



NOAA

Integrated Ocean Observing System (IOOS) Program Office

Data Integration Framework (DIF)

Functional Requirements Document

Version 1.0

November 19, 2007

Revisions

Version	Description of Version	Date Completed
Draft 0.1	Initial Draft released to IOOS Program Staff	09/17/2007
Draft 0.2	Second Draft released to IPT after incorporating initial IOOS project staff comments.	10/16/2007
Draft 0.3	Minor terminology changes from IOOS office staff	10/18/2007
1.0	Official release after comments from Functional Requirements working group of IPT	11/19/07
1.1		
1.2		
1.3		
1.4		
1.5		

Review & Approval

Review History

Reviewer	Version Reviewed	Signature	Date
Christopher Moore/IOOS			

Contents

Revisions.....	i
Review & Approval.....	ii
Contents.....	iii
1 Introduction.....	5
1.1 BACKGROUND AND OBJECTIVES.....	5
1.2 PURPOSE.....	6
1.3 DOCUMENT ORGANIZATION.....	6
2 Scope of the Data Integration Framework.....	8
3 Applicable Documents.....	10
4 IOOS DIF Overview.....	11
5 DIF Functional Requirements.....	13
5.1 PHASE I FUNCTIONAL REQUIREMENTS.....	13
5.1.1 <i>General Functional Requirements</i>	13
5.1.2 <i>Required Data Sets</i>	13
5.1.3 <i>Metadata Management</i>	14
5.1.4 <i>Quality Control</i>	15
5.1.5 <i>Data Transport and Access</i>	15
5.2 PHASE II FUNCTIONAL REQUIREMENTS.....	16
5.2.1 <i>General DIF Functional Requirements</i>	16
5.2.2 <i>Required Data Sets</i>	17
5.2.3 <i>Pre-Processing and Validation</i>	19
5.2.4 <i>Transformation and Assembly</i>	19
5.2.5 <i>Metadata Management</i>	20
5.2.6 <i>Quality Control</i>	21
5.2.7 <i>Data Storage</i>	22
5.2.8 <i>Data Transport and Access</i>	22
5.2.9 <i>IT Security</i>	24
5.2.10 <i>System Administration</i>	25
5.3 PHASE III FUNCTIONAL REQUIREMENTS.....	26
Appendix A Out-of-Scope Requirements.....	30

Appendix B Glossary and Definitions.....**Error! Bookmark not defined.**

1 Introduction

1.1 Background and Objectives

The National Ocean Service's (NOS) Integrated Ocean Observing System (IOOS) Program Office has been charged with developing a Data Integration Framework (DIF) for NOAA's Integrated Ocean Observing System.

The NOAA IOOS DIF project was proposed because currently there are no commonly accepted and applied standards for data format and transport to facilitate the ability to assemble data from diverse sources and meet the geographic coverage, vertical and horizontal resolution, measurement accuracy, timeliness requirements, and data processing needs of multiple NOAA ocean models, assessments or product owners.

This DIF is envisioned to address gaps in data management services that serve selected NOAA ocean models, assessments or products such that improvements in model outputs, assessments or products, and/or efficiencies in time and costs can be achieved and measured.

The project premise is that data integration and improved access to and management of mission-critical ocean-related data will increase the value and effectiveness of these data in supporting decision making tools/models, and will provide standards, best practices and other protocols of use to other IOOS partners. The DIF will not define how data providers' data holdings should be managed, but rather will define interfaces and specifications for how data should be delivered. The input parameters and customer models that have been selected are anticipated to provide a robust test of the potential of integrated data to improve model products and assessments within NOAA.

Specifically, the objectives of the DIF are to:

1. Validate the premise that integrated data and improved access to the data has value that can be measured. This premise will be tested using 5 IOOS core ocean variables, from NOAA and non-NOAA sources, and 4 specific NOAA decision support tools/models.
2. Utilizing the principles of IOOS DMAC, develop a methodology to improve upon existing ocean data integration efforts that will facilitate flexibility and extensibility to other variables, systems and decision support tools.
3. Achieve improved integration of and access to selected data sets by identifying, adopting, and adapting community-developed standards for data content, metadata, quality control, and transport and deploying these standards at selected data sources serving the 4 decision support tools.
4. Maintain the DIF for a minimum period of three years, from project inception, to allow for adequate performance monitoring and assessment.
5. Provide a set of lessons learned, draft standards, and other outputs that will allow the longer-term strategic ocean data integration efforts to leverage the DIF experience to the benefit of NOAA and the Nation.

To design, build and implement the DIF, the NOAA IOOS Program will utilize a combination of existing capacity and expertise resident in NOAA and contract resources. Project teams and affiliated working groups composed of cross Line Office and Goal Team representatives will design, carry out, or direct the technical work and building of DIF components, and will be involved in the testing and evaluation of the DIF. Specific DIF components to be built include pilot implementations at specific data provider and customer locations and reference implementations that can be distributed as software “toolkits” to facilitate expansion of participation in the DIF beyond the initial data provider and customer groups.

1.2 Purpose

The purpose of this document is to define the *functional* requirements of the DIF. It is not intended as a detailed specification or design document, but will serve as a vehicle for community agreement on the functionality of the DIF and to guide further system documentation, design and implementation work. Subsequent DIF technical documentation, including Concepts of Operations, designs for pilot implementations, and test and evaluation plans will reference these requirements.

