Guidance for Implementation of the Integrated Ocean Observing System (IOOS®) Data Management and Communications (DMAC) Subsystem

NOAA IOOS® Program Office White Paper (v1.0)*

March 12, 2010

This white paper is designed to advance community discussion about the national/regional IOOS® enterprise, specifically with a focus on implementation of DMAC subsystem elements.

Central to the success of IOOS® is the presence of a DMAC subsystem capable of delivering real-time, delayed-mode, and historical data for in-situ and remotely-sensed physical, chemical and biological observations, as well as, model-generated outputs, including forecasts, to IOOS® users and of delivering all forms of information to and from secure archive facilities. The DMAC design must be responsive to user needs. At a minimum, the DMAC subsystem must make data and products discoverable and accessible, and must provide essential metadata regarding information sources, methods and quality.

The overall objectives are for IOOS® partners to develop and maintain DMAC capabilities to:

- Deliver accurate and timely ocean observations and model outputs to a range of consumers including government, academic, private sector users, and the general public utilizing specifications common across all providers.

- Deploy the information system components (including infrastructure and relevant personnel) for full life-cycle management of observations from collection to product creation to public delivery, system documentation, and archiving.

- Establish robust data exchange that is responsive to variable customer requirements and routine feedback and not tightly bound to a specific application of the data or particular end-user decision support tool.

IOOS partners must therefore address the following DMAC-specific objectives:

- A standards-based foundation for DMAC capabilities: IOOS® partners must clearly demonstrate how they will ensure the establishment and maintenance of a standards-based approach for delivering their ocean observations data and associated products to users through local, regional and global/international data networks.

- Exposure of and access to coastal ocean observations: IOOS® partners must describe how they will ensure coastal ocean observations are exposed to users via a service-oriented architecture and recommended data services (see Section (d), below) that will ensure increased data interoperability including the use of improved metadata and uniform quality-control methods.

*Subsequent versions may augment or clarify sections of this guidance document without undermining progress established under the current or previous versions.

- Certification and governance of IOOS® data and products: IOOS® partners must present a description of how they will participate in establishing an effective IOOS® governance process for data certification standards and compliance procedures. This objective is part of an overall accreditation process which includes the other IOOS® subsystems (observing, modeling and analysis, and governance).

In addition, IOOS® partners must demonstrate how the DMAC Subsystem component will be implemented and sustained based on the following guiding principles:

a) Open Data Sharing: IOOS®, being a part of the Global Earth Observing System of Systems (GEOSS), ascribes to the GEOSS data sharing principles. The precise wording of those principles [see http://www.earthobservations.org/geoss_dsp.shtml] is still under development but the basic intent is for full and open exchange of data and metadata with minimum possible cost, delay and restriction. In particular, observations gathered with IOOS® funding are to be freely available. In some instances, IOOS® data providers may act as a "clearinghouse" for other lower capacity or smaller scale providers.

b) Provision of Data to WMO GTS: The World Meteorological Organization (WMO) Global Telecommunications System (GTS), and an emerging next generation system - the WMO Information System (WIS), - disseminates data in near-real-time to operational weather and ocean forecasting centers. NOAA IOOS® is committed to ensuring that all relevant U.S. coastal ocean observations will be contributed in near real time to the global GTS network. IOOS® partners can fulfill this requirement by entering into an agreement with NOAA's National Data Buoy Center (NDBC) to post data onto the GTS in near real time. NOAA IOOS® will provide specific guidance to partners on data priorities.

c) Service-Oriented Architecture (SOA): DMAC employs a service-oriented architecture (SOA). IOOS® partners must participate in highly distributed system of compatible components that includes:

   - Establishing standardized services that enable access to the information they wish to expose to IOOS® (see Section (d), below, for IOOS® service types);
   - Registering their service instances in a Service Registry to be established by IOOS® for the use of all participants;
   - Utilizing IOOS®-approved mechanisms for notifying users of service outages and enhancements;
   - Permitting basic monitoring, by IOOS®, of service availability and response time; and
   - Reporting summary data usage statistics to IOOS® on a monthly basis based on web server logs. This aggregate information should include basic statistics such as average daily requests during the month, and fraction of requests from various top-level domains (.com, .gov, .edu, etc.). Additional information is encouraged.

d) Recommended Data Access Services: The basic data access services currently used by IOOS® DMAC are listed below. IOOS® data partners are expected to establish, register, and maintain services which correspond to their data types (legacy services may also be maintained, if desired).

