


NOAA | Office of Response and Restoration | Assessment and Restoration Division 


DIVER Data Warehouse and Query Tools: Samples & Contaminant Chemistry Focus

Environmental Disasters Data Workshop

September 16-17

Ben Shorr
NOAA's Office of Response and Restoration


1

NOAA | Office of Response and Restoration | Assessment and Restoration Division 

Overview

- Discuss “Data Warehouse” approach**
Actual framework and processes; flexible and scalable
- Common Data Models**
Overview of data models and standards; focus on samples and chemistry and related information/data
- Data Query and Delivery**
Requirements that drive development of data discovery, query, reporting and export tools

9/24/2014 2

NOAA | Office of Response and Restoration | Assessment and Restoration Division 

Data Warehouse and Business Intelligence

Data Warehousing concept and reality

- **Default** to existing tools and processes; databases and data sources with faults and inefficiencies
- The earlier field collected and lab processed data streams are integrated, the better connections and management. **UP FRONT EFFORT** pays big dividends
- **Data Warehouse and Data Vaulting*** concepts
 - Ideally combine *data* beyond high level metadata
- **Business (Environmental) Intelligence**

9/24/2014 3

NOAA | Office of Response and Restoration | Assessment and Restoration Division 

Data Warehouse and Business Intelligence

- Use Industry standard tools
 - Collect and manage structured and unstructured information
- DWH Damage Assessment managing data with an Agile development approach
 - Evolve to meet data and development needs
 - Frequent brief video conference enhances accountability; minimizes silos; creates “team”

9/24/2014 4

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Data Warehouse and Business Intelligence

- Common Data Models – flexible and scalable
 - Core fields across datasets
 - Collect all digital possible (structured and unstructured) with key connections and hierarchy
 - ETL (Extract, Transform, Load)

9/24/2014 5

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Data Warehouse and Business Intelligence

Collate Source Data → Apply BI / ETL Methods → DIVER Data Warehouse → DIVER EXPLORER

Data Integration

Visualization, Exploration, and Reporting

DIVER's Common Data Model

Other Databases/Warehouse/Portals


Visualization

Steps include:

1. Define the common model
2. Accommodate additional data
3. STANDARDIZE **

Our Approach: Promoting Common Data Models


9/24/2014 6

NOAA | Office of Response and Restoration | Assessment and Restoration Division 

Common Data Model Examples (schemas)

- **Samples:** chemistry, biological+
- **Oceanographic:** cruise-collected sensor data
- **Observations:** shoreline, marsh, birds and mammals
- **Telemetry:** whales, dolphins, turtles, tuna
- **Photography:** keywords
- **Restoration data:** potential and implemented; budget and activities

9/24/2014 7

NOAA | Office of Response and Restoration | Assessment and Restoration Division 

Common Data Model Examples (schemas)

- **Core Fields**
Higher level across data models e.g. Analysis Type, Data Source, Status, Spatial
- **Data Specific**
Results, Methodology, Units
- **Related Information**
raw data, field information, source data packages, unstructured documents (reports, graphs, charts etc...)

9/24/2014 8

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Common Data Model: Contaminant Chemistry

Samples

- Used existing data standard and data processing (Query Manager) for contaminant chemistry
- Electronic Data Deliverables (lab templates)
- Work with data providers (owners)
- Use existing standards and nomenclature; expand and standardize when necessary
- Metadata, metadata, metadata

