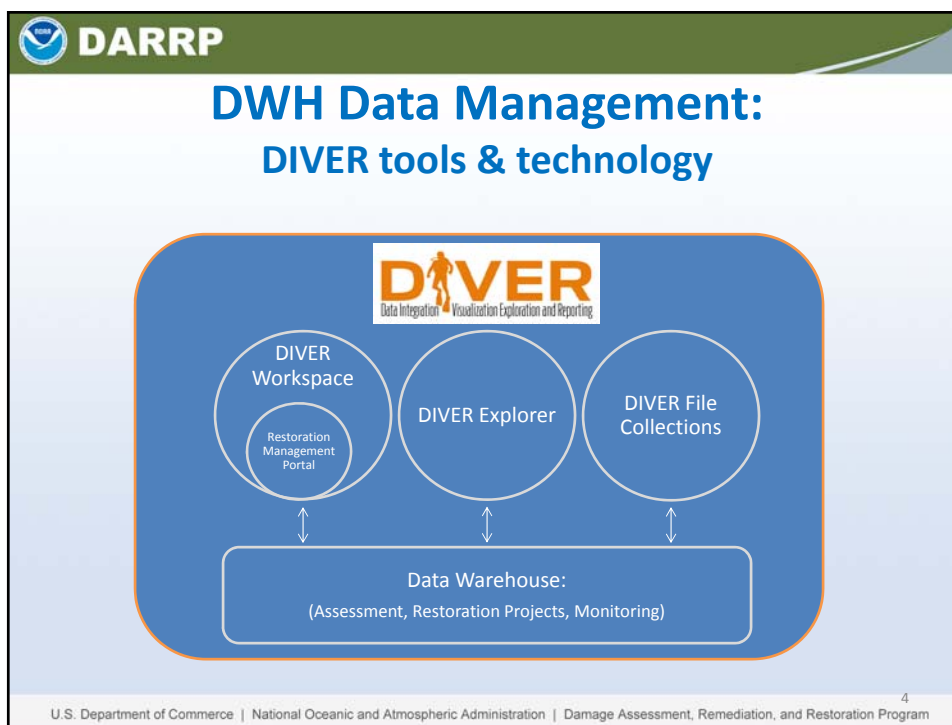
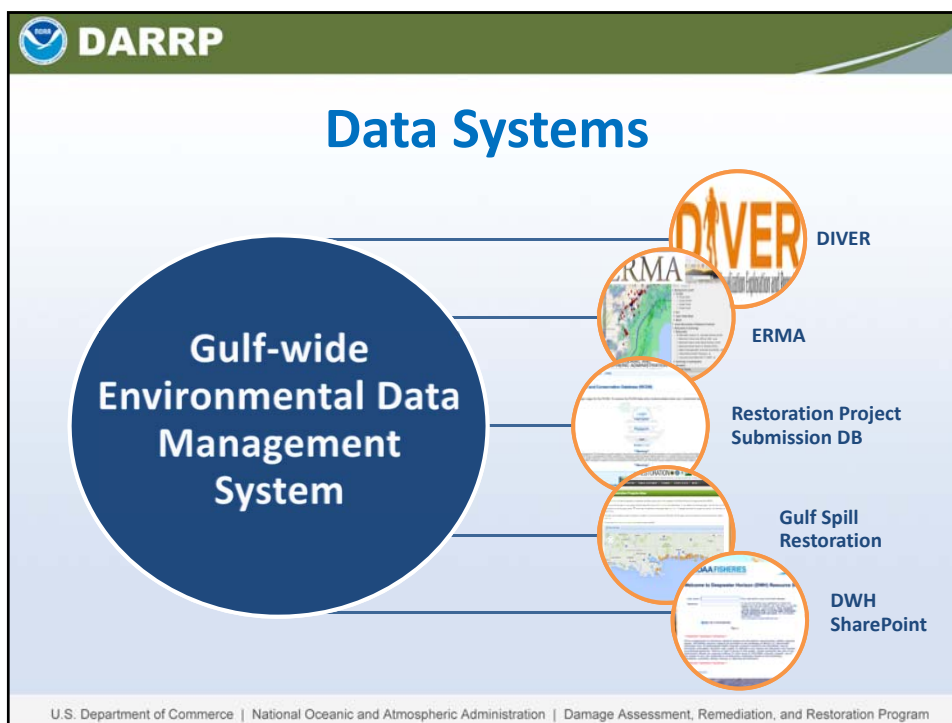
DARRP

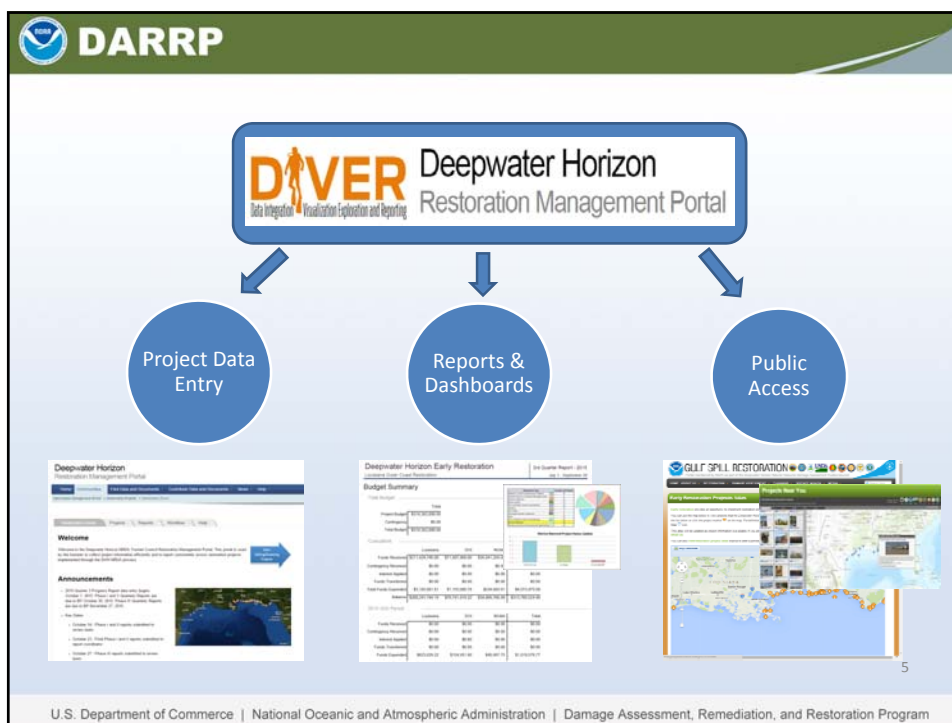
DAMAGE ASSESSMENT, REMEDIATION, & RESTORATION PROGRAM

Long-Term Data Management
Post-Deepwater Horizon

Amy A. Merten, Ph.D.
 Office of Response and Restoration
 June 7th, 2017








The screenshot shows the DIVER Restoration Management Portal interface for the Louisiana Restoration TIG. The top navigation bar includes links for Home, Data Overview, Reference Library, Activities Database, and a DIVER Explorer button. The main content area is titled "Louisiana Restoration TIG" and includes a "Dashboard" with three charts (Dashboard Chart A, B, and C), a "RECENT PROJECT ACTIVITY" section with a list of projects, a "DOCUMENTS" section with a table of documents, and a "CALENDAR" section showing dates and events. The footer includes the DIVER Explorer logo and the text "LOUISIANA RESTORATION TIG".

- Trustee Council and TIG homepages and dashboards
- TIG-level access permissions
- Flexible publishing to outreach products
- Reporting by TIG and restoration type
- Data query tools with seamless integration of assessment, restoration and monitoring data

U.S. Department of Commerce | National Oceanic and Atmospheric Administration | Damage Assessment, Remediation, and Restoration Program

**DARRP**

Workshop Objectives

- Foster collaboration among the Gulf of Mexico partners with respect to data management and integration for restoration planning, implementation and monitoring.
- Identify standards, protocols and guidance for long term data management being used by these partners for DWH NRDA, restoration, and public health efforts.
- Obtain feedback and identify next steps for the work completed by the Environmental Disasters Data Management (EDDM) Working Groups.
- Work towards best practices on public distribution and access of this data.

U.S. Department of Commerce | National Oceanic and Atmospheric Administration | Damage Assessment, Remediation, and Restoration Program

**DARRP**

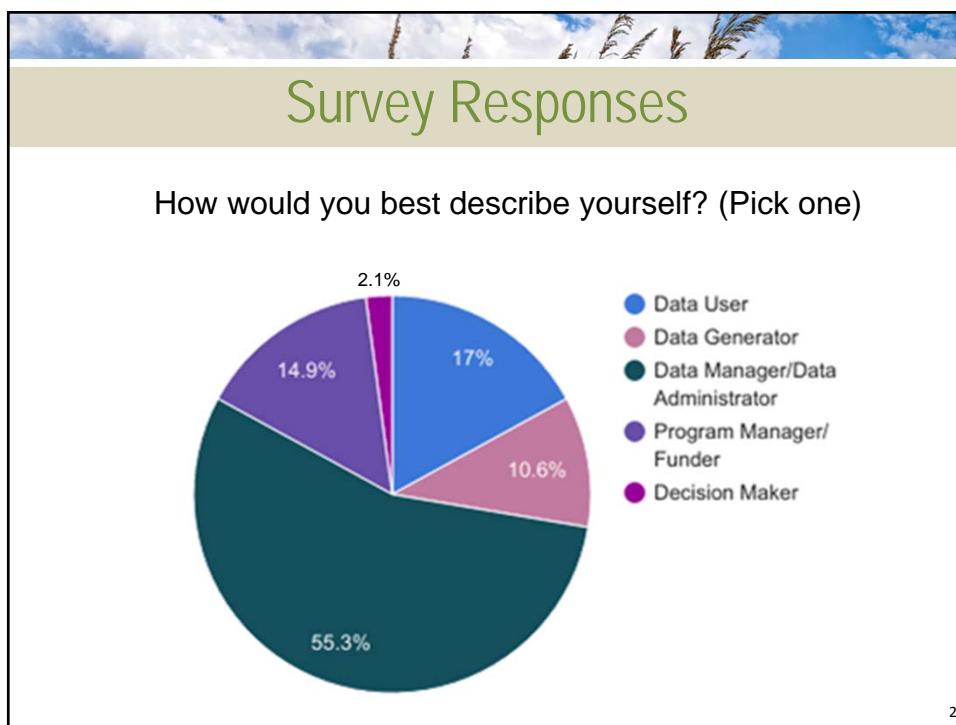
Questions & Discussion

U.S. Department of Commerce | National Oceanic and Atmospheric Administration | Damage Assessment, Remediation, and Restoration Program



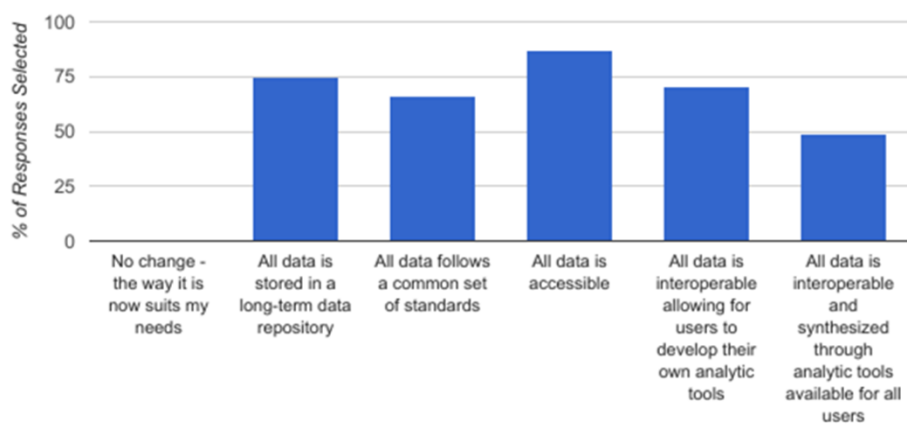
Visions for Long Term Data Management in the Gulf

DWH Long Term Data Management Workshop Participant Survey Responses



Survey Responses

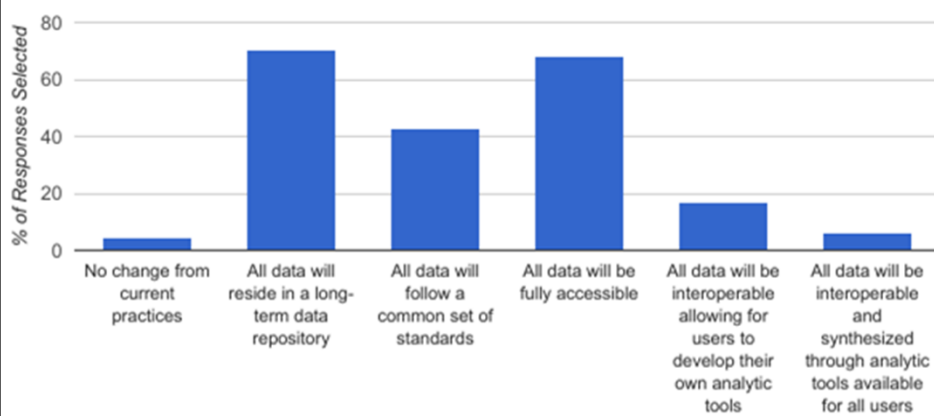
What do you want from Gulf research/monitoring data 15 years from now? (Pick all that apply)



3

Survey Responses

Realistically, what do you think Gulf research/monitoring data will have achieved in 15 years? (Pick all that apply)



4

Survey Responses

What do you see as the biggest challenge to data management in the Gulf?

"One group driving the bus!"

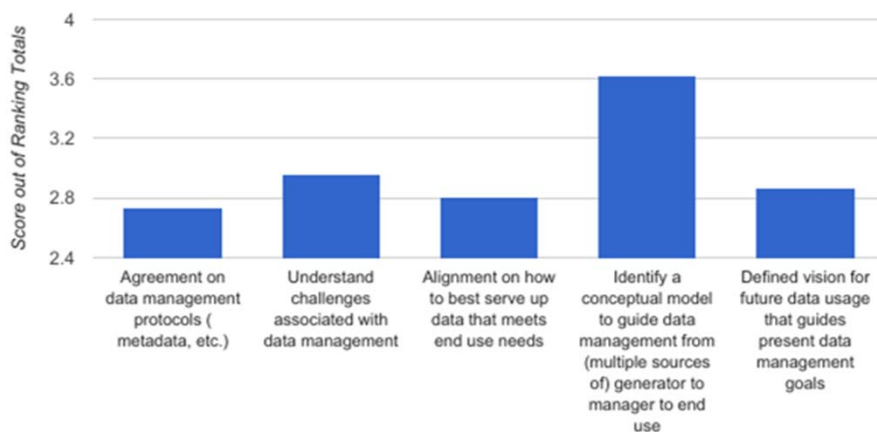
"Identifying how we can tailor data management towards the eventual use of the data on decision making through models, synthesis, etc. Connecting data management and data utilization."

"Move forward with collaboration despite remaining uncertainty. Take a calculated risk that existing data systems can expand to encompass common goals, and will be improved with greater engagement."

7

Survey Responses

What would be the most beneficial outcome of this workshop? (Rank 1-5)



8



Thank you!!

Up Next:

9:45 ***Break***

10:00 PLENARY: Overview of Data
Generation [*Michele Jacobi*]

General Counsel for Natural Resources • Office of Habitat Conservation • Office of Response and Restoration

DARRP

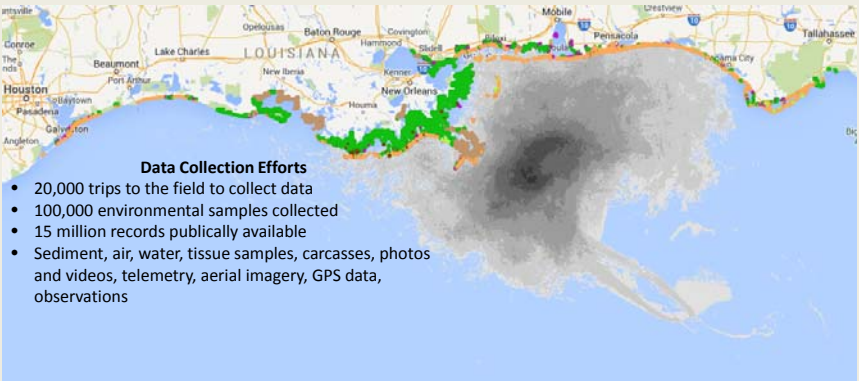
DAMAGE ASSESSMENT, REMEDIATION, & RESTORATION PROGRAM

Overview of Data Generation of the Deepwater Horizon Oil Spill

Michele Jacobi
Office of Response and Restoration
June 7th 2017

NOAA | Office of Response and Restoration

A massive spill, a massive response, a massive NRDA, & a massive research opportunity

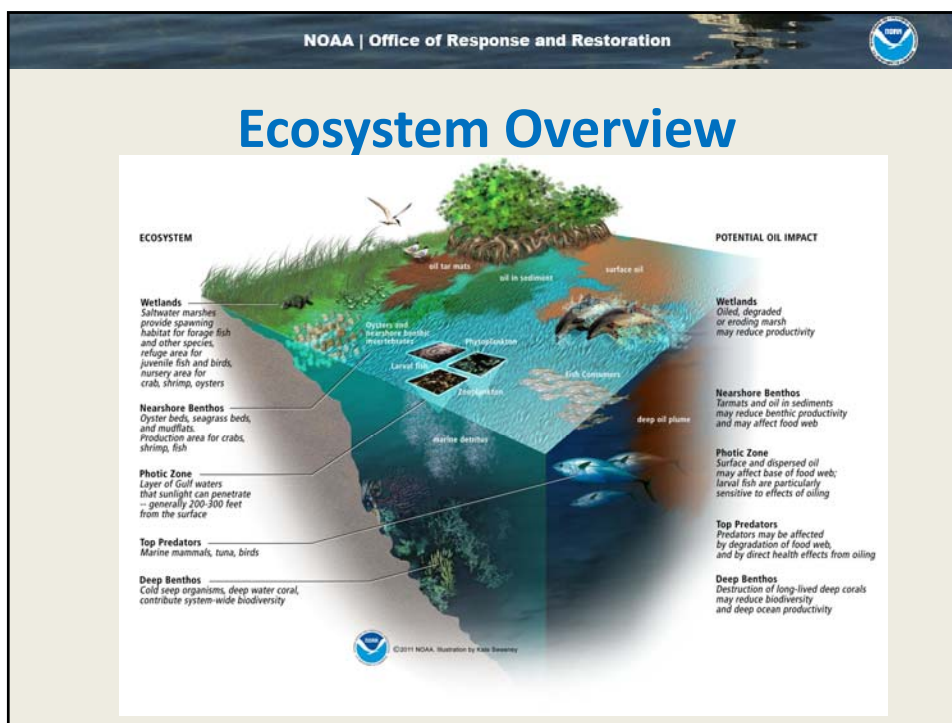


Data Collection Efforts

- 20,000 trips to the field to collect data
- 100,000 environmental samples collected
- 15 million records publicly available
- Sediment, air, water, tissue samples, carcasses, photos and videos, telemetry, aerial imagery, GPS data, observations


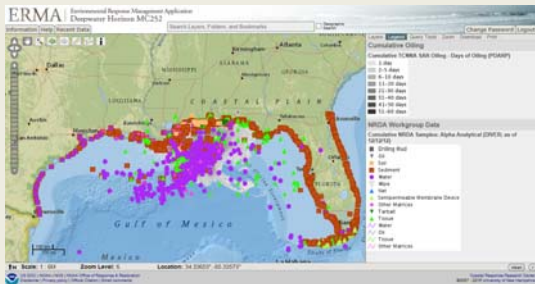
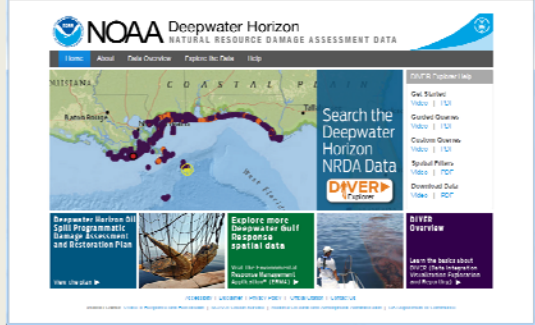
Shoreline Habitat Categories (Draft; 2014-09-29)

- Other
- Beach
- Mainland Herbaceous Marsh
- Back-barrier Herbaceous Marsh
- Delta Phragmites Marsh
- Mangrove



NOAA | Office of Response and Restoration

Data Managed


General Counsel for Natural Resources • Office of Habitat Conservation • Office of Response and Restoration

DARRP

DAMAGE ASSESSMENT, REMEDIATION, & RESTORATION PROGRAM

Response Data & NRDA Data

- Source oil
- SCAT data used to estimate shoreline injury from the quantity of affected shoreline
- Wildlife response data used to understand and quantify animal injuries
- Closures and advisories inform recreational affects
- Photos provide evidence of exposure, and document extent of spill



Following the Deepwater Horizon oil spill, numerous dolphins were documented encountering oil, such as those in this photo from July 2010.

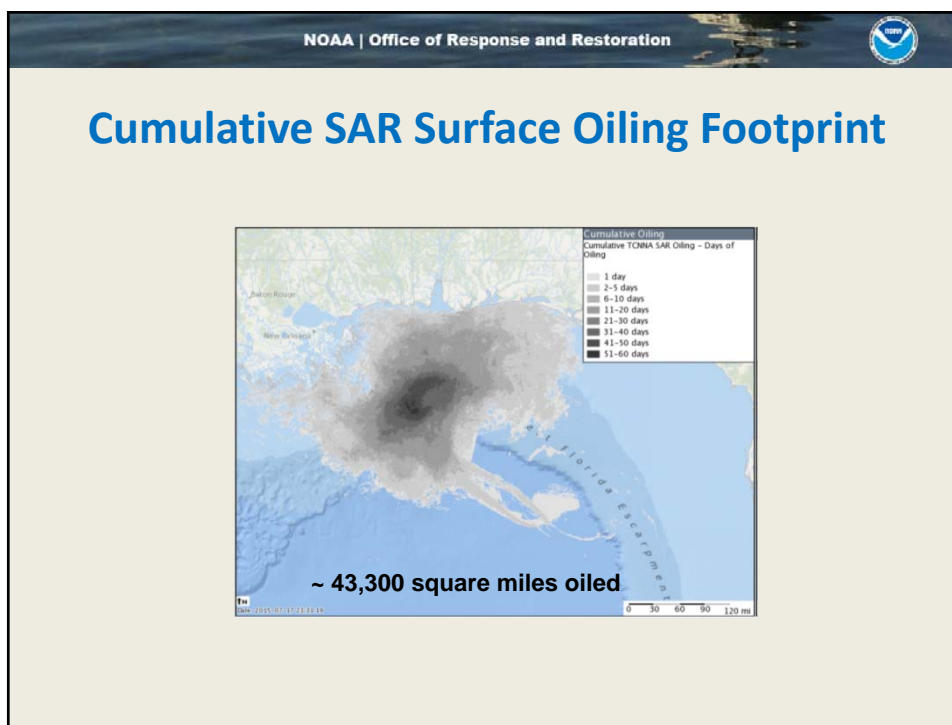
General Counsel for Natural Resources • Office of Habitat Conservation • Office of Response and Restoration

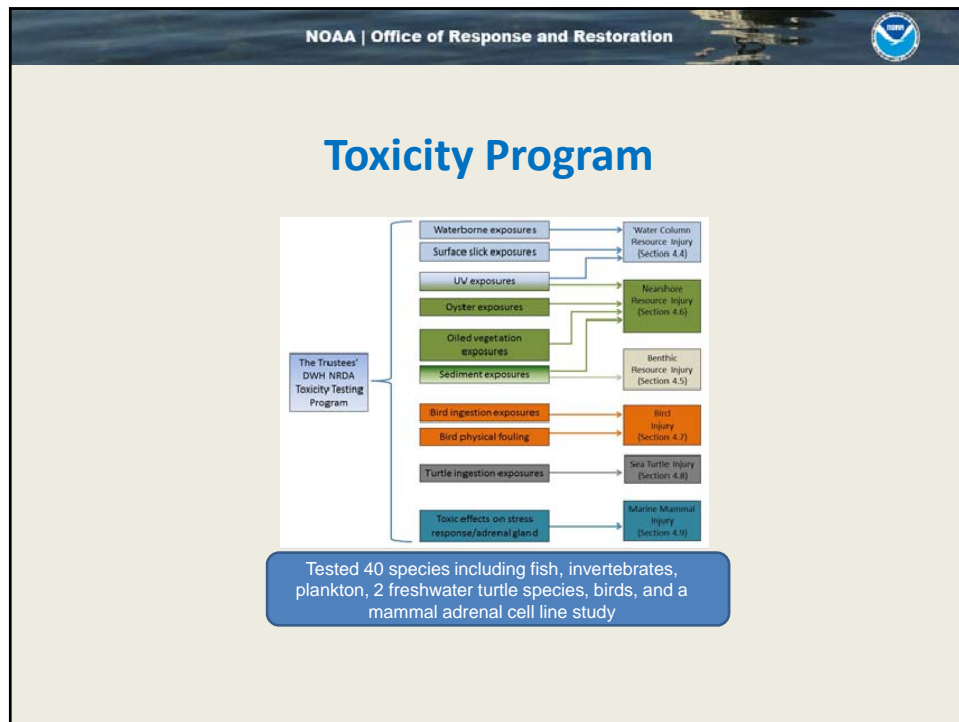
DARRP

DAMAGE ASSESSMENT, REMEDIATION, & RESTORATION PROGRAM

NRDA Activities During a Response

- Collection of ephemeral data
 - Conditions before oil reaches shoreline
 - Fish kills, bird carcasses
 - Fingerprinting of oil on shorelines, in water column
- Studying changes to recreational use & socio economics impacts
- Evaluating older projects and existing monitoring programs-
(access, comparability/ corrections)



General Counsel for Natural Resources • Office of Habitat Conservation • Office of Response and Restoration

DARRP

DAMAGE ASSESSMENT, REMEDIATION, & RESTORATION PROGRAM

Restoration Types and Response Data

Habitat	Recreational	Wildlife
SCAT	Closures/Advisories	Wildlife

<http://www.restoration.noaa.gov/dwh/storymap/>



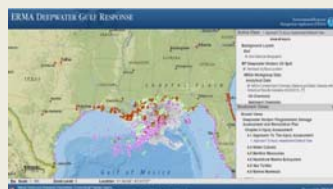
Restoration Monitoring Data

- Trustee Implementation Groups (TIGs) generally consist of both State and Federal trustees
- May leverage existing (historical/on-going) monitoring data
- May leverage data integration capabilities developed for assessment data
- Incorporate QA/QC and data validation
- Data can be publicly accessible through existing gulf-wide environmental data infrastructure

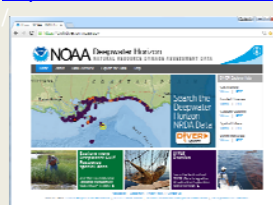


Where can I find the DWH Response and NRDA data?

<http://gomex.erma.noaa.gov>



<https://www.diver.orr.noaa.gov>





NOAA | Office of Response and Restoration

DWH NRDA publications

50+ peer reviewed publications and counting.....

- Deepsea corals and benthos
- Dolphins
- Fish Toxicity
- Sea Turtles
- Oil in the environment
- Forensic chemistry
- Nearshore
- Special theme volume Mar. Ecol. Prog. Ser.

Publications available to public:
<http://response.restoration.noaa.gov/deepwater-horizon-oil-spill/noaa-studies-documenting-impacts-deepwater-horizon-oil-spill.html>

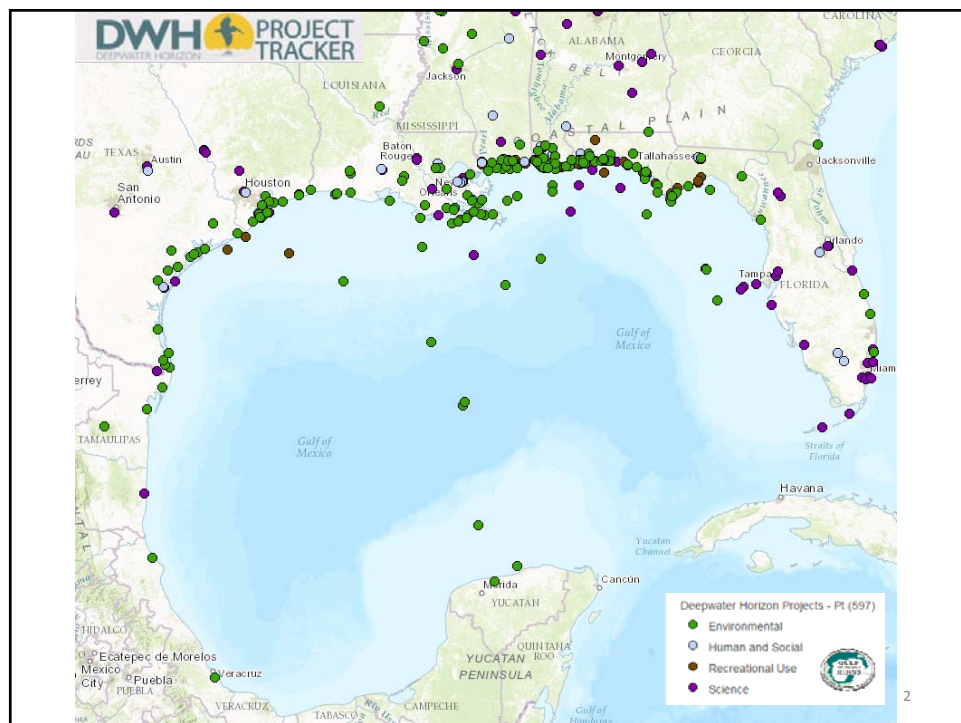
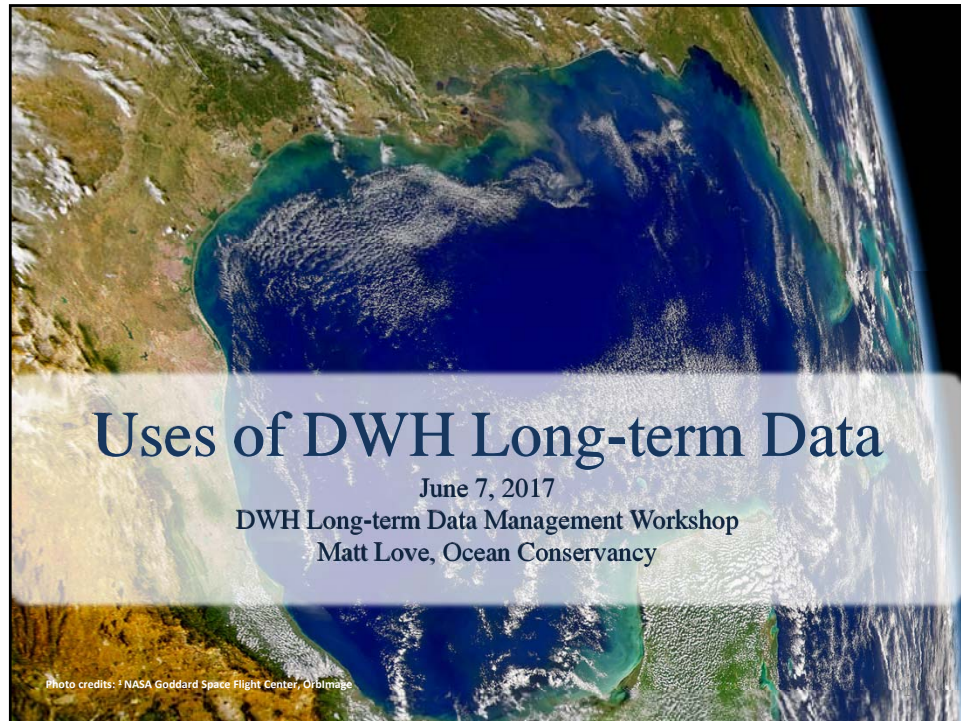
NOAA | Office of Response and Restoration

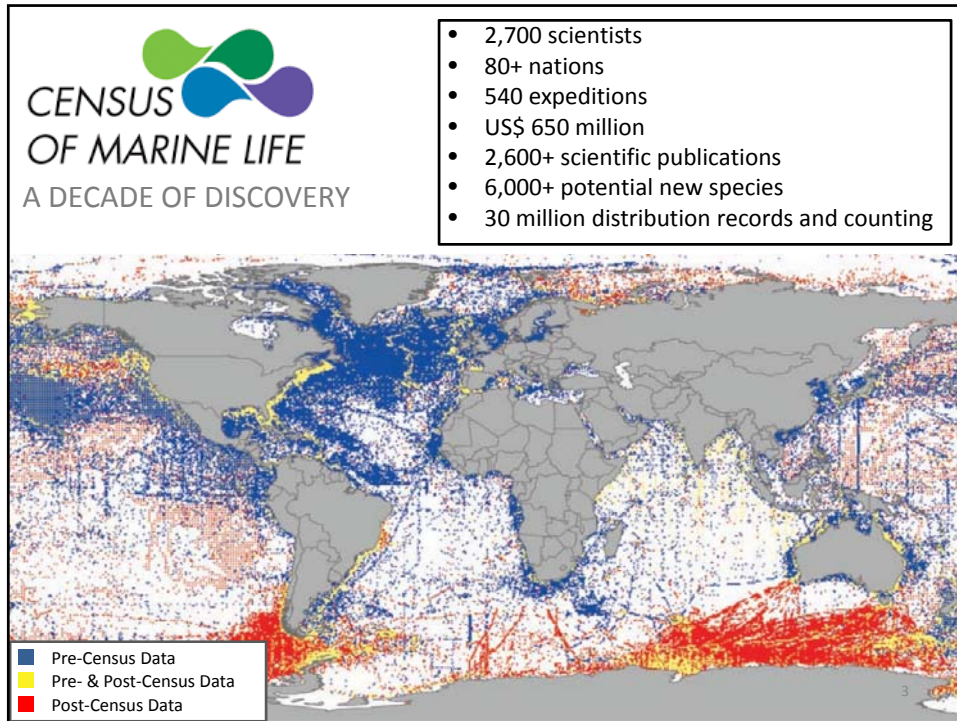
Overarching themes to Remember

Think before you collect!!!

- Consider the area and contaminant in sampling design
- Integrate design and metrics across disciplines to get to appropriate questions and experimental design
- Combine lab and field techniques
- Clearly define your data objective relative to action or decision
- Have a Data Management Plan & BUDGET so the data can be used effectively
- Keep agency data requirements and standards in mind
- Talk to you partners and friends

14







Response

Data needs:

- Common Operational Picture
- Decision Support

Use examples

- Coast Guard Search & Rescue
- Oil Spill Response
- Wildlife Rescue/rehab

Photo credits: U.S. Coast Guard—Reuters/Landov. NOAA ORR

Restoration

1

Data needs:

- Ecosystem function → Multiple Scales
- Decision support

Use examples

- Identify Restoration Need
- Project Level Assessment
- **Ecosystem Scale Evaluation**

2

Photo credits: ¹DWH NRDA Trustees, ²doi:10.1038/ngeo1620.