- OPeNDAP Data Access Protocol (DAP) and/or Open Geospatial Consortium (OGC) Web Coverage Service (WCS) for access to gridded data and model outputs (for example a THREDDS data server – Thematic Real-time Environmental Distributed Data Services – is one possible approach for implementing DAP and WCS services);
- OGC Sensor Observation Service (SOS) for access to in situ observations (e.g., observations at a point, profiles, trajectories, etc.);
- OGC Web Map Service (WMS) for access to georeferenced image data; and
- Other service types, in particular for event-based or bulk subscriptions, are under consideration, and IOOS® data partners are invited to participate in the adoption of such services. The overall objective of this approach, regardless of which specific services are adopted, is the participation in an ongoing community standards processes that result in common data model and/or common data taxonomies. Specific versions of each recommended technical specification will be stipulated to IOOS® partners.

e) Common Data Formats: IOOS® DMAC has selected several preferred data formats, and is researching others in an effort to maximize usefulness. IOOS® partners are expected to offer data in IOOS® formats (legacy formats may also be maintained, if desired). Formats currently in use include:

- Binary format for gridded data: Network Common Data Format (NetCDF) with Climate and Forecast (CF) conventions;
- Text format for in situ data: comma-separated value (CSV) with CF and IOOS-specific conventions; and
- XML format for in situ data: Open Geospatial Consortium (OGC) Geography Markup Language (GML) with IOOS application schema.

In an effort to maximize usefulness, other formats under consideration include OGC Sensor Web Enablement (SWE) Common format, Google Keyhole Markup Language (KML), and Javascript Object Notation (JSON). IOOS® partners are expected to offer data in one or more approved IOOS® formats (legacy formats may also be maintained, if desired). NOAA IOOS® will provide specific guidance to partners on versions for common data formats.

f) Common Vocabularies and Identifiers: IOOS® is presently adopting or researching shared vocabularies for terminology such as names of observed properties, units of measure, coordinate reference systems, animal species, etc. IOOS® is also adopting or defining conventions for identifiers for sensors, platforms, cruises, models, etc. IOOS® data partners are expected to support, and if appropriate to help define, such vocabularies and identifiers, as well as cross-walks (i.e., semantic mediation) between IOOS® and existing local or community-of-interest vocabularies and identifiers.

g) Metadata: Descriptive information about datasets, sensors, platforms, models, analysis methods, quality-control procedures is essential for the long-term usability and reuse of information. IOOS® is developing conventions for use of metadata formats such as ISO 19115 for dataset and collection-level metadata and Sensor Model Language (SensorML) for instrument and procedure metadata. IOOS® partners are expected to ensure relevant metadata is produced, accessible and compliant with IOOS® conventions as adopted, and to participate as appropriate in the development of such conventions.

h) Storage and Archiving: IOOS® partners must describe how they will:

- Provide initial stewardship; and
- Ensure permanent archiving of their observations, model outputs, metadata and derived information products of archival quality. NOAA's National Oceanographic Data Center (NODC) is the nominal archive center for IOOS® data. Applicants are expected to provide for storage of data, metadata and other supporting documentation and algorithm descriptions, to establish data recovery mechanisms, and to perform off-site storage of backups until the data and metadata are submitted to NODC for archiving. NOAA IOOS® will facilitate execution of NODC partnerships with IOOS® partners.

i) Data Management Planning and Coordination: IOOS® partners should maintain the capacity to and demonstrate how they will:

- Participate in all Quality Assurance of Real-Time Oceanographic Data (QARTOD) workshops; and actively participate in cross-regional data management policy and implementation plan development;
- Attend at least one national workshop/meeting per year to facilitate a closely coordinated national approach to IOOS® data management; and
- Participate in national data management committees and forums such as the Interagency Working Group on Ocean Observations' (IWGOO) DMAC Steering Team, the Federation of Earth Science Information Partners (ESIP), Smart Ocean Sensor Consortium (SOSC), among others.

j) IOOS® Maturity Levels and Certification Standards: IOOS® is presently developing capability maturity levels for particular areas in data management such as use of standardized data access services, provision of metadata, system reliability, and others. IOOS® will assess and display maturity levels for IOOS® data and service providers. IOOS® partners shall permit and assist with such assessment, and will be encouraged to achieve as high a level as practical in each category and to describe requirements or obstacles to moving to higher maturity levels if applicable.

k) Consideration for Long-term Operations: The IOOS® observing, data management, and modeling capacities being developed will, ideally, persist as the overall enterprise matures. IOOS® partners should include a discussion of potential plans for maintaining such persistence as part of normal IOOS® operations (e.g., by automating as many activities as possible, implementing operational procedures).