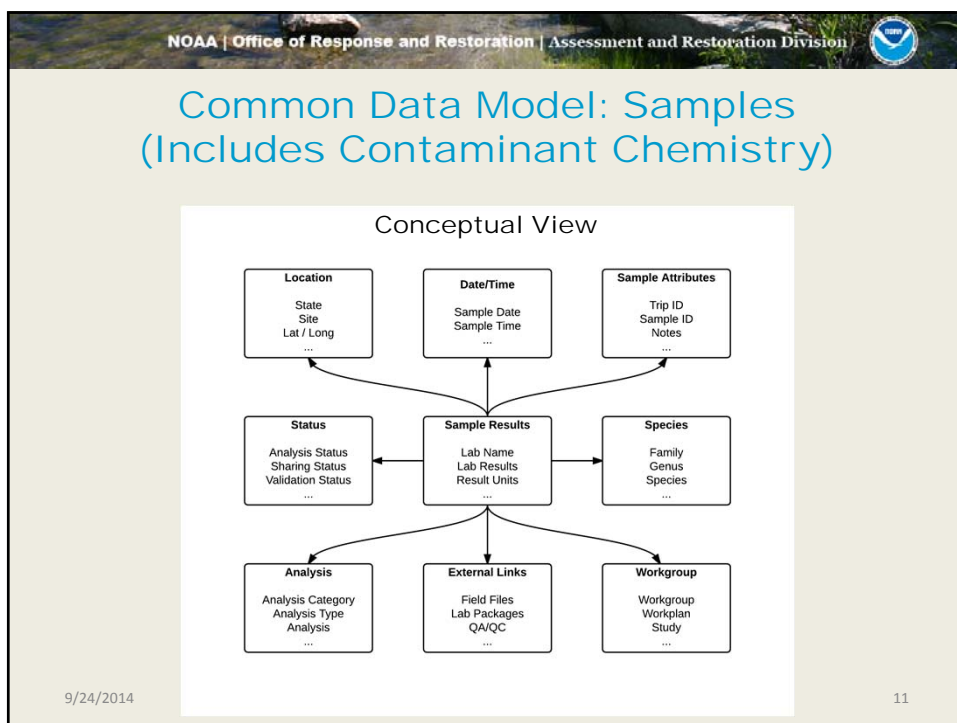
9/24/2014 9

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Common Data Model: Contaminant Chemistry

- Ingest into Data Warehouse
 - Contaminant chemistry source databases include:
 - **Historical Contaminant Chemistry** (Query Manager)
 - **DWH Response** collected (EPA ETL → NOAA QA/QC)
 - **BP NRDA** provided
 - Audit source data and queries
 - Integrate with other data streams (e.g. additional field collection information, related field and lab documents and raw packages, “value added” analysis like oil source fingerprinting)

9/24/2014 10



NOAA | Office of Response and Restoration | Assessment and Restoration Division

Common Data Model: Samples (Includes Contaminant Chemistry)


Sample Fields			
Age (years)	Crude Log	Latin Name: Class	QM Reporting Standard
Analysis	Crude Name	Latin Name: Family	QM Sample Details
Analysis Category	Data Category	Latin Name: Genus	QM Sample ID
Analysis Detail	Data Classification	Latin Name: Kingdom	QM Site ID
Analysis Method	Data Source	Latin Name: Order	QM Station ID
Analysis Result	Date	Latin Name: Phylum	QM Study ID
Analysis Result Unit	Day or Night Sample	Latin Name: Species	Qualifier Code
Analysis Status	Depth Category	Latin Name: Subphylum	Record ID
Analysis Type	Detection Extent	Length (cm)	Reporting Limit
Analyzer Detection Details	Detection Limit	Link to Lab Data Files	Result
Case/Activity	DV/ER Dataset	Location, Geom	Result (0.5 DL)
ChemCode	DV Qualifier Reason	LOSDMS Workplan ID	Result (0 DL)
Client ID	DV Qualifier Reason Code	Measurement Basis	Result (1 * 0.5 DL)
COC ID	DV Qual Reason	MeasurementLevel	Result (1 * Full ASA)
Collection Form	DV Qual Reason Code	Minutes of Hour	Result (1 * Full DL)
Collection Matrix	End Latitude	Month	Result (Full DL)
Collection Method	End Longitude	Month (Numerical)	Result Notes
Collection Study Name	File Collection ID	Month Short	Result Type
Collection Workplan	Fingerprint Class	NRSA Grid	Review Status
Common Name: Class	Fingerprint Class Source	Number Below Detection Limit	Sample Delivery Group
Common Name: Family	Guilford Workplan Name	Number in Composite	Sample Depth Unit
Common Name: Genus	Habitat Type	Number Measured	Sample ID
Common Name: Kingdom	Hour of day	Oil Presence	Sample Lower Depth
Common Name: Order	Image ID	Oil Presence Screening Class	Sample Notes
Common Name: Phylum	Lab ID	Oil Presence Screening Source	Sample Size
Common Name: Species	Lab Name	Percent Lipid	Sample Size Units
Common Name: Subphylum	Lab Replicate	Photo URL - Midsize	Sample Type
Composite/Part Sample	Lab/Result Matrix	Photo URL - Original	Sample Upper Depth
Composite Sample ID	Lab/Result Matrix Detailed	Photo URL - Thumbnail	Sex
	Last Update Date	QM Matrix	Sharing Status
			Source Type
			Species
			Start Latitude
			Start Longitude
			Station/Date
			Survey Notes
			Tissue Code
			Tissue Type
			Total Organic Carbon (pct)
			Trip End Date
			Trip ID
			Trip Start Date
			Validation Level
			Weight (g)
			Workgroup
			Year