The audience for this document includes potential users, or “customers”, of the system, data providers, and the various DIF project teams involved in the design and implementation of the DIF components.

It is expected that this document will be updated periodically through the design and implementation as functional requirements are refined and additional requirements are discovered.

1.3 Document Organization

The document is organized by DIF system functional area with the functional requirements presented in tabular form. The requirements tables include the following attributes for each requirement:

- Requirement ID number
- General subject
- Requirement description
- Source of the requirement
- Customers’ priority of the requirement (reference Section 2 for description of priorities)
- DIF project priority of the requirement (reference Section 2 for description of priorities)

The requirements in Section 5 are subdivided according to the phase in which they will be implemented.

In addition to the core functional requirements listed in the requirements tables, this document includes, in Appendix A, additional, out-of-scope, requirements that were identified during interactions with the various target customers and other DIF team members. These requirements are listed separately because they are well beyond the scope of the DIF project

and are considered non-achievable within project scope and resources. Their inclusion in this document is intended to act as a record of all requirements discovered and documented during the DIF functional requirements definition.

2 Scope of the Data Integration Framework

The IOOS DIF is being considered as a risk-reduction project whose objective is to validate that the performance of certain predictive models can be improved by providing access to input data in a more integrated and homogenous way. The IOOS DIF is intended to be a pilot implementation project to measure performance improvements that integrated data¹ provides to model performance over non-integrated data. The results of the DIF project will be communicated and shared with NOAA and the IOOS community and will be used to support options to NOAA management for continued contribution to the overall IOOS.

The four NOAA decision support tools/models included in the DIF scope are:

1. Harmful Algal Bloom (HAB) Operational Forecast System and Analyses
2. Integrated Ecosystem Assessments (IEA) Models and Analyses
3. Hurricane Intensity Models - Hurricane Weather Research and Forecasting (HWRF)
4. Coastal Inundation Models
 - a. Lake, and Overland Surges from Hurricanes (SLOSH)
 - b. Extratropical Storm Surge (ETSS) Model

The DIF effort will be limited to data containing at least one of the following five “core variables” for ocean observations:

1. Temperature
2. Salinity
3. Currents
4. Sea Level
5. Ocean Color

These data will also be limited to data sets and products that are in the form of time series, profiles, and/or gridded data. The data sources to be used in the DIF will be based on customer requirements, the five core variables, and ease of integration.

All requirements captured in Section 5 are considered within the scope of the DIF. The DIF will be developed in phases over the course of the project; Section 5 outlines which requirements will be implemented in which phase. Subsequent phases will build on the capability of previous phases until the full project scope and functionality is achieved. Requirements from previous phases apply to future work, typically applying to a larger scope in terms of data, data providers, etc. The DIF priorities in this document will be used to define

¹ Data integration refers to the process of combining data residing at different sources and providing the user with a unified access to these data. It involves the extraction, consolidation, and management of data from disparate systems to achieve broader capability by (functionally or technically) relating two or more data streams for the purposes of manipulation, analysis, and distribution. (http://en.wikipedia.org/wiki/Data_integration)

what functionality is implemented during which phase. The method for determining requirements' priority is described below.

Customers' priorities were identified from the perspective of the customers' mission fulfillment using the following priority definitions:

1. Mission Critical / Cannot meet operational mission objectives without this functionality.
2. Mission Optimal / Functionality not critical but would provide significant improvement to operational capability.
3. Mission Enhancing / Needed to enhance state of knowledge / assess potential for operational capability.

DIF priorities were based on anticipated performance impact on the customer models and whether or not the requirement is within the DIF scope (reference Section 2). DIF priorities are described as follows:

- H – High / Highly valuable to demonstrate project success; achievable at low risk and within project scope and/or resources.
- M – Medium / Moderately valuable to demonstrate project success; achievable at medium risk, outside project scope and/or resources.
- L – Low / Minimally valuable to demonstrate project success; not achievable or achievable at high risk, outside project scope and/or resources; may relate to non-core variable or data type.

3 Applicable Documents

- RD1 NOAA Integrated Ocean Observing System (IOOS) Program Office *Data Integration Framework (DIF) Master Project Plan*
- RD2 Data Management and Communications Plan for Research and Operational Integrated Ocean Observing Systems, The National Office for Integrated and Sustained Ocean Observations Ocean.US Publication No. 6
- RD3 IOOS Functional Requirements Baseline Document (produced by LMI)
- RD4 DIF As-Is Baseline Systems Document (DRAFT Version 0.6) September 19, 2007
- RD5 Harmful Algal Bloom Modeling – Draft Requirements for the IOOS Data Integration Framework, January 19, 2007
- RD6 Harmful Algal Bloom Customer Meeting Summary, July 19, 2007
- RD7 Hurricane Intensity Modeling Draft Report Summary, February 12, 2007
- RD8 Hurricane Intensity Customer Meeting Report Summary, July 18, 2007
- RD9 Coastal Inundation Modeling Draft Report Summary, January 29, 2007 and February 12, 2007
- RD10 Coastal Inundation Customer Meeting Report Summary, July 13, 2007
- RD11 Integrated Ecosystem Assessments – Draft Requirements for the IOOS Data Integration Framework, February 16, 2007
- RD12 IOOS and Integrated Ecosystem Assessments – Data Integration Framework (DIF) Follow-up Meeting Notes, July 11, 2007

4 IOOS DIF Overview

The DIF is a set of components and standards that will facilitate more efficient transfer of data products from a varied set of data providers to a number of data customers' models and analysis tools.