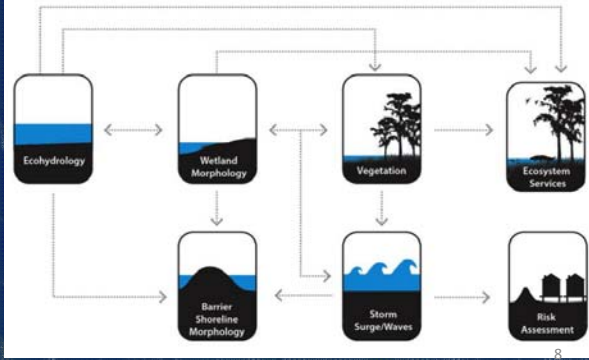


DWH Restoration

Scale of Restoration = Scale of Injury

- Collaboration: Data managers + Data generators + Research + Restoration/Management
- Integration of data types from many sources
- Ecosystem scale modeling

Louisiana Coastal Master Plan Predictive Models



```

graph TD
    Ecohydrology <--> WetlandMorphology[Wetland Morphology]
    WetlandMorphology <--> Vegetation
    Vegetation <--> EcosystemServices[Ecosystem Services]
    EcosystemServices <--> RiskAssessment[Risk Assessment]
    RiskAssessment <--> StormSurgeWaves[Storm Surge/Waves]
    StormSurgeWaves <--> BarrierShorelineMorphology[Barrier Shoreline Morphology]
    BarrierShorelineMorphology <--> Ecohydrology
    Ecohydrology <--> BarrierShorelineMorphology
    WetlandMorphology <--> StormSurgeWaves
    Vegetation <--> StormSurgeWaves
    EcosystemServices <--> RiskAssessment
  
```

8



Research

Data needs:

- Data Discoverability
- Data Access

Use examples

- System-wide Status & Trends
- Ecosystem Scale Evaluation

Photo credits: NOAA ORR



NOAA Technical Memorandum NMFS-SEFSC-706

2017 ECOSYSTEM STATUS REPORT UPDATE FOR THE GULF OF MEXICO

Mandy Karnauskas, Christopher R. Kelble, Seann Regan, Charline Quenée, Rebecca Allee, Michael Jepson, Amy Freitag, J. Kevin Craig, Cristina Carollo, Leticia Barbero, Neda Trifunova, David Hanisko, and Glenn Zapfe

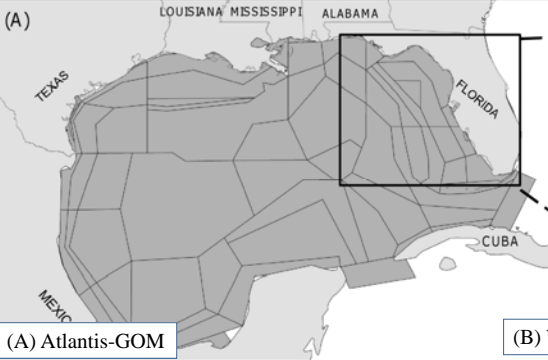
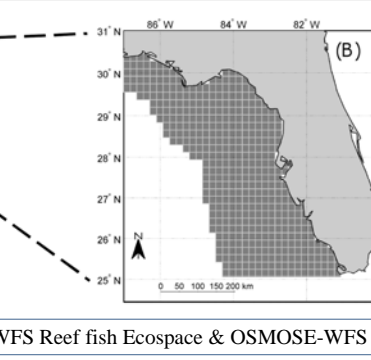


U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Southeast Fisheries Science Center
75 Virginia Beach Drive
Miami, Florida 33149

9

Enhanced Data Applications

- Analytics & Decision Support Tools
- Ecosystem-based Fisheries Management
- Challenges: Data compilation

(A) Atlantis-GOM

(B) WFS Reef fish Ecospace & OSMOSE-WFS

Communication With Data

- Derived Data Products - Information Synthesis

*EVOS Trustee Council
Injured Species List*



- Recovered
- Recovering
- Very Likely Recovered
- Not Recovering
- Recovery Unknown



Goals

Food Provision
Artisanal Fishing Opportunities
Natural Products
Carbon Storage
Coastal Protection
Coastal Livelihoods & Economies
Tourism & Recreation
Sense of Place
Clean Waters
Biodiversity

Foundations in Monitoring



- 20 year program initiated 2012
- Consistent scientific data to detect ecosystem change

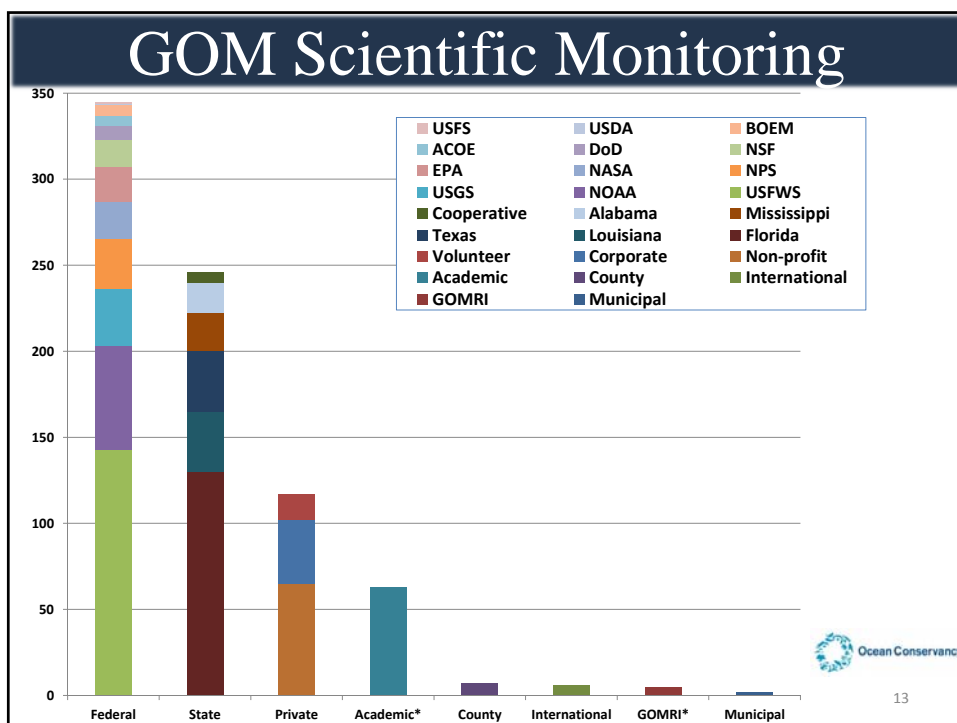
Ecosystem Monitoring Foundation

- Environmental Drivers
- Nearshore Ecosystems
- Pelagic Ecosystems
- Linger Oil



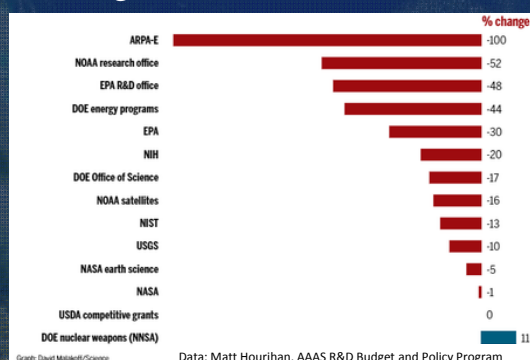
Data Users

Management Agencies
Scientific Research Community
General Public



Data Value Increases With Use

- Every observation is an investment in our understanding
- Collaborative science is the new norm
- Era of defunding science
- We can no longer afford loss of data



Long-term Vision – 15+ years

- Gulf restoration is an opportunity in collaboration
- Successful restoration and management based on science requires open, accessible data
- Need to consider uses of data beyond direct application
- Innovation in science and management requires an integrated information infrastructure



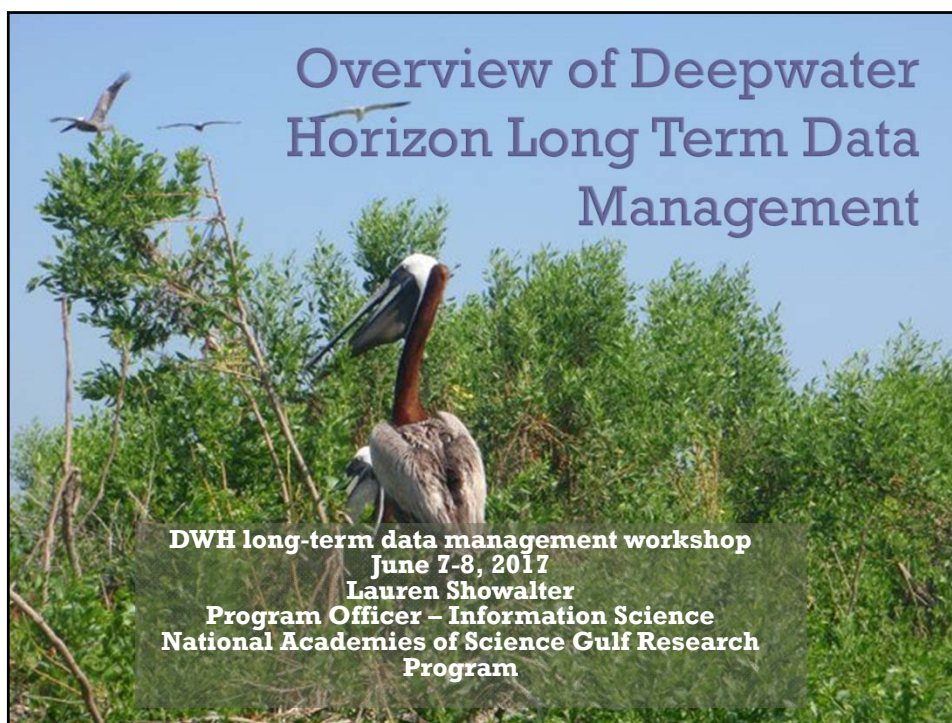
Photo credits: Apollo 11/NASA

15

Discussion Questions

1. What are key constraints or considerations in effectively engaging users in the development of data products?
2. Do you agree with the data users and uses described in this presentation? What types of users do we have at this workshop?

16



DWH LTDM

- **Goals:**

- To ensure the legacy of the science from the DWH disaster is accessible to researchers studying future disasters in the region and around the world
- To improve the quality of science coming out of the funds from the DWH disaster and ensure the research products are well documented and in stable formats

Data Archives

In the field of data management, the terms "archive" and "repository" often are used interchangeably. Within the Federal government, however, the term "archive" is specific to the mission and activities of the National Archives and Records Administration (NARA). Only NARA, or a Federal entity officially delegated by NARA for the long-term curation of specific products, should be referred to as an "archive."

Federal Archival

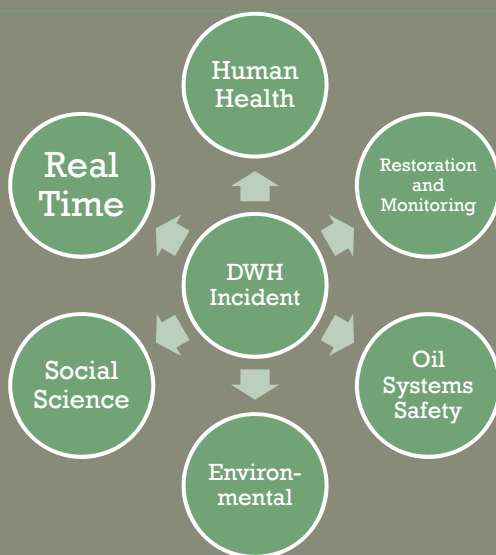
- From Open Archival Information System (OAIS)

...an archive, consisting of an organization, which may be part of a larger organization, of people and systems that has accepted the responsibility to preserve information and make it available for a designated community. It meets a set of such responsibilities as defined in this International Standard, and this allows an OAIS archive to be distinguished from other uses of the term "archive". (from iso.org)

- from NOAA and U.S. National Archives and Records Administration (NARA)

The NOAA National Data Centers are tasked with storing environmental data and making this data available to researchers, scientists, and anyone else that has a need for it, as well as in support of NOAA's mission. Destroy/delete 75 years after cutoff upon approval by NOAA and NESDIS stakeholders. A longer retention may be necessary for research purposes.

Data Types



Baselines

- How do we know we are restoring to previous conditions?
 - What information is available to know that
- Comparable metrics and baselines for monitoring and restoration activities
- What is the new baseline for the GoM since DWH?
- Want to be able to look back at DWH data to answer questions for future spills

Metadata

- Documentation of data is essential to ensure that future users understand how the data was collected and who to contact with questions
- ISO 19115 standard- should be adopted as much as possible
 - This is what the federal government is using
 - Other standards should be able to be transformed into ISO
- Darwin Core could be considered for biological data, for specific repositories

Metadata Creation

ISO 19115-2 Metadata Editor

Load from File Load from Submitted Dataset Save to File Clear Form Check and Save to File Help

Dataset Contact Dataset Information Keywords Data Extent Distribution Info Distribution Contact Metadata Contact

NOTE: Fields with * are required.

Dataset Contact

This section collects contact information about the person primarily responsible for the creation and maintenance of the dataset. Provides future researchers the means to contact the dataset originator if additional information is needed.

*Name	The name of the individual responsible for the creation of the majority of the dataset.
*Organization	Responsible individual's organization, generally this is the University and Department where this person resides. This information can be found at research.gov.i.org
*Position	Position of the responsible individual within their organization, e.g., Assistant Professor.

Resource Provider: Party that supplies the resource.
Custodian: Party that accepts accountability and responsibility for the data and ensures appropriate care and

XML Metadata

```

<gmd:citation>
  <gmd:CI_Citation>
    <gmd:title>
      <gco:CharacterString>2013-2014 NOAA PAH concentrations and granulometry data</gco:CharacterString>
    </gmd:title>
    <gmd:alternateTitle>
      <gco:CharacterString></gco:CharacterString>
    </gmd:alternateTitle>
    <gmd:date>
      <gmd:CI_Date>
        <gco:Date>2016-07-08</gco:Date>
        </gmd:CI_Date>
      </gmd:date>
      <gmd:dateType>
        <gmd:CI_DateTypeCode codeSpace="002" codeListValue="publication"
          codeList="http://www.isotc211.org/2005/resources/CodeList/gmxCodeList.xml#CI_DateTypeCode">creation</gmd:CI_DateTypeCode>
        </gmd:dateType>
      </gmd:date>
    </gmd:CI_Citation>
  </gmd:citation>
  <gmd:abstract>
    <gco:CharacterString>This dataset has CTD and positional data from sediment grabs taken on the NOAA ship Pisces (2013) and Gordon Gunter (2014). Two shoreline collections from Louisiana and Alabama are included in the 2014 data. All sediment grabs on the NOAA cruises were made with a multicorer, unless a Shipex (B) grab is indicated. Samples from the multicorer tubes or the Shipex grab were subsampled to a depth of 5 cm. The two shoreline collections were taken to a depth of 25cm (Shoal Louisiana, shore collection) or to a depth of 5cm (Dauphin Island, hand core). Sediment samples were sent to the Louisiana State University Ag center for granulometry analysis (2013 sediment). The 2014 sediment granulometry was analyzed at Troy University. PAH data was determined at Troy University. The data were collected at the same time as dataset R2.x22.600-0004, which reports trace metals from these same samples.</gco:CharacterString>
  </gmd:abstract>
  <gmd:purpose>
    <gco:CharacterString>To determine PAH concentrations and sand fraction concentrations in sediment collected along the continental shelf in the Gulf of Mexico.</gco:CharacterString>
  </gmd:purpose>
  <gmd:status>
    <gmd:MD_ProgressCode codeSpace="001" codeListValue="completed"
      codeList="http://www.isotc211.org/2005/resources/CodeList/gmxCodeList.xml#MD_ProgressCode">completed</gmd:MD_ProgressCode>
    </gmd:status>
  <gmd:pointOfContact>
    <gmd:CI_Responsibility>
      <gmd:individualName>
        <gco:CharacterString>Stephen Landers</gco:CharacterString>
      </gmd:individualName>
      <gmd:organisationName>
        <gco:CharacterString>Troy University</gco:CharacterString>
      </gmd:organisationName>
      <gmd:positionName>
        <gco:CharacterString>Professor</gco:CharacterString>
      </gmd:positionName>
      <gmd:contactInfo>
        <gmd:CI_Contact>

```

Metadata Training

NOAA HOME WEATHER OCEANS FISHWISER CHARTING SATELLITES CLIMATE RESEARCH COASTS CAREERS

NOAA NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Formerly the National Coastal Data Development Center (NCDDC)... more on NCEI

Search GO

you are here: home → metadata standards → metadata training

NCEI Metadata Training

Metadata is integral to data management and data archive activities. The National Centers for Environmental Information (NCEI) provides metadata training to help data providers and data managers accomplish the goal of providing discovery-level, access-level and understanding-level metadata for their geospatial and environmental data.

NCEI's metadata training focus is on the development of ISO 19115-2 and related standards in accordance with NOAA's Data Documentation Directive. Other metadata topics can be addressed as needed.

"Intro to ISO" Online Metadata Training

This course presents the concept, principles, and value of metadata utilizing the International Organization for Standardization (ISO) 19111-2 metadata in several online sessions. The "Intro to ISO Metadata" course will focus on how the ISO 19115-2 metadata standard is organized and will demonstrate how the different levels of metadata content (discovery, access, and understanding) are expressed in the ISO metadata standards.

The course consists of six separate one-hour online modules using a GoToWebinar format.

Course Modules

- Introduction to Metadata, Current Metadata Policies
- Intro to XML and Intro to UML
- Discovery Level ISO 19115-2 Metadata, Overview of Data Catalogs
- Metadata Content for Access
- Metadata Content for Understanding
- ISO Metadata Creation Methods - Tools and Templates

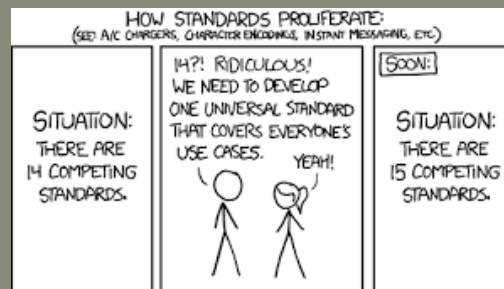
Please contact us at NCEI.Info@noaa.gov with questions.

Point of Contact
NCEI@noaa.gov
866.732.2382

OCEAN FACTS

Data Standards

- Identification of standards early in the process and get community buy in
- Standards need to be adequately communicated to data collectors



Federal Mandates

- The Digital Government Strategy and Open Data Policy were developed for the Government to better deliver information (data) and services.
- Federal agencies are under certain mandates that could inhibit data from being accepted if not properly formatted and documented
- *Common Framework for Earth-Observation Data*, March 2016, Office of Science and Technology Policy
- These standards need to be properly communicated to the data collectors

Data Sharing

- Length of time from collection to sharing
 - Real time data – data that is shared as soon as it is collected
 - Cruise data
 - Satellite data
 - Buoy data
 - Other data is shared depending on:
 - Funder/publisher requirements
 - Federal or state mandates
 - Requirements of other collaborators (foreign, private, industry, etc.)

Data Holds

- If data is to be held for any reason the documentation of that data should begin before it is submitted for public access.
- Groups that start documentation before the data is collected have a leg up when the data is ready for publication
- Tracking of data from project onset is essential



Data Citation

- support proper attribution and credit
- support collaboration and reuse of data
- enable reproducibility of findings
- foster faster and more efficient research progress
- provide the means to share data with future researchers

When using this data, please cite the original publication:

Portney DL, Puritz JB, Hohenbeck CM, Gotschke J, Chapman D, Gold JR (2015) Selection and size based dispersal in a coastal shark: the influence of phylogeny on adaptive variation. *Molecular Ecology* 24(23): 5977–5989. <https://doi.org/10.1111/mec.13441>

Additionally, please cite the Dryad data package:

Portney DL, Puritz JB, Hohenbeck CM, Gotschke J, Chapman D, Gold JR (2015) Data from: Selection and size based dispersal in a coastal shark: the influence of phylogeny on adaptive variation. Dryad Digital Repository. <https://doi.org/10.5061/dryad.7k6c1>

[Cite](#) | [Share](#)

Download the data package citation in the following formats:

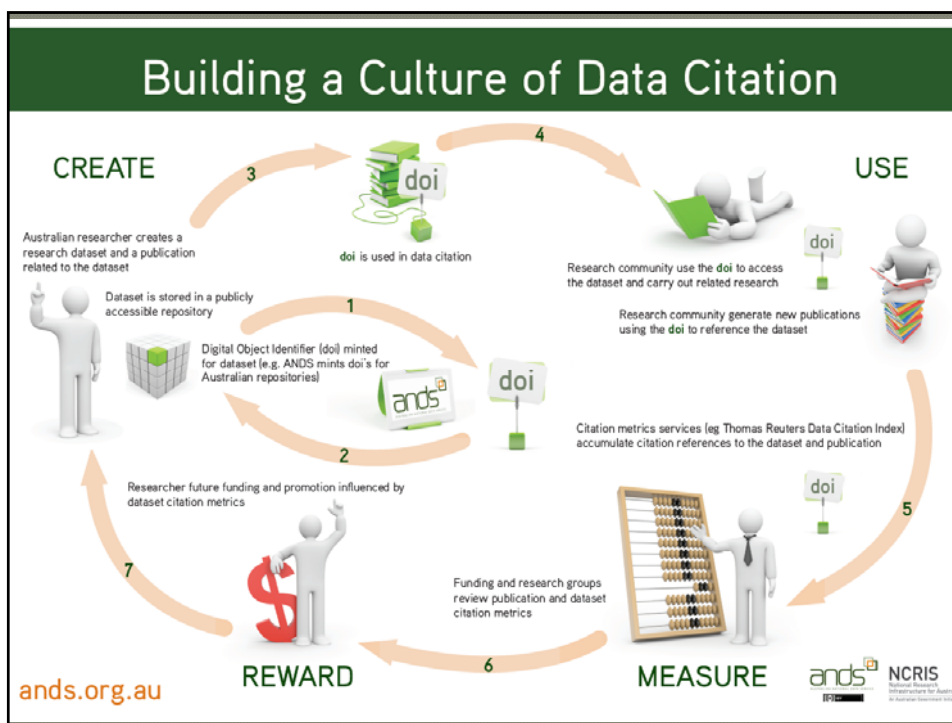
RIS (compatible with EndNote, Reference Manager, ProCite, RefWorks)
BibTex (compatible with BibDesk, LaTeX)

DOI: 10.5061/dryad.7k6c1
Title: R.V. Adams (ALVIN (DRV-2): AZ20-11 Hydrographic (CTD) Data, Northern Gulf of Mexico, March 11 - April 23, 2014
Author(s): Susanna Jeps
Metadata: 10.5061/dryad.7k6c1
File name: 10.5061/dryad.7k6c1
Estimated Download Time: 19 seconds (based on your current connection speed)
Detailed Description:
alt:
doi:
doi:10.5061/dryad.7k6c1

Citation suggestion:
Susanna Jeps (2017) R.V. Adams (ALVIN (DRV-2): AZ20-11 Hydrographic (CTD) Data, Northern Gulf of Mexico, March 11 - April 23, 2014. Distributed by: Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC), Data Research Institute, Texas A&M University-Corpus Christi. doi: 10.7267/252177

Persistent Identifiers

- A Digital Object Identifier (DOI) is a commonly used type of identifier that is used to link to digital objects.
- Use of a persistent identifier makes data search and accessibility easier for future users
- Open Researcher and Contributor ID (ORCID) is a persistent identifier for researchers



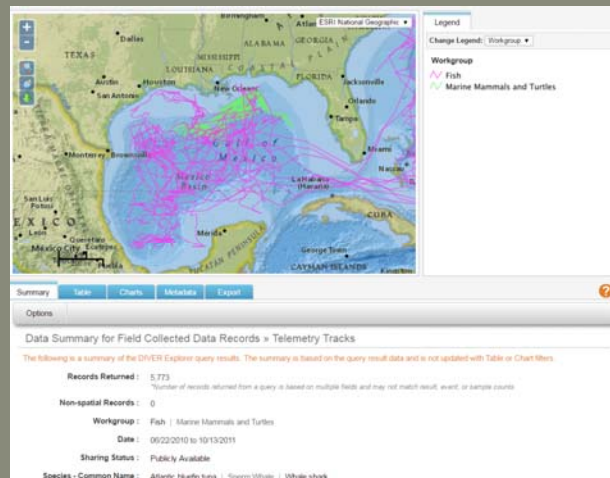
Data Access

- Providing services that allow users to retrieve data for exploration, analysis, or decision making
- Rely on sets of common standards and protocol (e.g. OPeNDAP, WMS, WCS, ERRDAP, FTP, SOS)
- Often community-driven
- Need for both human access and machine-to-machine access

Distributed Data Frameworks



Telemetry Data



Visualization examples

- <https://story.maps.arcgis.com/apps/MapSeries/index.html?appid=597d573e58514bdbbeb53ba2179d2359>
- <https://storm.pps.eosdis.nasa.gov/storm/cesium/EventViewer.html?position=-84.50,20.44,1199998&view=6.28,-1.00,6.28&fname=2A.GPM.DPR.V620160118.20170524-S141109-E144108.V04A.RT-H5>

General Counsel for Natural Resources • Office of Habitat Conservation • Office of Response and Restoration

DARRP

DAMAGE ASSESSMENT, REMEDIATION, & RESTORATION PROGRAM


DIVER and ERMA: Data and Visualization

DWH Long Term Data Management June 7-8, 2017

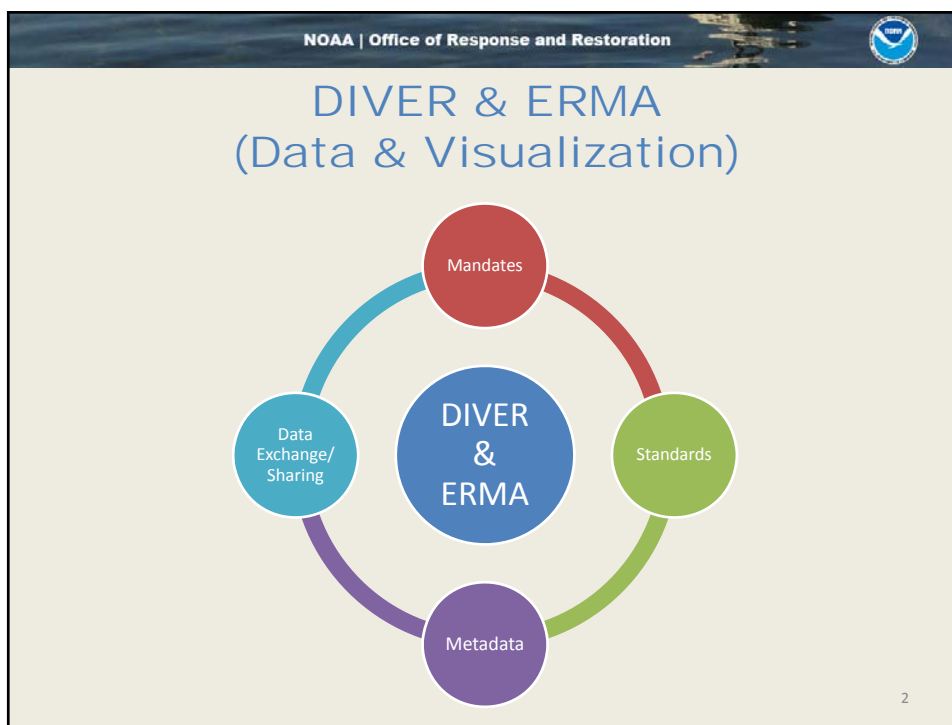
Ben Shorr (presenting)

Dr. Amy Merten, Marti McGuire, Mike Peccini, Jamey Redding
Nick Eckhardt, Jay Coady, Michele Jacobi, George Graettinger

NOAA - NOS - Office of Response & Restoration
Assessment & Restoration Division: Spatial Data Branch



1





Overview

DIVER

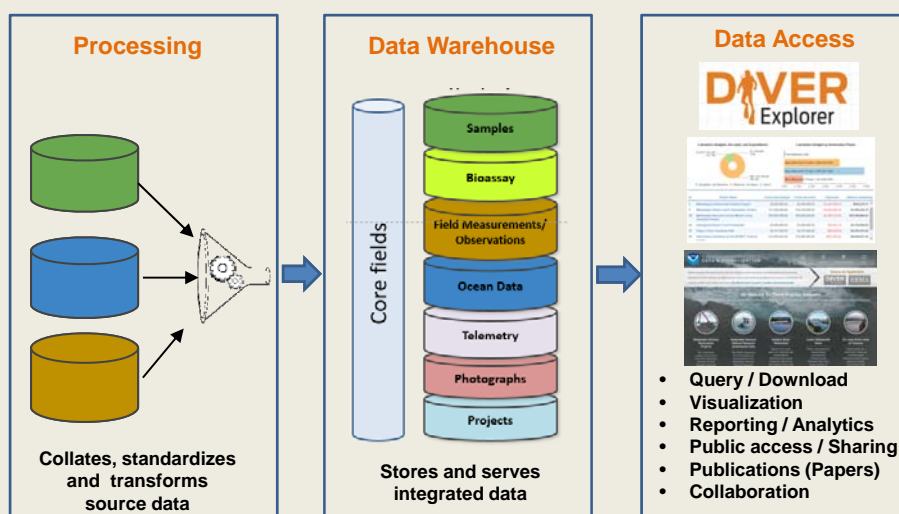
(Data Integration, Visualization, Exploration, and Reporting)

DIVER is a **data warehouse, query tool, and collaboration** application. The DIVER approach **integrates standardized** datasets so users can query across data holdings and download **information and results**.

3



What is DIVER?



4

NOAA | Office of Response and Restoration

DIVER: Key Capabilities

Data Integration and Sharing

- **Integrate** and **standardize** data from multiple sources (e.g. field-collected data, laboratory data, monitoring data, analysis)
- **Query, Export, Reporting** tools
- Federal **data sharing requirements** (e.g. Open Data Policy; PARR; ISO Metadata)


Cloud-based Infrastructure

Secure

- Federal **IT Security** Requirements

Scalable/Flexible

- Evolving application to meet evolving needs



DIVER Explorer is a querying tool that pulls together a broad array of environmental and restoration project data from numerous sources. The tool allows users to search, filter, and download data of interest.

Go Directly to a Regional DIVER Explorer

5

NOAA | Office of Response and Restoration

DWH Damage Assessment Data

- **20,000** trips for field data collection
- **1 million** field data forms and related electronic files
- **100,000** water, tissue, oil and sediment samples
- **15 million+** database records
- **30 terabytes** of data
- Data Referenced in many Publications/Journals

Shoreline Data

Marsh Assessment

Oyster Collections

Telemetry Data

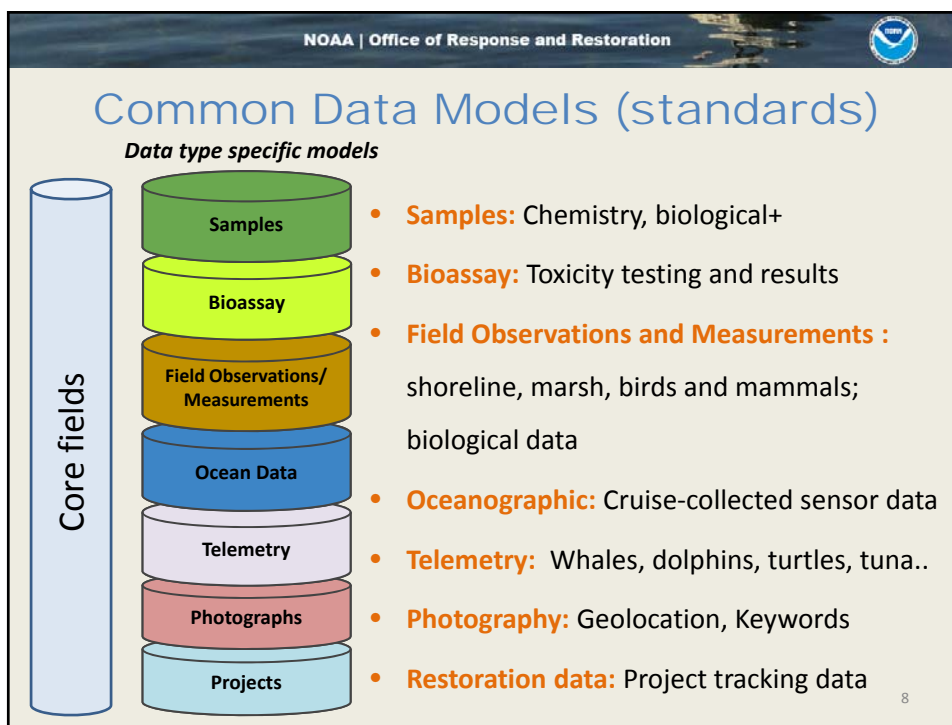
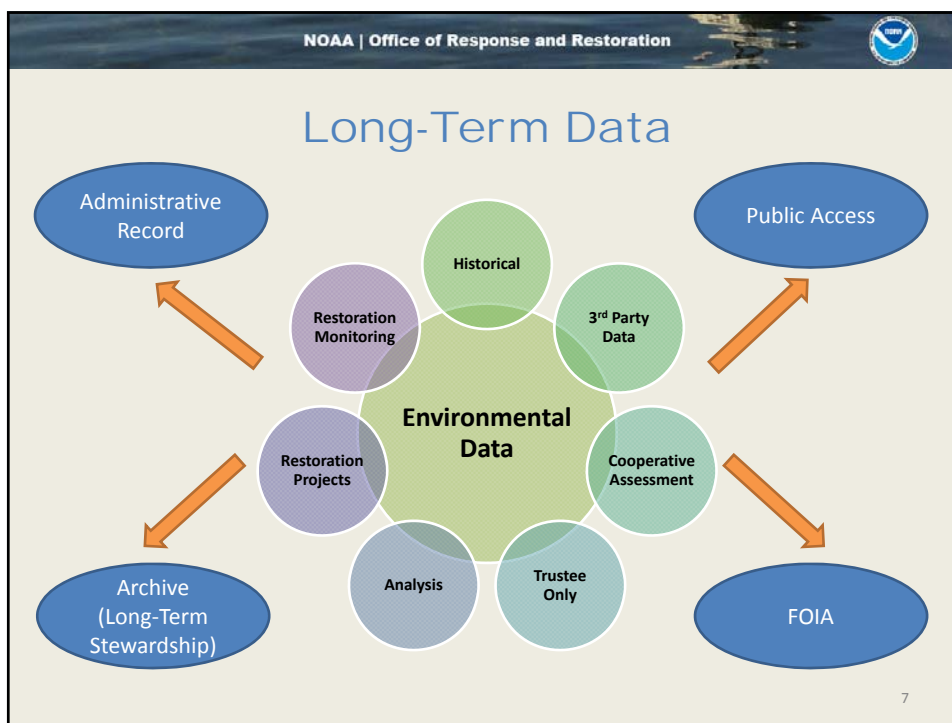
Toxicity Data

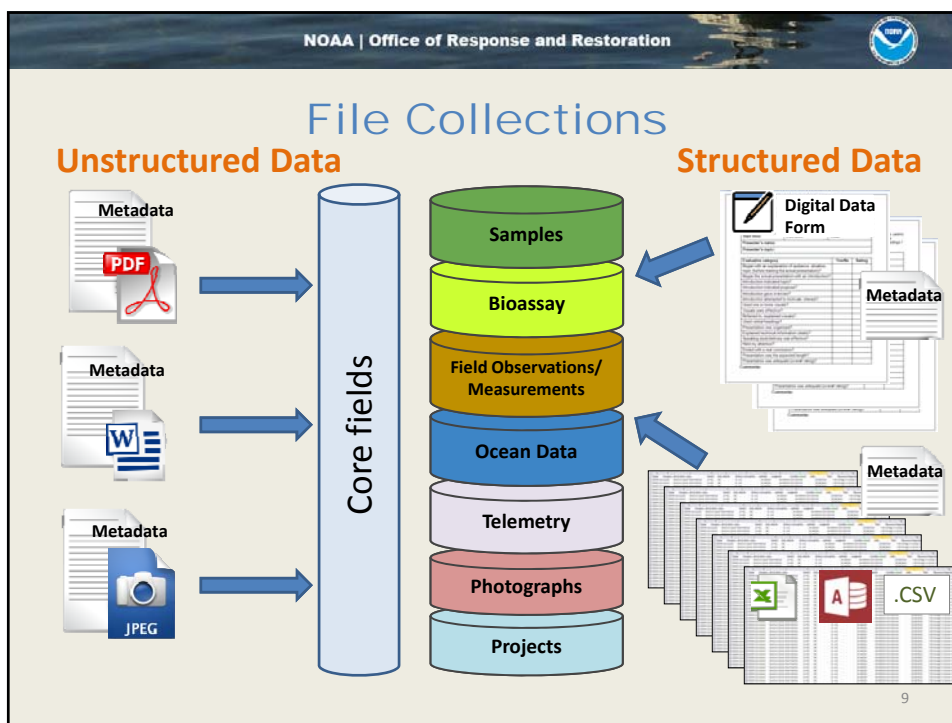
Water Column

Seafood Safety

Marine Mammal & Turtle Assessment

6





NOAA | Office of Response and Restoration

National DIVER Portal Overview

The screenshot shows the National DIVER Portal interface. At the top, there's a navigation bar with 'Home', 'About the Data', 'Help', and 'Manage' links. Below this is a 'SELECT A WORKSPACE' section with a list of favorites including 'Great Lakes', 'Northwest', 'Northwest & Arctic', 'Southwest', 'A Bay Long Incident LA', 'A Deepwater Horizon', 'Alabama TIG', 'DWH File Collections', 'DWH Remediation Hazards', 'Equipment and Sample Forms', 'Florida TIG', 'Louisiana TIG', 'Mississippi TIG', 'Open Ocean TIG', 'Regionwide TIG', 'Restoration Plan Tracking', 'Texas TIG', 'Toxicity Data', 'Treasure Coast', 'Demonstration', 'A Green Canyon 248', 'A Hattley Canyon 919 (HRC919)', 'A SE Brown Tracking', 'A South Alabama Mercury NODAR', 'A Taylor Energy', 'User Workspaces', and 'Southwest & Pacific Islands'. The main content area has a welcome message: 'Welcome to the DIVER Portal. Select an option in the workspace navigation menu to access collaboration features and files specific to your region or activity. For additional information on using the site and functions, please see the help materials located in the About the Data and Help menus.' Below the welcome message are three sections: 'BACKGROUND' (describing the data resource), 'DIVER EXPLORER' (describing the query tool), and 'RELATED TOOLS' (describing ERMA). Each section includes a map of the United States with a highlighted region.

10

NOAA | Office of Response and Restoration

National DIVER Portal Overview

Login environment; data sharing status

Help documents, data details and data management tools including OR&R Activities Database

Quick Access to query data by Regions

Quick Access to ERMA, online mapping and visualization tool

The Portal is organized into

- Regions
- Activities and Workspaces.

Regional pages provide DIVER Explorer filtered to regional data and case-specific data based on user access. Workspaces can contain File Collections, collaboration and data entry tools.

<https://portal.diver.orr.noaa.gov>

11

NOAA | Office of Response and Restoration

DIVER Portal: Restoration

Technical Implementation Groups (TIGs) have customized content and group specific permissions

Project Details

Overview | Budget | Actions | Monitoring | Exit Compliance | Contacts | Documents

Lake Hermitage Marsh Creation - NRDA Early Restoration Project

Project ID: 4 | Project Status: In Progress

Description:
The Lake Hermitage Marsh Creation - NRDA Early Restoration Project involves the creation of marsh within a project boundary located in the Lake Hermitage Marsh Creation Project, developed for and funded through the Coastal Wetlands Planning, Protection and Restoration Act (CWPRA) Program. This project substitutes approximately 104 acres of created marsh for approximately 104 acres of wetland loss that would otherwise have been constructed within the CWPRA project boundary.

The Lake Hermitage Marsh Creation - NRDA Early Restoration Project will create approximately 104 acres of marsh in two of the 7,300 linear feet of eastern levees that was included in the final design of the Lake CWPRA project. This additional marsh area will be constructed entirely within the Lake CWPRA project's levee boundary. Sediment will be hydraulically dredged from a borrow area in the Mississippi River, and dumped via pipeline to create new marsh in the project area. Over time, natural dewatering and compaction of dredged sediments should result in elevations within the marsh design which would be conducive to the establishment of emergent marsh. The 104-acre area will be planted with native marsh vegetation to accelerate benefits to be realized from this project.

The Lake Hermitage Marsh Creation - NRDA Early Restoration Project is located within the Barataria Hydrologic Basin in Plaquemines Parish, Louisiana, to the west of the community of Pointe à la Pêche, and northwest of the community of Magnolia. This basin was identified as a priority area for coastal restoration, and has been the focus of extensive study and project design and implementation.

Trustee Implementation Group(s):
• Louisiana

Implementing Trustee(s):
• Louisiana Coastal Protection and Restoration Authority (CPRA)

Restoration Type(s):
• Wetlands
• Coastal and Nearshore Habitats

Restoration Phase:
Early Restoration Phase 1

Site Location(s):
• Louisiana

12

NOAA | Office of Response and Restoration

Public DIVER website

Public website for Assessment & Restoration data

<https://www.diver.orr.noaa.gov>

13

NOAA | Office of Response and Restoration

DIVER Explorer: Query Result

DIVER Explorer is a query tool. Filter, map and download. Dashboard display.