...Just a few favorite fields

9/24/2014 12

NOAA | Office of Response and Restoration | Assessment and Restoration Division

DIVER (Data Integration Visualization Exporting and Reporting): Explorer Tool



- Objectives & Requirements
 - Flexible query and export of all data including NRDA collected and external datasets
 - Documented lineage and connections to data holdings
 - Metadata, Metadata, Metadata
 - Export for analysis, visualization and processing

13

NOAA | Office of Response and Restoration | Assessment and Restoration Division

DIVER (Data Integration Visualization Exporting and Reporting): Explorer Tool

- Queries: Guided, Custom & Saved
- Download Data Packages
- Map & Legend
- Charts
- Data Tables
- Photos
- Metadata
- Study Notes
- Export



NOAA | Office of Response and Restoration | Assessment and Restoration Division

Guided Queries

Select a Guided Query

Expand All | Collapse All

Contaminant Chemistry Results > Single Contaminant Results > Sediment

Edit Query Filters

Guided Query: Contaminant Chemistry Results > Single Contaminant Results > Sediment

Show Query Summary

Field List: Search by field or source name

Values List: Enter text in the field to quickly find values

Selected Values (1): Click on a value to remove from the query

1-Methylbenzothiophene (Sampled)

1-Methylnaphthalene (Sampled)

1-Methylphenanthrene (Sampled)

2,3,5-Trimethylnaphthalene (Sampled)

2,3-dihydrodione (Sampled)

2,6-Dimethylnaphthalene (Sampled)

2-Methylanthracene (Sampled)

2-Methylnaphthalene (Sampled)

2-Methylphenanthrene (Sampled)

Back Date Range : No range selected Depth Range : No range selected Cancel Run Query

15

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Query Filters

Choose from Field List

Choose value(s)

Add Filters

Specify Date (or Depth Range)

Selected Value(s)

Edit Query Filters

Guided Query: Contaminant Chemistry Results > Single Contaminant Results > Sediment

Show Query Summary

Field List: Search by field or source name

Values List: Enter text in the field to quickly find values

Selected Values (1): Click on a value to remove from the query

1-Methylbenzothiophene (Sampled)

1-Methylnaphthalene (Sampled)

1-Methylphenanthrene (Sampled)

2,3,5-Trimethylnaphthalene (Sampled)

2,3-dihydrodione (Sampled)

2,6-Dimethylnaphthalene (Sampled)

2-Methylanthracene (Sampled)

2-Methylnaphthalene (Sampled)

2-Methylphenanthrene (Sampled)

Back Date Range : No range selected Depth Range : No range selected Cancel Run Query