Data providers collect observational data from various platforms and sensors, assemble and process these data, and provide various output data products. Examples of data providers and the products they offer are shown in the table below.

Center/System	Product(s)
National Data Buoy Center (NDBC)	Wind, Currents, Temperature
Center for Operational Oceanographic Products and Services (CO-OPS) National Water Level Observation Network (NWLON)	Water Level
CO-OPS Physical Oceanographic Real-Time System (PORTS [®])	Salinity
National Environmental Satellite, Data, and Information Service (NESDIS) CoastWatch MODIS Aqua	Ocean Color
Temperature and Salinity Profile Program (GTSP)	Temperature, Salinity

The DIF may also collect and integrate data from selected models and forecasting systems as required by individual customers.

The concept of the DIF is to allow the identified NOAA customers to more easily and efficiently access oceanographic and other meteorological data from the data providers for input into their modeling and analysis systems. The goal of the DIF is to provide these customers with a uniform set of methods to access to the multiple data sets that they ingest, and to make these data sets available in standardized formats. Currently these customers tend to access these data sets using specialized automation processes and through manual retrieval over the Internet from various web sites and FTP servers. The DIF will allow users to implement more homogenous, automated processes for retrieving and using the data. This will be done through the selection, specialization, application and testing of data content, data transport, metadata and quality standards. These standards are intended to be used by DIF data providers to formalize a common data sharing infrastructure.

The DIF will interface with the various data providers to retrieve the data sets defined in Section 5.1. The DIF will ensure that these data are consistent with a common data model and data content standards when transported to the customer. Data in the common format will be made available to the user community using standard transport protocols. It is anticipated that the DIF will be a distributed system, with components implemented at data providers, software gateways, customer locations or all of these. The functionality required to be provided by the various components is described in this document. How are where that

functionality is implemented will be addressed in the design and implementation phases of the project.

The DIF is likely to be developed in phases over the course of the project; the initial phase may, for example, result in adaptation of some standards and integration of one or two variables. Subsequent phases will build on that capability until the full project scope and functionality is achieved. The DIF priorities in this document will likely be used to define what functionality is implemented during which phase.

5 DIF Functional Requirements

5.1 Phase I Functional Requirements

This section contains requirements for implementation Phase I of the DIF.

5.1.1 General Functional Requirements

This section outlines the general functional requirements for Phase I of the DIF.

ID	Subject	Requirement	Source	Customer Priority *	DIF Priority**
GEN045	Common Data Model	All DIF data, including model outputs, defined in Section 5.2 shall adhere to a common data model or set of models and conventions to ensure consistency of content and format. The common data model(s) and conventions shall adhere to DIF-approved standards.	IOOS Program	N/A	H

5.1.2 Required Data Sets

This section specifies the data sources that are required for integration and access through the DIF.

ID	Variable/ Platform/ Provider	Requirement	Source	Customer Priority *	DIF Priority**
RDS001	Currents / ADCP / NDBC	The DIF shall provide MMS ADCP Currents data from NDBC.	HAB Draft Reqs (1/19/07); 2nd customer meeting req'ts	3	H
RDS035	Currents / TABS / NDBC	The DIF shall provide Texas Automated Buoy Systems (TABS) Currents data from NDBC.	HAB Draft Reqs (1/19/07); 2nd customer meeting req'ts	3	H

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

ID	Variable/ Platform/ Provider	Requirement	Source	Customer Priority *	DIF Priority**
RDS040	Currents / COMPS / USF (NDBC?)	The DIF shall provide Coastal Ocean Monitoring and Prediction System (COMPS) Buoy Currents data from University of South Florida.	HAB Draft Reqs (1/19/07); 2nd customer meeting req'ts	2	H
RDS100	Currents / HF Radar (CODAR) / NDBC?	The DIF shall provide Gulf Currents data from HF Radar (CODAR) in NetCDF format. GAP: Verify data source provider.	HAB Draft Reqs (1/19/07)	2	H
RDS115	Currents / ADCP / NDBC?	The DIF shall provide Current data from shipboard ADCP profiles. GAP: Verify data source provider.	IEA Draft Reqs (2/16/07)	1	H

5.1.3 Metadata Management

ID	Subject	Requirement	Source	Customer Priority *	DIF Priority**
MTD001	Metadata Standards	Each DIF logical dataset or data service shall have at least one FGDC or ISO metadata record.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H
MTD002	Metadata Standards	DIF metadata standards shall adequately define data attribute details (e.g. unit of measure, reporting convention, precision, code definitions)		N/A	H
MTD015	Metadata, Public Access	The DIF shall encode all metadata in valid XML and make it available for public access.	DS Working Group	N/A	H
MTD015	Metadata Reporting	When publishing data utilizing a service interface, the DIF shall report metadata for the service (eg using the GetCapabilities record for an OGC service).	IOOS DIF Proposed Funcs/Capabilities (4/3/07), modified by WG	N/A	H