14

NOAA | Office of Response and Restoration

Dashboards

DIVER Portal

SELECT A WORKSPACE

Home

Southwest

Red drum (*Sciaenops ocellatus*); embryo; 24 hr
Slick A HEWAF
Test 201

NOAA's Deepwater Horizon Trustee Toxicity Testing Program Results

Export Data

Species: Common = "Red drum"

Select Download Types ☐ all ☐ none

Results

- ☒ Treatment mean results for each available time point and endpoint (bio_mean.csv)
- ☐ Individual replicate results for each available time point and endpoint (bio_rep.csv)
- ☒ Effect and Lethal Concentrations (ECLC) from modeled concentration-response curve (bio_stat.csv)
- ☐ Details of modeled concentration-response curve (bio_stat_detail.csv)
- ☒ Matched bioassay and analytical chemistry results (bio_chem.csv)
- ☐ Analytical chemistry and calculated results (full_chem.csv)

Supplemental Data

- ☐ Field-collected sample information (field_sample.csv)
- ☐ Test conditions (test_cond.csv)
- ☐ UV exposure doses (uv_doses.csv)
- ☐ Water quality measurements (water_quality.csv)
- ☐ Chemistry notes (chem_notes.csv)

Metadata

All exports contain the following files for selected studies: Study Notes (study_notes.txt), Study References (study_ref.csv), Field Definitions (field_definitions.csv), Qualifier Definitions (qual_definitions.csv), and Metadata (DWH_Tox_ISO19115.html and DWH_Tox_ISO19115.xml).

EXPORT DATA

NOAA | Office of Response and Restoration

Data Access & Sharing

- **Explorer Query Tool**
 - Environmental and Projects data
 - New Search tool
- **Data Services**
 - ERDDAP
 - NCEI Archive
- **DIVER Data Specification**
 - Common Data Models
 - Data Templates

17

NOAA | Office of Response and Restoration

DIVER Explorer: Export Packages

The diagram illustrates the export capabilities of the DIVER Explorer tool. It shows a screenshot of the tool's interface, which includes a map of the Gulf of Mexico, a legend, and a data summary table. Arrows indicate the flow from the interface to four main export categories:

- Mapping/GIS**: Includes icons for Shapefile, GeoTIFF, and KML, along with the ERMA logo.
- Metadata & Study Notes**: Includes the ISO logo and a document icon.
- Data Packages**: Includes icons for a folder of documents, the ISO logo, and a document icon.
- Expanded Data Services**: Includes the ERDDAP logo and the text "Easier access to scientific data".

NOAA | Office of Response and Restoration

Restoration Data

- **\$8.8 billion** settlement paid out over **15 years**
- Funds allocated to **7 Trustee Implementation Groups (TIGs)** across **15 resource types**
- Projects will be implemented by **17 trustee agencies**
- Commitment to **data-driven adaptive management**

Project details, activities & status

Locations

Monitoring

Budget, receipts & expenditures

Environmental compliance

Accomplishments

NOAA | Office of Response and Restoration

DIVER Restoration Portal

Technical Implementation Groups (TIGs) have customized content and group specific permissions

Project Details

Lake Hermitage Marsh Creation - NRDA Early Restoration Project

Project ID: 4 | Project Status: In Progress

Description:
The Lake Hermitage Marsh Creation - NRDA Early Restoration Project involves the creation of marsh within a project footprint located in the Lake Hermitage Marsh Creation Project, developed for and funded through the Coastal Wetlands Planning, Protection and Restoration Act (CWPRA) Program. This project substitutes approximately 104 acres of created marsh for approximately 104 acres of wetlands lost due to erosion within the CWPRA project boundary.

The Lake Hermitage Marsh Creation - NRDA Early Restoration Project will create approximately 104 acres of marsh in two of the 7,300 linear feet of eastern levees that was included in the final design of the Lake CWPRA project. This additional marsh area will be constructed entirely within the Lake CWPRA project's levee boundary. Sediment will be hydraulically dredged from a borrow area in the Mississippi River, and dumped via pipeline to create new marsh in the project area. Over time, natural revegetation and compaction of dredged sediments should result in elevations within the marsh that are suitable for the establishment of emergent marsh. The 104-acre area will be planted with native marsh vegetation to accelerate benefits to be realized from this project.

The Lake Hermitage Marsh Creation - NRDA Early Restoration Project is located within the Barataria Hydrologic Basin in Plaquemines Parish, Louisiana, to the west of the community of Pointe à la Pêche, and northwest of the community of Magnolia. This basin was identified as a priority area for coastal restoration, and has been the focus of extensive study and project design and implementation.

Trustee Implementation Group(s):
• Louisiana

Implementing Trustee(s):
• Louisiana Coastal Protection and Restoration Authority (CPRA)

Restoration Type(s):
• Wetlands
• Coastal and Nearshore Habitats

Restoration Phase:
Early Restoration Phase 1

Site Location(s):

Implementation Timeline(s):
Project Component (Date input in progress)
• Biological Evaluation Exam

NOAA | Office of Response and Restoration

Gulf Spill Restoration Website (Public)

Gulf Spill Restoration Projects
Restoration projects in the Gulf States approved by the Deepwater Horizon MDA Trustees.

Designated Nonpoint Source Trustee Councils: DEQ EPA MDA USFWS AL FL LA MS TX

All Projects FL AL MS LA TX Open Ocean Regionwide

Project ID: 4 | Project Status: In Progress

Lake Hermitage Marsh Creation - NRDA Early Restoration Project

Description:
The Lake Hermitage Marsh Creation - NRDA Early Restoration Project involves the creation of marsh within a project budget known as the "Lake Hermitage Marsh Creation Project" developed for and funded through the Coastal Wetlands Planning, Protection and Restoration Act (CWPRA) Program. This project subsidizes approximately 104 acres of created brackish marsh for approximately 5-6 acres of earthen terraces that would otherwise have been constructed within the CWPRA project boundary.

The Lake Hermitage Marsh Creation - NRDA Early Restoration Project will create approximately 104 acres of brackish marsh in lieu of the 7,300 linear feet of earthen terraces that was included in the final design of the base CWPRA project. This additional marsh area will be constructed entirely within the base CWPRA project's terrace boundary. Sediment will be hydraulically dredged from a borrow area on the Mississippi River, and pumped via pipeline to create new marsh in the project area. Over time, natural dewatering and compaction of dredged sediments should result in elevations within the intended range which would be conducive to the establishment of emergent marsh. The 104-acre RI area will be planted with native marsh vegetation to accelerate benefits to be realized from this project.

The Lake Hermitage Marsh Creation - NRDA Early Restoration Project is located within the Barataria Hydrologic Basin in Plaquemines Parish, Louisiana, to the west of the community of Plaquemine, and northward of the community of Magnolia. This basin was identified as a priority area for coastal restoration, and has been the focus of extensive study and project design and implementation.

Trustee Implementation Group(s):
• Louisiana

Implementing Trustee(s):
• Louisiana Coastal Protection and Restoration Authority (CPRA)

Restoration Type(s):
• Wetlands
• Coastal and nearshore habitats

Restoration Phase:
Early Restoration Phase 1


Site Location(s):
• Lake Hermitage
• Lake Hermitage Marsh Creation

English (United States)

NOAA | Office of Response and Restoration

ERMA (Environmental Response Management Application)

The Environmental Response Management Application (**ERMA**) is an online geographic information system (GIS) and visualization tool that allows you to **view** response, assessment, and restoration **mapping layers in context** with other environmental information



22

NOAA | Office of Response and Restoration

Gulf of Mexico ERMA

- Standard layers; specialized by region/state
- Public access; Login to view privileged data and additional tools

Legend | Query | Download

ERMA® Environmental Response Management Application
Gulf of Mexico

Map Tools

Search

Layers

Bookmarks

Scale: 1 : 12,330,070 Zoom Level: 5 Location: 27.0422°, -84.5861°

23

NOAA | Office of Response and Restoration

Gulf of Mexico ERMA

Environmental data focused on the Gulf of Mexico specifically in support of Response, Assessment and Restoration efforts

ERMA® Environmental Response Management Application
Gulf of Mexico

Active View: Gulf of Mexico GAPs Analysis & Long-term Monitoring

Layers

Background Layers

Esri

Restoration

Ocean Conservancy GAP Analysis

Oysters

Shallow and Mid-water Corals

Deepwater Horizon

Scale: 1 : 6,166,950 Zoom Level: 6 Location: 22.7992°, -80.1037°

NOAA | Office of Response and Restoration

Gulf of Mexico ERMA

- Response spatial data
 - Common Operating Picture
- Natural Resource Damage Assessment data
 - Programmatic Damage Assessment & Restoration Plan
- Environmental resources and habitat
- Monitoring and Restoration information
 - Oceans Conservancy Data Gaps Analysis

25

NOAA | Office of Response and Restoration

Gulf of Mexico ERMA

ERMA® Environmental Response Management Application
Gulf of Mexico

Information | Help | Recent Data | Admin | Upload

Search Layers, Folders, and

Scale: 1 : 6,166,950 Zoom Level: 6 Location: 22.7992°, -80.1037°

USDOC | NOAA | NIOS | NOAA Office of Response & Restoration
Disclaimer | Privacy policy | Official Citation | Contact

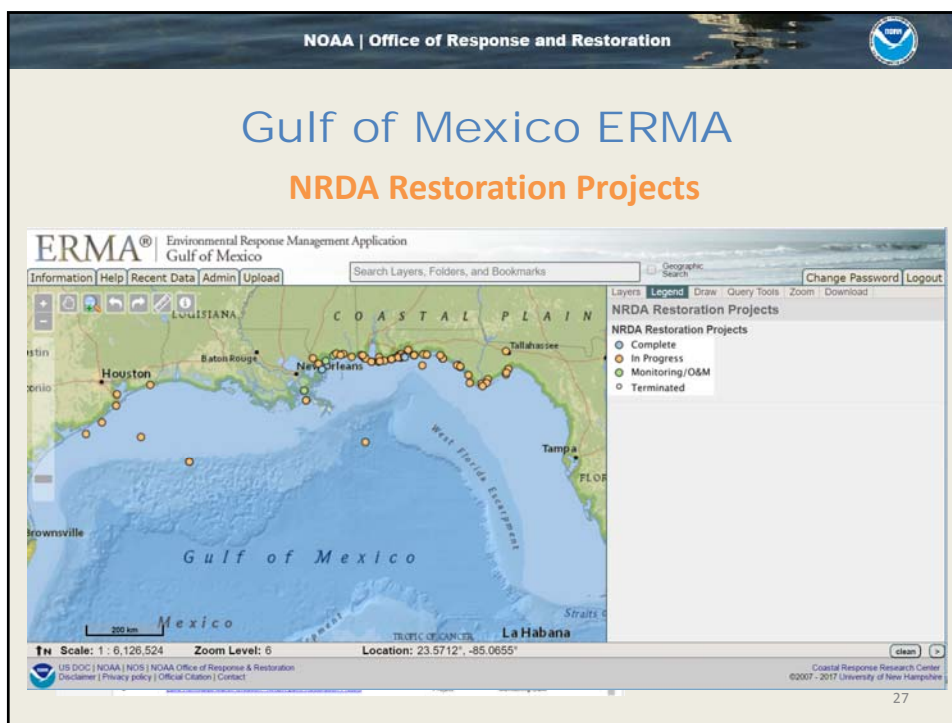
Ocean Conservancy
CHARTING THE GULF
Analyzing the Gaps in Long-term
Monitoring of the Gulf of Mexico

Bookmark Views: new Expand

clean

Coastal Response Research Center
©2007 - 2017 University of New Hampshire

26



NOAA | Office of Response and Restoration

ERMA Data Exchange

- Opportunity to display and distribute spatial data (**projects, environmental, restoration**) in Gulf of Mexico ERMA:
 - State, Federal, Tribal, Non-governmental organizations, Academic
- Provisional (in process, under review) and Public data

28

NOAA | Office of Response and Restoration

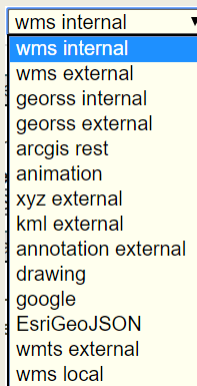
Gulf of Mexico ERMA

Standards

- Web Mapping services and
- spatial data files (shapefile)

Metadata

- ERMA ❤️ Metadata.
- Flexibility to include notes (summary metadata) and FGDC or ISO Metadata files



A screenshot of a dropdown menu with the following items: wms internal, wms internal (highlighted), wms external, georss internal, georss external, arcgis rest, animation, xyz external, kml external, annotation external, drawing, google, EsriGeoJSON, wmts external, and wms local.

29

NOAA | Office of Response and Restoration

Thanks!

Questions at the end?

30

NOAA Restoration Center – DWH Long Term Data Management Systems

On behalf of the DWH Trustee Council:



DIVER

- NRDA restoration project tracking
- Restoration project monitoring data

NRDA Public Submissions Database

- Public project ideas used for restoration planning



U.S. Department of Commerce | National Oceanic and Atmospheric Administration | NOAA Fisheries

Mandates

- **DWH Consent Decree** - ...establish, populate, manage, and maintain a Gulf-wide environmental data management system that shall be readily accessible to all Trustees and the public.
- **DWH Trustee Council SOPs**
- **DWH Monitoring and Adaptive Management Manual**
- **What ever happened to that \$8.8 billion?**
- **Oil Pollution Act (OPA)**

Data Standards

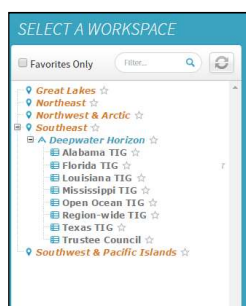
Project tracking – Project tracking data structure driven by Trustee Council reporting needs

Monitoring – Looking to adopt or coordinate with existing standards where possible

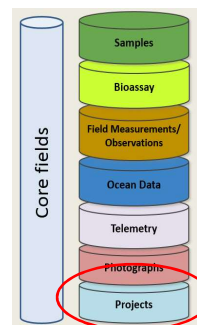


U.S. Department of Commerce | National Oceanic and Atmospheric Administration | NOAA Fisheries | Page 2

DIVER Portal - Project Tracking Module



- Descriptive overview information
- Financial information
- Project activities
- Environmental compliance
- Accomplishments
- Monitoring
- Locations



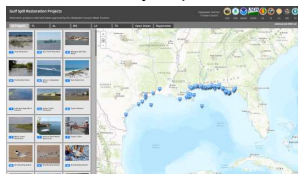
Shareability & Interoperability

Project Tracking – Primary focus has been on public accessibility

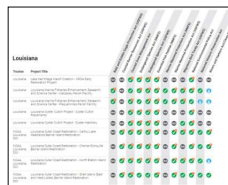
Project information pages



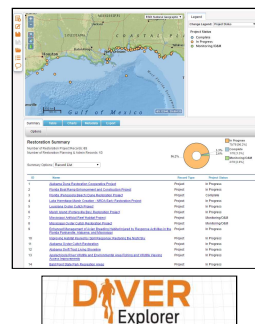
Story Maps



Data Dashboards



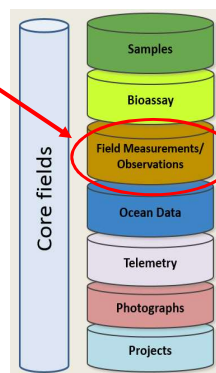
Search and query tools




Shareability & Interoperability

Project monitoring data

- Across NRDA projects, TIGs, states, agencies etc.
- With DIVER assessment field measurements
- Maximize interoperability across Gulf data







NOAA National Centers for Environmental Information (NCEI)

Dr. Rost Parsons
Acting Chief, Oceanographic Sciences Branch


Kirsten Larsen and Julie Bosch
Coastal Sciences Branch

June 2017




 National Oceanic Atmospheric Administration | NOAA Satellite and Information Service | www.ncei.noaa.gov

1



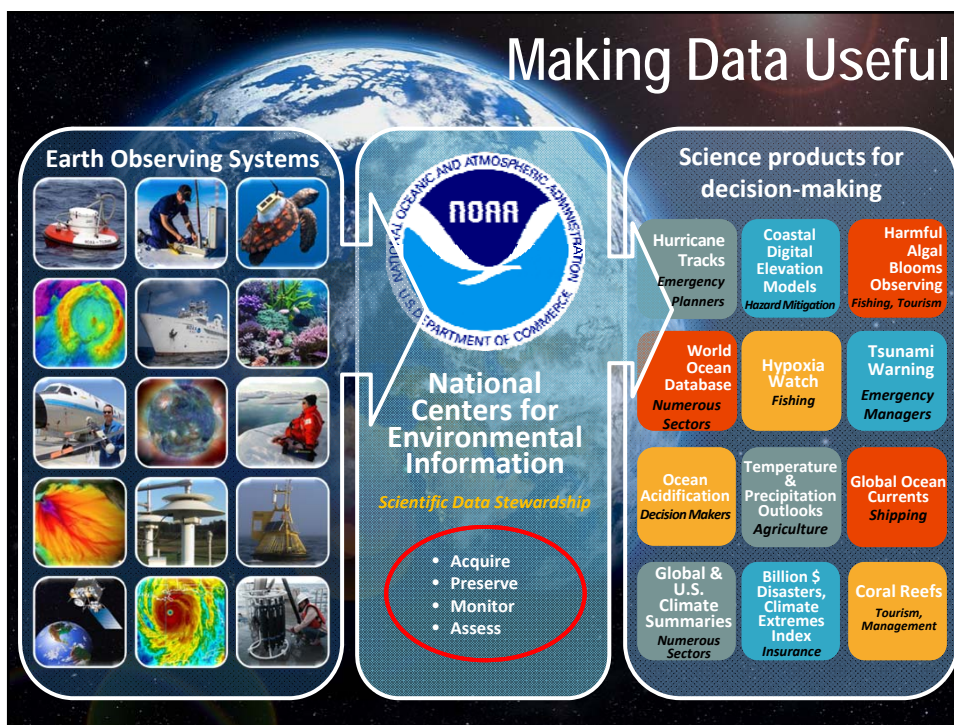
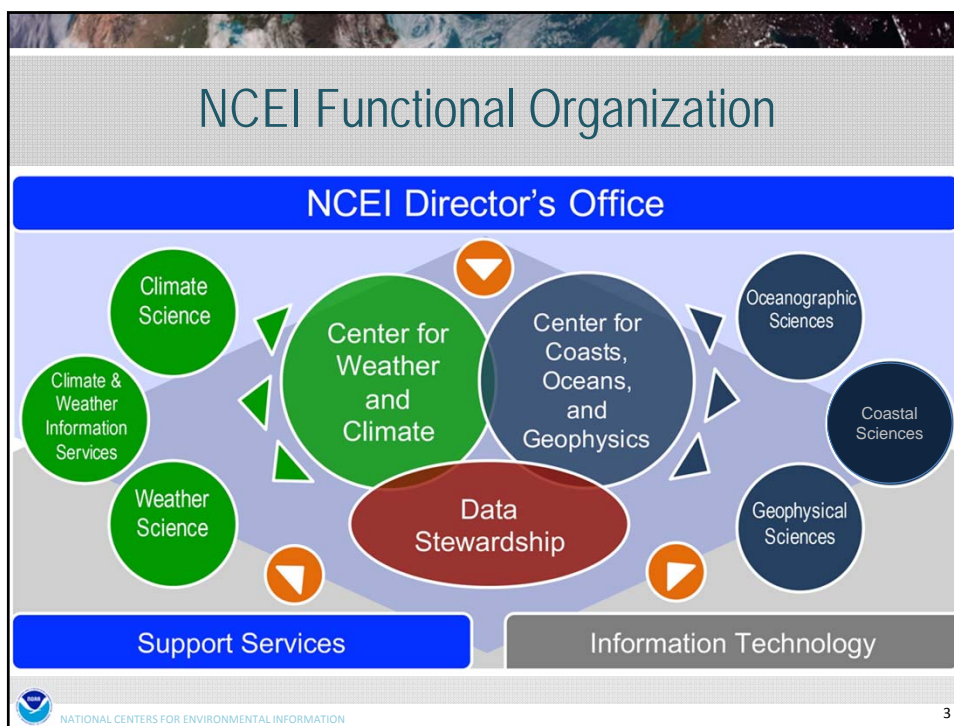
The National Centers for Environmental Information (NCEI) is the official data management entity for oceanographic, geophysical, and climatological information with the United States...

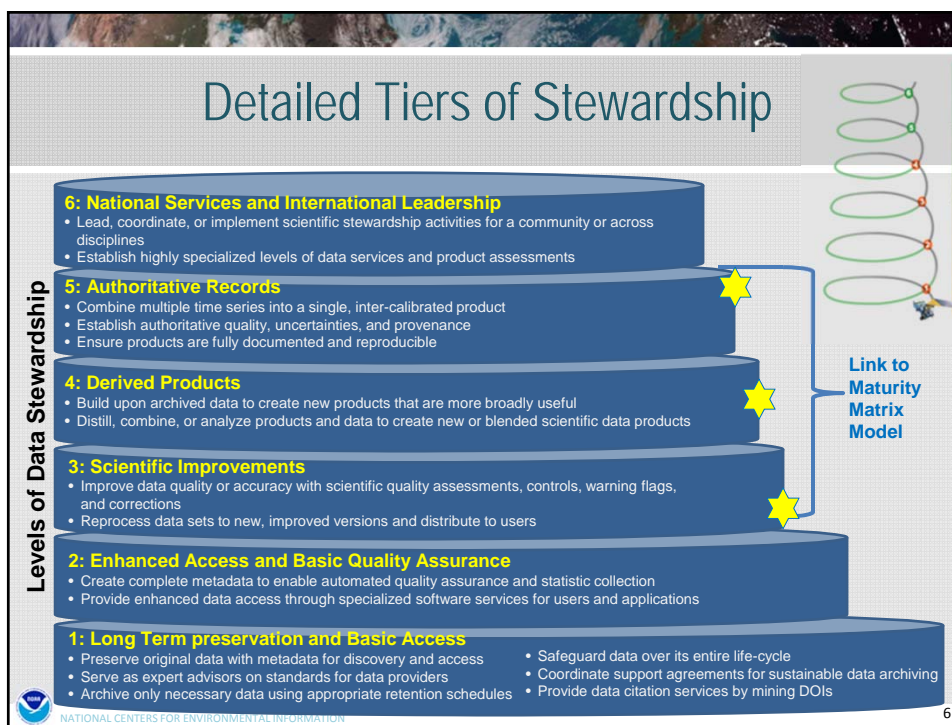
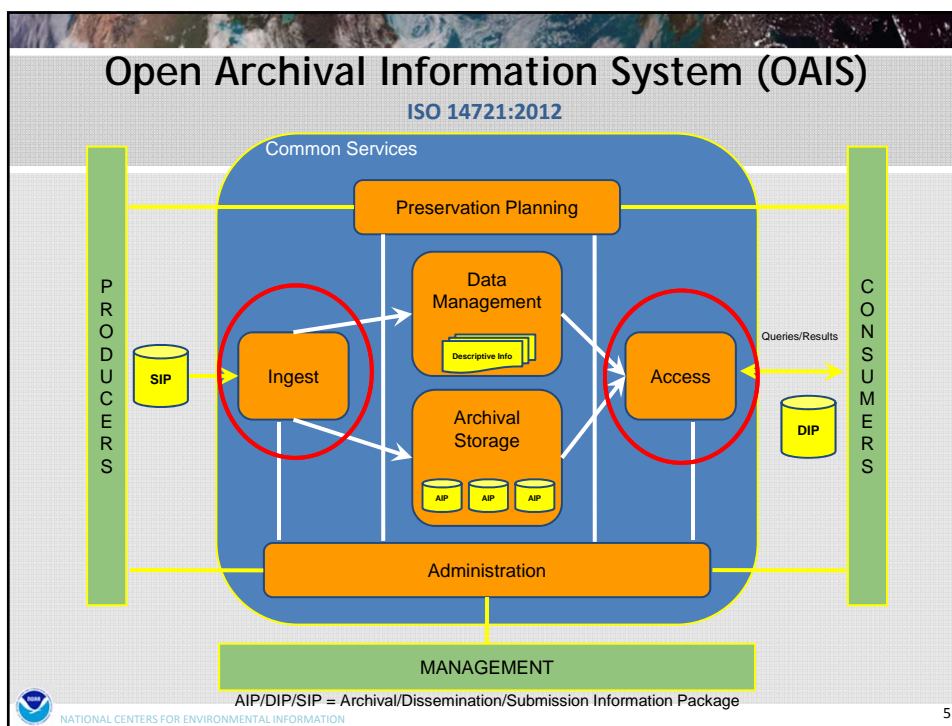
—FY2017 Omnibus Appropriation
Original Language in House Report 114-605



NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION

2





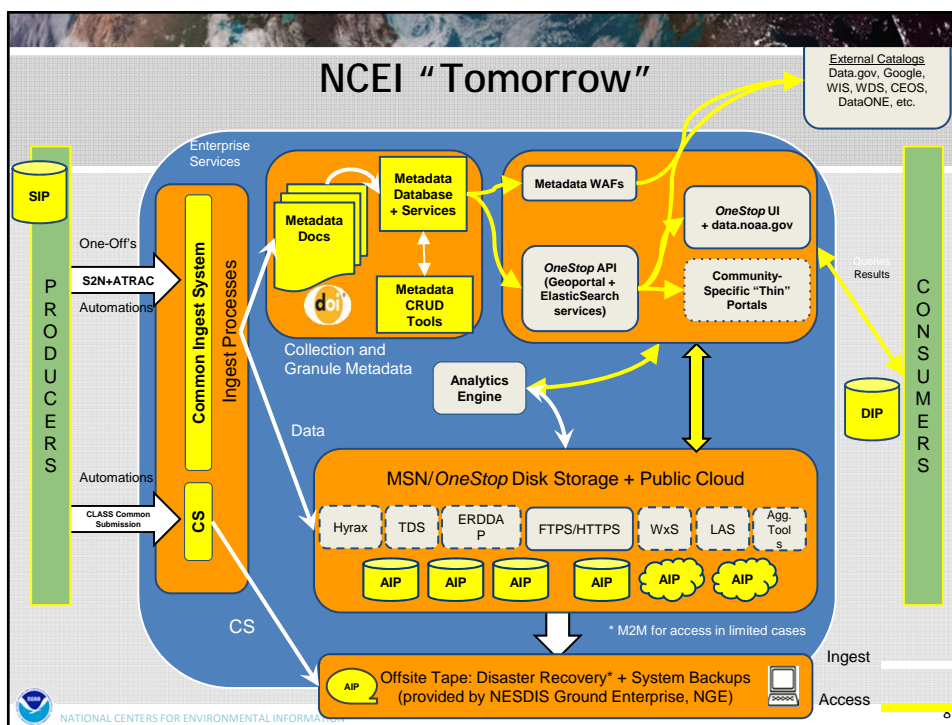
Questions to Answer

- Metadata
 - ISO 19115 et al.
 - Transforms from other ... e.g. Darwin Core
 - That required for Archive ...
- Standards
 - Data archive formats (commonly follow Library of Congress)
 - Common or managed vocabularies
 - NCEI Standards Section
- Mandates
 - NOAA Administrative Order 212-15 (*Management of Environmental Data and Information*) cites overarching mandates from Federal Records Management Act to NARA to Agency Directives
- Shareability / Interoperability
 - Access efforts – focused on online data services ... One Stop, DataOne Node, etc.
 - NCEI Higher Levels of Stewardship ... World Ocean Database



NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION

7





Ensuring a data and information legacy that promotes continual scientific discovery and public awareness of the Gulf of Mexico.

GRIIDC Serves Data Life Cycle Plan – Track – Document – Archive – Disseminate

Report Issue | Suggest Improvement | LOGIN

GULF OF MEXICO RESEARCH INITIATIVE
Investigating the effect of oil spills on the environment and public health.

GRIIDC

SEARCH DATA ▾ SUBMIT DATA ▾ TRACKING & STATS ▾ ABOUT US ▾ HELP ▾ RESEARCH

Ensuring a Data Legacy

SEARCH DATA SUBMIT DATA MONITOR DATA

626 DATASETS 250 INSTITUTIONS 219 PROJECTS 2,751 RESEARCHERS

GRIIDC Data Stories, News, & Updates
Showcasing how sharing data benefits the scientific community & data management activities in the Gulf of Mexico.

GRIIDC Is a GoMRI Legacy

- Committed to serving GoMRI data until at least 2030
- Expanding services beyond GoMRI-funded research
- Integration with other repositories (e.g., NCEI, DataOne)
- Harte Research Insititute committed to expanding a data sharing culture through GRIIDC

GOMAportal

<http://www.gomaportal.org>

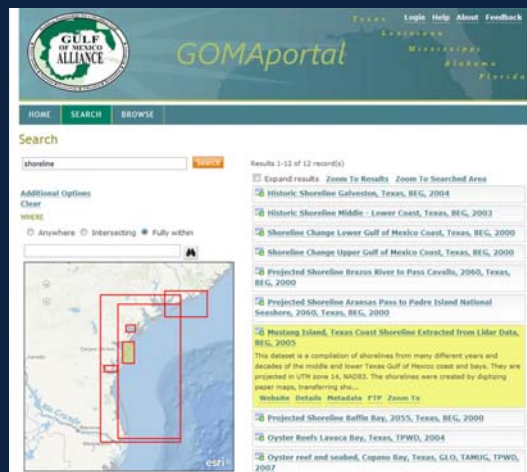
William Nichols - william.nichols@tamucc.edu
 Harte Research Institute for Gulf of Mexico Studies
 DWH Long Term Data Management Workshop
 Mobile, AL - June 7-8, 2017

1

GOMAportal.org

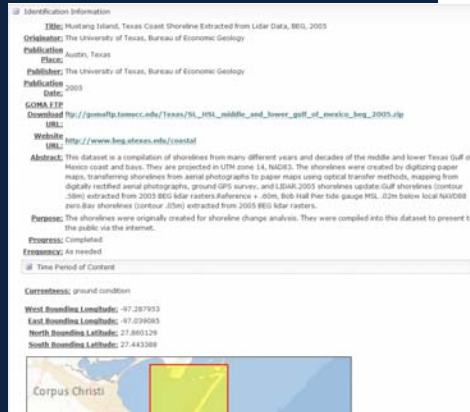
GOMAportal.org is a data catalog and repository that provides data discovery and access to Gulf of Mexico geospatial datasets

- Metadata catalog
- Data repository
- Browse by Topic
- Keyword Search
- Spatial Search
- View Information
- Download via FTP



GOMAportal.org

GOMAportal.org is driven by metadata. GOMA state partners harvest and upgrade metadata records to meet FGDC standards



```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<metadata>
  <idinfo>
    <citation>
      <origin>
        The University of Texas, Bureau of Economic Geology
      </origin>
      <pubdate>2005</pubdate>
    </citation>
    <title>
      Mustang Island, Texas Coast Shoreline Extracted from Lidar Data
    </title>
    <geoform>vector digital data</geoform>
  </idinfo>
  <pubinfo>
    <pubplace>Austin, Texas</pubplace>
  </pubinfo>
  <publish>
    The University of Texas, Bureau of Economic Geology
  </publish>
  <onlink>http://www.beg.utexas.edu/coastal</onlink>
</metadata>
```

Currently over 800
geospatial datasets
covering Texas,
Mississippi, Louisiana,
and Alabama

History of GOMAportal

- AP II - Ecosystem Integration and Assessment Priority Issue Team
- Identify, collect, document 'orphaned' datasets as identified by state partners
 - Metadata
 - Data
- Initial work completed 2011 – 800 datasets
- Continued to add datasets
 - HCRT – SLAMM
 - NOAA ECSC – Worldview 2 Imagery
- Currently 900 datasets

GOMAportal 2017

- Gulfstar award to update and enhance
- March – December 2017
- Move to new / better server
- Update to 1.2.7
- Enhance with new features

5

Metadata

- Currently FGDC
- Will support ISO 19115-2 and 19119 (web services) and Dublin Core for non geospatial data types

Standards

- Open formats for data, primarily geospatial formats
- Complete and valid metadata

6

Mandates

- Maintain repository of GOMA data products that do not have long-term archives

Interoperability

- Based on open-sourced Esri geoportal platform
- Federated search from other geoportals or via any CS-W client

7

Questions?

8

DWH Project Tracker



www.dwhprojecttracker.org

Presented by
Laura Bowie
Gulf of Mexico Alliance



DWH Project Tracker

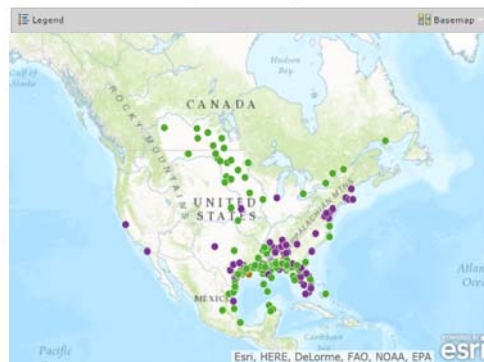


www.dwhprojecttracker.org

HOME - PROJECTS SUMMARIES - ABOUT - FAQ - DOWNLOADS CONTACT US LAUNCH MAP VIEWER

DEEPWATER HORIZON PROJECT TRACKER

Map of All Projects



What is the Deepwater Horizon Project Tracker?

The Deepwater Horizon Project Tracker provides an easy and comprehensive way to track restoration, research, and recovery projects resulting from the 2010 Deepwater Horizon oil spill.

Quick Summaries

Summary by State
Summary by Category
Summary by Funding Program

Launch Searchable Map Viewer

**currently
597 projects**

DWH Project Tracker

How Complete is the Project Tracker?

FUNDING PROGRAM	FUNDERS	PROJECTS AWARDED DWH FUNDS	PROJECTS SENT TO TRACKER	PROJECTS IN THE TRACKER	% IN THE TRACKER
Berms to Barriers Fund	California Coastal Protection and Restoration Authority (CCPRA)	2	2	2	100%
Early Restoration Phase I	Natural Resource Damage Assessment (NRDA) Trustees	8	8	8	100%
Early Restoration Phase II	Natural Resource Damage Assessment (NRDA) Trustees	2	2	2	100%
Early Restoration Phase III	Natural Resource Damage Assessment (NRDA) Trustees	44	44	44	100%
Early Restoration Phase IV	Natural Resource Damage Assessment (NRDA) Trustees	10	0	0	0%
Gulf Environmental Benefits Fund	National Fish and Wildlife Foundation (NFWF)	26	24	24	92%
Gulf Region Health Outreach Program	Gulf Region Health Outreach Program	8	8	8	100%
Gulf Research Program	National Academy of Sciences (NAS)	33	33	33	100%
MOEX Clean Water Act Penalties	Various Entities	8	8	8	100%
MOEX Supplemental Environmental Projects	Various Entities	8	8	8	100%
National Institute of Health Year One Block Grants	Gulf of Mexico Research Initiative (GoMRI)	4	0	0	0%
North American Wetlands Conservation Act	United States Fish and Wildlife Service (USFWS)	47	47	38	81%
Recovered Oil Fund for Wildlife	National Fish and Wildlife Foundation (NFWF)	48	48	48	100%
RTT & Community Grants	Gulf of Mexico Research Initiative (GoMRI)	8	8	8	100%

Funding Programs (13):

Berms to Barriers
 NRDA Phases I-V
 NFWF GEBF
 Gulf Region Health Outreach Program
 NAS GRP
 MOEX Settlements
 NAWCA
 NFWF Recovered Oil for Wildlife
 GoMRI
 RESTORE Bucket 1
 RESTORE Bucket 2
 RESTORE Bucket 4
 RESTORE Bucket 5

DWH Project Tracker



HOME - PROJECTS - SUMMARIES - ABOUT - FAQ - DOWNLOADS - CONTACT US - LAUNCH MAP VIEWER



Project Types:

- Environmental
- Human and Social/Planning
- Recreational Use
- Science/Research

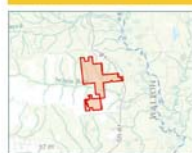


Summary by Category

PROJECT CATEGORY	DWH PROJECTS	FUNDING FROM DWH SOURCES (\$USD)	KNOWN LEVERAGED FUNDS (\$USD)
Environmental	210	1,258,925,916	111,080,314
Human and Social	32	107,826,201	3,400,890
Recreational Use	39	276,246,782	0
Science	316	361,457,754	7,864,940

Seven Runs Creek Easement



Implementing Organization:	The Trust for Public Land	DWH Project Funding*: \$ 2,400,000
		Total Project Funding*: \$ 2,400,000
* No value or a zero means the project is not completed, or the total is unknown.		
Location:	Florida	
Project Category:	Environmental	
Project Actions:	Land Acquisition / Protection, Habitat Restoration and Enhancement	
Targeted Resources:	Wetlands / Marshes / Estuaries	

Conservation / Acquisition Status:

Landowner	Acres Purchased	Year Purchased	Conservation Type	Conservation Purpose	Public Access	Easement Holder	Easement Holder 2	Easement Duration
	2,336	2012	Easement	Environmental System	Open	Florida Department of Environmental Protection - Division of State Lands		

Project Funders:

Name	Florida Department of Environmental Protection	Known Leveraged Funds
Amount Awarded	\$ 2,400,000	\$ 0
Year Awarded	2012	
Funding Program	MOEX Supplemental Environmental Projects	
Funding Type	Civil Settlement	

Project Description:

This project will benefit groundwater resources, surface water quality and wildlife. Also part of military buffer lands strategy.

For the complete project description, please visit:
http://www.dep.state.fl.us/deepwaterhorizon/files/Fact_Sheet_settlement.pdf

Project Contact:

Settlement Source DWH Controlling Document MOEX Consent Decree Between The United States and MOEX Offshore 2007 LLC	Name: Ashley M. Williams Phone Number: 850-245-2197 Email: Ashley.M.Williams@dep.state.fl.us
--	---

This report was created by The Gulf of Mexico Alliance and its partners on April 28, 2015. Information on the map is for discussion and visualization purposes only.