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Query Shapes

Draw and Edit Shape Load/Manage Shapes

Units: Miles Geometry: Buffer - Geodesic Area: 60,867.10 sq mi Radius: 139.19 mi Update Cancel Run Query

[illegible]

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Query Results (Table)

Requirements and Functionality:

- Present tabular results
- Integrated with map
 - Selected row highlighted in map
 - Select in map creates filtered table
- Link to source data files, related data and information (e.g. documents, photographs)

19

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Query Results (Table)

Query Name: Contaminant Chemistry Results - Single
Owner: Global Query
Last Modified: 28/10/2014

Description: For sediment samples, create a table with one row for each sample. Data available, method, and other sample specific information are available.

Query Filters:
Analysis Category = "Contaminant Chemistry"
QM Matrix = "Sediment"
Matrix Type = "Residue"
Depth Category = "Surface Sediment"
Sampling Station = "Surface Sampling Station"
Analysis = "Sediment Chemistry"

Display Results for: Concentration

Legend: Concentration - Benzo(a)pyrene (PPB)
 ○ Below Detection Limit (DL = 1.22)
 ○ Below Detection Limit (1.22 < DL = 9.33)
 ○ Below Detection Limit (DL = 9.33)
 ● Concentration < 1.22
 ● 1.22 < Concentration < 9.33
 ● Concentration > 9.33

#	Starting Station	Workgroup	Collection Workgroup	Collection Study Name	QM Study ID	QM Sample ID
1	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
2	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
3	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
4	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
5	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
6	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
7	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
8	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
9	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
10	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
11	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
12	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
13	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
14	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
15	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
16	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001
17	Publicly Available	Chemicals	Research CR	Chemical Research CR Sampling 01	CR	0001

5,882 of 5,882 rows displayed

20

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Charts

Requirements and Functionality:

- Provide overview summary of query results
- Interactive - click on charts to show filtered data
- Flexible - built to handle new information

21

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Charts

Query Name: Contaminant Chemistry Results - Single
Owner: Default Query
Last Modified: 09/17/2014

Display Results By: Concentration

Records by Collection Media

Collection Media	Percentage
Water	67.53%
Sediment	30.82%
Soil	0.23%
Other	0.07%
Unknown	0.05%

Records by Workgroup

Workgroup	Percentage
Emergency Response	44.03%
Other	3.24%
Other	3.24%
Other	3.24%
Other	3.24%
Other	3.24%
Other	3.24%
Other	3.24%
Other	3.24%
Other	3.24%

Records by Concentration

Concentration	Percentage
Concentration < 0.33	0.46%
Concentration > 0.33	9.33%
Concentration > 0.33	9.33%
Concentration > 0.33	9.33%
Concentration > 0.33	9.33%

22

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Export

Requirements and Functionality:

- Spreadsheet and GIS formats (CSV, Shapefile, KML)
- Include metadata and related study notes (contaminant chemistry)
- Export results
- Export related data (additional fields and collection forms)

23

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Display Results By: QM Matrix