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

5.1.4 Quality Control

ID	Subject	Requirement	Source	Customer Priority *	DIF Priority**
QC015	Provide data quality information in metadata	The DIF shall document quality procedures in the metadata record.	HI Draft Reqs (2/12/07); HABs 2 nd customer meeting (7/19/07)	1	H

5.1.5 Data Transport and Access

ID	Subject	Requirement	Source	Customer Priority *	DIF Priority**
XPT010	Data Transport – Real Time	The DIF shall provide the capability for the transmission of DIF data sets to customers and archive centers in real-time using DMAC standards and protocols.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H
XPT015	Data Transport – Delayed Mode	The DIF shall provide the capability for the transmission of DIF data sets to customers and archive centers in delayed mode using DMAC standards and protocols.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H

5.1.5.1 Data Aggregation and Filtering

ID	Subject	Requirement	Source	Customer Priority *	DIF Priority**
FILT010	Aggregation	DIF data providers shall assemble compatible distributed data sources, along geographic and temporal boundaries, into seamless logical collections differentiated by variable, to allow users to exploit data from distributed sensor arrays.	Func Req's Working Group (11/16/07)		H

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

5.2 Phase II Functional Requirements

This section contains the requirements for implementation Phase II of the DIF.

5.2.1 General DIF Functional Requirements

This section outlines the general functional requirements for the DIF.

ID	Subject	Requirement	Source	Customer Priority *	DIF Priority**
GEN001	Core Variable – Temperature	The DIF shall retrieve, process, and deliver to end users data containing the core variable of Temperature.	IOOS Program	N/A	H
GEN005	Core Variable – Salinity	The DIF shall retrieve, process, and deliver to end users data containing the core variable of Salinity.	IOOS Program	N/A	H
GEN010	Core Variable – Currents	The DIF shall retrieve, process, and deliver to end users data containing the core variable of Currents.	IOOS Program	N/A	H
GEN015	Core Variable – Sea Level	The DIF shall retrieve, process, and deliver to end users data containing the core variable of Sea Level.	IOOS Program	N/A	H
GEN020	Core Variable – Ocean Color	The DIF shall retrieve, process, and deliver to end users data containing the core variable of Ocean Color.	IOOS Program	N/A	H
GEN025	Real-time data	The DIF shall support access and delivery (transport) of real time data, as applicable, for the data sets defined in Section 5.2.	NOAA IOOS Project Req'ts Document (per LMI, modified)	N/A	H
GEN026	Delayed mode data	The DIF shall support access and delivery (transport) of delayed mode data, as applicable, for the data sets defined in Section 5.2.	NOAA IOOS Project Req'ts Document (per LMI, modified)	N/A	H
GEN030	Data Type – Time Series	The DIF shall support data in the form of time series.	IOOS Program	N/A	H
GEN035	Data Type – Profiles	The DIF shall support data in the form of profiles.	IOOS Program	N/A	H
GEN040	Data Type – Gridded	The DIF shall support data in the form of gridded data.	IOOS Program	N/A	H

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

5.2.2 Required Data Sets

This section specifies the data sources that are required for integration and access through the DIF.

ID	Variable / Platform / Provider	Requirement	Source	Customer Priority	DIF Priority
RDS005	Salinity / C-MAN / NDBC	The DIF shall provide CTD offshore Salinity data from NDBC buoys.	HAB Draft Reqs (1/19/07)	2	M
RDS020	Ocean Color / SeaWiFS / CoastWatch	The DIF shall provide near real-time SeaWiFS Ocean Color data from CoastWatch servers.	HAB Draft Reqs (1/19/07)	1	M
RDS025	Ocean Color / MODIS Aqua (NASA) / CoastWatch	The DIF shall provide near real-time MODIS Aqua Ocean Color data from CoastWatch servers.	HAB Draft Reqs (1/19/07)	1	M
RDS045	Salinity (Estuarine) / NERRS / NERRS	The DIF shall provide estuarine Salinity data from NERRS.	HAB Draft Reqs (1/19/07)	2	M
RDS050	Salinity (Estuarine) / PORTS [®] / CO-OPS	The DIF shall provide estuarine Salinity data from the CO-OPS PORTS [®] data sets.	HAB Draft Reqs (1/19/07)	2	M
RDS055	Salinity (Estuarine) / NEPs / EPA	The DIF shall provide estuarine Salinity data from the EPA National Estuary Program (NEP)	HAB Draft Reqs (1/19/07)	2	H
RDS180	Temperature / CO-OPS	The DIF shall provide Water Temperature data from CO-OPS systems.			M
RDS185	Temperature / NDBC	The DIF shall provide Water Temperature data from NDBC.			M
RDS180	Temperature / NERRS	The DIF shall provide Water Temperature data from NERRS.			M

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

ID	Variable / Platform / Provider	Requirement	Source	Customer Priority	DIF Priority
RDS180	Temperature/ /NEP	The DIF shall provide Water Temperature data from NEP systems.			M
RDS185	Water Level/ /CO-OPS	The DIF shall provide Water Level data from CO-OPS systems.			M
RDS185	Water Level/ /NDBC	The DIF shall provide Water Level data from NDBC systems.			M
RDS190	Sea Level/Water Level Gauges/USGS (via NHC)	The DIF shall provide Sea Level data from USGS portable water level gauges via NHC has data.	CI Draft Rpt Summary (2/12/07)	2	M