DWH Project Tracker

HOME - **PROJECTS** - SUMMARIES - ABOUT - FAQ - DOWNLOADS - CONTACT US - LAUNCH MAP VIEWER

All Projects

Project Name: State: Project Category: Funding from DWH Sources: Funding Program:

DWH Source:

Copy PDF CSV Print Excel

PROJECT NAME	STATE	PROJECT CATEGORY	FUNDING FROM DWH SOURCES	FUNDING PROGRAM	DWH SOURCE
Application of Passive Samplers to Monitor PAHs Concentrations in Water, Sediment Porewater, Sediment, and Commercially Important Organisms in the Gulf of Mexico in Order to Quantify	FL, MD, TX	Science / Research	95,000	GoMRI RFP-III: Bridge Grant	BP
The Influence of Oil Dispersants on the Biodegradation of South Louisiana Crude Oil	LA, TX	Science / Research	311,561	GoMRI Year One Block Grant	BP

DWH Project Tracker

EXPORT

SEARCH

206.169.56.48/DWH/

DWH PROJECT TRACKER

Search:

Logical Operator:

Query Value:

Query Conditions:

- Primary Project Action = Species Restoration
- AND
- Project Category = Environmental

Use And or Or Between Conditions: ☒ And ☐ Or

Run Query

Layers

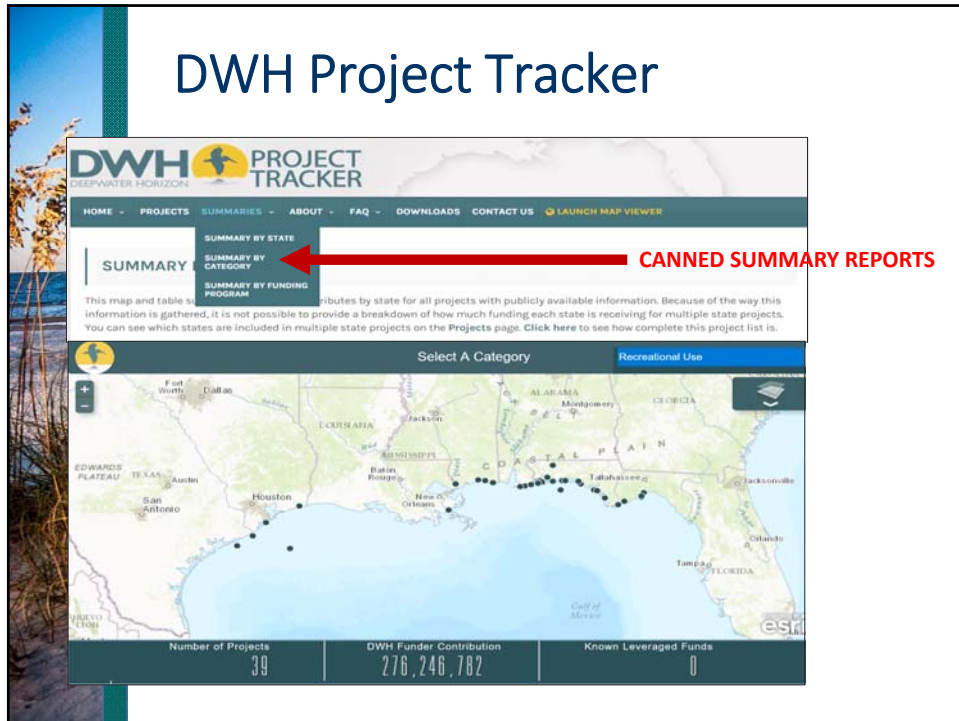
Map View:

DWH Project Query [1]

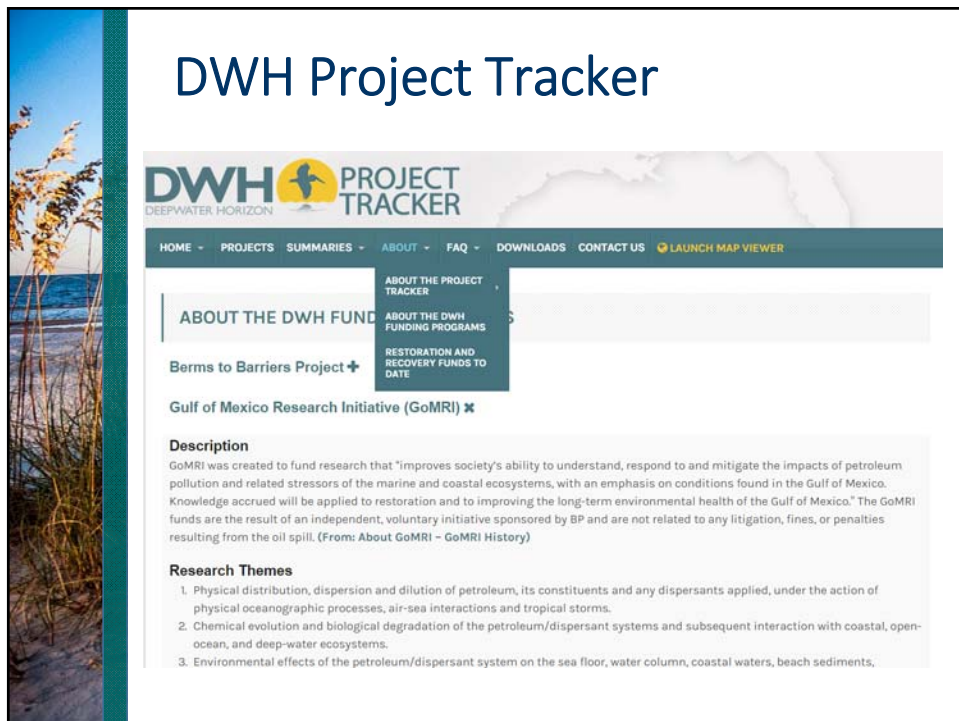
Project Name	Project Category	Year Funded	Deepwater Horizon Funding Amount	Funding Program	Deepwater Horizon Funding Source	Deepwater Horizon Funding Origin
Restoration and Enhancement of Oyster Reefs in Alabama	Environmental	2013	3,750,000	Gulf Environmental Benefit Fund (GEBF)	BP and Transocean	Criminal Settlement

1 - 10 of 10 results

DWH Project Tracker



DWH Project Tracker



DWH Project Tracker



Stats*

- 3,300 visits in last 12 months
- Over 10,900 page views
- 31% increase from year before
- Most popular uses:
 - All Projects List
 - Summary by Category
 - Summary by State
 - GIS downloads

Usage Snapshot:

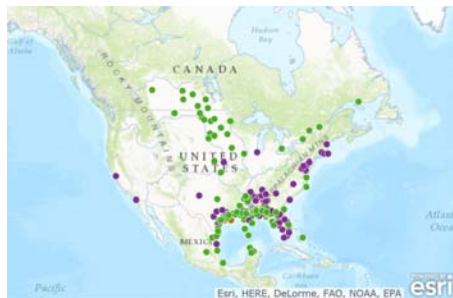
- Public
- Restoration Programs
- Media

**From Google Analytics in December 2016*

Questions?

Contact Information:

Laura.Bowie@gomxa.org



Gulf of Mexico Coastal Ocean Observing System (GCOOS.ORG)

Dr. Matthew Howard – GCOOS DMAC Lead
Deep Water Horizon Long-Term Data Management Collaboration Workshop
Mobile, Alabama
6-8 June 2017



Hierarchy

- Global Earth Observation System of Systems (GEOSS)
- Global Ocean Observing System (GOOS)
- U.S. Integrated Ocean Observing System (IOOS)
- 11 Regional Coastal Ocean Observing Systems (RCOOS')
- Gulf of Mexico Coastal Ocean Observing System

**We are in our third 5-year cycle
Renewed to 2021 ~\$1.5M/yr
~3 Data-type FTEs**

Vision Statement

“Develop and maintain an automated largely-unattended interoperable system of systems which delivers high-quality data, metadata and products from sensors to desktops in preferred formats.”



—Matthew Howard

DMAC Scope

- Metadata Management
 - Data Discovery
 - Uniform On-line Browse
 - Data Access and Transport
 - Data Archive
-
- Web Services Sensor Observation Service (SOS), CSV
 - ERDDAP/TDS (NetCDF)



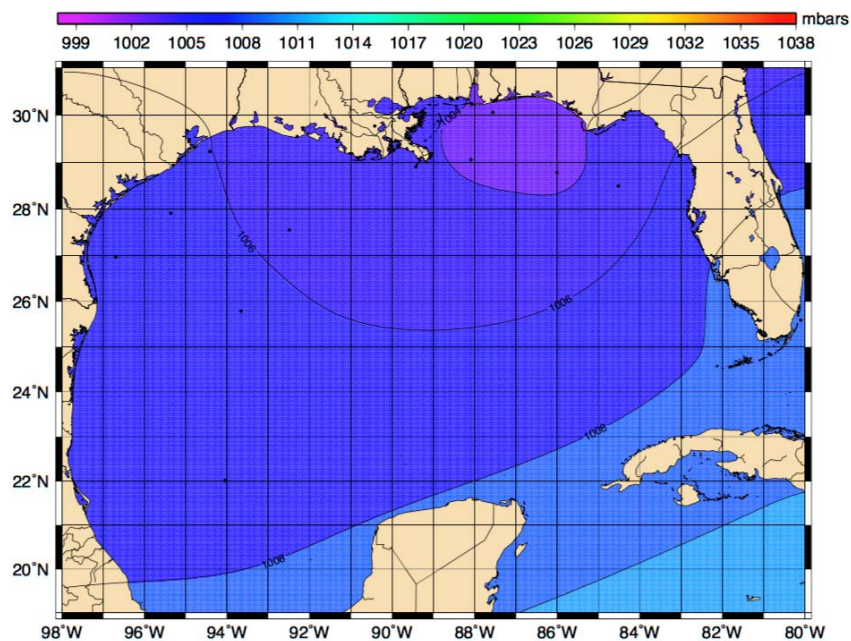
Data Types (fixed, mobile, remote sensing)

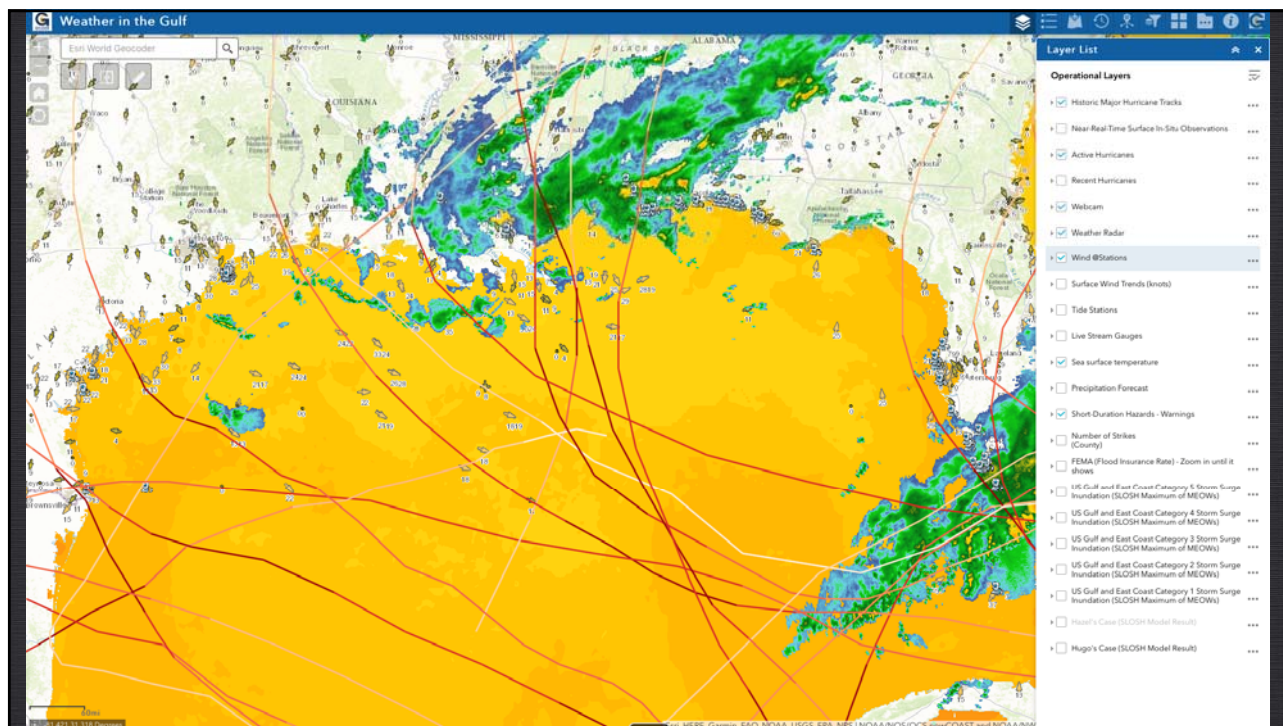
- Physical Oceanographic (T, S, Currents, Water level, River discharge ...)
- Marine Meteorological (Winds, Temperature, Pressure ...)
- Biogeochemical (nutrients, dissolved oxygen, pH, ...)
- Biological (plankton, fish, ...)
- Near real-time
- Numerical Model Forecasts (winds, currents)
- Historical Data (near real-time, field cruises, reanalysis, ...)
- Climatologies (Temperature, Salinity, ...)
- Static (bathymetry, coastlines)



Barometric Press.

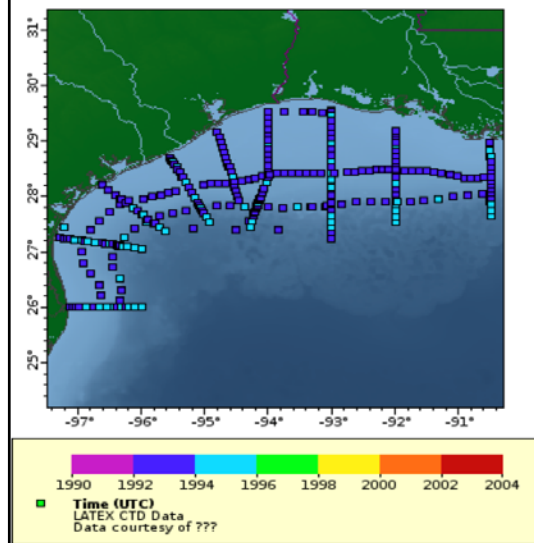
Local Date/Time: 17060618 UT Date/Time: 17060623



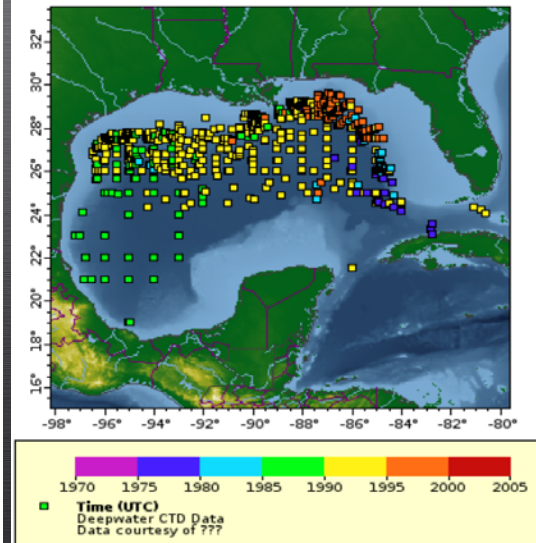


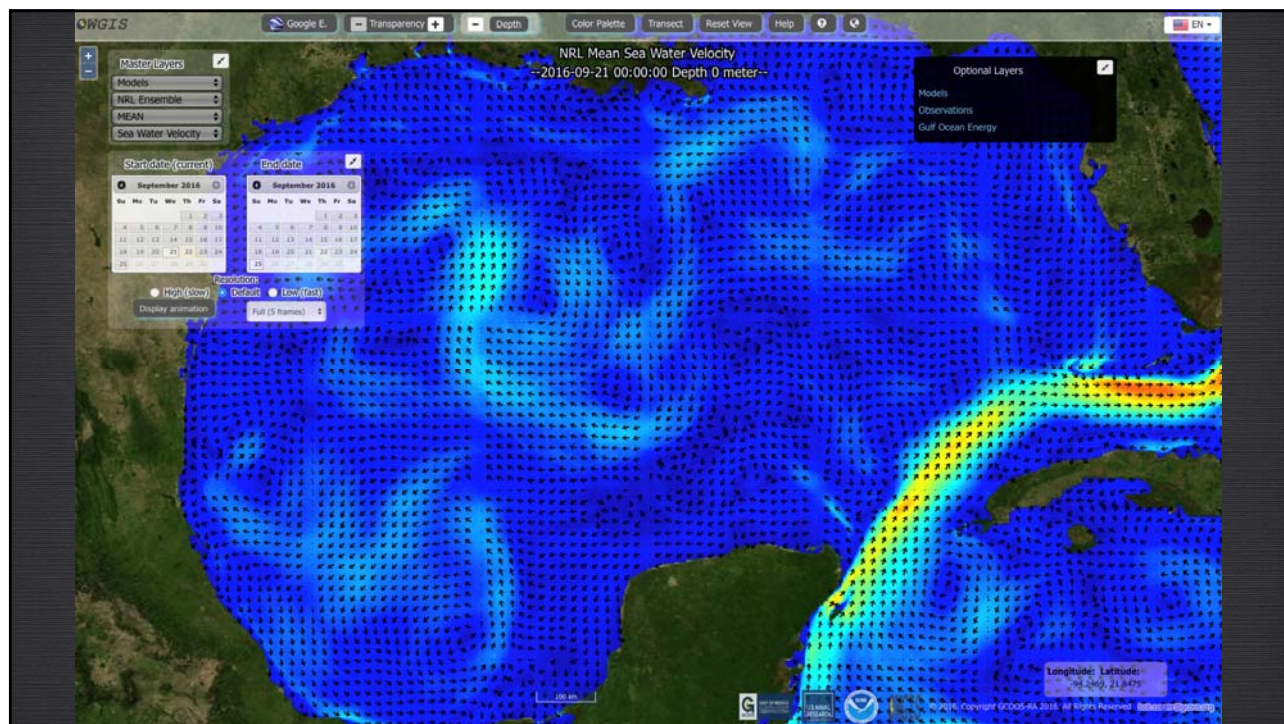
Historical Data (LATEX & Deepwater Reanalysis)

Map of Distinct Data [\(Refine the map and/or download the image\)](#)
Optional: Click on the map to select the closest data.



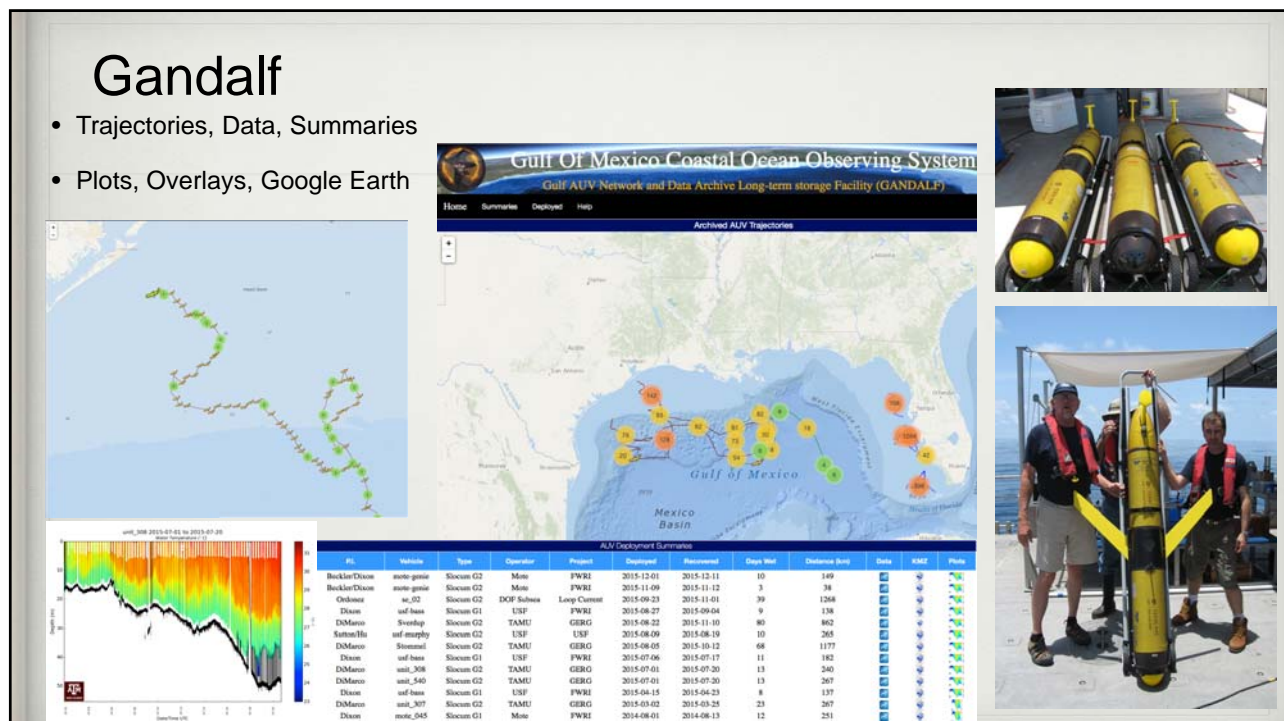
Map of Distinct Data [\(Refine the map and/or download the image\)](#)
Optional: Click on the map to select the closest data.

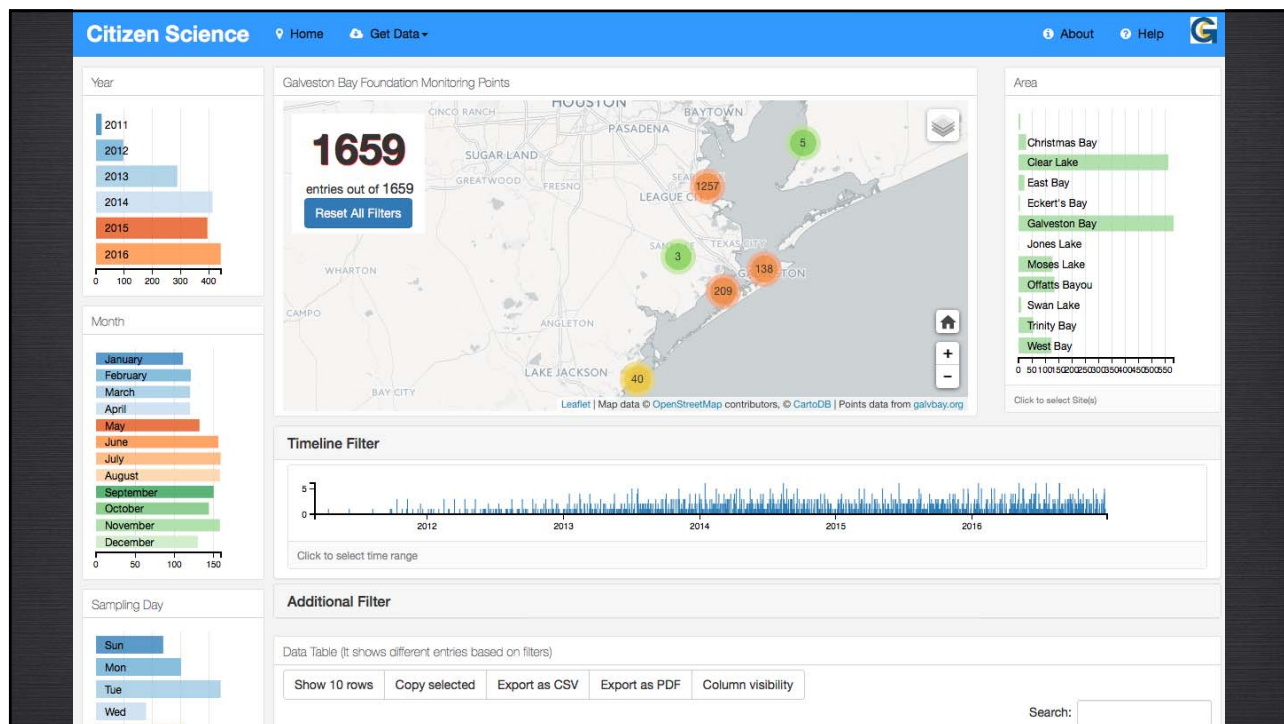




Gandalf

- Trajectories, Data, Summaries
- Plots, Overlays, Google Earth







The Louisiana
**Coastal
Information
Management
System**

Craig Conzelmann
U.S. Geological Survey
conzelmannc@usgs.gov

Ed Haywood
LA Coastal Protection and
Restoration Authority
edh@la.gov

for DWH Long Term Data Management Coordination Meeting – Mobile, AL



Standards | Metadata | **Mandate** | Share-ability | Interoperability



What is CIMS?

Suite of data driven applications
and tools used to manage,
visualize, share and analyze
coastal data.



Mandate: Internal La CPRA policy



Standards | Metadata | **Mandate** | Share-ability | Interoperability



What kind of data can I find in CIMS?

Tabular

- Hydrographic
 - Real-time
 - Discrete
- Vegetation
 - Emergent marsh
 - Forested swamp
- Soil
 - Accretion
 - Sediment elevation
 - Soil properties

Spatial

- Project boundary
- Monitoring stations
- Infrastructure
- Master Plan projects, features, flood modeling
- Habitat analysis
- Modeling results
- Bathymetry
- Sediment core
- Deposit/Borrow

Unstructured (binary library)

- Documents
- Photos
- Videos
- Non-Standard data files such as ADCP or LISST

NOTE: These are the publicly available data in CIMS.



Standards | Metadata | Mandate | Share-ability | Interoperability



Standards

Geophysical



Ecological



Metadata



CSDGM: Content Standard for Digital Geospatial Metadata



- 191* Suite:

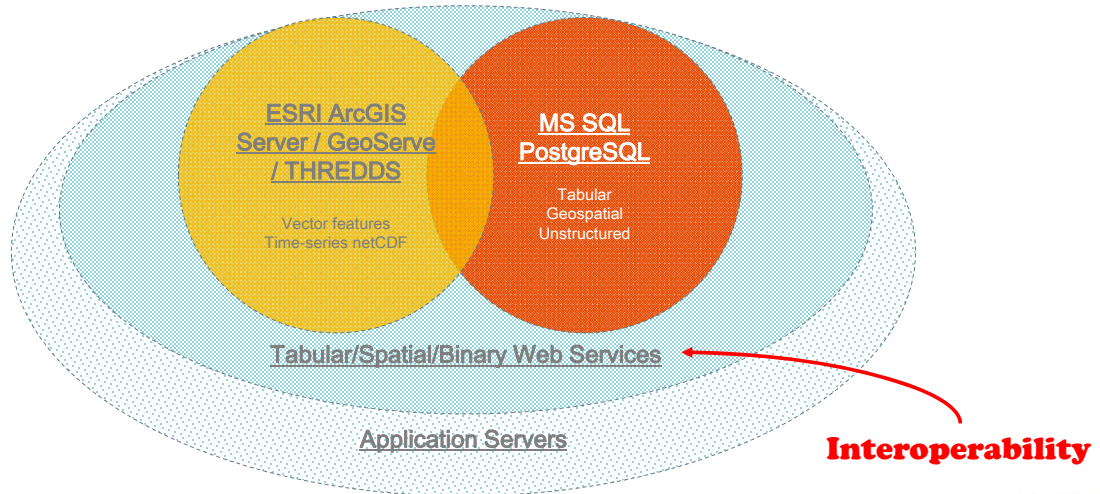
<https://cims.coastal.louisiana.gov/RecordDetail.aspx?Root=0&sid=12362>
<https://cims.coastal.louisiana.gov/RecordDetail.aspx?Root=0&sid=11504>



Standards | Metadata | Mandate | Share-ability | Interoperability



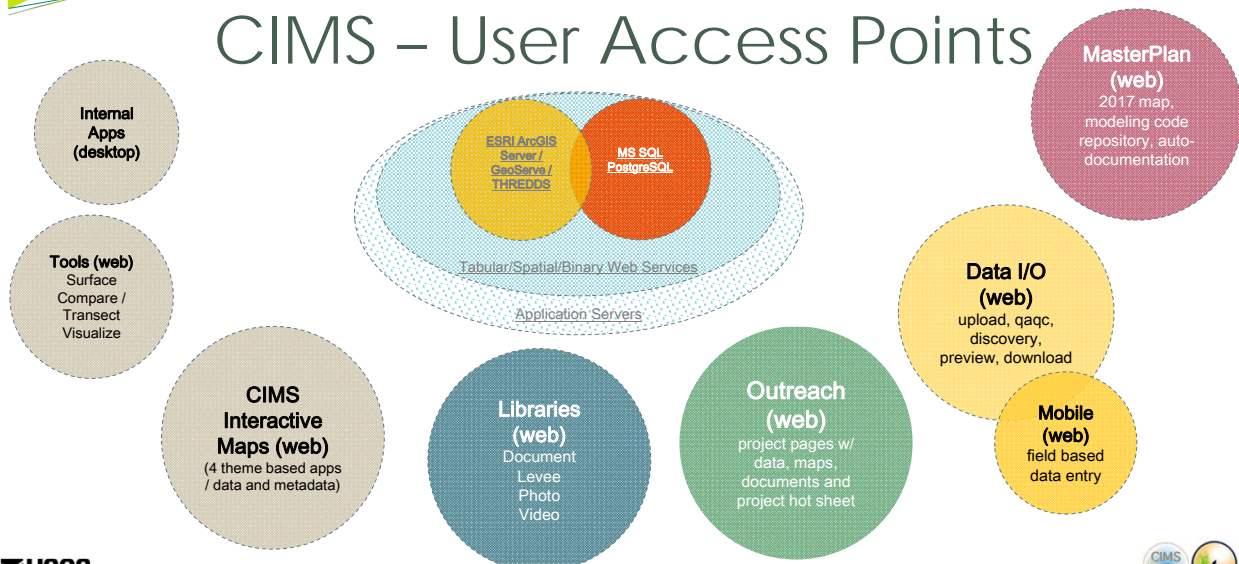
CIMS – Behind the Scenes



Standards | Metadata | Mandate | Share-ability | Interoperability



CIMS – User Access Points



Standards | Metadata | Mandate | Share-ability | Interoperability



CIMS Main Interface



Easy access to the 3 main public modules:
Maps, Data,
Document Library

<https://cims.coastal.la.gov>

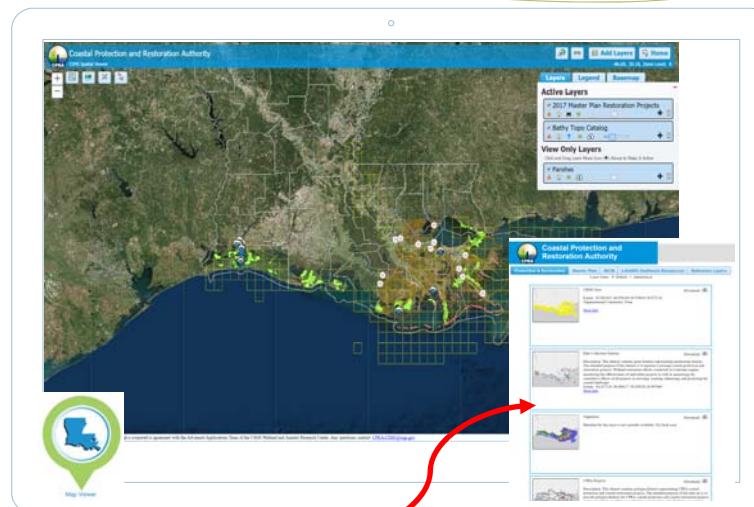


Standards | Metadata | Mandate | Share-ability | Interoperability



CIMS Main Spatial

Full featured viewer allowing user to layer, filter, extract and compare data.



Building on the same data services, customized maps are also supported.



Outreach Focused



Master Plan Focused

<https://cims.coastal.la.gov>



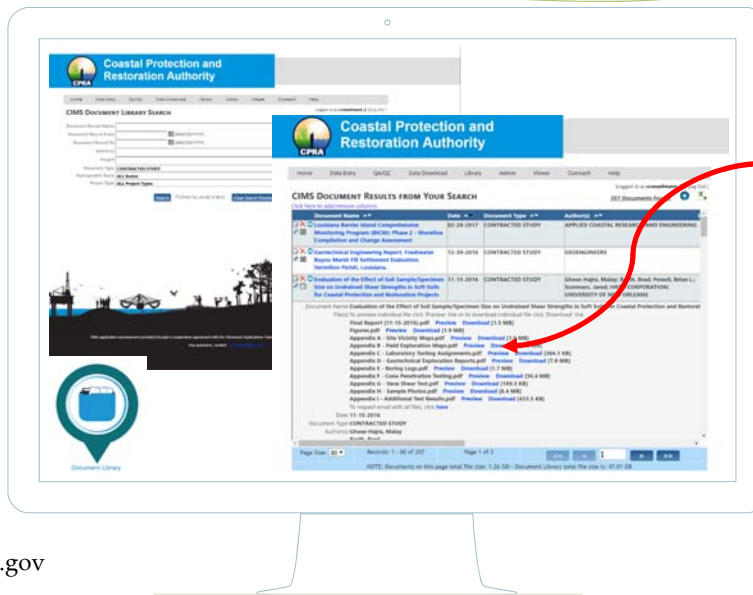
Share-ability

Standards | Metadata | Mandate | Share-ability | Interoperability



CIMS Main Library

All documents tagged with projects, parish or hydro-basin enabling service level access to other CIMS modules



Share-ability

Building on the same database structures and data services, customized libraries are also supported.



Photographs

<https://cims.coastal.la.gov>

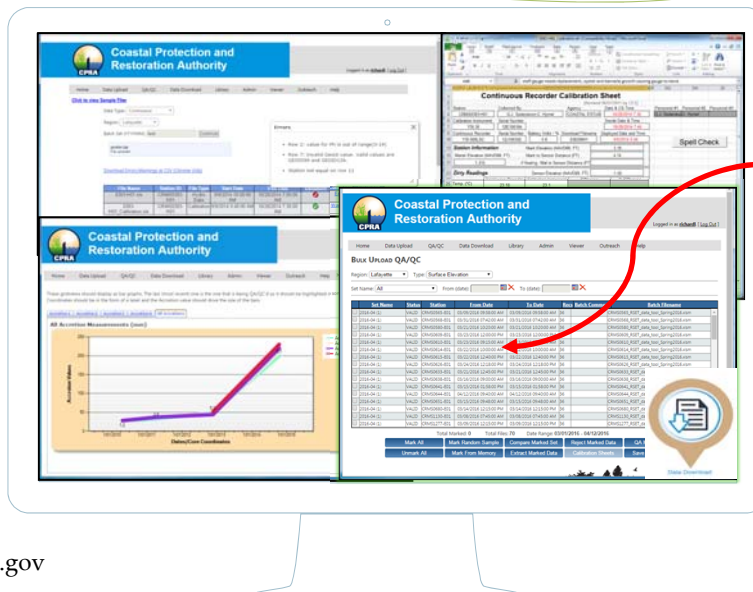


Standards | Metadata | Mandate | **Share-ability** | Interoperability



CIMS Data – Load, QC, Review, Download

All CIMS data is rigorously reviewed BEFORE being made available for download or service enabled for other consumers.



Share-ability

<https://cims.coastal.la.gov>

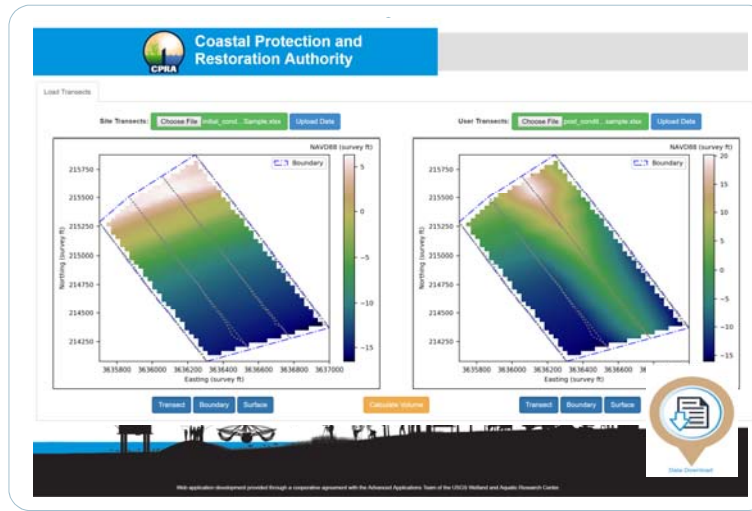


Standards | Metadata | Mandate | **Share-ability** | Interoperability



CIMS Tools Surface and Volume Comparison

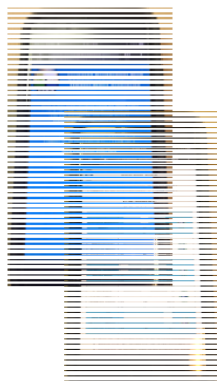
Planning and engineering tool allowing user to upload initial and desired elevation data, create surfaces then compare to get volume change information.



<https://cims.coastal.la.gov>



Standards | Metadata | Mandate | Share-ability | Interoperability



Field Data Entry

Contractors perform data entry on SmartPhone
....CIMS shows boat bay status to the public

<https://cims.coastal.la.gov>



Standards | Metadata | Mandate | Share-ability | Interoperability



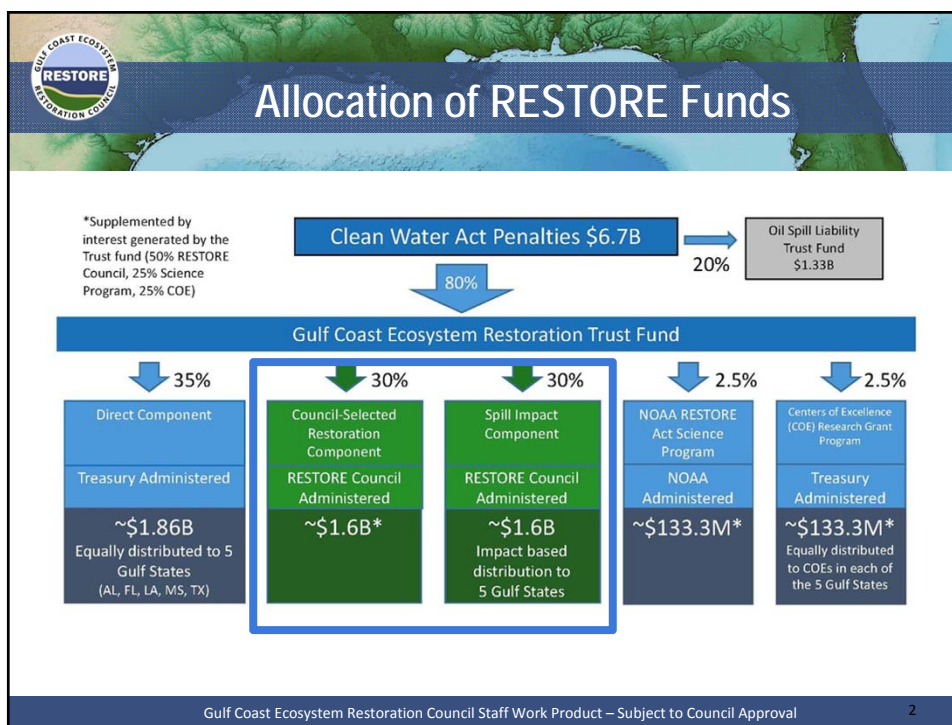


Questions?