Legend

QM Matrix

ERMA Environmental Response Management Application
Deepwater Horizon MC252

Information Help Recent Data Admin Upload Incident Find

Layers Legend Query Tools AOI Labels Zoom Download Print

NRDA Workgroup Data

Sharing Status	Workgroup	Collection Workplan	Collection Study Name	QM Station ID	QM Sample ID	QM Matrix	Date	VER)
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 03 2010	ALAK46249	S001	Sediment	7/21/2010	VER)
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 03 2010	ALAK49333	S001	Sediment	8/4/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 03 2010	MSAK41053	S001	Sediment	7/1/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 03 2010	MSAK42177	S001	Sediment	7/1/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	ALAK47189	S001	Sediment	7/29/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	ALAK47197	S001	Sediment	7/23/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	ALAK47200	S001	Sediment	8/1/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	ALAK48184	S001	Sediment	8/1/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	LAAL40011	S001	Sediment	8/9/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	LAAL40035	S002	Sediment	8/15/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	LAAG39185	S001	Sediment	6/18/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	LAAR37091	S001	Sediment	8/24/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	LAAR37092	S001	Sediment	8/24/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	LAAR38145	S001	Sediment	8/20/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	LAAR42047	S001	Sediment	7/2/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	LAAR42048	S001	Sediment	7/2/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	MSAJ43039	S001	Sediment	7/7/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	MSAJ44070	S001	Sediment	7/27/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	MSAJ44074	S001	Sediment	7/7/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	MSAJ44075	S001	Sediment	7/8/2010	
Publicly Available	Chemistry	Forensic Oil	Chem--Forensic Oil Sampling 04 2010	MSAK44087	S001	Sediment	7/27/2010	

100 km 50 mi

Scale: 1: 3M Zoom Level: 7 Location: 27.04611°, -86.71761°

US DOC | NOAA | NOS | NOAA Office of Response & Restoration
Disclaimer | Privacy policy | Official Citation | Email comments

Coastal Response Research Center
©2007 - 2014 University of New Hampshire

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Metadata

Requirements and Functionality:

- “Lite” version with key information (HTML)
- FGDC compliant metadata (XML and HTML)
 - moving to ISO 19115
- Query Details: fields and values chosen
- Data Details: when were datasets updated?
- Data Caveats: notes about data
- Field Definitions

25

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Summary Table Charts **Metadata** Study Notes Export

Show Section Download

Query Details

Analysis Category = "Contaminant Chemistry"
 QM Matrix = "Sediment"
 Result Type = "Result"
 Depth Category = "Surface Sediment"
 Analysis Status = "Results Available"
 QM Reporting Standard = "Standard Re"
 Source Type = "DWH Trustee NRDA"
 Analysis Type = "PAH"
 Analysis = "Benzo(a)pyrene"

Data Details

DIVER Dataset

Samples - DWH NOAA NRDA

Chemical Definitions

Chemical Code	Chemical
BAP	Benzo(a)py

Data Caveats

Category	Descripti

Metadata "Lite"

FGDC Metadata

Metadata:

- Identification Information
- Data Quality Information
- Spatial Reference Information
- Temporal and Accuracy Information
- Distribution Information
- Metadata Reference Information

Identification Information:

Citation:

Creation Information:

Originator: NOAA's Office of Response and Restoration
 Publication Date: 20100523
 Title: DIVER Engine Data Export 20100523
 Publication Information:
 Publication Place: Seattle, WA
 Publisher: NOAA Office of Response and Restoration
 Online_Linkage: www.noaa.gov
 Online_Linkage: www.noaa.orl
 Online_Linkage: www.gulfipollution.noaa.gov

Description:

Abstract:
 The datasets contained within the DIVER Data Workspaces represent DIVER, Response, and historical data gathered in support of the Natural Resource Damage Assessment (NRDA) process. These datasets are subject to revision based on violations and QA/QC.

Purpose:
 These datasets represent information gathered to support the Natural Resource Damage Assessment (NRDA) for the Deepwater Horizon Oil Spill in the Gulf of Mexico. NOAA and Supplemental Information:
 For more information and documentation related to the Deepwater Horizon oil spill event and NRDA, see NOAA DIVER: www.gulfipollution.noaa.gov.

Time Period of Content:

Time Period Information:

Range of Dates Times:
 Beginning Date: 20100429
 Ending Date: 20120116

Currentness Reference: Publication date

DIVER is intended to provide users with access to the available data as well as information on the status of all samples and

NOAA | Office of Response and Restoration | Assessment and Restoration Division

Summary Table Charts Metadata **Study Notes** Export

Show Section Download

Atlantis Cruise Dec 4-15 2010 [AJ]

DATA SOURCE

Data were compiled from surveys conducted in the Gulf of Mexico. Data were compiled from NewFields Environmental Forensics Practice, LLC (Alpha) lab electronic data. The following SDGs (QC Batches) have been incorporated into the database: 1012123, 1012125, 1012132, 1012134, 1012135, 1012137, 1012138, 1012140, 1108135, 1108136. The data sets were for samples collected from Atlantis Cruise Dec 4-15 2010.