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

5.2.3 Pre-Processing and Validation

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
VAL005	Remove duplicate records	The DIF shall filter and delete duplicate records for data sets collected from multiple sources. GAP: Identify precise method for determining which records are duplicates.	IEA Draft Reqs (2/16/07)	1	H

5.2.4 Transformation and Assembly

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
TRAN001	Data transformation	The DIF shall provide the capability and tools to transform data, as needed, into the format consistent with the DIF common data model and content standards.	IOOS Program	N/A	H
TRAN020	Aggregation by data type (VDAC)	The DIF shall provide mechanism for aggregation of data of the same data type from multiple data providers. Such aggregation shall be provided over useful geographic and temporal boundaries to allow users to exploit data of the same type from multiple sources.	CONOPS Working Group 11/14/07		M

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

5.2.5 Metadata Management

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
MTD003	Metadata Standards	DIF metadata standards shall adequately convey data file formats or structures.		N/A	H
MTD004	Metadata Standards	DIF metadata standards shall adequately capture data quality information (e.g. QC tests applied, QC flags and flag definitions)		N/A	H
MTD005	Metadata Standards	DIF metadata standards shall convey all transformations (e.g. unit conversions, format conversions, sub-setting) that have occurred to data from the entry point to the DIF to the output/delivery to the data user.		N/A	H
MTD010	Data-Metadata Linkage	The DIF shall ensure that the linkages between data and metadata are maintained with great reliability.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H
MTD020	Metadata Tools	The DIF shall provide tools to enable end users and data providers to increase their capability in metadata generation and management.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H
MTD025	Consistent Terminology	The DIF shall establish a standard glossary for use by customers and data providers.	IEA Draft Reqs (2/16/07); was XPT035	1	H
MTD030	Central Metadata Registry	The DIF shall provide a central metadata registry.	Func. Req's Working Group (11/16/07)	N/A	M

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

5.2.6 Quality Control

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
QC001	QC Partnership	The DIF shall ensure that quality control operations are a partnership among data observation/collection components, processors, analysts, other users, and the DMAC.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H
QC002	QC Standards	The DIF shall identify, adopt, and adapt existing standards for quality control.	DIF IPT	N/A	H
QC005	Known/documented Quality	The DIF shall provide a mechanism for ensuring that data are of known and documented quality.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H
QC010	QC Documentation	The DIF shall ensure that all DIF data products should include readily accessible documentation of sensor characteristics, processing, calibration, provenance, quality control, accuracy, precision, and other data and information required to satisfy user needs.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H
QC020	QC Transport	The DIF shall transport QC and error characteristics, flags, through from data provider to data consumer. Gap: will need to identify all error and QC data per integrated data set.	HABs 2nd customer meeting (7/19/07); modified by WG	1	H
QC030	Data QC	The DIF shall accurately report lat/long and time of measurement.	CI Draft Rpt Summary (2/12/07)	1	H
QC035	Flag Missing Data	The DIF shall consistently flag missing data.	CI 2 nd round?	1	H
QC040	QC of Sea Level products	The DIF shall provide evaluated QC'd Sea Level products and elevation contours from data sources. GAP: Identify data sources.	IEA Draft Reqs (2/16/07)	2	M
QC050	Data quality flags	The DIF shall implement data quality flags consistently across the data community.	IEA Draft Reqs (2/16/07)	1	H
QC055	Data quality flags	The DIF shall provide raw data with data quality flags rather than cleaned data. GAP: Identify which variables and data sources.	IEA Draft Reqs (2/16/07)	1	M

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
QC060	Deliver Sea Level data of known quality	The DIF shall deliver data of known quality, possibly using NWLON QC procedure, for all Sea Level data. GAP: Identify which providers specifically need improved QC.	HI Draft Reqs (2/12/07)	1	H
QC065	Access and QC of ADCP Current data	The DIF shall provide QC of Current data from shipboard ADCP profiles.	IEA Draft Reqs (2/16/07)	1	H

5.2.7 Data Storage

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
DSA005	Contribution to archive	The DIF shall facilitate contribution of standards-based data to NODC archive from ecosystem community.	IEA Draft Reqs (2/16/07)	1	L - Not in DIF scope

5.2.8 Data Transport and Access

5.2.8.1 Data Transport

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
XPT001	Transport, Machine-to-Machine	The DIF shall support transport protocol(s) to allow machine-to-machine data transport.	IOOS Program	N/A	H
XPT005	Transport, End-User	The DIF shall support transport protocol(s) to allow data retrieval manually by a user.	IOOS Program	N/A	M

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
XPT020	Metadata Transport	The DIF shall provide for the transport of metadata associated with all data transported.		N/A	H
XPT025	Data Transport Standards	Data transport standards applicable to the DIF shall be identified and adapted as needed to support DIF functional requirements.		N/A	H
XPT040	Automated Push Delivery	The DIF shall provide automated delivery, "push", of data to a customer model data input.		N/A	M
XPT045	Manual Delivery	The DIF shall allow users to manually select data for delivery.		N/A	M
XPT050	Integrated Access Functions	The DIF data and metadata access functions shall be seamlessly integrated.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H