**Gulf Coast Ecosystem Restoration Council
Data Management Framework Development**

Jessica Henkel, PhD
jessica.henkel@restorethegulf.gov





Gulf Coast Ecosystem Restoration Council


Council Composition:

- Governors of Alabama, Florida, Louisiana, Mississippi and Texas
- Secretaries of Agriculture, Army, Commerce, Homeland Security, Interior, Administrator of the EPA
- Chair: Agriculture



Gulf Coast Ecosystem Restoration Council Staff Work Product – Subject to Council Approval

3



Gulf Coast Ecosystem Restoration Council


Funding Recipients

- State Agencies:
 - TX CEQ
 - LA CPRA
 - MS DEQ
 - AL DCNR
 - FL DEP
- Federal Agencies:
 - USDA, Army, DOC, DOI, EPA
 - Bureaus under each agency
- Sub-recipients of each agency/bureau



Gulf Coast Ecosystem Restoration Council Staff Work Product – Subject to Council Approval

4




Council Monitoring and Assessment Program

Council Monitoring and Assessment Workgroup (CMAWG):

- 1 primary/1 alternate per Council Member
- Funded on Initial FPL
- Lead by NOAA & USGS
- Council staff representation
- Coordination of, and reach-back to, available monitoring and data management capabilities and info
- Generate recommendations to the Council


Gulf Coast Ecosystem Restoration Council Staff Work Product – Subject to Council Approval 5



Current Data Requirements


Currently, projects or programs are required to include an:

- ***Observational Data Plan (ODP)*** - information relevant to project ***data collection and compilation***
- ***Preliminary Observational Data Management Plan (DMP)*** - information relevant to project ***data management and delivery***



Interim Guidance: Data Management Plans

- Recipients are responsible for providing all project-related data to the Council
- Current data requirements:
 - Digital
 - Machine-readable
 - Non-proprietary formats (publicly available)
 - Appropriate metadata
 - Compliance with all federal laws and policies



Interim Guidance: Data Management Plans

- Project name, sponsoring agency, project phase, and an estimated budget for data management
- Contact information for one or more Data Stewards
- Estimated data collection period (*start and end dates*)
- A short description of the project location & data collection
- Description of each of the data types generated by the project
- GIS information (*if known and applicable*)
- **Organization's data management and metadata capacities and how the organization intends to store, archive, and disseminate project data**



Next steps: 2017

- Metadata Standards
 - CMAWG consensus on recommendation of adoption of ISO metadata standard for RESTORE Council funded projects to the Steering Committee
 - Council Staff Investigating Open Source Metadata Tool
- Draft Data Management Framework
 - Will work with CMAWG to develop options for Council consideration

Gulf Coast Ecosystem Restoration Council Staff Work Product – Subject to Council Approval 9

Thank You!

For More Information on Council Data Activities:

Jessica Henkel (jessica.henkel@restorethegulf.gov), Alyssa Dausman
(alyssa.dausman@restorethegulf.gov)



DWH NRDA Restoration Data Management Framework

Definition of Monitoring Data and Info (TC SOP 10.6.1)

Monitoring data include, but are not limited to,

- Datasets or model results collected, compiled, or utilized as part of DWH NRDA restoration
- Generated during any phase or component of restoration
- Project-specific monitoring or non-project specific data collection

Monitoring and Data Standards^(TC SOP section 10.6.2)

- Established by the Cross-TIG MAM work group
- More in-depth list and description will be in the MAM Manual
- **Monitoring standards** will include, but are not limited to, parameters/metrics, performance criteria, and data collection protocols (further described in the MAM Manual).
- **Data standards** will include, but are not limited to, FGDC/ISO standard metadata, acceptable units, measurement precision (number of digits), a QA/QC process, a data dictionary, and a readme file.
- Coordination with other programs

MAM Plan: Data Management Sections

All MAM plans will include a description of how the monitoring data will be managed (i.e., QA/QC procedures, metadata, data sharing, and storage).

MAM Data Management Steps Defined in TC SOP

- MAM Data Review and Clearance (TC SOP 10.6.4)
- MAM Data Storage and Accessibility (TC SOP 10.6.5)
- MAM Data Sharing (TC SOP 10.6.6)
- MAM Data Analysis and Synthesis (TC SOP 10.6.7)

MAM Data Review and Clearance

- Data should go through the appropriate QA/QC process in accordance with the data management section of the monitoring plan.
- Implementing Trustees will verify and validate MAM data and information
- Submitting Trustee will provide other TIG members time to review the data before data becomes public

MAM Data Storage and Accessibility

- DIVER (Restoration Portal) is the central repository and facilitates public access to restoration MAM data.
 - Trustees may also maintain records on other platforms.
 - Must explain data origin and long-term management and archiving
- MAM data stored and accessible within a year from when collected.
 - If not possible, explanation needed MAM plan

MAM Data Sharing

- The Trustees will follow standards and protocols set forth in the Open Data Policy*
- Throughout the calendar year, MAM data and information may be added to the Restoration Portal (DIVER) and made publicly available via the Trustee Council website

[*<https://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government->](https://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-)

MAM Data Analysis and Synthesis

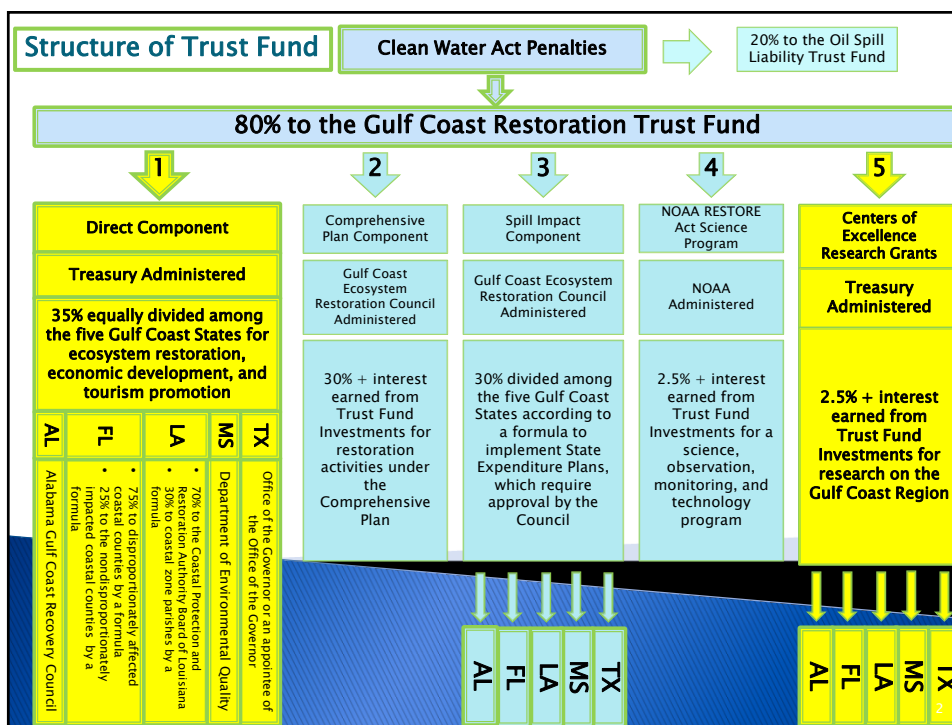
- Data from outside sources may be incorporated into analysis
 - Outside sources need adequate metadata and meet minimum QA/QC standards.
- TIGs share MAM data aggregation and analysis responsibilities with each other, especially when Restoration Types overlap with geographic areas

Next Steps

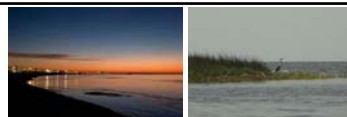
MAM Manual Version 1.0

- Data Management section that starts to outline in more specificity some of our data standards and data management procedures

Building the capacity and functionality within DIVER for our MAM data



Direct Component



- ▶ Grants to Gulf Coast states, 20 Louisiana Parishes, and 23 Florida counties for:
 - *Restoration and protection of the natural resources, ecosystems, fisheries, marine and wildlife habitats, beaches and coastal wetlands of the Gulf Coast region.*
 - *Mitigation of damage to fish, wildlife and natural resources.*
 - *Implementation of a federally approved marine, coastal, or comprehensive conservation management plan, including fisheries monitoring.*
 - *Workforce development and job creation.*
 - *Improvements to or on State parks located in coastal areas affected by the Deepwater Horizon oil spill.*
 - *Infrastructure projects benefitting the economy or ecological resources, including port infrastructure.*
 - *Coastal flood protection and related infrastructure.*
 - *Planning assistance.*
 - *Administrative costs.*
 - *Promotion of tourism in the Gulf Coast region, including recreational fishing.*
 - *Promotion of the consumption of seafood harvested from the Gulf Coast region.*

3

Status of Centers of Excellence Grants

- ▶ Awarded 4 Centers of Excellence Research Grants



Grant Awarded	Texas, Florida, Louisiana, Mississippi
Selection Started	Alabama

4

Centers of Excellence

Gulf Coast Region Science, Technology, Monitoring

Disciplines	FL	MS	LA	TX	AL
Coastal and deltaic sustainability, restoration and protection, including solutions and technology that allow citizens to live in a safe and sustainable manner in a coastal delta in the Gulf Coast Region		✓	✓	✓	
Coastal fisheries and wildlife ecosystem research and monitoring in the Gulf Coast Region	✓	✓	✓	✓	
Offshore energy development, including research and technology to improve the sustainable and safe development of energy resources in the Gulf of Mexico		✓		✓	
Sustainable and resilient growth, economic and commercial development in the Gulf Coast Region		✓	✓	✓	
Comprehensive observation, monitoring, and mapping of the Gulf of Mexico	✓	✓	✓	✓	

5

RESTORE Act Grant Requirements for Performance Reporting

- ▶ Summarize any significant findings or events, including any **data compiled, collected, or created**, if applicable.
- ▶ Describe any activities to disseminate or publicize results of the activity, project, or program, **including data and its repository and citations for publications** resulting from this Award.

Treasury RESTORE Act Standard Terms & Conditions

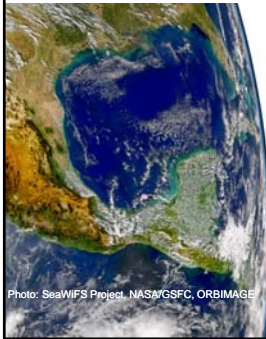
6

Centers of Excellence Data Management

Data Management Approach	FL 8 Centers	MS	LA	TX One Gulf	TX Subsea Systems	AL
Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC)	✓	✓		✓	✓	
Gulf of Mexico Coastal Ocean Observing System (GCOOS)				✓		
National Centers for Environmental Information (NCEI)		✓	✓			
Inter-University Consortium for Political and Social Research			✓			
DataOne Dash			✓			
Make data available within 2 years, after QA/QC, using community-accepted standards and protocols			✓		✓	

GULF RESEARCH PROGRAM

Data Framework



DWH LTDM Workshop
Lauren Showalter
Program Officer
Information Science

The Gulf Research Program

- A \$500 million, 30-year program (until 2043) managed by the National Academies. Funds grants, fellowships, and other activities
- Directed to operate in three areas:
 - Oil system safety
 - Human health
 - Environmental resources
- Directed to work via three mechanisms:
 - Research & development
 - Education & training
 - Environmental monitoring
- Guided by *Strategic Vision* (2014)
- and 20+ member Advisory Board



Photo credits: Background photos ©iStock

Program Initiatives



- Reducing risk in offshore oil and gas operations



- Observation and monitoring for healthy ecosystems and coastal communities



- Planning and action for healthy and resilient coastal communities



- Building capacity to address cross-boundary challenges

Photo credits (from top to bottom): ©iStock/nielubieklon; NASA image courtesy Norman Kuring, Ocean Color Team; ©iStock/stretchc; Photograph by Kelly M. Darnell

Data Timeline

- Pre-award
 - Data management plans are required for all proposals. I review all DMPs and provide those reviews to the review committee
- Project duration
 - I work with all grantees to identify what data or information products will be generated from the project and where they should be made available
- Post-award
 - Data must be submitted to one of the GRP recommended repositories (under development) within one year of project end date
 - Data catalog will be developed to record and point to locations of all data and information products

Data availability and accessibility

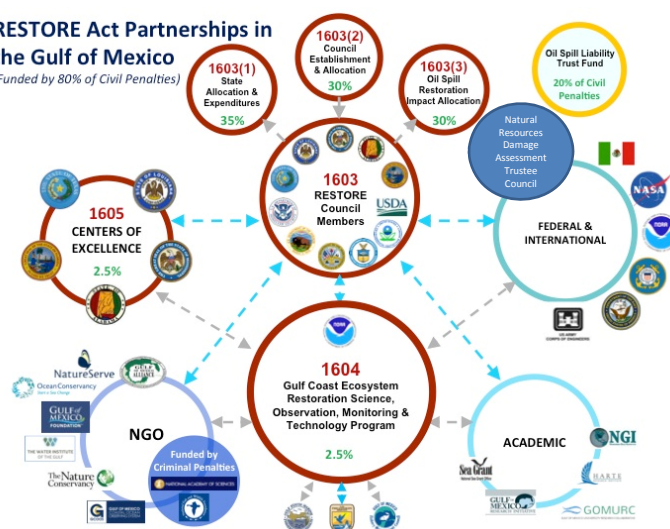
- The GRP will provide data producers with a list of recommended repositories from a variety of disciplines.
 - The GRP Advisory Board recently agreed to a list of repositories that will be contacted to set up agreements to accept GRP data
 - In order for a grantee to submit to a repository not listed they will have to provide a written justification to the grants management team for approval
- The GRP will have a data and information product catalog that will describe and point to all funded projects

Monitoring Community Data Standards/Guidance/Challenges

Greg Steyer
U.S. Geological Survey

Monitoring/Data Acquisition Community

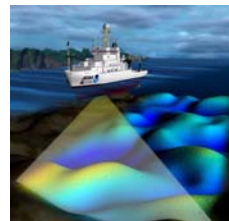
**RESTORE Act Partnerships in
the Gulf of Mexico**
(Funded by 80% of Civil Penalties)



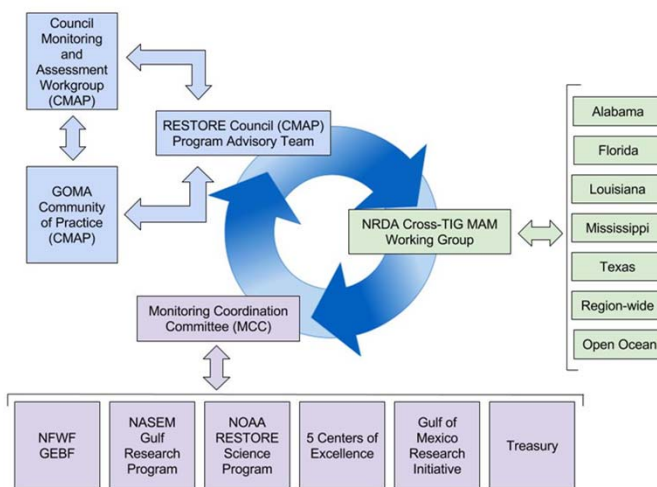
Approach

Use and build on the numerous existing monitoring activities & programs and science in the Gulf

- Identify, catalogue, and understand historic and ongoing monitoring activities and associated data
 - Measurements taken
 - Location
 - Timing
 - Methods/Protocols
- Improve coordination/leverage regional capabilities
- Develop and ensure consistent methods and protocols
- Develop data quality, management, and accessibility standards
- Monitor at different scales (project, basin, state, Gulf-wide)
- Identify and address information gaps
- Utilize science-based decision support tools and adaptive management applications – design to learn



Monitoring and Data Coordination



Guidance

- Coordinated data management system to aggregate, quality assure, store, and disseminate environmental data for the Gulf
- Build an integrated, standards-based, largely virtual system that will support web-based discovery of and access to data streams for diverse end users
- Utilize existing capabilities (web portals, catalogues, archives) where possible, adding new capabilities as necessary
- Common standards for data description, formats, and services (for catalogue queries, Web mapping and data access) should be employed to promote interoperability
- Establish clear and consistent data management, monitoring, adaptive management, and science delivery policies as part of its overarching strategy

Biggest Challenges

- Monitoring and data management communities working together from inception to develop integrated processes
- Communicating and coordinating across both DWH and non-DWH programs
- Designing to the needs of users while meeting the mandates of agencies
- Clearly articulating measurable objectives from project to programmatic scales and common sets of questions we want the monitoring and data management programs to address
- Adoption of common data standards
- Tweaking designs of long-term monitoring and data management programs
- Responsibilities for following minimum monitoring standards & data requirements
- Governance across programs

Big Challenges...but Achievable



National Institute of Environmental Health Sciences
Your Environment. Your Health.



Environmental Disaster Data Management (EDDM) Working Group Update

Field Protocols Working Group

June 7, 2017

National Institutes of Health • U.S. Department of Health and Human Services



National Institute of Environmental Health Sciences
Your Environment. Your Health.

Objectives

- Inventory existing resources for field data collection
- Inventory existing equipment, monitors, devices, and monitors for field data collection
- Apprise academics and NGOs of sampling protocols they should use to get data included

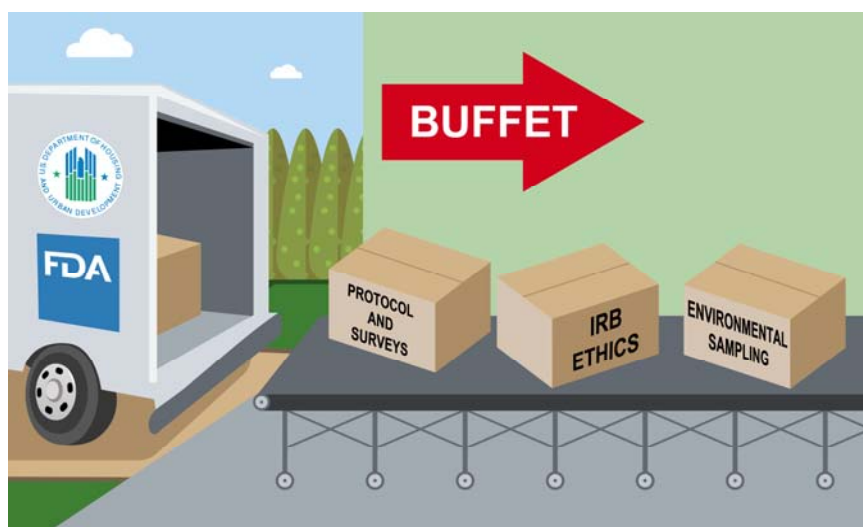
National Institutes of Health
U.S. Department of Health and Human Services



The future...bringing in tools/platforms across USG



Current tools and new tools being added soon





National Institute of Environmental Health Sciences
Your Environment. Your Health.

Providing tools to empowering researchers to quickly assemble comprehensive EHS protocols

DISASTER RESEARCH CAFÉ

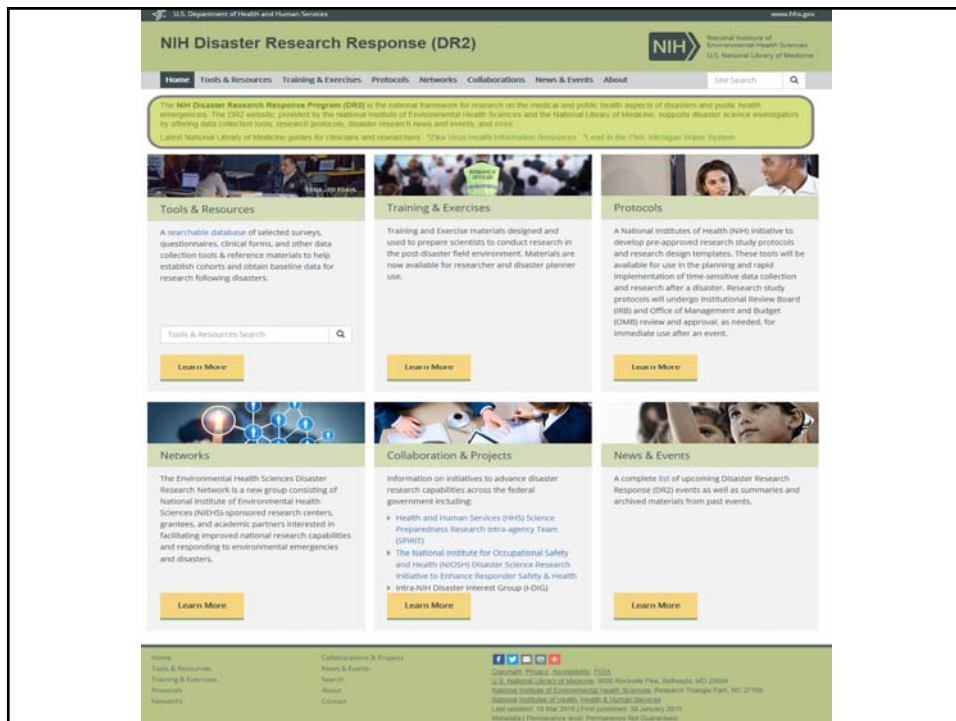


National Institute of Environmental Health Sciences
Your Environment. Your Health.

Empowering others to perform research that includes EHS components

GOVERNMENT WORKERS COMMUNITY ACADEMIC





U.S. Department of Health and Human Services www.hhs.gov

NIH Disaster Research Response (DR2)

National Institute of Environmental Health Sciences
U.S. National Library of Medicine

Home Tools/Resources Training/Exercises Planning/Protocols/Funding Networks Goals/Projects News/Events About Site Search

Data Collection Tools Search

From Disaster LiR®, a database of the NLM Disaster Information Management Research Center.

Refine Your Results

Data Collection on

- ☐ Environmental Exposure(s) (16)
- ☐ Lifestyle and Quality of Life (78)
- ☐ Mental Health and Cognitive Function (139)
- ☐ Occupational Health (29)
- ☐ Preparedness (13)
- ☐ Social Support and Resilience (33)
- ☐ Specific Body Systems (36)
- ☐ Specific Disasters (42)

Source

- ☐ Ashenbach System of Empirically Based Assessment (1)
- ☐ Action Against Hunger (1)
- ☐ Agency for Toxic Substances and Disease Registry (3)
- ☐ American Academy of Orthotists and Prosthetists (1)

[+ See More](#)

Displaying records 1 - 20 of 209

Page 1 of 15 Results/ Page 10 Sort: Newest First

1. ☐ **Youth Risk Behavior Surveillance System (YRBSS)**

Source: Centers for Disease Control and Prevention (CDC)

Format: Text

Date Published: 6/9/2016

Annotation: The Youth Risk Behavior Surveillance System (YRBSS) monitors six types of health-risk behaviors that contribute to the leading causes of death and disability among youth and adults: behaviors that contribute to unintentional injuries and violence; sexual behaviors that contribute to unintended pregnancy and sexually transmitted diseases, including HIV infection; alcohol and other drug use; tobacco use; unhealthy dietary behaviors; and inadequate physical activity. YRBSS also measures the prevalence of obesity and asthma among youth and young adults. Questionnaires are available for both high school and middle school aged participants. This resource was identified by the NIH Disaster Research Response Program (DR2) for researchers looking for pre- and post-disaster data collection instruments.

Ease of Use in Disaster Setting: ● Moderate

Flesch-Kincaid Reading Level: 6.1

Population: Adults and teens

Length: The number of questions varies from middle to high school questionnaires and from year to year. The 2015 standard high school questionnaire contains 50 questions, while the middle school questionnaire contains 40 questions.

Administered by: Self Administered/Self Report

Language(s): English [en]

Data Collection on: Lifestyle and Quality of Life

URL: <https://www.cdc.gov/healthyyouth/data/yrb/index.htm>

Type: Guidelines/Assessment Tool

Access Notes: Free/Publicly Available

Citation(s):
Ying L, Chen C, Lin C, Greenberger E, Wu X, Jiang L. The relationship between post-traumatic stress symptoms and suicide ideation among child survivors following the Wenchuan earthquake. *Suicide Life Threat Behav* 2015 Apr;45(2):230-42. doi: 10.1111/sltb.12118. Epub 2014 Sep 10. PubMed PMID: 25166443.
<https://www.ncbi.nlm.nih.gov/pubmed/25166443>. Subscription required.

Zhou X, Wu X, An Y, Fu F. Longitudinal relationships between post-traumatic stress symptoms and sleep problems in adolescent survivors following the Wenchuan earthquake in China. *PLoS One*. 2014;9(8):e104470. doi: 10.1371/journal.pone.0104470. Epub 2014 Aug 12. PubMed PMID: 25090826. PubMed Free Full Text.

Results from:

Disaster LiR The Resource Guide for Disaster Medicine and Public Health.

[View additional search results from Disaster LiR](#)

PubMed Database of biomedical journal citations and abstracts.

[View search results from PubMed](#)

MedlinePlus Health information resources for the general public.

[View search results from MedlinePlus](#)

U.S. Department of Health and Human Services www.hhs.gov

NIH Disaster Research Response (DR2)

National Institute of Environmental Health Sciences
U.S. National Library of Medicine

Home Tools/Resources Training/Exercises Planning/Protocols/Funding Networks Goals/Projects News/Events About Site Search

Data Collection Tools Search

From Disaster LiR®, a database of the NLM Disaster Information Management Research Center.

Refine Your Results

Data Collection on

- ☐ Environmental Exposure(s) (13)
- ☒ Environmental Sampling (1)
- ☐ Lifestyle and Quality of Life (44)
- ☐ Mental Health and Cognitive Function (60)
- ☐ Occupational Health (25)
- ☐ Preparedness (8)
- ☐ Social Support and Resilience (17)
- ☐ Specific Body Systems (23)
- ☐ Specific Disasters (37)

Source

- ☐ U.S. Geological Survey (1)

Displaying records 1 - 1 of 1

Page 1 of 1 Results/ Page 10 Sort: Newest First

1. ☐ **TEST: National Field Manual for the Collection of Water-Quality Data**

Source: U.S. Geological Survey (USGS)

Format: PDF

Date Published: 1900

Annotation: Provides continuously updated protocols for the proper collection of environmental data: water, sediment, and microbiological. Information in the chapters of this SOP are continuously being updated to reflect the latest technology and sample collection methods and analysis—check back often to be sure the most up-to-date protocol is being followed. U.S. Geological Survey, variously dated. National field manual for the collection of water-quality data. U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chaps. A1-A9, available online at <http://pubs.water.usgs.gov/twri/9A>. This resource was identified by the NIH Disaster Research Response Program (DR2) for researchers looking for pre- and post-disaster data collection instruments.

Sampled media: Water, Sediment

Sampling parameter: Mold, Other chemicals

Administered by: Hydrologist [less]

Data Collection on: Environmental Sampling

URL: <https://water.usgs.gov/owq/fieldManual/>

Type: Instructional/Training Material

Access Notes: Free/Publicly Available

Includes Research Tools: Yes

ID: 13060. From Disaster LiR®, a database of the U.S. National Library of Medicine.

Results from:

Disaster LiR The Resource Guide for Disaster Medicine and Public Health.

[View additional search results from Disaster LiR](#)

PubMed Database of biomedical journal citations and abstracts.

[View search results from PubMed](#)

MedlinePlus Health information resources for the general public.

[View search results from MedlinePlus](#)

Common Data Model

EDDM Working Group

1

The Working Group

- Great support/facilitation by Nancy, Kathy, Laura Belden, and Whitney Hauer
- Members - Ben Shorr, Steve Delgreco, Dan Hudgens, Mike McCann, Mark Stenzel, Scott Thompson, Stephanie Sneyd, Fred Sparks, Joe Schaefer, Lauren Showalter

2

Common Data Models: Objectives

- Objective 1: Document what specific data models, portals (data sets), and web services people are using across different disciplines and compile details regarding each one (portal name, description, type of data accessible, data base compatibility, url, key contacts).
 - Outcome: Spreadsheet of data systems pertinent to environmental disasters

3

Data Systems Examined

- Initially looked at 24 different data systems
- Focused on 7 for initial analysis and gathered information on:
 - Purpose of system
 - Update Frequency
 - Use Restrictions
 - Contacts
 - Category of Data Included (e.g., Weather, Environmental, Operations, Human Dimensions)

4

EDDM Common Data Model: Objective 1 - Data Model						
Portal Name	Location	Purpose	Category (e.g., human health, weather, shipping traffic)	Frequency of Data Updates (e.g., real time, hourly, daily, monthly, yearly)	url	EDDM Contact
Climate Data Online	Asheville	This system is a publicly accessible web-based and web services system that allows the public to request and subset data from various different datasets. Datasets available are GPCPv2 GPCPv2 Monthly Summaries, Radar, Annual Climate Summaries, Precipitation Data (15 minute and hourly), Marine Data, Global Hourly, Global Summary of the Day, and Normals. This system allows selection by either station-level or geographic (map) interfaces. The customer is allowed to query by date ranges and data exists for the most popular datasets from 1700 to present.	Weather and Climate	Hourly, daily, monthly yearly (depends on dataset)	http://cds.ndbc.noaa.gov	Steve Deigne
Storm Events	Asheville	Provides access to historical storm events data (data, fatalities, narratives) provided to NCE-Asheville by NWS on a monthly basis. Data historically goes back to 1950.	Weather and Climate	Monthly, yearly	http://www.ndbc.noaa.gov/stormevents/	Steve Deigne
Severe Weather Data Inventory	Asheville	The Severe Weather Data Inventory (SWDI) is an integrated database of severe weather records for the United States. The records in SWDI come from a variety of sources in the NCEC archive. SWDI provides the ability to search through all of these data to find records covering a particular time period and geographic region, and to download the results of your search in a variety of formats. The formats currently supported are Shapefile (for GIS), KMZ (for Google Earth), CSV (comma-separated), and XML.	Weather and Climate	Monthly, yearly	http://www.ndbc.noaa.gov/sewd/index	Steve Deigne
DIVER (Data Integration/Visualization Exploration and Reporting)	Amazon Cloud East/West (FedRAMP)	Web-based access to National Environmental data warehouse and query tools. Data download with metadata. Details on common data models including: Samples, Photographs, Visual Observations, Telemetry, Instruments.	Environmental	Monthly; As-Needed	https://diver.veron.noaa.gov	Ben Short
Environmental Response Management Application (ERMA)	Amazon Cloud East/West (FedRAMP)	Online mapping tool that integrates both static and real-time data, such as Environmental Sensitivity Index (ESI) maps, ship locations, weather, and ocean currents. In a centralized, easy-to-use format for environmental responders and decision makers. Ability to display static spatial data (e.g. shapefiles, KMZ) and services (e.g. WMS, GeoRSS, other services).	Multiple	As-Needed (Daily)	http://response.restoration.noaa.gov/maps-and-spatial-data/s/national-environmental-response-management-application.html	Ben Short
		TONNET - HSDR contains the following types of information by chemical or chemical family: human health effects, emergency medical treatment, animal toxicity, metabolism/pharmacokinetics, pharmacology, environmental fate and				

5

Common Data Models: Objectives

- Objective 2: Crosswalk existing data models to find similar elements.
 - Outcome: Identify redundancy, compatibility across data models

6

Objective 2 -- Crosswalking

- Recognize importance of “Federated” Data
 - not one system, but connected systems
- Cross-walking to facilitate information sharing
- Develop common vocabularies
- Example: SCRIBE ↔ DIVER cross walking

7

Example Data

SCRIBE FIELDS

Sample #
 CLP Sample #
 Location
 Matrix
 Lab Matrix
 Analysis
 Analyte
 Result
 Units
 Test Type
 Qualifier
 Lab Qualifier
 MDL
 MDL Units
 Lab COC No
 Lab Batch No
 QC Type
 Event
 Lab_Location_ID
 Date_Collected
 Lab_Name
 Lab_Samp_No

DIVER Field Name

Analysis
 Analysis_Category
 Analysis_Detail
 Analysis_Method
 Analysis_Result
 Analysis_Result_Unit
 Analysis_Type
 Case-Activity
 Collection_Form
 Collection_Matrix
 Collection_Method
 Collection_Study_Name
 Collection_Workplan
 Common_Name_Class
 Common_Name_Family
 Common_Name_Genus
 Common_Name_Kingdom
 Common_Name_Order
 Common_Name_Phylum
 Common_Name_Species
 Common_Name_Subphylum
 Composite_Sample_ID

8

Step 1: Field Matches

PART 1: FIELD TO FIELD MAPPING			
SOURCE - SCRIBE FIELD NAME		TARGET -- DIVER FIELD NAME	Notes
Sample #	maps to	Sample_ID	
CLP Sample #			[no direct mapping]
Location		Station-Site	
Matrix		Collection_Matrix	Scribe Uses codes (e.g., S for sediment) that need to be translated on ingest (See Example)
Lab Matrix		Lab-Result_Matrix	Scribe Uses codes (e.g., S for sediment) that need to be translated on ingest
Analysis		Analysis_Type	Scribe Uses codes (e.g., P for PAH) that need to be translated on ingest (See Example)
Analyte		Analysis	Would need to establish lookup table to make sure same conventions used
Result		Analysis_Result	
Units		Analysis_Result_Unit	
Test Type			(data used to establish mapping had no values in Scribe for this field)
Qualifier		Qualifier_Code	Merge Qualifier and Lab Qualifier to Qualifier Code?
Lab Qualifier		Qualifier_Code	Merge Qualifier and Lab Qualifier to Qualifier Code?

9

Step 2: Lookups and Constants

PART 2: DIVER AUTOMATED FIELDS	
DIVER "Calculated" Fields	Explanation
Analysis_Category	Lookup based on Analysis Type
Analysis_Detail	Automatic calculation to create unique key
Hour_of_day	Calculates based on time within date field
Minutes of Hour	Calculates based on time within date field

Part 3: DATASET SPECIFIC CONSTANTS THAT WOULD RECOMMEND SETTING	
DIVER Field Name	Constant to Apply
Case-Activity	?
Collection_Form	= "Scribe"
Collection_Workplan	?
Data_Category	= "Sample"
Data_Classification	?
Data_Source	= "Scribe"
Review_Status	?
Sharing_Status	= "Publicly Available"

10

Common Data Models: Objectives

- Objective 3: At all levels (field collection, synthesis, analysis) inventory/identify existing ways to be interoperable.
 - Outcome: Make recommendations where we can leverage approaches to interoperability and security.
 - Schedule: We hope completion date for this objective will be decided at this workshop

11

12

Written Summary also Available

SIMPLE EXAMPLE OF "CROSS-WALKING" IN NOAA DIVER APPLICATION

The concept of cross-walking focuses on a simple aspect of mapping data fields, typically as part of a larger ETL (Extract, Transform, Load) process. For the example below (taken from DIVER), the focus is on mapping data fields from a data source (table in this case) to the data fields in the DIVER Data Warehouse. This is a simple example – a more complex scenario may include transforming the data so that the structure of the source dataset matches the structure of the destination database, standardizing values, and building a routine to repeat that task.

DIVER implements a "Dataset Template" concept, where a user can build the actual mapping from source data fields to destination dataset fields. In the DIVER Data Template construct, **Fields** contain data of a specific type. **Constants** are values that are applied across all records and are dataset specific – meaning a user can specify them when they load an individual file (see below for more information). The "Cross-Walking" occurs when the Data Manager reviews the structure and field definitions of both the source dataset ("from") and the destination dataset ("to") and creates the mapping. In DIVER, a data manager can create these mappings and then data contributors can use them for a specific project.

FROM FIELD	TO FIELD	MAPPING	CONSTANT
1	1		
2	2		
3	3		
4	4		
5	5		
6	6		
7	7		
8	8		
9	9		
10	10		
11	11		
12	12		
13	13		
14	14		
15	15		
16	16		
17	17		
18	18		
19	19		
20	20		
21	21		
22	22		
23	23		
24	24		

13

Full Data Systems

- Looked at 24 different data systems to start:

Climate Data Online, Storm Events, HDSS Access System, Integrated Surface Data, Geoportal, Severe Weather Data Inventory, Climate Data Records (CDR) Website, NOMADS, Geospatial Services, Earth Observations from Space, Marine Geology & Geophysics, Natural Hazards, Ocean Archive System, World Ocean Database Select (WODselect), DSCRT, MDICH, MOSS, DIVER (Data Integration, Visualization, Exploration, and Reporting), Inter-university Consortium for Political and Social Research (ICPSR), National Environmental Public Health Tracking Network, Environmental Response Management Application (ERMA), Climate Reference Network, Marine Cadastre, Toxicology Data Network (TOXNET) - Hazardous Substance Data Bank (HSDB), Toxicology Data Network (TOXNET) - Integrated Risk Information System (IRIS), Toxicology Data Network (TOXNET) - Toxics Release Inventory (TRI), Toxicology Data Network (TOXNET) - TOXMAP®, Shoreline Cleanup Assessment Technique (SCAT), CAFE: Chemical Aquatic Fate and Effects database, Economic Impact Data, NIOSH Pocket Guide to Chemical Hazards, NIOSH Health Hazard Evaluations (HHEs), National Health and Nutrition Examination Survey (NHANES), Agency for Toxic Substances and Disease Registry (ATSDR), American Conference of Governmental Industrial Hygienists (ACGIH), United States Census Bureau APIs, Integrated Public Use Microdata Series (IPUMS_USA), NASA Socioeconomic Data and Applications Center (SEDAC), WebEOC, EPOC.org_Scribe, VIPER, CAMEO Chemicals

14

EDDM - Gold Standard Working Group

Julie Bosch

DWH Long Term Data Management Coordination Workshop, Mobile, AL

June 7, 2017

Gold Standard Working Group

- Julie Bosch, NOAA NCEI
- Linda Cook, Exponent
- Felimon Gayanilo, Harte Research Institute/GOMRI
- James Gibeaut, Harte Research Institute/GRIIDC
- Matt Howard, GCOOS/GOMRI/GRIIDC
- Ann Jones, Industrial Economics, Inc
- Ben Shorr, NOAA ORR ARD, Spatial Data Branch
- Trish Stewart, Stewart Exposure Assessments, LLC
- Jason Weick, Coastal Waters Consortium/LUMCON
- Kyle Wilcox, Axiom Consulting AOOS Team
- Sarah Wright, Locus Technologies

Gold Standard Working Group

Objective: Identify the functionality needed for information management and decision support tools for different disaster types and where these functionalities are located (e.g., IPAC, HAZUS, ERMA) or missing (gaps).