DATA COLLECTION PURPOSE

Natural Resource Damage Assessment

DATA USE QUALIFICATION

Values for concentration and detection limit should be interpreted to 3 significant figures. Values for reporting limits should be interpreted to 1 significant figure.

STUDY

This study includes the following data: Water chemistry (SLURP filter samples and one water [floculant] sample); Sediment surface and subsurface chemistry (cores); and Tissue chemistry. There was also one 'slurped fluff' sample reported in smptar/chemtar tables.

STATION

StationIDs are based on the Station IDs recorded in the NOAA Field Sampling Information database.

SAMPLES AND REPLICATES


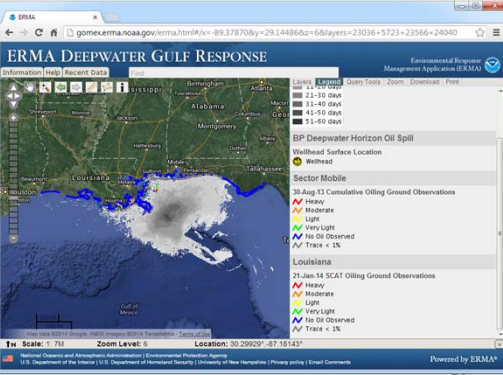
The collection depth of water samples in the fields UDepth and LDepth are reported in meters. The collection depth of sediment samples in the fields UDepth and LDepth are reported in centimeters as measured from the sediment water interface. The original SampleIDs reported by the lab from the Chain-of-Custody is stored in the ExSampleID field. Core samples were given a SampleID starting with "S" for sediment, followed by the Core letter and then a depth identifier. For example, S006-2 was a core sample collected from the second depth interval of Core 6. SLURP samples are coded with a prefix of "SL" on the SampleID. Samples analyzed as filter (Matrix = RS) or filter/liquid (Matrix = PT) have a suffix of "P." Water samples run through multiple filters have the filter letter as part of the sampleID. For example, SampleID SL002P-C is the second Slurp sample run through filter "C" and analyzed as a particulate (filter). Slurp samples were reported at the water depth collected. Two Slurp samples (GU2888-A1208-OE301A, and 1B) were reported as oil were analyzed both as solid and as a filter sample; these were stored in smptar/chemtar. One sample (ExSampleID GU2888-A1209-OE301) was defined as top water collected on push core 3, and thus defined as a flocculant sample. Flocculant (floc) samples were given the prefix "E." Floc samples analyzed as total water have a "W" at the end of the


Study Notes

NOAA | Office of Response and Restoration | Assessment and Restoration Division

www.gulfspillrestoration.noaa.gov & ERMA Gulf Response

Validated public NRDA data available at these websites:

NOAA | Office of Response and Restoration | Assessment and Restoration Division 


Next Steps & Challenges

- More widely available DIVER tools
 - Gulf of Mexico and Great Lakes
- Enhanced data search functionality
- Create flexible & scalable national approach
 - Ability to ingest digital field collected data and unstructured information

Internet Security

Discussion? Happy Hour?

29

NOAA | Office of Response and Restoration | Assessment and Restoration Division 

Acknowledgements

DIVER Data Management Team: Mike Jackson, Dan Hudgens, Jim Anderton, Ann Jones, Amy Merten, Ben Shorr, Kevin Kirsch
NOAA: Amy Merten, Kevin Kirsch, Jay Coady
IEc: Kate Doiron, Ann Jones, Jess Fydenkevez, Lena Flannery, Neal Etre, Amy Anderton
Sirius Solutions: Vincent Luzzo, Nicole Williams, Brian Thompson