5.2.8.2 Data Aggregation and Filtering

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
FILT001	Subsetting, Server side	The DIF shall allow for server-side subsetting of data for all relevant data.	NOAA IOOS Project Requirements Document (per LMI), modified by IPT WG	2	H
FILT002	Subsetting, Data Field	The DIF shall allow subsetting of data based on any field in the relevant data set.	NOAA IOOS Project Requirements Document (per LMI), IEA Draft Reqs (2/16/07), modified by IPT WG	2	H

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

5.2.9 IT Security

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
SEC001	NOAA Req'ts	The DIF shall implement applicable NOAA requirements for IT security.	IOOS Program	N/A	H
SEC005	External Interfaces	The DIF shall ensure that interfaces to external systems are secure from unauthorized access.	NWS/SEC ITSO Guidance	N/A	H
SEC010	User and System Functional Separation	The DIF shall separate user functionality (including user interface services) from information system management functionality.	NWS/SEC ITSO Guidance	N/A	H
SEC015	Security Function Separation	The DIF shall isolate security functions from non-security functions.	NWS/SEC ITSO Guidance	N/A	H
SEC020	Protect Secure Information	The DIF shall protect the confidentiality of security-relevant system information and integrity of all transmitted information.	NWS/SEC ITSO Guidance	N/A	H
SEC025	Network Connection Termination	The DIF shall terminate network connections at the end of each session or after a configurable period of inactivity.	NWS/SEC ITSO Guidance	N/A	H
SEC030	Session Authenticity	The DIF shall provide mechanisms to protect the authenticity of communications sessions.	NWS/SEC ITSO Guidance	N/A	H
SEC035	Virus Protection	The DIF shall incorporate antivirus protections with automatic updates.	NWS/SEC ITSO Guidance	N/A	H
SEC040	Input Validation	The DIF shall check information inputs for accuracy, completeness, and validity.	NWS/SEC ITSO Guidance	N/A	H
SEC045	Source Authentication	The DIF shall authenticate that data sets received at client locations are from a valid DIF data source	IOOS Program	N/A	H

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

5.2.10 System Administration

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
ADM001	System Monitoring	The DIF shall provide a mechanism for continually monitoring results and publication of those results as a method for evaluating improvements in performance as a result of data integration.	DMIT Mtg 7/24/07	N/A	H
ADM005	Fault Detection and Correction	The DIF shall provide a mechanism for detection of status of DIF data provider interfaces and report the status externally.	IOOS DIF Proposed Funcs/Capabilities (4/3/07), mod Func Reqs WG (11/16/07)	N/A	H
ADM010	System Extensibility	The DIF shall provide a mechanism for system extensibility.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H
ADM015	Data Availability Policies	The DIF shall establish and publicize policies for data availability.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H
ADM020	User Feedback	The DIF shall provide a mechanism for soliciting and responding to user feedback.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H
ADM025	International Linkages	The DIF shall establish and maintain international linkages.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	H

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
ADM030	Service Registry	The DIF shall provide a registry or other means for determining what data are available within the DIF based upon queries that may be issued by others or by other machines.	IOOS DIF Proposed Funcs/Capabilities (4/3/07)	N/A	M

5.3 Phase III Functional Requirements

The table below identifies additional data sets to be included in Phase III of the DIF.

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
RDS010	Real-time Winds / C-MAN / NDBC	The DIF shall provide C-MAN real-time Wind Speed and Direction data from NDBC.	HAB Draft Reqs (1/19/07)	1	L- Not core variable
RDS011	Real-time Winds / NWLON / NDBC	The DIF shall provide NWLON real-time Wind Speed and Direction data from NDBC.	HAB Draft Reqs (1/19/07)	1	L- Not core variable
RDS012	Real-time Winds / Buoys / NDBC	The DIF shall provide NDBC Buoy real-time Wind Speed and Direction data from NDBC.	HAB Draft Reqs (1/19/07)	1	L- Not core variable
RDS030	Currents / TGLO 3-D / Tx A&M (NDBC?)	The DIF shall provide TGLO 3-D circulation model Currents data from Texas A&M. NOTE: this is not deemed an operational system by NOAA.	HAB Draft Reqs (1/19/07); 2nd customer meeting req'ts	3	L – Not operational
RDS060	Winds (Forecasted) / NAM / NCEP	The DIF shall provide NAM model Forecasted Winds data from NCEP sites .	HAB Draft Reqs (1/19/07)	1	L- Not core variable

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
RDS065	Winds (Forecasted) / NWS / NWSTG – GTS	The DIF shall provide Forecasted Winds data from NWS regional marine text forecasts. GAP: Verify source provider of forecast data.	HAB Draft Reqs (1/19/07)	1	L- Not core variable
RDS070	Cell Counts / Field surveys / State & local partners	The DIF shall provide Cell Count data from various state partners including, but not limited to, Florida Fish and Wildlife Research Institute, Mote Marine Laboratory, Collier County Pollution Control and Prevention Department, Sarasota County Health Department, Alabama Department of Public Health.	HAB Draft Reqs (1/19/07)	1	L- Not core variable
RDS075	Bathymetry / Coastal Relief Model / NGDC	The DIF shall provide Bathymetry data from the NGDC digital elevation model (DEM) Coastal Relief Model.	CI Draft Rpt Summary (2/12/07)	1	L- Not core variable
RDS080	Bathymetry / Hydrographic Surveys / NOS	The DIF shall provide Bathymetry data from NOS Hydrographic Surveys.	CI Draft Rpt Summary (2/12/07)	1	L- Not core variable
RDS085	RESERVED				
RDS090	Unknown / Unknown / USGS	The DIF shall provide USGS data to replace current web scraping. GAP: Determine variables, platforms, sources and formats.	HABS 2nd customer meeting (7/19/07)	2	L
RDS095	Currents (Offshore) / AUVs / Unknown	The DIF shall provide offshore Currents and other data from existing AUVs. GAP: Verify data source provider.	HAB Draft Reqs (1/19/07)	3	L-data may not be readily available
RDS105	Topography / National Elevation Database / USGS	The DIF shall provide Topography data from the USGS National Elevation Database.	CI Draft Rpt Summary (2/12/07)	1	L- Not core variable