Outcome: Completed [table](#) including a series of matrices of tool vs. disaster type for different disaster scenarios

- ID functionality & purpose
- Where it exists
- Gaps
- Key data types examples
- Type of disasters
- Summary

Function: Analysis - routine statistical analysis and output

Why: A common platform for viewing analysis or value added to data and/or observations is critical to provide decision makers with raw or observation data in context with thresholds or guidelines.

Does it exist: NOAA DIVER Explorer presents queries for sediment contaminant chemistry that compare to thresholds and guidelines.

NOAA's legacy Query Manager application has an expanded capability for comparison to tissue and water guidelines- NOAA is working to bring these guidelines/thresholds into DIVER Explorer

Gap and Significance: There is a gap in updating guidelines/thresholds and making them available in context of integrated data. In an emergency situation, integrating data from multiple sources and comparing to guidelines is very challenging.

Gold Standard Working Group

Objective: Identify criteria to evaluate data and procedures (for QA/QC, data transport, security, and data use analytics) that can be considered a Gold Standard.

Outcome: Developing a [list of criteria](#), subdivided depending on types of data, methodology, disaster. Develop an evaluation worksheet – of criteria and ranking/result.

- Data type category & data type - *Laboratory Based Measurement - chemical analyses (water, sediment, tissue, blood, oil, other)*
- QC criteria - *Method specified QA/QC criteria for instrument calibration and QC analyses.*
- Current QA/QC procedure
 - 1) *US EPA National Functional Guidelines for Data Review and Validation.*
 - 2) *Professional judgement based on method requirements*
- Responsible party - *Independent third-party data validators*
- Suggestions for QA/QC improvements & efficiencies

Require use of a consistent Standard Reference Material (SRM) or released source material (i.e., control oil) within a program to allow for accurate assessment of inter- and intra-laboratory variability.

Gold Standard Working Group

Objective: Identify critical data types for baseline data for different environments and types of disasters

Outcome: [Listing](#) of critical data types and recommended authoritative sources.

- Critical data types for baseline data
- Parameters
- Media and category
- Recommended resource
- >170 parameter/media identified

Extreme events for coastal environments	Environmental data	Toxicology
Water level	pH	Human toxicology
Water	Sediment/soil	Biologic tissues
NOAA, COOPS; USGS	USDA Natural Resources Conservation Service	International Toxicity Estimates for Risk (ITER)

Gold Standard Working Group

Objective: Identify definitions of terms (data dictionaries).

Outcome: [Listing](#) of different data dictionaries as a function of environmental disaster type and provide access to them.

- Data dictionary name
- Links
- Critical data types
- 56 vocabularies listed

Chemical Carcinogenesis Research Information System (CCRIS)	https://toxnet.nlm.nih.gov/
Climate and Forecast (CF) Conventions	http://cfconventions.org/standard-names.html
Coastal Marine Ecological Classification Standard (CMECS)	http://mmisw.org/orr/#http://mmisw.org/ont/noaa/cmeecs
Darwin Core	http://rs.tdwg.org/dwc/
DIVER	https://dwhdiver.orr.noaa.gov/data-overview
GENE-TOX: Genetic Toxicology Data Bank	https://toxnet.nlm.nih.gov/
Integrated Risk Information System (IRIS)	https://toxnet.nlm.nih.gov/
National Health and Nutrition Examination Survey (NHANES)	https://toxnet.nlm.nih.gov/

DWH LONG-TERM DATA MANAGEMENT WORKSHOP

CRRC, NOAA ORR, NOAA NMFS RC, NOAA NCEI, and GoMRI

June 7 & 8, 2017

Mobile, Alabama

Larry Langebrake



CONNECT SIX LLC

1301 Country Trails Drive
Safety Harbor, FL 34695

Of the hundreds (or thousands) of things we *could* do, what *should* we do?



*“This system is great – we can get exactly the **information** we need, when we need it...”*

Workshop Objectives:

- Foster collaboration among the Gulf of Mexico partners with respect to data management and integration for restoration planning, implementation and monitoring.
- Identify standards, protocols and guidance for long term data management being used by these partners for DWH NRDA, restoration, and public health efforts.
- Obtain feedback and identify next steps for the work completed by the Environmental Disasters Data Management (EDDM) Working Groups.
- Work towards best practices on public distribution and access of this data.

Workshop Objectives:

- Work towards best practices on public distribution and [broad] access of this data.

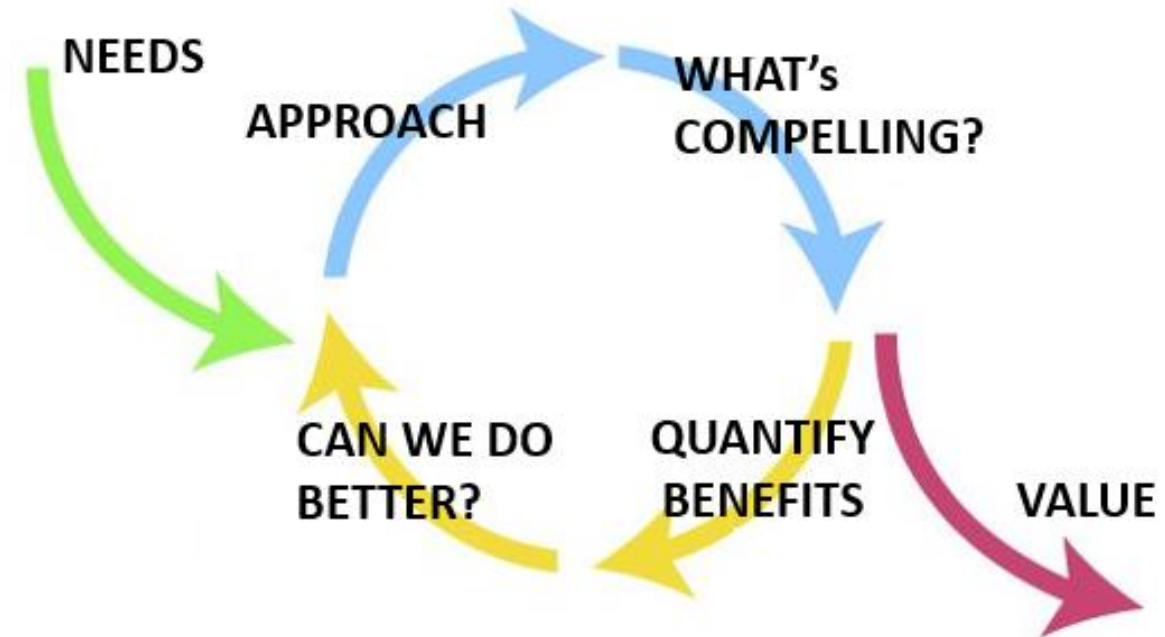
Work towards best practices ...

Value Creation is a **best practice** that produces an optimum and compelling outcome.



The value creation process has three main components:

1. Identifying and quantifying customer *need(s)*;
2. Iterating on an *approach*; and
3. Quantifying *benefits* and *cost* then contrasting those against *alternatives*.



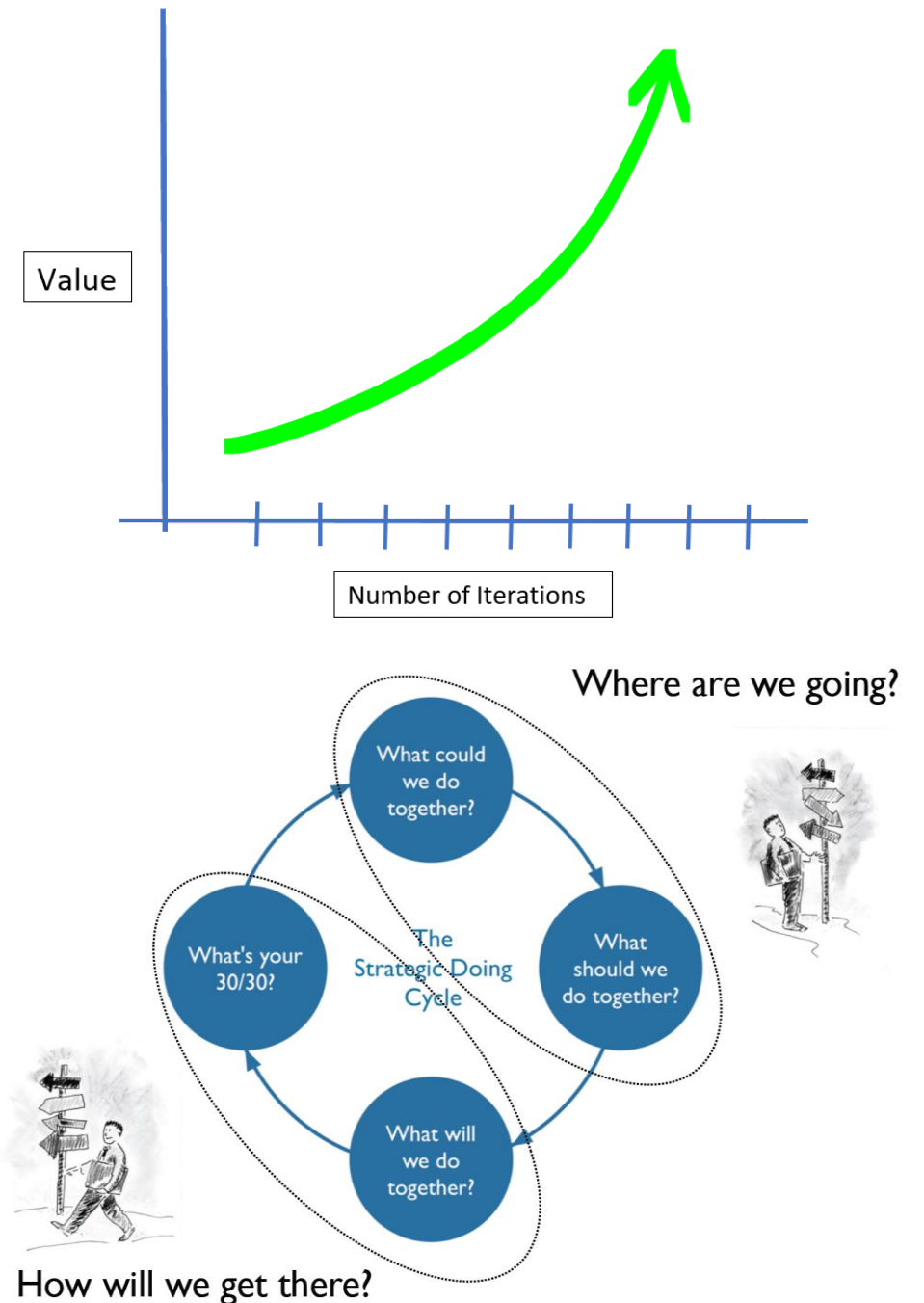
Value creation is an *iterative* process.

$$\text{Value} = \text{Benefit} / \text{Cost}$$

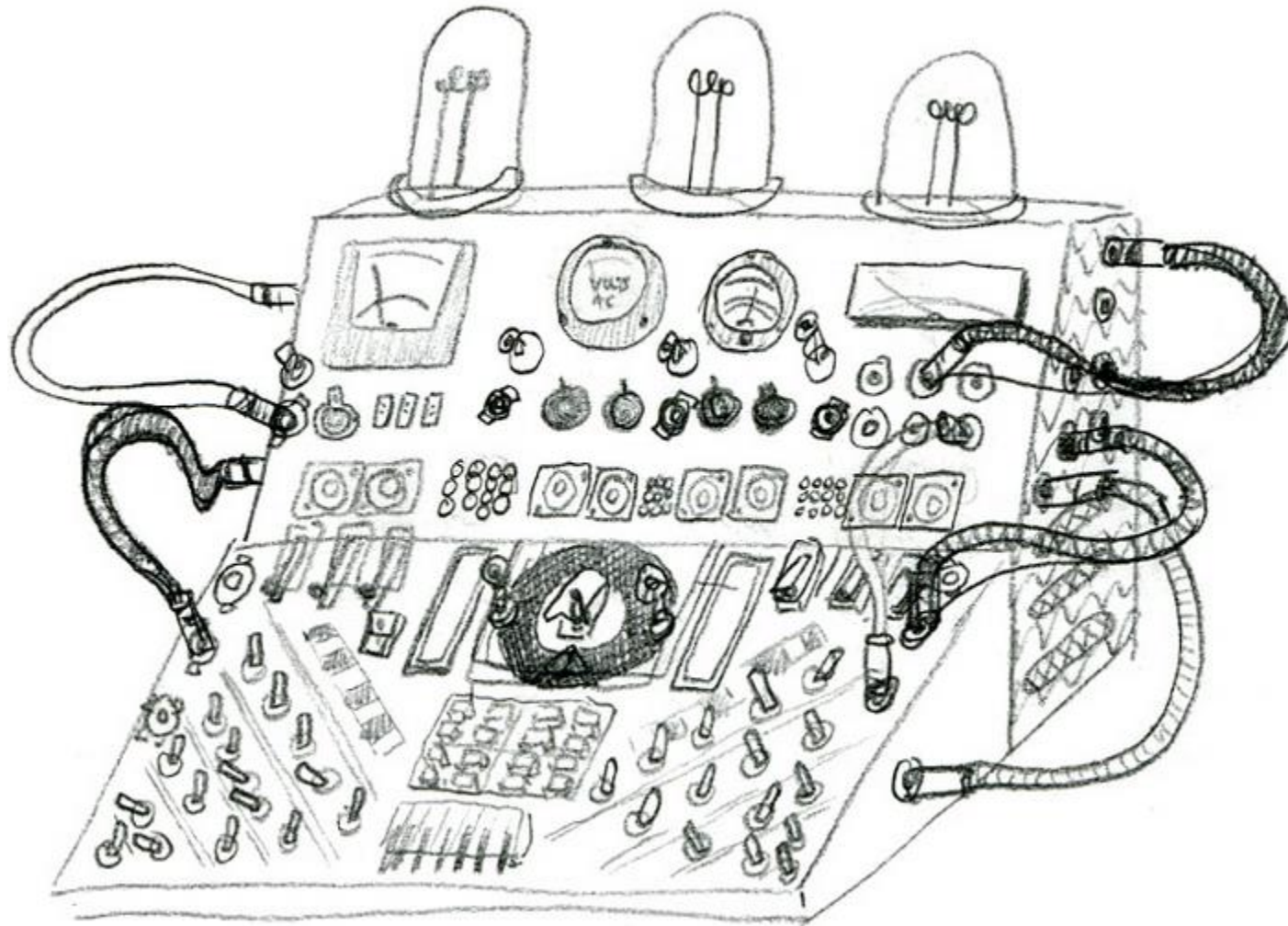


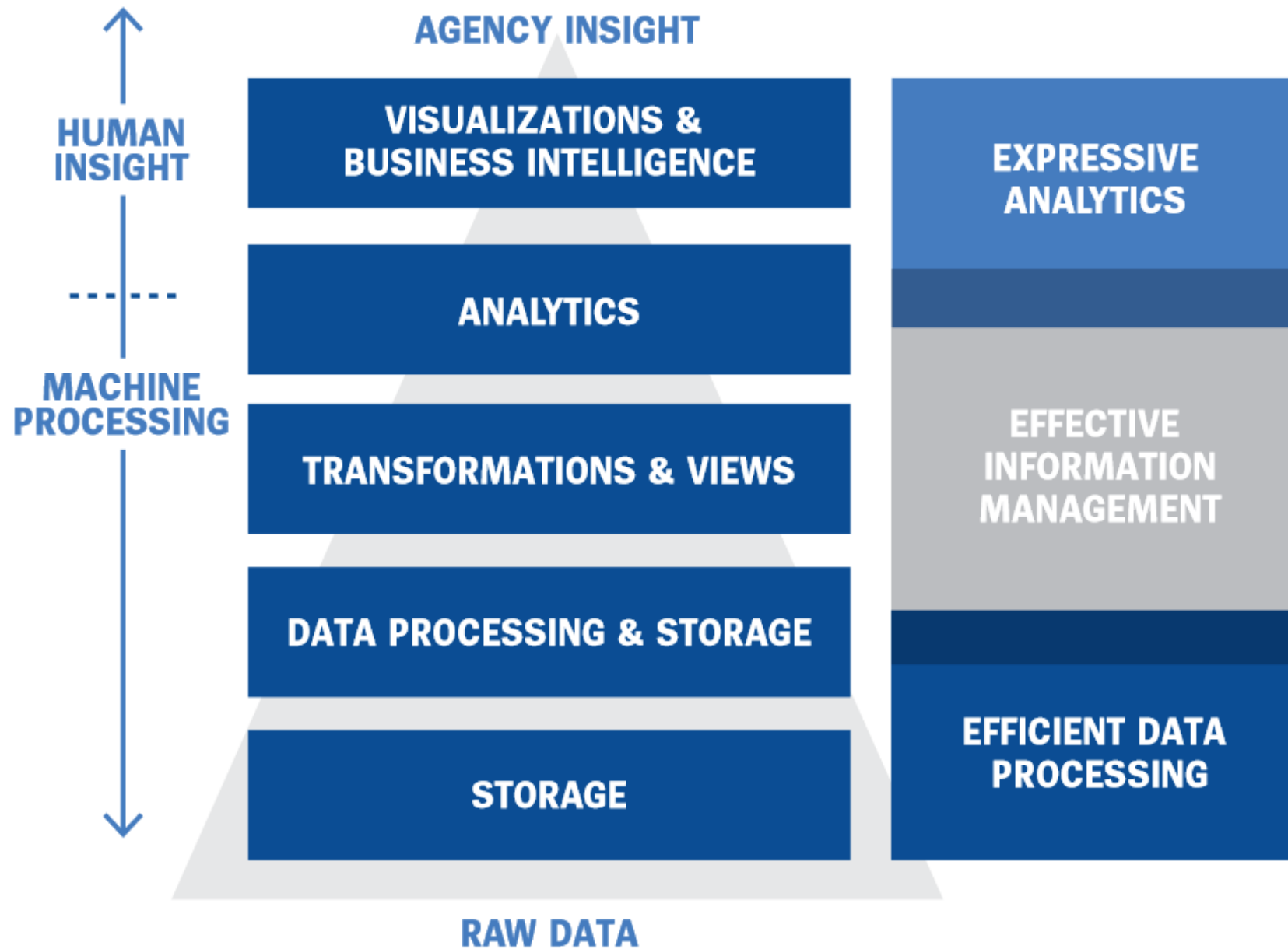
Some elements of Value Creation:

1. Common language
2. Iteration to a compelling solution (includes divergent & convergent thinking)
3. Champions
4. Alignment
5. Use of subject matter experts!



A brief look at how others are
addressing big-data...





The “value chain” of data and information management:

Data ➔ Information ➔ Insight ➔ Inspiration

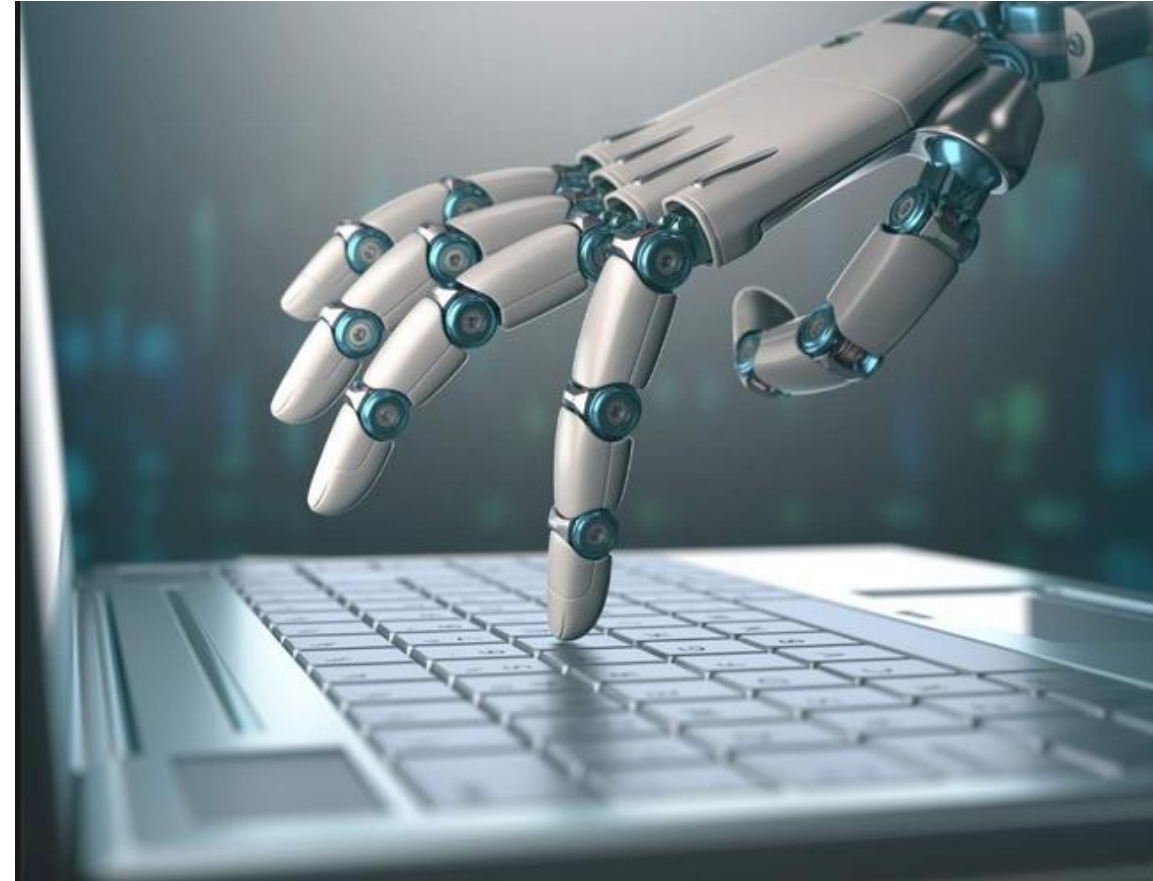


Increasing value

How is industry responding?

- HPE: “The Machine”, 160TB to implement “memory-driven-computing” → 4090 yotta bytes
- Google: knowledge graph, semantic web
- IBM: “cognitive computing”, Watson
- Microsoft: data to insight to... (inspiration?)
- SRI: SOA and beyond
- Amazon: plumbing, to measurement, to content, to...loop (recommender engines)
- D-Wave, IBM: quantum computing
- AI

... and there's no shortage of search tools, cloud resources, analytical services, more...



In short, there are no constraints in technology but typically there are in its implementation. However, usually those constraints are either financial or expertise-related. In both cases, the root of such constraints stem from policy. And, policy stems from culture. A benefit of the value-creation process is the opportunity for culture change (and thus policy change).

... back to the value creation process

First step – who are the data/information customers?

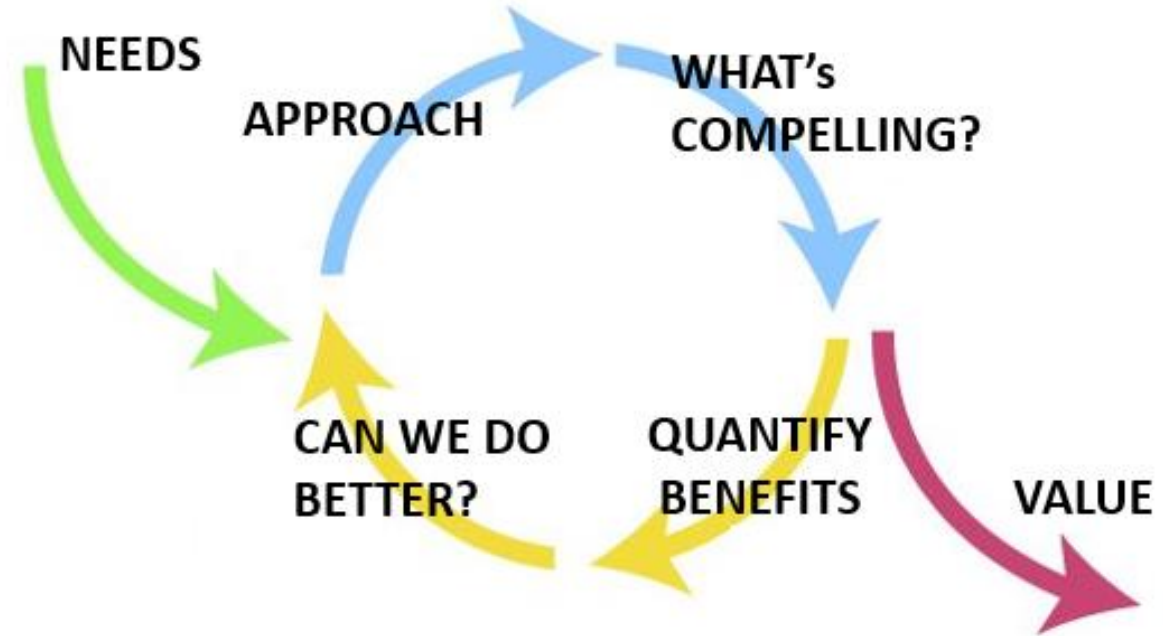
- Researchers/Scientists
- Disaster responders
- Coastal Communities
- The Medical Community-of- interest
- Policy Makers
- Students
- Industry
- Commercial and Rec. Fishing



An example: Researcher/Scientist

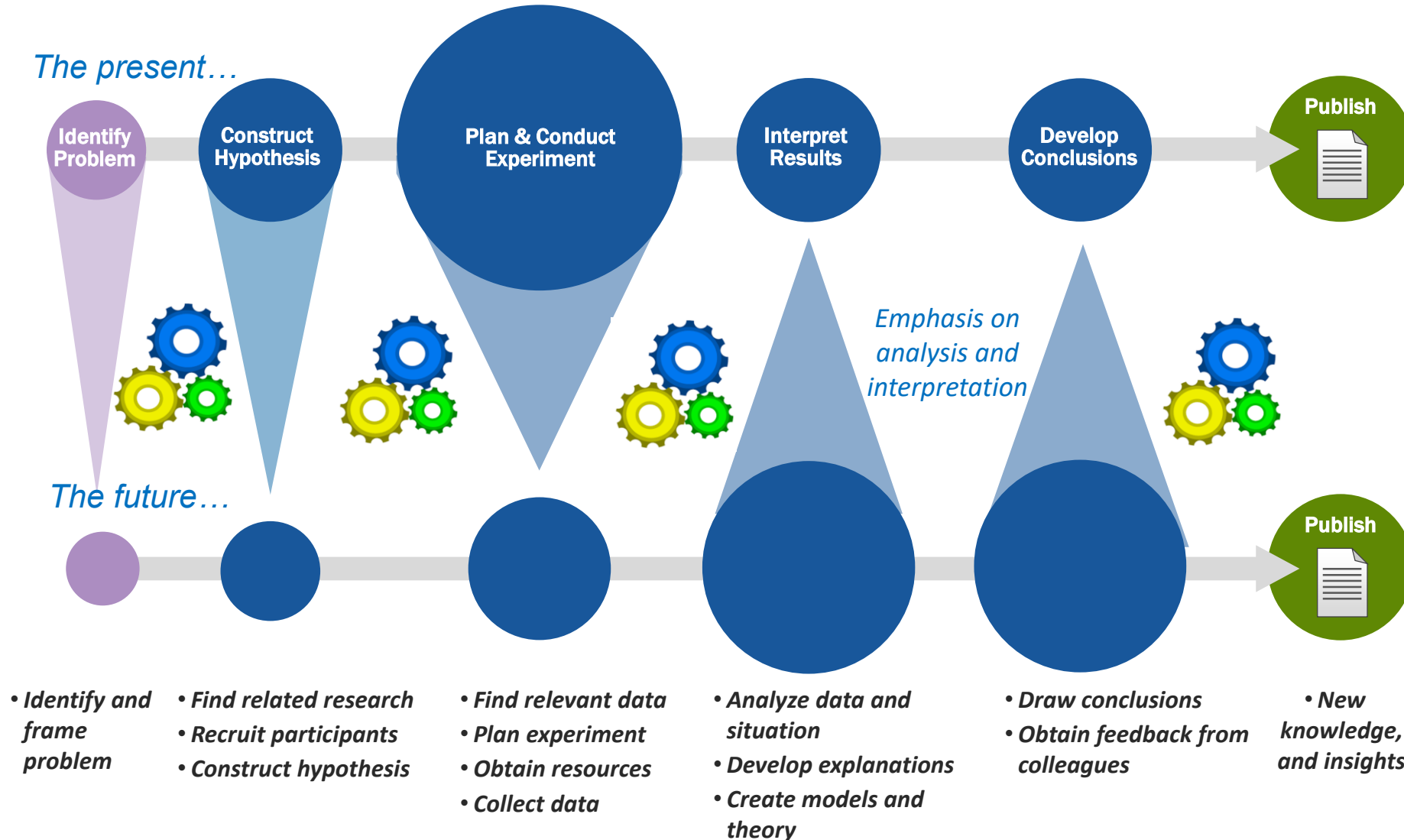
1. Identifying and quantifying customer need(s) .

- Reduce time to identify important problems
- Reduce time needed for preliminary research
- Reduce requirement for new data
- Automate analysis
- Automate publishing



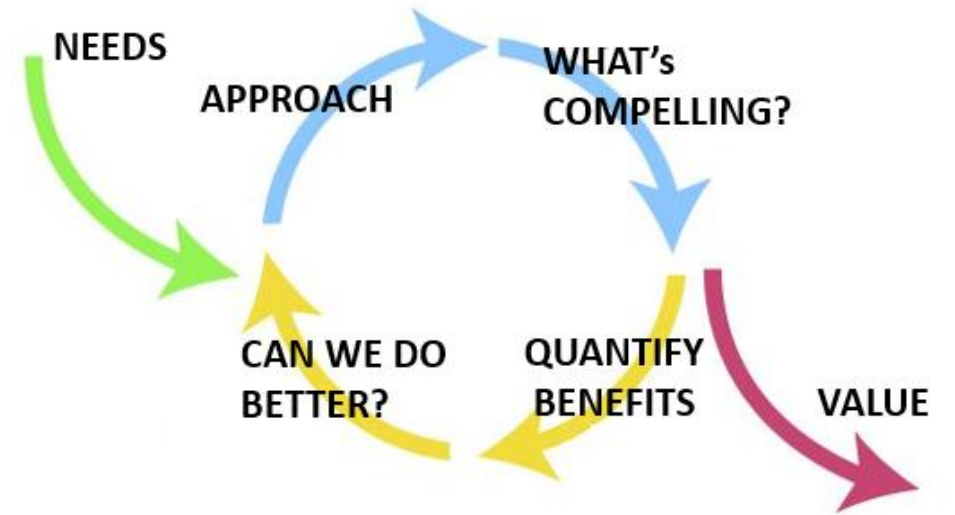
Value Creation for the scientist/researcher

(in a data and information management context.)

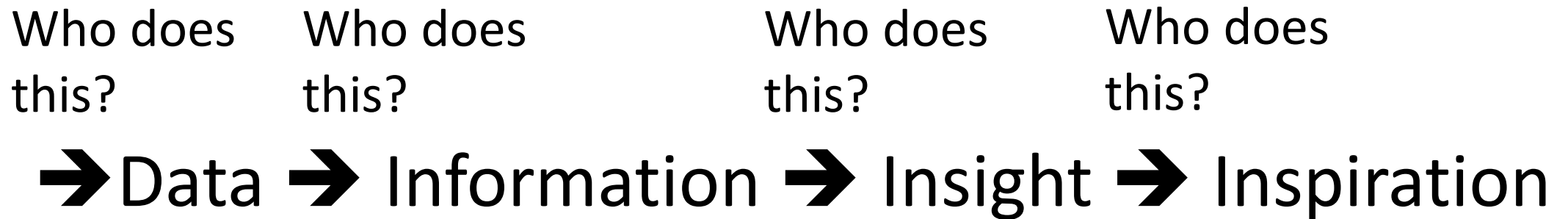


Converging on the main points for the workshop:

1. Consider first: who is the customer and what are their **important needs**.
2. Consider the **benefits and costs** of an **approach** – **plan to iterate with others & SME's**
3. Adopt and use a **common language**
4. The solution will need a **passionate champion** - consider that when identifying the approach.
5. **Alignment is crucial**, there **must** be a team, organization or dedicated collaboration for a viable approach.



The “value chain” of data and information management:



Where does the value creation process lead the conversation? What could we do ➔ what should we do? We know the needs – what tools (or solutions) fit best? Does the customer gain value? ***What are the important needs?***

Welcome to Kaggle Competitions

Challenge yourself with real-world machine learning problems



New to Data Science?

Get started with a tutorial on our most popular competition for beginners, [Titanic: Machine Learning from Disaster](#).



Build a Model

Get the data & use whatever tools or methods you prefer to make predictions.



Make a Submission

Upload your prediction file for real-time scoring & a spot on the leaderboard.

✕ Dismiss



15 active competitions

Sort by Prize

Active All Entered

All Categories

Search



Zillow Prize: Zillow's Home Value Prediction (Zestimate)

Can you improve the algorithm that changed the world of real estate?

Featured - 8 months to go

\$1,200,000

467 teams



Intel & MobileODT Cervical Cancer Screening

Which cancer treatment will be most effective?

Featured - 20 days to go

\$100,000

719 teams



Google Cloud & YouTube-8M Video Understanding Challenge

Can you produce the best video tag predictions?

Featured - a day to go

\$100,000

659 teams



Planet: Understanding the Amazon from Space

Use satellite data to track the human footprint in the Amazon rainforest

Featured - 2 months to go

\$60,000

341 teams



Instacart Market Basket Analysis

Which products will an Instacart consumer purchase again?

Featured - 2 months to go

\$25,000

418 teams



Mercedes-Benz Greener Manufacturing

Can you cut the time a Mercedes-Benz spends on the test bench?

Featured - a month to go

\$25,000

675 teams



Sberbank Russian Housing Market

Can you predict realty price fluctuations in Russia's volatile economy?

Featured - a month to go

\$25,000

2,444 teams



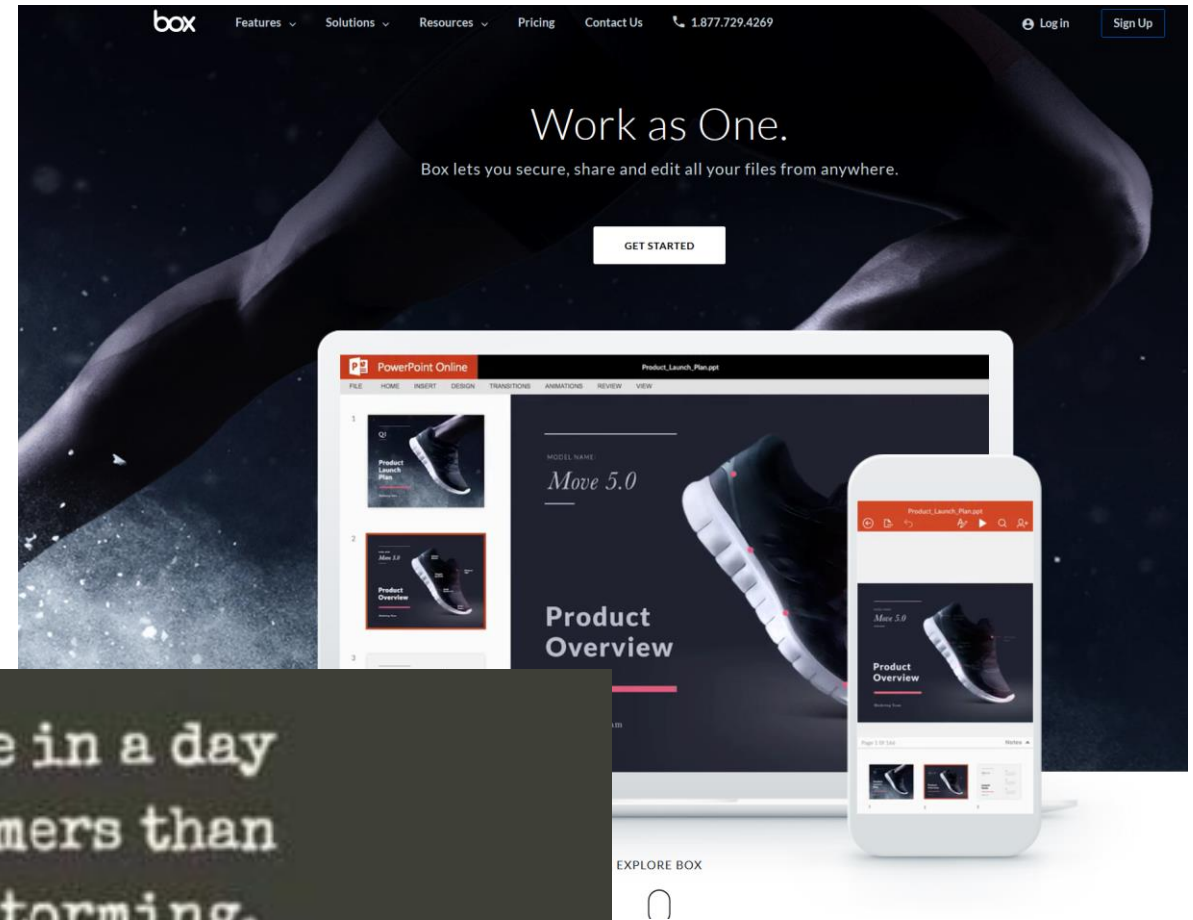
NOAA Fisheries Steller Sea Lion Population Count

How many sea lions do you see?

Featured - a month to go

\$25,000

207 teams



You'll learn more in a day
talking to customers than
a week of brainstorming,
a month of watching competitors,
or a year of market research.

Aaron Levie, Box.com

Appendix D: Pre-workshop survey

Visions for Long Term Data Management in the Gulf

*** 1. How would you best describe yourself? (Pick one)**

- ☐ Data User
- ☐ Data Generator
- ☐ Data Manager/Data Administrator
- ☐ Program Manager/Funder
- ☐ Decision Maker

Other or Additional details about your selection:

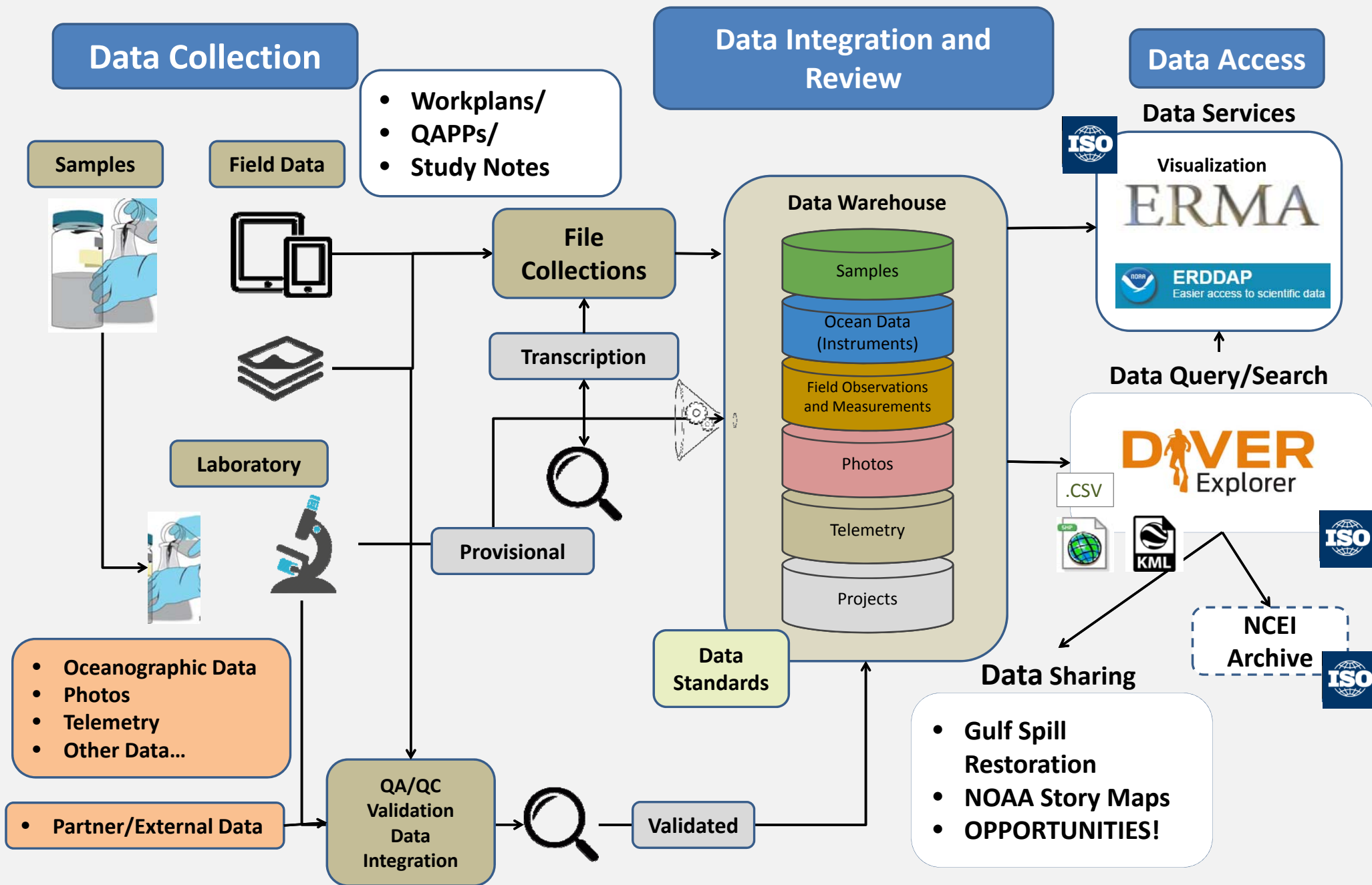
*** 2. What do you **want** from Gulf research/monitoring data 15 years from now? (Pick all that apply)**

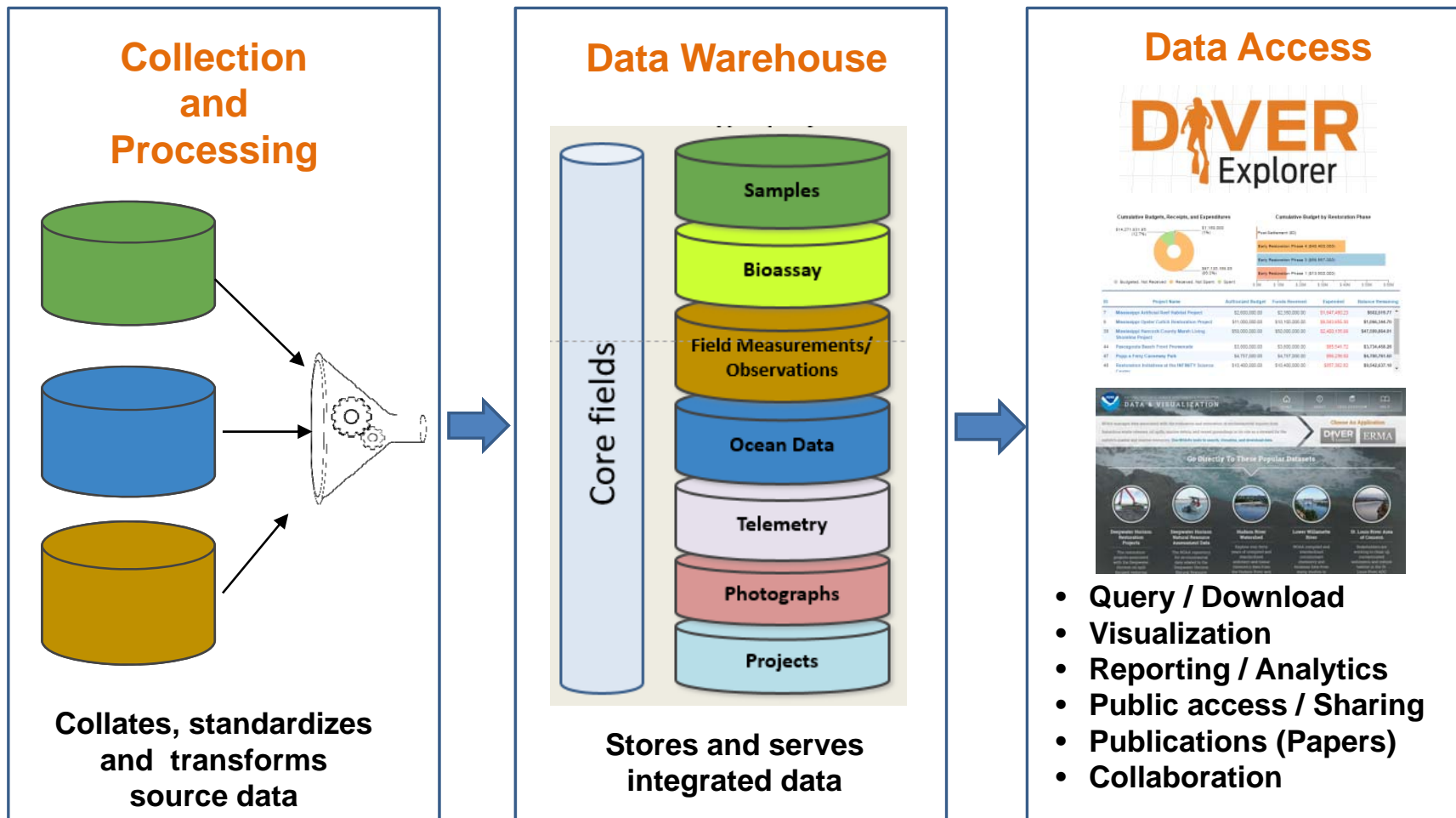
- ☐ No change - the way it is now suits my needs
- ☐ All data is stored in a long-term data repository
- ☐ All data follows a common set of standards
- ☐ All data is accessible
- ☐ All data is interoperable allowing for users to develop their own analytic tools
- ☐ All data is interoperable and synthesized through analytic tools available for all users

*** 3. Realistically, what do you **think** Gulf research/monitoring data will have achieved in 15 years? (Pick all that apply)**

- ☐ No change from current practices
- ☐ All data will reside in a long-term repository
- ☐ All data will follow a common set of standards

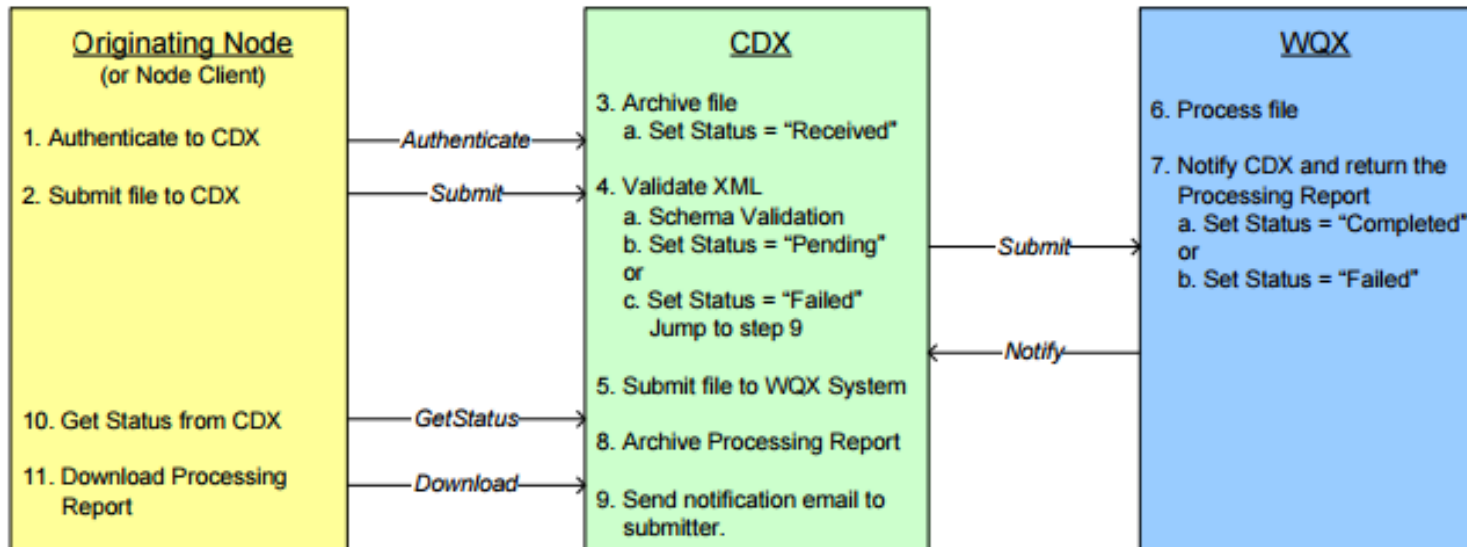
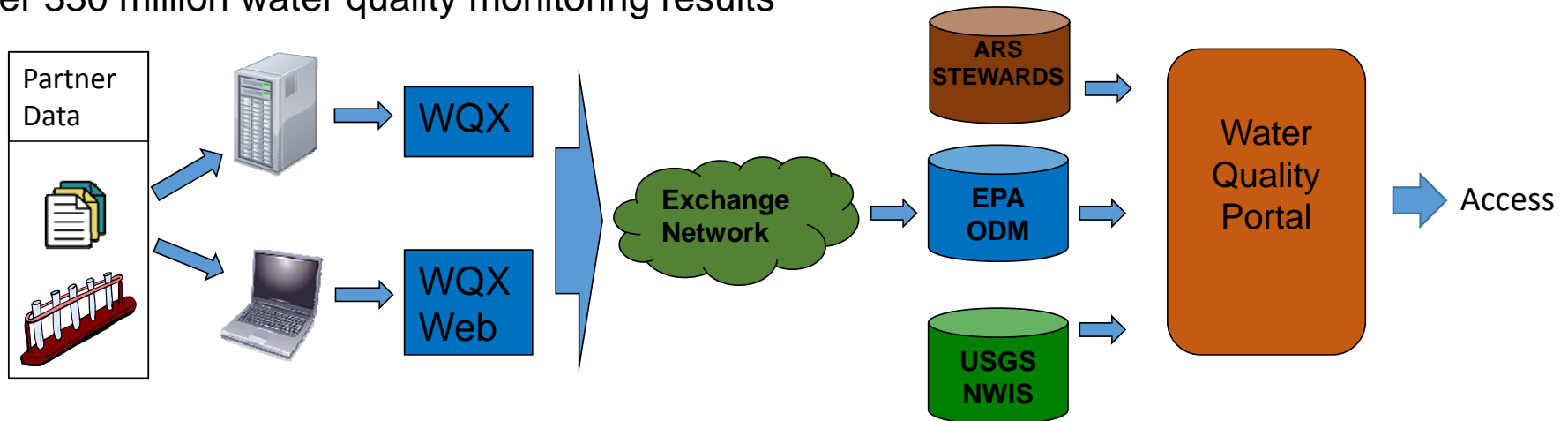
Appendix E: Flow diagrams

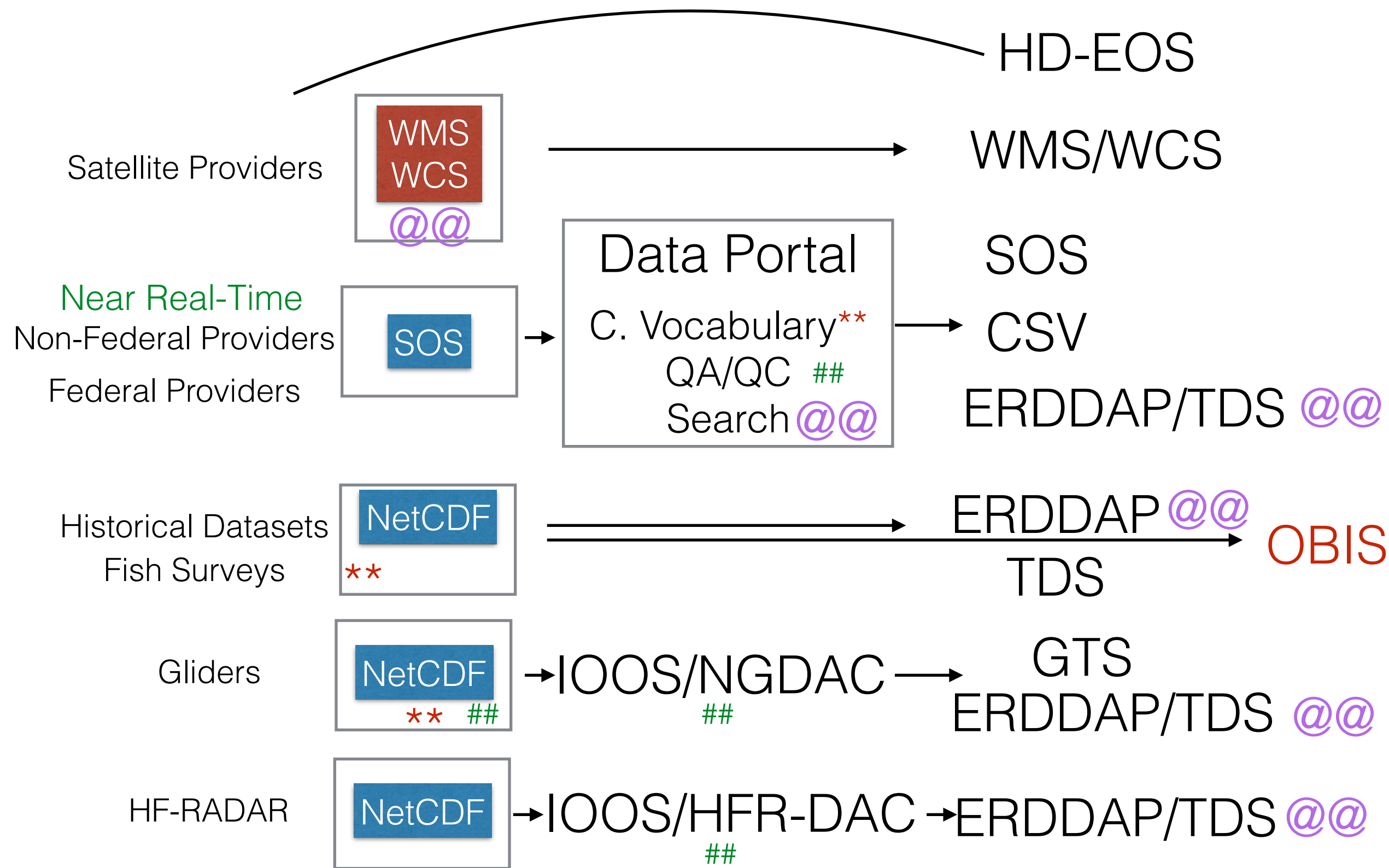




For more information on the portal see: www.waterqualitydata.us
 Currently over 330 million water quality monitoring results

States, tribes,
 other feds,
 local groups





Sources

FO-Real-Time

FO-Human Collected

SP-Human Generated

SP-Model Generated

QC

QA – Automated /
Subject Matter Expert

QA – Automated /
Subject Matter Expert

Interop

Spatial DBs

Tabular DBs

Services ☺

Discoverability

Access / Usability

Applications
Public AND Private

Synthesis...NA

Appendix F: Breakout session templates

Breakout Session I: Group B: Data Generator

1. Reactions to plenary sessions

-

2. Participation in long-term data collaborative

Requirements	Incentives:	Challenges:	Cost (high, med, low)	Priority (high, med, low)
1. Interoperability				
2. Ease of Discovery/Searchability				
3. Data Access				
4. Data Synthesis				
5. Data Usability				
6. Metadata/Data Documentation				
7. Other?				

	Incentives:	Challenges:	\$ Cost (optional) (high,med,low):
Individual agency requirements			
1.			
2.			
3.			
Common set of metadata standards across community			
Maintenance of datasets in perpetuity			
Timely availability to user community			
Accessibility (use/reuse) of data in future			

Breakout Session I: Group C: Data Managers/Governors

1. Workload Effort

Requirements	Challenges:	Priority (high, med, low)
1. Interoperability		
2. Ease of Discovery/Searchability		
3. Data Access		
4. Data Synthesis		
5. Data Usability		
6. Metadata/Data Documentation		
7. Other?		

2. Long term repository

Challenges:	Costs:

3. Definition of success

	Characteristics/ Metrics
Data management system	
Data repository	

4. Barriers

	Barriers
Getting generators to submit data	
Finding data	

Breakout Session I: Group C: Data Managers/Governors

Using data	
Enforcing data policy	
Data security	

Day 2: Breakout Group Session II: Group D: Data Synthesis

Challenges	Solutions (including from other long term data management disciplines)
<i>Listed here from day 1 in priority list</i>	

Data management requirements for successful data synthesis
<ul style="list-style-type: none">• xx

Guidance/ best practices on data synthesis for funders/data generators
<ul style="list-style-type: none">• xx

Breakout Session III

Draw end-to-end process / Flow Diagram

Day 2: Breakout Group Session II: Group D: Data Usability

Challenges	Solutions (including from other long term data management disciplines)
<i>Listed here from day 1 in priority list</i>	

Data management requirements for successful data usability
<ul style="list-style-type: none">• xx

Guidance/ best practices on data usability for funders/data generators
<ul style="list-style-type: none">• xx

Breakout Session III

Draw end-to-end process / Flow Diagram

Day 2: Breakout Group Session II: Group F: Metadata / Data Documentation

Challenges	Solutions (including from other long term data management disciplines)
<i>Listed here from day 1 in priority list</i>	

Data management requirements for successful metadata/data documentation
<ul style="list-style-type: none">• xx

Guidance/ best practices on metadata/data documentation for funders/data generators
<ul style="list-style-type: none">• xx

Breakout Session III

Draw end-to-end process / Flow Diagram

Appendix G: Breakout group participant list

Group A: Data User	Group B: Data Generator	Group C: Data Manager/Governor
Lead: Mark Miller	Lead: Dan Hudgens	Lead: Rost Parsons
Recorder: Melissa Gloekler	Recorder: Megan Verfaillie	Recorder: Kathy Mandsager
Holly Binns	Courtney Arthur	Jonathan Blythe
Melody Chimahusky	Jennifer Bauer	Julie Bosch
Jay Coady	Ryan Druyor	Laura Bowie
Linda Cook	Neal Etre	Craig Conzelmann
Alyssa Dausman	Carl Ferraro	Steve Delgreco
Nic Eckhardt	Shawn Fisher	Sandra Ellis
Jim Gibeaut	George Graettinger	Lei Hu
Jessica Henkel	Mark Howard	Christina Hunnicutt
Amy Hunter	Dan Hudgens	Michele Jacobi
Helga Huntley	Ann Jones	Steve Jones
JB Huyett	Kirsten Larsen	Laurie McGilvray
Syed Khalil	Matt Love	Marti McGuire
Barb Kirkpatrick	Kate McClure	Mike Peccini
Julien Lartigue	Amy Merten	Dave Reed
Gareth Leonard	Tamay Ozgokmen	Denise Reed
Nadia Martin	Steve Ramsey	Angela Schrift
William Nichols	Rick Raynie	Greg Steyer
Jon Porthouse	Jamey Redding	Tom Strange
Lauren Showalter	Ben Shorr	Kevin Suir
Danny Wiegand	Hugh Sullivan	Jason Weick
Eric Weissberger	Carrie Wall	Caitlin Young
Dwane Young		

Group A: Data User	Group B: Data Generator	Group C: Data Manager/Governor
Lead: Mark Miller	Lead: Dan Hudgens	Lead: Rost Parsons
Recorder: Melissa Gloekler	Recorder: Megan Verfaillie	Recorder: Kathy Mandsager
Holly Binns	Courtney Arthur	Jonathan Blythe
Melody Chimahusky	Jennifer Bauer	Julie Bosch
Jay Coady	Ryan Druyor	Laura Bowie
Linda Cook	Neal Etre	Craig Conzelmann
Alyssa Dausman	Carl Ferraro	Steve Delgreco
Nic Eckhardt	Shawn Fisher	Sandra Ellis
Jim Gibeaut	George Graettinger	Lei Hu
Jessica Henkel	Mark Howard	Christina Hunnicutt
Amy Hunter	Dan Hudgens	Michele Jacobi
Helga Huntley	Ann Jones	Steve Jones
JB Huyett	Kirsten Larsen	Laurie McGilvray
Syed Khalil	Matt Love	Marti McGuire
Barb Kirkpatrick	Kate McClure	Mike Peccini
Julien Lartigue	Amy Merten	Dave Reed
Gareth Leonard	Tamay Ozgokmen	Denise Reed
Nadia Martin	Steve Ramsey	Angela Schrift
William Nichols	Rick Raynie	Greg Steyer
Jon Porthouse	Jamey Redding	Tom Strange
Lauren Showalter	Ben Shorr	Kevin Suir
Danny Wiegand	Hugh Sullivan	Jason Weick
Eric Weissberger	Carrie Wall	Caitlin Young
Dwane Young		

Appendix H: Breakout group notes

Breakout Session I: Data Users

1. Reactions to plenary sessions
 - Larger than DWH- what are the bounds of the endeavor?
 - i. Ideas discussed here could be expanded to a broader arena
 - Look at data pre DWH for restoration purposes
 - Participation (donate data) not funded by DWH
 - i. How do you engage a wider community of researchers?
 - ii. Generated data that is not represented in this room (e.g., county data)
 - iii. Identify other data generators and how it relates to monitoring/restoration
 - Baseline: data suitable to help with restoration assessments
 - Requirements placed on data generators by funders;
 - i. Standard set of recommended repositories
 - ii. repository would have a standard data format
 - Identify different types of standards and what is important to the data user.
 - What are we losing when focusing on standards?
 - i. E.g., Title, geographic location, date of location
 - ii. There are data standards regulated by statute
 - iii. Quality standards vs. type of information required (this depends on the data user, allow them to make the decision)
 - iv. Data that does not follow certain guidelines cannot be endorsed by government agencies.
2. Define the term “user” purpose and/or people with varying levels of domain knowledge
 - Perform domain analysis to understand users and their needs
 - Cannot develop repository for entire range of users, determine users and then develop the repository
 - Data can be in repository at the highest quality, but user decides what level it is drawn out at.
3. Types of Data Users
 - High Domain:
 - Researchers
 - Resource Managers (e.g., fish and game)
 - Decision Maker
 - Regulator
 - Government
 - Funders
 - Low Domain:
 - Resource Users (e.g., fishermen)
 - Media
 - Public
 - Litigation
4. Outcomes of this work should be applied to a broader situation, share with and engage with other agencies.

5. List of requirements from the user community:

Requirements	Features/Needs	Challenges	Priority (high, med, low)
1. Interoperability			
2. Ease of Discovery/Searchability	<ul style="list-style-type: none"> • Data are catalogued across systems • Geographic search • Temporal search • Parameter search 	<ol style="list-style-type: none"> 1. Recommended set of repositories 2. Periodic review to add repositories 3. Match anticipated use with key words 4. Characteristics of a good repository 	4.
3. Data Access	<ul style="list-style-type: none"> • Access to data system to input data • Open (unrestricted) access • Download data or subset • Getting a preview of the data 	<ol style="list-style-type: none"> 1. Funding 2. IT security 3. Confidentiality of data 4. In order to allow subset download, repository must have detailed knowledge of data 5. Lack of network technology 	1., 2., 3.
4. Data Synthesis	<ul style="list-style-type: none"> • Visualization/ simple analysis tools • Do not stop at making data available; develop synthesized products based on user needs 	<ol style="list-style-type: none"> 1. Apply resources to prepare the interpreted/synthesized products 2. Anticipate user needs 3. Anticipate diversity of user needs 	2., 3.
5. Data Usability	<ul style="list-style-type: none"> • Data standards so data can be shared in a format 	<ol style="list-style-type: none"> 1. Provide sufficient information to allow users to compare data 	1.
6. Metadata/Data Documentation	<ul style="list-style-type: none"> • Must support discoverability 	<ol style="list-style-type: none"> 1. Very robust metadata takes a lot of work 	1.
7. Longevity		<ol style="list-style-type: none"> 1. Sustain a repository 2. Magnitude of data 	1., 2.

Breakout Session I: Data Generators

1. Reactions to plenary sessions

- Crosswalk between these applications. What is located in each place and where you can find it? What is the turnout in terms of report, metadata, etc.?
- Major data systems are very different. Package of information versus integrated data that has been standardized. Unreasonable to expect that finding integrated data will be easy. What are the goals from each of the different groups?
- Is a federated data system a goal? If we can get all systems built on common framework it will make it more valuable.
- Minimum data quality standards and how the data has been collected. Set number of required data that are comparable in quality.
- Integration is the highest level goal we are trying to achieve, but also one of the most challenging. Requires investment in logistics.
- Tools and analysis are often geared towards users. What tools exist for data generators? For example, quality standards. Connecting data generators in order to increase the investment in quality data.

2. Participation in long-term data collaborative

Requirements	Incentives:	Challenges:	Cost (high, med, low)	Priority (high, med, low)
1. Interoperability	<ul style="list-style-type: none">• Participating more may create more funding and visibility• Allows you to articulate your project, what other ways can this data be used (thanks to its interoperability)• Working in a system outside of other generators (mutual benefit)	<ol style="list-style-type: none">1. May not have understanding or funding that allows you to meet a minimum standard2. How to make this part of everyday collection efforts3. Primarily a user concern, how to make it more relevant to generators4. Ensuring clear metrics As an agency, trying to establish themselves as an expert in this area in order to educate others (also applies to academia)		2. Highest 1. High

2. Ease of Discovery/ Searchability	<ul style="list-style-type: none"> Inputting data gives an opportunity to apply for more funds 	<ol style="list-style-type: none"> Level of budget available to hire people to monitor the imputing and sharing of data by generators Data users should work with generators to determine the needs/search preferences 	7. Highest	6. 7. High
3. Data Access	<ul style="list-style-type: none"> Funding incentives from funding organizations 	<ol style="list-style-type: none"> Ownership of data and delay of access 		
4. Data Synthesis	<ul style="list-style-type: none"> Citation requirements for academia (recording # of citations/ downloads) or set timeline for private data that will then become public Share information about projects more broadly so that recognition is being given for existing work Making sure that the data systems are tracking DOIs 	<ol style="list-style-type: none"> Separate funding for integration as an incentive Combining information from one component to find commonalities (structured and unstructured) 		
5. Data Usability	<ul style="list-style-type: none"> Flexibility to work different QAPs 	<ol style="list-style-type: none"> Single use versus reuse by others All data should be of known quality (must be enough information to judge quality) Ensuring that data is entered in a way that it may be used for visualization by users Need a set system for determining data maturity 	20. Potential for leveraging data maturity or other models	20. Highest 21. High

		5. Ensure that the methods for contributing data are easy to use (current exchange used by data generators) 6. Making it easier to know what is required by QAC		
6. Metadata/Data Documentation	<ul style="list-style-type: none"> • Templates that generate metadata (& other documentation) marketed as time and money savers 	1. Create templates/minimum standards from the beginning of data collection (data management plans, etc.)	25. Medium	25. Highest

Common set of metadata standards across community	<ul style="list-style-type: none"> • Will benefit all if single set can be developed 	
Maintenance of datasets in perpetuity	<ul style="list-style-type: none"> • How do we ensure that machine readable standards evolve with technology • Getting data to the archives that store/manage 	
Timely availability to user community	<ul style="list-style-type: none"> • QA, especially when there are no common standards • Ownership of data/"I want to publish" 	
Accessibility (use/reuse) of data in future	<ul style="list-style-type: none"> • Having the resources to make it accessible/ease of transfer to upload and share 	

Breakout Session I: Data Managers

1. Reactions to Plenary Session

- Interoperability (datasets or metadata) – how to match due to different system requirements.
- Need good definition of terms.
- Data redundancy; different QCs for different systems. Transparency and version control is needed.
- What is the process for data correction (data owner to data repository), versioning of the datasets. Suite of products may have been developed from earlier versions, not the corrected, updated data.
- DOI assignment is needed.
- Note the funding agency's mission for creating database; to better understand the data. Define funders in the metadata.

2. Workload Effort

Requirements	Challenges:	Priority (high, med, low)
1. Interoperability	1. Scale of interoperability down to the metadata record 2. Selective technical data standard	Med to high depending on dataset & scale
2. Ease of Discovery/Searchability	1. Lack of common vocabulary 2. Multiple portals with easy find for internet searches 3. Web page with different “themes” to help drill down for needed data 4. Search by “keywords” 5. Design for user experience is difficult	High
3. Data Access	1. Paying for infrastructure 2. Data volume 3. Common API interfaces as standard 4. Number/amount of people accessing data (infrastructure behind access) 5. Restrictions & sensitivity & patents & security; level 2 product can be accessed but not the raw data	Med
4. Data Synthesis	1. Better interoperability feeds into better synthesis	Very high (due to the variability of what comes into the system and output synthesis)

	2. Multiple data synthesis (human vs environmental; timescale and granularity) 3. Data ambiguity/ biased 4. If automated/computer synthesis 5. Funding 6. Defining synthesis	
5. Data Usability	1. Accuracy, resolution, level of confidence, fitness of uses 2. Sufficient recorder level data 3. Known quality 4. Can't control user usability 5. Versioning	Low
6. Metadata/Data Documentation	1. Lack of common metadata vocabulary (units, time range, scale) 2. Multiple portals with easy find for internet searches 3. Consistency of implementing the standards 4. Training of what metadata is (data description) 5. Maintaining 6. Versioning	Med to low
7. Other: Communication to the user base	1. Data nerds can't do this 2. Common data management strategies (priority of the program) 3. Time	Med

7. Long term repository

High - \$1M+; Med - \$100,000+; Low- less than \$100k; Micro-low-less than \$5k

Challenges:	Costs:
Resources (funding)	High
Changing Technology	Med to High
Changing Standards & Sources	Med to High
Leadership (not organization; recognize importance of data)	Micro
Security	High

8. Definition of success

	Characteristics/ Metrics
Data management system (transactional system; access, discovery portals; queries)	<ul style="list-style-type: none"> Happy customer; data user, funder, data provider

	<ul style="list-style-type: none"> • Reuse/publication of data
<p>Data repository (warehouse; central system; stewardship management; preservation)</p> <p>Combine these 2 into data management program</p>	<ul style="list-style-type: none"> • Happy customer; data user, funder, data provider • Reuse/publication of data • Adequate growth of archives • Usage analytics • Open services • Third party reuse

9. Barriers

	Barriers
Getting generators to submit data	<ul style="list-style-type: none"> • Time, money, ease, willingness to share, training, difficulty to submit, guidance & mandates • Data security embargoes
Finding data	<ul style="list-style-type: none"> • Natural language (from user perspective) • Flexible search engine • Usability of the interface • Create search indexes that allow for quick search • Score data maturity • Not sharing (dark data)
Using data	<ul style="list-style-type: none"> • Lack of documentation • Tools • Good data tools encourage data sharing • Sample design
Enforcing data policy	<ul style="list-style-type: none"> • “No stick” • Labor intensive (enforced by funding) • Too many guidance options; conflicting data policies; staff change over
Data security	<ul style="list-style-type: none"> • Ever evolving • Changes between agencies

Breakout Session II: Interoperability

1. What does interoperability mean to us:

- The ability to choose data and streams
- Pass information to different systems with ease
- Added analysis or synthesized data
- Different datasets
- Blending data across domains and sectors (economic, scientific, etc.)
- Take data and share (communicate, talk to it) in something else. Understanding how data translates to other systems or data. Vocabulary translation. Also types and formats
- System exposes data and works across different machines or types of data
- Aggregate data across different sources. Could come down to nomenclature. Consistent way to pull data together.
- Difficult to get data communities to work with others. Need a backbone or infrastructure to allow for communities to be interoperable.
- Interoperability is only a part of the entire process.

Challenges	Solutions
<i>Listed here from day 1 in priority list</i>	
1. Who is interoperability important to? (the importance of interoperability)	<ul style="list-style-type: none"> • Communications to different users or sectors (example: using climate data to cross over to health sector to answer public health questions) • Providing use cases and success stories for data interoperability • Catalogue of existing frameworks for data interoperability; no need to re-invent the wheel. • Being aware of who is collecting what: data catalogue (scope of data and systems for interoperability)
2. Why is interoperability important?	<ul style="list-style-type: none"> • Reducing duplication of efforts • Creatively using data to answer questions in a variety of sectors • By making data interoperable it can be used in a very powerful and robust way; interoperability supports synthesis, data discovery, access, dissemination, archive, etc. • Showing value through combining different seemingly divergent data streams. • Data itself does not do much; interoperability of systems is required to answer complex questions • Interoperability enhances machine to machine capabilities
3. How does interoperability happen? Who is responsible?	<ul style="list-style-type: none"> • Standards for the entire life of the data stream • Standards for systems; need a known set of functions (example: OGC) • Standards for data collection • Data management plans that would include collection through archive • Government agencies (State and Federal) may be the backbone or foundation for interoperability and standards creation.

	<ul style="list-style-type: none"> • Communication of standards through consortiums • People who are enforcing standards should also be responsible for communicating those standards • Feedback loop between government, academia, NGOs, and industry creates innovation. This creates and refines standards. • Data owners have the responsibility to conform to standards. • Automation of data streams to standards regardless of the collector. Providing tools to support interoperability. • Generally tools to support interoperability need to be open source • Belonging to system of systems (example: IOOS to GEOS or DataOne)
--	---

What are the data management requirements for successful interoperability?
<ul style="list-style-type: none"> • Clear plans that follow standards (data format, metadata, quality, access, etc.) • Enticement to follow and enforcement of standards. • Training • Proper resources • Communication between systems trying to achieve interoperability • A common vision across organizations that translates to priority within individual organizations • Catalogue of their existing frameworks for data interoperability; no need to re-invent the wheel.

What are best practices and guidance for interoperability?
<ul style="list-style-type: none"> • Homogenous standards • Develop standards • Adhere to standards • Refine standards as systems and technology advance • Documentation and preservation of legacy standards or systems that may be important for future users • When new systems are developed they must converge or support existing standards

Breakout Session II: Ease of Discovery/Searchability

Challenges	Solutions
<i>Listed here from day 1 in priority list</i>	
1. What are the characteristics of a good repository- in terms of ease of discovery/searchability (e.g., keyword search, themes).	<ul style="list-style-type: none">• Abundant Keywords with a Common Vocabulary• Semantic Search• Canned Search Reports• Searchable data (columns and values)• Themes (tailored data packaging depending on user) or different versions of repositories• Findable repository
2. How do we ensure metadata quality?	<ul style="list-style-type: none">• Investment in human resources (management working with generators)• Early focus on complete/accurate metadata• Training generators on ensuring metadata quality
3. How do we ensure user needs are met?	<ul style="list-style-type: none">• Know the user (public, academic, etc.)

What are the data management requirements for successful ease of discovery/searchability?
<ul style="list-style-type: none">• Early involvement by data management team and definition of user needs/questions• Editing of metadata to make it searchable

What are the best practices and guidance for ease of discovery/searchability?
<ul style="list-style-type: none">• Federated database (kayak, hotels.com)• Funding mandates• Identification of end user at the beginning of process

Breakout Session II: Data Access

1. What does data access mean?

- There are subtleties data access is more complex not just availability. High domain awareness in finding it not understanding.
- Successful endpoint to data discovery
- User can get to the data that they need
- Data access vs interoperability. We usually want more data across data providers without a lot of extra work.
- Electronic data access vs handwritten letter by carrier pigeon
- Being able to obtain data easily that you need or are looking for
- Variety of ways to access not just downloads (both types and sizes)
- Data has to be in the system; system open to adding data. Interface has to be able to find the data. The presentation has to be usable by the user.
- Data access doesn't always mean data download; view only

Challenges	Solutions
<i>Listed here from day 1 in priority list</i>	
1. How do you deal with data that is restricted/sensitive?	(Definition: this refers to the actual sensitivity of the data itself (human, archeological, embargoed data, dark data, tribal data)) <ul style="list-style-type: none"> • Limit the amount of restricted data • Summarized forms; GIS obfuscation; built into the system (by computer code) using summarized forms • Data owners/generators agree to this process <ul style="list-style-type: none"> ○ Educate data owners that their data if put into the system in a protected mode • Rolled up so "individuals" can't be identified built into the access location • Granularity/credentials/differential access in login access roles
2. How do you deal with data security?	(Definition: broad interpretation to data integrity and system security) <ul style="list-style-type: none"> • Granularity/credentials/differential access in login access roles • Meeting common IT security requirements; different systems have different oversight guidance • Require data centers to be certified • Credentialed login
3. What are your options for the data volume challenge? (i.e., high volume: large data set or many hits on one data set)	<ul style="list-style-type: none"> • Scalable cloud (size of pipeline) <ul style="list-style-type: none"> ○ The challenge is expense ○ Procurement challenges • Subsetting data – design system or queried services to access useable bytes • Low res previews & downloads (summarized query results) • Multiple methods of access (ftp, direct cloud download, cold storage for very large datasets for shipment)

	<ul style="list-style-type: none"> • Leverage already managed by private industry and make our data accessible (google earth engine)
--	---

What are the data management requirements for successful data access?
<ul style="list-style-type: none"> • Common summarization approaches (common unit reporting such as county) • Complete documentation (explain why this is the requirement) on data generator end as well as user end; robust metadata (system, project or record levels) • Bolster public accessibility is priority • Effective user interface (human) and services (machine)

What are best practices and guidance for data access?
<ul style="list-style-type: none"> • The system should ensure data integrity • The system has to be designed for public access as well as credentialed logins (for data generators as well as data users). • Single login across all platforms (i.e., EPA exchange network, university single login). • Ensure identification of data owners/generators and give credit • Sharing data in enticing way. Communicate with story maps (“bites, snacks, and meals”).