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
RDS110	Topography / LiDAR Data / Unknown	The DIF shall provide Topography data LiDAR. GAP: Verify data source providers, formats.	CI Draft Rpt Summary (2/12/07)	1	L- Not core variable
RDS120	Water Pressure / Unknown / Unknown	The DIF shall provide subsurface Pressure data from available sources. GAP: Identify platforms, data source provider(s), formats.	IEA Draft Reqs (2/16/07)	3	L - Not core variable
RDS125	Salinity / regional observing systems / Unknown	The DIF shall provide sub-surface and near-shore Salinity data from regional observing systems. GAP: Verify data source providers, access methods, formats.	HI Draft Reqs (2/12/07)	2	L
RDS130	Currents / TBD / TBD	Increase spatial coverage of current data, or access to in situ currents. GAP: Identify specific sources.	HAB Draft Reqs (1/19/07)	3	L
RDS140	Temperature / TBD / TBD	Obtain additional surface and subsurface Temperature data at various resolutions. GAP: Determine data availability and sources.	IEA Draft Reqs (2/16/07)	2	L
RDS145	Temperature / TBD / TBD	The DIF shall provide access to Temperature interpolations for Sea Surface and Subsurface Temperatures from models. GAP: identify models and source provider.	IEA Draft Reqs (2/16/07)	2	L
RDS150	TBD / TBD / GTS	The DIF shall provide access to historical data for all time series data provided through the GTS. GAP Identify variables and observation platforms.	IEA Draft Reqs (2/16/07)	2	L
RDS155	Temperature / TBD / TBD	The DIF shall provide Subsurface Temperature data to address the Gulf of Mexico data gaps. GAP: Identify observation platforms and data source providers.	HI Draft Reqs (2/12/07)	2	L
RDS160	Salinity / TBD / TBD	The DIF shall provide Salinity data to address Gulf of Mexico data gaps. GAP: Identify observation platforms and data source providers.	HI Draft Reqs (2/12/07)	2	H

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

ID	Subject	Requirement	Source	Customer Priority	DIF Priority
RDS165	Winds (Forecasted) / TBD / Regional sources	The DIF shall provide Forecasted Winds from regional sources. GAP: Need to identify specific sources.	CI Draft Rpt Summary (2/12/07)	2	L- Not core variable
RDS170	High Water Mark / Storm gauges / USGS	The DIF shall provide post-storm High Water Mark data from USGS storm gauges.	CI Draft Rpt Summary (2/12/07)	2	L - Not core variable
RDS175	Ocean Heat / TBD / Tropical Prediction Center	The DIF shall provide the Ocean Heat content product produced at Tropical Prediction Center.	HI Draft Reqs (2/12/07)	2	L

* 1-Mission Critical; 2-Mission Optimal; 3-Mission Enhancing.

** H-High; M-Medium; L-Low.

Appendix A Out-of-Scope Requirements

This section tracks requirements that came to light during the functional requirements definition phase of the DIF project, but are out of scope of the DIF project for one reason or another. They are listed here to keep a record of all documented requirements.

ID	Subject	Requirement	Source
OOS 001	Ocean Color Data from SeaWiFS	Continue to purchase SeaWiFS data (NOAA) beyond September 2007.	HAB Draft Reqs (1/19/07)
OOS 005	Offshore Current data from AUVs	Deploy additional AUVs to detect offshore currents and other data.	HAB Draft Reqs (1/19/07)
OOS 010	Consistent Cell Count data	Establish a standardized format for Cell Count data contribution in all Gulf region states (currently only TX and FL).	HAB Draft Reqs (1/19/07)
OOS 015	Sea Level Data Collection During Storm Events	Increase availability of Sea Level data during storm events, through hardening of gauges.	CI Draft Rpt Summary (2/12/07)
OOS 020	Backup power for observing systems	Provide backup power for observing systems to ensure data collection during storm events.	CI Draft Rpt Summary (2/12/07)
OOS 025	Data Latency	Reduce latency of data delivery from observing systems to data processing centers.	CI Draft Rpt Summary (2/12/07)
OOS 030	Data integration	Allow users to graphically overlay multiple observations with forecast data. NOTE: this is likely outside the scope of DIF.	HAB 2nd customer meeting (7/19/07)
OOS 040	Forecasted Winds, NDFD	The DIF shall provide Forecasted Winds data from NDFD. It has been stated that the NDFD has been built with oceanographic data fields but that these data fields are generally not being populated with data.	HABS 2nd customer meeting (7/19/07)
OOS 045	Ocean Color temporal resolution	Provide Ocean Color from SeaWiFS at a temporal resolution of 3 hours. Determine availability of data at this resolution. GAP: Need source for data.	HAB Draft Reqs (1/19/07)
OOS 050	Ocean Color spatial resolution	Provide Ocean Color at a spatial resolution of 300 meters. GAP: Need source for data.	HAB Draft Reqs (1/19/07)
OOS 051	Integrate regional data sources into NOAA products	Provide integrated access to data from regional networks through NOAA data centers; output products should be consistent with quality control and quality assurance. GAP: Identify regional networks, the products desired from each, and the source data formats.	CI Draft Rpt Summary (2/12/07)