Breakout Session II: Data Synthesis

1. Data Synthesis definition:

- Bringing together different data sets to do comparisons and analyses
 - Tied to interoperability
 - Human guided activity
 - Putting analysis in context
 - Multi-disciplinary
 - Using data beyond the original purpose to answer new questions
 - Original data and not derived data products
- Examples:
 - Looking at restored oyster reef and examining the economic and ecosystem services benefits from the restoration.

Challenges	
<i>Listed here from day 1 in priority list</i>	
1. Who is data synthesis important to? How to make data discoverable and accessible for unknown future users?	<ul style="list-style-type: none"> • Interoperability <ul style="list-style-type: none"> ○ Metadata, key wording, standards, crosswalks ○ Facilitating machine discovery ○ Standard archiving formats • Optimizing for web searchability (from a google like search) <ul style="list-style-type: none"> ○ Cover all users we don't know about ○ Marketing the data
2. Why is data synthesis important?	To answer questions such as (DWH NRDA Restoration Example): <ul style="list-style-type: none"> • Has the resource recovered? • Did what I do help the resource recover? • Was there a benefit to the ecosystem as a whole?
How does data synthesis happen? Who is responsible? 3. How does data management facilitate data synthesis?	<ul style="list-style-type: none"> - Through the use of transformation and analytical tools - Effective archive and dissemination - Building the capacity for the researchers to do the work and teach them to use the accepted tools. - Coordinate with synthesis centers <p>Data managers at various levels are responsible (It takes a village)</p>
4. Future Use	Effective discovery of archived data

What are the data management requirements for successful data synthesis?
<ul style="list-style-type: none"> • Direct access to the data • Searchability • Consistent metadata • Knowledge and training on both repositories and their tools

What are best practices and guidance for data synthesis?
<ul style="list-style-type: none">• Talk to experts, ask questions• Defining the question• Knowing the audience

- Talk to experts, ask questions
- Defining the question
- Knowing the audience

Breakout Session II: Data Usability

1. Data usability definition

- Little effort to execute task (e.g., units, can actually open file, use it, turn it into info with little cleanup/prep/transformation)
- Knowing what dataset is – understanding how it was collected, why, biases included, to really understand context of the info, quality of data
- Attributes, characteristics of data, project sufficient to be useful to a user (e.g., metrics that will be useful in designing future restoration projects)
- Making sure others can access the exact same dataset you're using to make interpretations (i.e., "version") – ensure explanations attributed to the data are well defined
- What, when, why, purpose, context, quality, versions, confidence in the data / Info at the right place, right time, right format, known value so can be used for many purposes
- Usability at dataset level – and across datasets (interoperability) – confidence in each dataset to ensure interoperability, etc.
- Can I build off of info to answer my questions / what question are you trying to answer with the data (hopefully built into the dataset as it was developed)
- Understand the usage restrictions – should be clearly defined in the metadata - and possibly overcome restrictions where possible

Challenges	Solutions
<i>Listed here from day 1 in priority list</i>	
1. How do you get generators to define quality, resolution & accuracy of the data?	<ul style="list-style-type: none">• Can be required by funder in contract (require generator to define quality plan in advance).• Funder can even include specific standards, basic requirements, explicit instructions• Need to make it easy for generator – point to existing methods, automated process, templates (and make sure it starts early on)• Funder can require approval of quality plan, metadata standards, etc. – to allow for review, feedback• Remove reasons for not doing it – funding, basic tools, difficulties,• If there are existing repositories for which these requirements (templates, tools, best practices already built in) – and we require/recommend generators to use these repositories• Including disclaimers in templates or models that generators are using / providing info on "known quality"• Generator needs to document methods of data collection (which would provide info on accuracy and resolution) – well defined source• Generator needs to provide "quality plan" that was / was not followed – whether qa/qc was conducted, how it was conducted, status, is data raw

2. How do you convey the quality, accuracy, resolution of the data to the users?	<ul style="list-style-type: none"> • Disclaimer (that includes what data was used for, purpose, why it was created, context) • Need quality of data documented in dataset descriptions, well defined source, context of data collection – all in “metadata” • Repositories – need to document their purpose or focus of the repository, in addition to specifics of the dataset
3. How do you assess and report data quality, accuracy & resolution?	<ul style="list-style-type: none"> • Identify and use authoritative data sources (knowing the right system to find the dataset that meets your needs) • How do you standardize the assessment of quality? How do you identify the “gold standard”? Peer review can help; some agencies have identified tiered data sources; • Can we identify our own quality indicators / can we require data generators to identify the quality level • Identify data user, ensure we engage data users to determine their needs (for resolution, accuracy, etc.) to ensure data generator meets their needs • Communities of practice – learning from data users
4. OTHER	<ul style="list-style-type: none"> • Define key policy questions • Do we have the right metrics, data to answer the bigger picture questions from policy makers – did we make a difference in GOM from all of this restoration post DWH? • Coordination within DWH community to enable common metrics that can be aggregated to answer those key policy questions • Difficulty in attributing increases/benefits to specific restoration projects or programs • Unknown baselines

What are the data management requirements for successful data usability?
<ul style="list-style-type: none"> • Common data descriptors – common knowledge and understanding <ul style="list-style-type: none"> ○ Robust metadata and data descriptions (including mini report card of dataset characteristics and detailed level of metadata) ○ Ensure resources are available to review metadata, etc. ○ Engaging data users early ○ Ensuring acceptance by data user community • Machine readable metadata that can be pulled into data systems • Preview of data • Ease of system and User Testing – funder or data manager should go through the system, use it, fill it out, figure out how long it takes, how time intensive.

What are best practices and guidance for data usability?
<ul style="list-style-type: none"> • Developing communities of practice – follow consensus approach across many communities? How do we bridge gaps between diverse communities with diverse needs? • Data fits to common data model and meets minimum data standards

- Developing summary info, indicator of quality (consumer report) – mini report card, easy way to convey info on dataset (e.g., could have levels of metadata, various levels of detail)
- Interface/tool – between data/metadata and the user (based on user needs)
- Interfaces – should not have just “text” but graphical representations, color coding for various levels of data quality – intuitive ways to communicate quality
- Clearly conveying date
- User training webinars (to help data users) – e.g., CPRA has a training center related to usability of the data and other aspects – applied training session
- Repositories could develop trainings
- Identify baselines

Breakout Session II: Metadata/Data Documentation

1. Scope: NRDA restoration/ restoration Act (primary data of discussion); there will be other end users of this data, and therefore need quality metadata
2. Overall: Greater value given to archived data sets related to the research to be funded

Challenges	Solutions
<i>Listed here from day 1 in priority list</i>	
1. How do you establish and ensure consistency of geospatial metadata standards?	<ul style="list-style-type: none"> • Mandate from funding agency (existing directives for each of these funding agencies that metadata should be collected) • Agreement of what tags are for and their supplemental information, required and understood definitions for those fields • Training on metadata to understand what core fields • Call “data set description” instead of metadata • Templates • Automate metadata creation • Engage communities who are doing work to agree on protocols/standards/procedures that can be drawn from • Get Monitoring and Adaptive Management (MAM) & Trustee Implementation Group (TIGs) • Accountability: written in contract- grant requirements. Must have champions on the ground.
2. How do you implement controlled vocabulary into metadata?	<ul style="list-style-type: none"> • Automation • Templates- limited number of choices • Community-agree upon vocabulary • When applying- here are the tools/templates for data collection • Use community of practice to establish controlled vocabulary (constrain list/check fields) • Easy access to “data documentation” tool creations • Require metadata prior to receiving data • Value: time saving, how often your data is being used, your data is being called from more, •
3. How do you ensure that the metadata are complete enough for uses of that data set in the future?	<ul style="list-style-type: none"> • Implement a review process, both human and automated • Rubric for completeness as part of your metadata tool • Templates and guidance of completeness • Understand what was fundamentally done (meets minimum required standards, to provide maximum value of the data set for the funder)
4. How do you reduce the burden of metadata on generators?	See above

What are the overarching data management system requirements for successful metadata/data documentation?

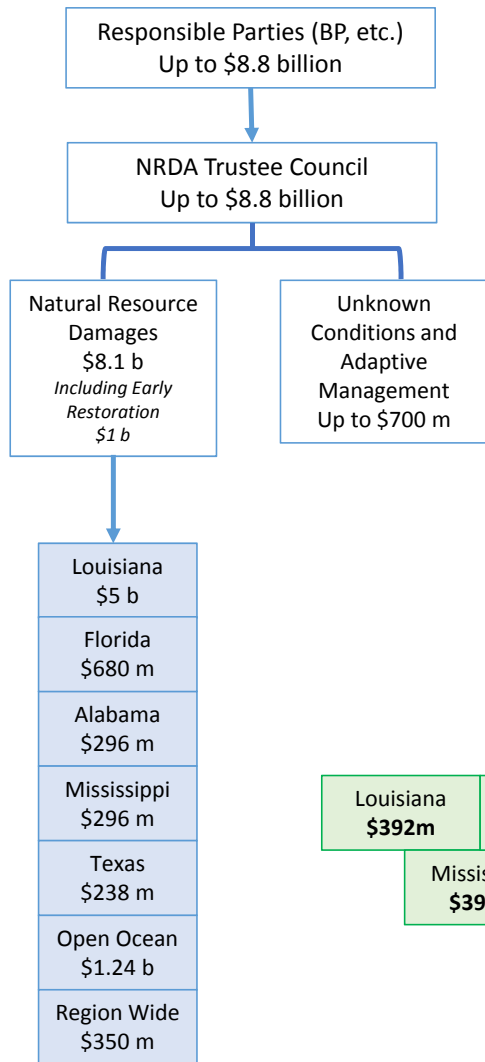
- | |
|---|
| <ul style="list-style-type: none">• Generate complete standard metadata (all stakeholders)• Machine readable• Discoverable data• Connected to source data• Transferable/movable• Version controlled• Adaptively managed |
|---|

What are best practices and guidance for metadata/data documentation?
--

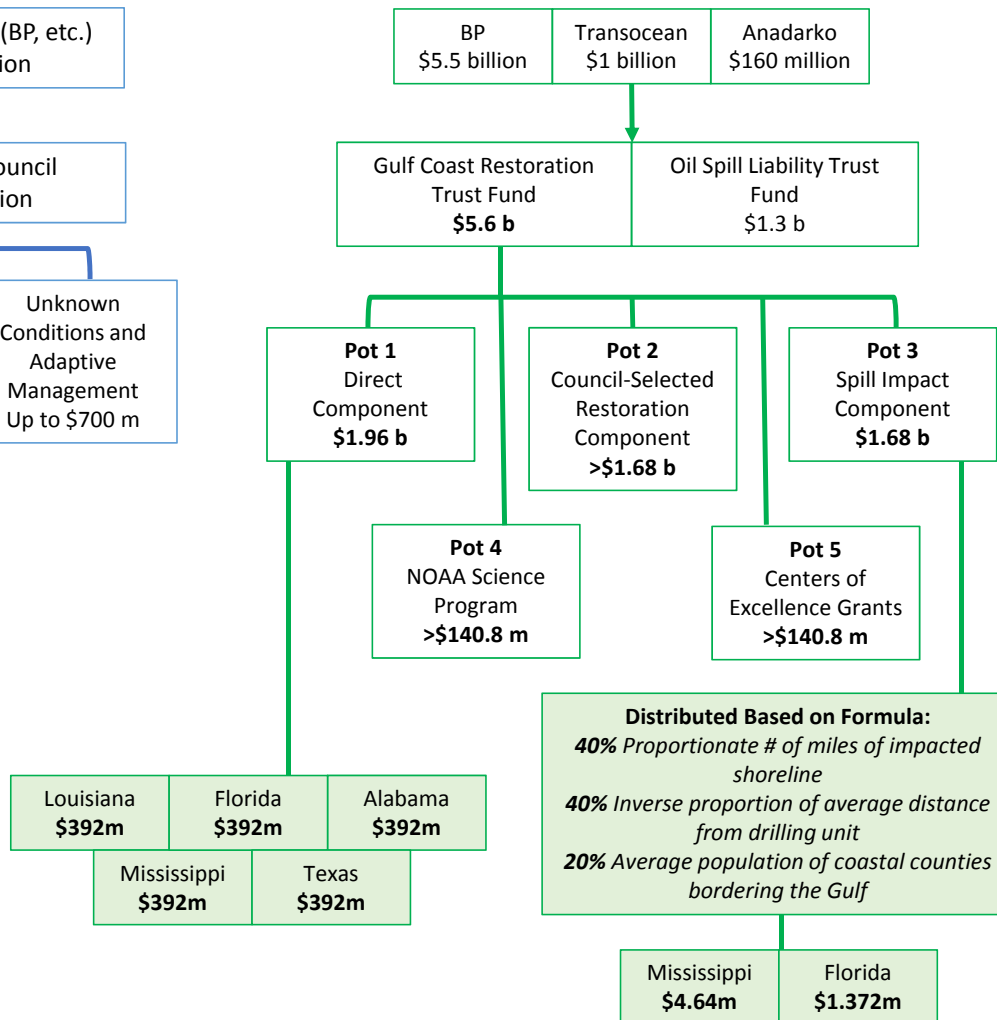
- | |
|--|
| <ul style="list-style-type: none">• Done by communities of practice and stakeholders• Generate as you go• Start metadata as start to collect• Implement existing workflows (e.g., rolling deck to repository) |
|--|

Appendix I: Funding Diagram

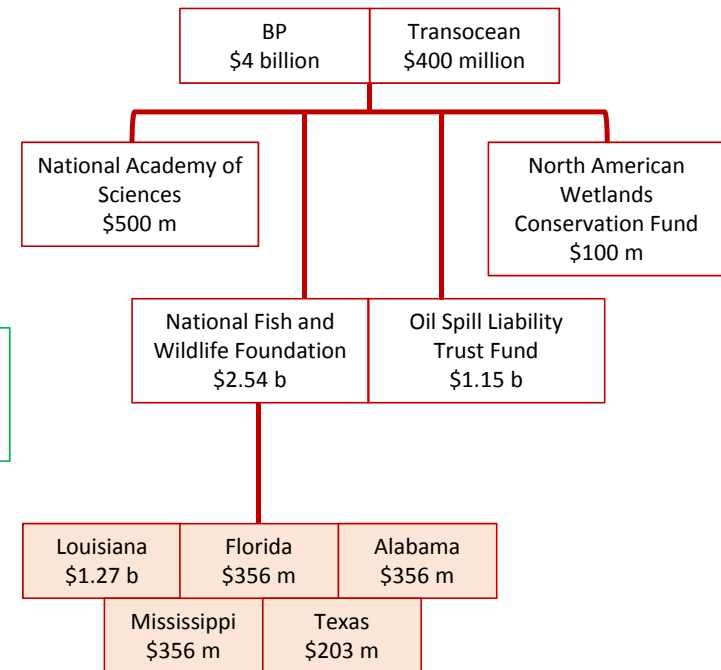
Natural Resource Damages



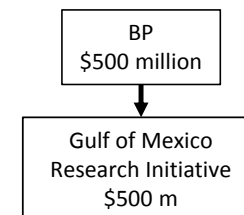
Civil Penalties



Criminal Penalties



Others



Each box links to the slide for the state or organization listed

Louisiana

Florida

Alabama

Mississippi

Texas

National Fish and
Wildlife Foundation
(NFWF)

National Academy
of Sciences (NAS)

North American
Wetlands
Conservation Fund

Oil Spill Liability
Trust Fund

NRDA Trustee
Council

Gulf of Mexico
Research Initiative
(GOMRI)

<http://www.gulfofmexicoalliance.org/learn-more/gulf-restoration/>
<http://eli-ocean.org/gulf/updates/>

Each box links to the slide for the state or organization listed

Gulf Coast Restoration Trust Fund

Pot 1:
Direct Component

Pot 2:
Council Selected
Restoration
Component

Pot 3:
Spill Impact
Component

Pot 4:
NOAA Science
Program

Pot 5:
Centers of
Excellence Grants

Louisiana Restoration Area

NRDA Representation:

- [Louisiana Trustee Implementation Group](#)
 - State Trustees:
 - Coastal Protection and Restoration Authority (Principal)
 - Oil Spill Coordinator's Office
 - Department of Environmental Quality
 - Department of Wildlife and Fisheries
 - Department of Natural Resources
 - Federal Trustees: NOAA, USDA, EPA, DOI
- Regionwide Trustee Implementation Group

State and Local Entities for Gulf Coast Restoration Trust Fund Dollars:

- Coastal Protection and Restoration Authority of Louisiana
 - CPRA is the state entity in Louisiana responsible for designating priorities and to guide development and implementation in order to establish a safe, sustainable coast to protect communities, infrastructure, and natural resources
 - [2017 Coastal Master Plan](#)
- Pot 1: CPRA and 20 Louisiana Parishes with individual Multiyear Implementation Plans (MYP) in various stages of development
- Pot 2: Gulf Coast Ecosystem Restoration Council (Council) Representation: Governor, with CPRA as Designee
- Pot 3: CPRA [Louisiana State Expenditure Plan \(SEP\)](#)
- Pot 5: CPRA establishes Centers of Excellence Research Grants Program

Louisiana

Louisiana Projects		
Program	Projected Funding Amount	Managed By
Natural Resource Damages NRDA Trustee Council	\$5 billion	Louisiana and Federal Trustees
Gulf Coast Restoration Trust Fund: Pot 1 Direct Component	\$392 million (20% of Pot) State: \$274.6 m (70%) 20 Coastal Parishes: \$117.7 m (30%)	Coastal Protection and Restoration Authority (CPRA) and 20 designated Louisiana Parishes/ Treasury
Gulf Coast Restoration Trust Fund: Pot 2 Council-Selected Restoration Component	Determined by Council Currently \$38.3 million	CPRA/ Council
Gulf Coast Restoration Trust Fund Pot 3 Spill Impact Component	\$581.7 million (34.59% of Pot)	CPRA/ Council
Gulf Coast Restoration Trust Fund Pot 5 Centers of Excellence Research Grants Program	\$28 million + interest (20% of Pot)	CPRA/ Treasury
Gulf Environmental Benefit Fund	\$1,272 million	NFWF

Florida Restoration Area

NRDA Representation:

- [Florida Trustee Implementation Group](#)
 - State Trustees:
 - Department of Environmental Protection (Principal)
 - Fish and Wildlife Conservation Commission
 - Federal Trustees: NOAA, USDA, EPA, DOI
- Regionwide Trustee Implementation Group

State and Local Entities for Gulf Coast Restoration Trust Fund Dollars:

- Pot 1: 23 named Florida Counties with individual MYPs in various stages of development
- Pot 2: Council Representation: Governor, with selected Designee
- Pot 3: A consortium of named 23 Florida Counties with a SEP in development
- Pot 5: Florida Institute of Oceanography (FIO)* establishes Centers of Excellence Research Grants Program
 - Supports wildlife ecosystem research and monitoring in the Gulf Coastal Region along with comprehensive observation, monitoring and mapping of the GOM

*FIO is an academic institution that serves as the state entity for Florida

Florida

Florida Projects		
Program	Projected Funding Amount	Managed By
Natural Resource Damages NRDA Trustee Council	\$680 million	Florida and Federal Trustees
Gulf Coast Restoration Trust Fund: Pot 1 Direct Component	\$392 million (20% of Pot) 15 Non-disproportionately Affected Counties: \$98 m (25%) 8 Disproportionately Affected Counties: \$294 m (75%)	Each of the 23 Florida Counties/ Treasury
Gulf Coast Restoration Trust Fund: Pot 2 Council-Selected Restoration Component	Determined by Council	Florida/ Council
Gulf Coast Restoration Trust Fund Pot 3 Spill Impact Component	\$308.7 million (18.36% of Pot)	Consortium of Florida/ Council
Gulf Coast Restoration Trust Fund Pot 5 Centers of Excellence Research Grants Program	\$28 million + interest (20% of Pot)	Florida Institute of Oceanography (FIO) /Treasury
Gulf Environmental Benefit Fund	\$356.16 million	NFWF

Alabama Restoration Area

NRDA Representation:

- [Alabama Trustee Implementation Group](#)
 - State Trustees:
 - Alabama Department of Conservation and Natural Resources (Alabama DCNR) (Principal)
 - Geological Survey of Alabama
 - Federal Trustees: NOAA, USDA, EPA, DOI
- Regionwide Trustee Implementation Group

State and Local Entities for Gulf Coast Restoration Trust Fund Dollars:

- Pot 1: Alabama Gulf Coast Recovery Council (AGCRC) with a MYP in development
 - Support restoration and protection of natural resources, ecosystems, fisheries, marine and wildlife habitats in the GC region, mitigate damages, implement a management plan, etc.
 - [A Roadmap to Resilience \(2011\)](#)
- Pot 2: Council Representation: Governor, with Alabama DCNR as Designee
- Pot 3: AGCRC with a SEP in development
- Pot 5: AGCRC establishes Centers of Excellence Research Grants Program

Alabama

Alabama Projects		
Program	Projected Funding Amount	Managed By
Natural Resource Damages NRDA Trustee Council	\$296 million	Alabama and Federal Trustees
Gulf Coast Restoration Trust Fund: Pot 1 Direct Component	\$392 million (20% of Pot)	Alabama Gulf Coast Recovery Council (AGCRC)/Treasury
Gulf Coast Restoration Trust Fund: Pot 2 Council-Selected Restoration Component	Determined by Council	Alabama DCNR/ Council
Gulf Coast Restoration Trust Fund Pot 3 Spill Impact Component	\$343 million (20.40% of Pot)	AGCRC/ Council
Gulf Coast Restoration Trust Fund Pot 5 Centers of Excellence Research Grants Program	\$28 million + interest (20% of Pot)	AGCRC/ Treasury
Gulf Environmental Benefit Fund	\$356.16 million	NFWF

Mississippi Restoration Area

NRDA Representation:

- [Mississippi Trustee Implementation Group](#)
 - State Trustees:
 - Mississippi Department of Environmental Quality (Principal)
 - Federal Trustees: NOAA, USDA, EPA, DOI
- Regionwide Trustee Implementation Group

State Entity for Gulf Coast Restoration Trust Fund Dollars:

- Mississippi Department of Environmental Quality (MDEQ)
 - Protecting the state's air, land, and water and safeguard the health, safety, and welfare of present and future generations. Eight areas of focus are eco-restoration, economic development, seafood, infrastructure, tourism, workforce development, small businesses, research and education
 - [GoCoast 2020](#)
- Pot 1: MDEQ with a MYP in place
- Pot 2: Council Representation: Governor with MDEQ as Designee
- Pot 3: MDEQ [Mississippi State Expenditure Plan](#)
- Pot 5: MDEQ establishes Centers of Excellence Research Grants Program

Mississippi

Mississippi Projects		
Program	Projected Funding Amount	Managed By
Natural Resource Damages NRDA Trustee Council	\$296 million	Mississippi and Federal Trustees
Gulf Coast Restoration Trust Fund: Pot 1 Direct Component	\$392 million (20% of Pot)	Mississippi Department of Environmental Quality (MDEQ)/Treasury
Gulf Coast Restoration Trust Fund: Pot 2 Council-Selected Restoration Component	Determined by Council	MDEQ/ Council
Gulf Coast Restoration Trust Fund Pot 3 Spill Impact Component	\$320.7 million (19.07%)	MDEQ/ Council
Gulf Coast Restoration Trust Fund Pot 5 Centers of Excellence Research Grants Program	\$28 million + interest (20% of Pot)	MDEQ/ Treasury
Gulf Environmental Benefit Fund	\$356.16 million	NFWF

Texas Restoration Area

NRDA Representation:

- [Texas Trustee Implementation Group](#)
 - State Trustees:
 - Texas Commission on Environmental Quality (Principal)
 - Texas General Land Office
 - Texas Parks and Wildlife Department
 - Federal Trustees: NOAA, USDA, EPA, DOI
- Regionwide Trustee Implementation Group

State Entity for Gulf Coast Restoration Trust Fund Dollars:

- Texas Commission on Environmental Quality (TCEQ): RESTORE the Texas Coast
 - Ecosystem restoration, economic recovery, and tourism promotion in the GC Region
 - [GoCoast 2020](#)
- Pot 1: TCEQ with a MYP in development
- Pot 2: Council Representation: Governor with TCEQ as Designee
- Pot 3: TCEQ with a SEP in development
- Pot 5: TCEQ establishes Centers of Excellence Research Grants Program

Texas

Texas Projects		
Program	Projected Funding Amount	Managed By
Natural Resource Damages NRDA Trustee Council	\$238 million	Texas and Federal Trustees
Gulf Coast Restoration Trust Fund: Pot 1 Direct Component	\$392 million (20% of Pot)	TCEQ/ Treasury
Gulf Coast Restoration Trust Fund: Pot 2 Council-Selected Restoration Component	Determined by Council Currently \$26.3 million	TCEQ/ Council
Gulf Coast Restoration Trust Fund Pot 3 Spill Impact Component	\$127.5 million (7.58% of Pot)	TCEQ/ Council
Gulf Coast Restoration Trust Fund Pot 5 Centers of Excellence Research Grants Program	\$28 million + interest (20% of Pot)	TCEQ/Treasury
Gulf Environmental Benefit Fund	\$203.52 million	NFWF

Direct Component (Pot 1)

About:

The Direct Component is funded by 35% of the DWH Clean Water Act Civil Penalties (estimated to total \$1.86 billion) to be deposited in the Gulf Coast Restoration Trust Fund. It is divided evenly among the five Gulf States. Each state has appointed state and local entities responsible for managing the funds and selecting projects. These projects support restoration and protection to ecosystems, mitigation of damage, monitoring plans, workforce development, and more.



Council-Selected Restoration Component (Pot 2)

About:

Managed by the Gulf Coast Ecosystem Restoration Council and funded by 30% (plus 50% earned interest) of the DWH Clean Water Act Civil Penalties (estimated to total \$1.86 billion) to be deposited in the Gulf Restoration Trust Fund. Approximately \$150-180 million is dedicated to these projects and programs. This component supports ecosystem restoration and protection based on the Council's Comprehensive Plan.

Spill Impact Component (Pot 3)

About:

This component is managed by the Gulf Coast Ecosystem Restoration Council and supported by 30% of the funds of the DWH Clean Water Act Civil Penalties (estimated to total \$1.86 billion) to be deposited in the Gulf Coast Restoration Trust Fund. Unlike the Council-Selected Restoration Component, these funds are invested in projects and programs identified by the five State Expenditure Plans. These plans must be approved by the Council and must align with the Comprehensive Plan objectives. Allocation of these funds to the five states was based on the DWH impact on each state. The distribution is as follows;

Louisiana: 34.59%

Florida: 18.36%

Alabama: 20.40%

Mississippi: 19.07%

Texas: 7.53%

(These percentages are out of the 30% allocated to the spill impact component via the Gulf Coast Restoration Trust Fund)

Centers of Excellence Research Grants Program (Pot 5)

About:

2.5% of the Clean Water Act Civil Penalties (estimated to total \$1.86 billion) to be deposited in the Gulf Coast Restoration Trust Fund and 25% of the interest is dedicated to the Centers of Excellence Research Grants Program. These grants fund science, technology, and monitoring related to Gulf restoration. The money is divided equally among the five [Gulf Coast Region eligible entities](#) designated by the RESTORE Act. They are as follows;

Louisiana: Coastal Protection and Restoration Authority of Louisiana (CPRA)

Florida: Florida Institute of Oceanography (FIO)*

Alabama: Alabama Gulf Coast Recovery Council (AGCRC)

Mississippi: Mississippi Department of Environmental Quality (MDEQ)

Texas: Texas Commission on Environmental Quality (TCEQ)

(FIO is an academic institution that serves as the state entity for Florida)

Gulf Coast Ecosystem Restoration Council

About:

The Gulf Coast Ecosystem Restoration Council was created in 2012 through the RESTORE Act. It is responsible for 60% of the Gulf Coast Restoration Trust Fund deposits, which includes the Council Selected Restoration Component and the Spill Impact Component. Its primary goal is to develop a Comprehensive Plan to “restore the ecosystem and economy of the Gulf Coast Region.”

Members include a designee and the governor of each Gulf Coast state along with representatives from the Department of Agriculture, Department of the Army, Department of Commerce, Environmental Protection Agency, Department of Homeland Security, and Department of the Interior.

<https://www.restorethegulf.gov/about-us>

National Oceanic and Atmospheric Administration (NOAA)

DWH Funding:

- NOAA RESTORE Act Science Program (>\$140.8 million)
- [NRDA Trustee Council](#): NOAA on behalf of the Department of Commerce

NOAA RESTORE Act Science Program

About:

The RESTORE Act Science Program funds research, observation, and monitoring to support the sustainability of the ecosystem, fish stocks, fish habitat in the Gulf of Mexico.

Science Plan:

- Support the science and coordination necessary to understand and manage the GOM ecosystem in order to support:
 - Healthy, diverse, sustainable, and resilient estuarine, coastal and marine habitats and living resources
 - Resilient and adaptive coastal communities

NOAA RESTORE Act Science Program

Funding:

- Gulf Coast Restoration Trust Fund
 - Bucket 4: NOAA Restore Act Science Program (>\$140.8 million)

Gulf Coast Restoration Trust Fund (Pot 4)		
Program	Funding Amount	# Projects Funded
NOAA Restore Act Science Program	\$2,659,200	7

NRDA Trustee Council

About:

The Oil Pollution Act of 1990 authorizes a group of federal agencies, states, and Indian Tribes to serve as Natural Resource Damage Assessment Trustees. These trustees study and evaluate the impacts of oil spills and assist in planning and implementing restoration. The DWH NRDA Trustee Council have improved standard operating procedures for the management of settlement funds. Decision making is usually on a consensus basis.

[Final Programmatic Damage Assessment and Restoration Plan \(PDARP\) and Final Programmatic Environmental Impact Statement \(PEIS\)](#)

Federal Trustees Include: NOAA, USDA, EPA, DOI

Trustee Implementation Groups (TIGs):

- [Louisiana](#): Trustees for Louisiana and Federal Trustees
 - [Louisiana Restoration Plan](#)
- [Florida](#): Trustees for Florida and Federal Trustees
- [Alabama](#): Trustees for Alabama and Federal Trustees
 - [Alabama Restoration Plan](#)
- [Mississippi](#): Trustees for Mississippi and Federal Trustees
 - [Mississippi Restoration Plan](#)
- [Texas](#): Trustees for Texas and Federal Trustees
 - [Texas Restoration Plan](#)
- [Open Ocean](#): Federal Trustees
- [Region Wide](#): All Trustees

Natural Resource Damage Assessment Trustee Council Projects

Program	Funding Amount	Status	# Projects Funded
NRDA Early Restoration Phase 1	\$56,216,508	Funded Some awaiting completion due to long term implementation or ongoing design/ engineering work	8
NRDA Early Restoration Phase 2	\$8,741,490		2
NRDA Early Restoration Phase 3	\$626,836,165		44
NRDA Early Restoration Phase 4	\$133,647,580		10
NRDA Early Restoration Phase 5	\$34,372,184		1

National Fish and Wildlife Foundation (NFWF)

About:

- Conduct or fund **projects**
- Remedy harm, or reduce or eliminate risk of future harm, to **Gulf Coast natural resources**
- Where there has been injury to, destruction of, loss of, or loss of use or resources **resulting from the Macondo Oil Spill**
- **Consult** with state resource agencies, USFWS, and NOAA
- **Maximize environmental benefits**

Priorities:

- Priority 1: Restore and Conserve Coastal Habitat
 - Coastal marsh
 - Barrier islands & beach/dune habitat
 - Coastal Bays and Estuaries
- Priority 2: Enhance Populations of Priority Living Coastal and Marine Resources
 - Oysters, Gulf Coast Birds, Red Snapper & Reef Fish, Sea Turtles, Marine Mammals

National Fish and Wildlife Foundation (NFWF)

Funding

- Criminal Penalties (\$2.54 billion)
- Funds are distributed by state

National Fish and Wildlife DWH Projects			
Program	Funding Amount	Purpose	# Projects Funded
Gulf Environmental Benefit Fund 2013-2018	\$1.272 billion	Barrier island and river diversion projects in Louisiana	101
	\$356 million	Natural resource projects in AL, FL, MS	
	\$203 million	Natural resource projects in Texas	

National Fish and Wildlife Foundation (NFWF)

	Payment (in millions of dollars)	Louisiana	Alabama	Florida	Mississippi	Texas
Apr. 2013	\$158.00	\$79.00	\$22.12	\$22.12	\$22.12	\$12.64
Feb. 2014	353.00	176.50	49.42	49.42	49.42	28.24
Feb. 2015	339.00	169.50	47.46	47.46	47.46	27.12
Feb. 2016	300.00	150.00	42.00	42.00	42.00	24.00
Feb. 2017	500.00	250.00	70.00	70.00	70.00	40.00
Feb. 2018	894.00	447.00	125.16	125.16	125.16	71.52
Totals	\$2,544.00	\$1,272.00	\$356.16	\$356.16	\$356.16	\$203.52

BP = \$2,394M; Transocean = \$150M

National Academy of Science (NAS): Gulf Research Program

About:

The National Academy of Science Gulf Research Program strives to improve oil system safety and protect communities and the environment by better understanding the interconnectivity of the region and its resources.

Strategic Vision:

1. Foster improvements to safety technologies and culture, and environmental protection systems associated with offshore oil and gas development
2. Improve understanding of the connections between human health and the environment to support the development of healthy and resilient Gulf communities
3. Advance Understanding of the GOM region as a dynamic system with complex, interconnecting human and environmental systems, functions, and processes to inform the protection and restoration of ecosystem services

National Academies of Science (NAS): Gulf Research Program

Funding:

- Criminal Penalties (\$500 million)
- Program is managed by [Advisory Board](#)

National Academies of Science Projects			HOME
Program	Funding Amount	Purpose	# Projects Awarded DWH Funds
Research and Development Grants 2017	TBD Fall 2017	Support research to understand risk leading to the release of oil and gas	TBD Fall 2017
Research-Practice Grants 2017	Up to \$10 million	Funds projects related to resilience to climate change and disasters	(3-6)
Capacity Building Grants 2016		Support enhancement of community networks that improve coastal environments	
Synthesis Grants 2016	\$2,120,000	Supports projects that use scientific synthesis to understand impacts of offshore oil and gas operations	3
Exploratory Grants 2015	\$4,571,000	Increase training for offshore oil and health professionals	21
Exploratory Grants 2016		Support innovative work on scenario planning to improve safety	
Data Synthesis Grants 2015	\$4,416,000	Grants for activities that synthesize existing GOM data	9

North American Wetlands Conservation Fund

(US Fish and Wildlife Service)

About:

The North American Wetlands Conservation Act (NAWCA) grants serve to increase bird populations and wetland habitat in addition to improving local economies and hunting, fishing, farming, etc.

Fund Mission:

- Protect, restore, or enhance wetland and associated habitats throughout the country
- Promote long-term protection of habitats for birds and other wetland-dependent species
- Catalyze conservation partnerships with federal, state, non-profit, and private organizations
- Support conservation of priority migratory bird species in the US

North American Wetlands Conservation Fund

(US Fish and Wildlife Service)

DWH Funding:

- Criminal Penalties (\$100 million)
- Managed by the North American Wetlands Conservation Council

Programs		
DWH Funding Amount	Leveraged Funds	# Projects Awarded DWH Funds
\$46,494,053	\$75,882,741	52

Oil Spill Liability Trust Fund

About:

Established under the Oil Pollution Act of 1990, this fund is meant to cover the cost of spill related damages or removal when the responsible party is unknown or noncompliant.

The Trust Fund is supplied by a five-cents per barrel fee on oil (ended in 1994), interest on existing funds, cost recovery from spills, and fines or civil penalties collected from responsible parties.

Oil Spill Liability Trust Fund

DWH Funding:

- Civil Penalties (\$1.3 billion)
- Criminal Penalties (\$1.15 billion)
- Funds are administered by US Coast Guard National Pollution Fund Center

Programs	
Funding Amount	Purpose
Fund can provide up to \$1 billion for any one oil pollution incident, including up to \$500 million for natural resource damage assessments	State access for removal actions Payment to trustees to carry out NRDA restorations Payment of claims for removal and damages Research and development

Gulf of Mexico Research Initiative (GoMRI)

Funding:

- BP Funded Independent Research Program (up to \$500 million)
- Funding decisions are made by a [Research Board](#) comprised of 20 marine scientists, education, and public health experts
- [GoMRI Administrative Unit](#) is an internal department of [the Gulf of Mexico Alliance \(GoMA\)](#)
- GoMRI shares data with [Gulf of Mexico Research Initiative Information and Data Cooperative \(GRIIDC\)](#)

Research Themes:

1. Physical distribution, dispersion, and dilution of petroleum and its constituents and associated contaminants, under the action of physical oceanographic processes, air sea interaction, and tropical storms
2. Chemical evolution and biological degradation of petroleum/dispersant systems and their interaction with coastal, open-ocean, and deep-water ecosystems
3. Environmental effects on the petroleum/dispersant system on the sea floor, water column, coastal waters, beach sediments, wetlands, marshes, and organisms; and the science of ecosystem recovery
4. Technology developments for improved response, mitigation, detection, characterization, and remediation associated with oil spills and gas releases
5. Impact of oil spills on public health including behavioral, socioeconomic, environmental risk assessment, community capacity, and other population health considerations and issues

GoMRI Programs			HOME
Program	Funding Amount	Purpose	# Projects Awarded Funds
Year-One Block Grants 2010	\$45 million	Provided by BP to determine baseline data	158
RFP I – Consortia Grant 2011	\$110 million	Awarded to 8 research consortia in 27 US States and 5 countries	8
RFP II – Investigator Grants 2012	\$18.5 million	Awarded to 19 efforts involving a PI and up to 3 co-PI's for 3 institutions	19
RFP III – Bridge Grants 2011	\$1.5 million	Awarded to 17 projects supporting observations and sampling	17
RFP IV – Consortia Grants 2015	\$140 million	Awarded to 12 research consortia	12
RFP V – Investigator Grants 2016	\$38 million	Awarded to individuals and teams studying the effects of oil in GOM	22