ID	Subject	Requirement	Source
OOS 055	Sea Level Temporal Frequency	Increase Sea Level temporal frequency (currently 6 min and 1 min during storm events). GAP: Identify source data products for higher temporal frequency.	CI Draft Rpt Summary (2/12/07)
OOS 060	Sea Level Spatial Resolution	Increase Sea Level coverage data for monitoring storm surge events. GAP: Identify additional data sources.	CI Draft Rpt Summary (2/12/07)
OOS 065	Improve vertical accuracy of data	Improve vertical accuracy of data along the Louisiana coast. GAP: Identify data sources, formats, and availability.	CI Draft Rpt Summary (2/12/07)
OOS 070	Relation between Water Level and GEO-IDE data	Provide relation between Water Level data and GEO-IDE data. GAP: Needs clarification.	CI Draft Rpt Summary (2/12/07)
OOS 075	Minimize occurrence of truncated data	Improve the data transmission process to minimize occurrence of truncated records. GAP: Could not validate issue in 2 nd round meetings.	CI Draft Rpt Summary (2/12/07)
OOS 080 (previously QC045)	QC of Current data	The DIF shall provide additional processing and QC of archived Current data. GAP: Identify data sources.	IEA Draft Reqs (2/16/07)
OOS 085 (previously TRAN010)	Multiple sources for ADCP data	The DIF shall compile and process raw ADCP data from multiple sources. GAP: Identify sources and desired data.	IEA Draft Reqs (2/16/07)

Appendix B Acronym List

ADCIRC	Advanced Circulation Hydrodynamic Model
ADCP	Acoustic Doppler Current Profiler
AUV	Autonomous Underwater Vehicles
AWIPS	Advanced Weather Interactive Processing System
CalCOFI	California Cooperative Oceanic Fisheries Investigations
CDMO	Centralized Data Management Office
CI	Coastal Inundation
C-MAN	Coastal-Marine Automated Network
CODAR	Coastal Ocean Dynamics Applications Radar
COMPS	Coastal Ocean Monitoring and Prediction System
CO-OPS	Center for Operational Oceanographic Products and Services
CSC	Coastal Services Center
CSDL	Coast Survey Development Lab
CTD	Conductivity, Temperature, Depth
DIF	Data Integration Framework
DMIT	Data Management Integration Team
EPA	Environmental Protection Agency
ETSS	Extratropical Storm Surge
FGDC	Federal Geographic Data Committee
FIPS	Federal Information Processing Standards
FTP	File Transfer Protocol
GOM	Gulf of Mexico
GSFC	Goddard Space Flight Center
GTS	WMO Global Telecommunications System
GTSPP	Global Temperature and Salinity Profile Program
HAB	Harmful Algal Bloom
HDF	Hierarchical Data Format
HF	High frequency (radar)
HI	Hurricane Intensity
HTTP	Hyper Text Transfer Protocol
HWRF	Hurricane Weather Research and Forecasting
HYCOM	Hybrid Coordinate Ocean Model
IDS	Required Data Sets
IEA	Integrated Environmental Assessments
IGOSS	Integrated Global Ocean Services System
IMS	Information Management System
IOC	Initial Operating Capability
IOC	Intergovernmental Oceanographic Commission
IODE	International Oceanographic Data and Information Exchange committee
IOOS	Integrated Ocean Observing System

IT	Information Technology
JCOMM	Joint Commission on Oceanography and Marine Meteorology
MEDS	Canada's Marine Environmental Data Service
MMS	Minerals Management Service
MODIS	Moderate Resolution Imaging Spectroradiometer
NAM	North American Mesoscale
NCCOS	National Centers for Coastal Ocean Science
NDBC	National Data Buoy Center
NEP	National Estuary Program
NERRS	National Estuarine Research Reserve System
NESDIS	National Environmental Satellite, Data, and Information Service
netCDF	Network Common Data Form
NGOM	Northern Gulf of Mexico
NIST	National Institute of Standards and Technology
NODC	US National Oceanographic Data Center
NOS	National Ocean Service
NWLON	National Water Level Observation Network
NWS	National Weather Service
NWSTG	NWS Telecommunications Gateway
OPeNDAP	Open-source Project for a Network Data Access Protocol
POM	Princeton Ocean Model
PORTS [®]	Physical Oceanographic Real-Time System [®]
QA	Quality Assurance
QC	Quality Control
RD	Reference Document
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
SLOSH	Lake, and Overland Surges from Hurricanes
TABS	Texas Automated Buoy System
TGLO	Texas General Land Office
USF	University of South Florida
USGS	United States Geological Survey
WMO	World Meteorological Organization