XL CENTRAL AND NORTHERN CALIFORNIA OCEAN OBSERVING SYSTEM (CENCOOS) DATA MANAGEMENT AND COMMUNICATIONS SYSTEM PLAN

1. INTRODUCTION

The U.S. Integrated Ocean Observing System (IOOS) is a nationwide effort to provide access to a wide variety of coastal oceanographic and environmental observations and data. IOOS is a program within the National Ocean Service of the National Oceanic and Atmospheric Administration (NOAA) and a Regional Alliance of the Global Ocean Observing System. The Central and Northern California Ocean Observing System (CeNCOOS) was established in 2004 by a consortium of partner agencies and research institutions to serve as the regional association for integrating coastal and ocean observing activities from Point Conception north to the California-Oregon border. The CeNCOOS region extends 200 nautical miles offshore of the coastline (the seaward extent of the Exclusive Economic Zone) and includes bays and estuaries. Along with 10 other regional associations (RAs), CeNCOOS is designated a Regional Coastal Observing System (RCOS) under the authority of the Integrated Coastal and Ocean Observation System Act of 2009 (ICOOS Act) and amended by the Coordinated Ocean Observations and Research Act of 2020 (COORA Act).

As a member of IOOS, CeNCOOS has a mandate to collect, organize, and provide access to regional oceanographic data. These data need to be quality reviewed, easily understandable, easily discoverable, electronically accessible, and well organized to allow researchers, policy makers, industry, and the general public to make well-informed decisions. To satisfy this mandate, CeNCOOS supports a web-based data portal for the entire region providing ocean, coastal, and relevant watershed environmental data and information products.

The goal of the CeNCOOS data management system is to curate multiple data streams from the sensors and models supported by CeNCOOS as well as from independent data providers, document the data using IOOS-approved metadata standards, provide these data to users via standard services, and archive the data in appropriate long-term archives. The CeNCOOS Data System is based on a service-oriented architecture that employs interoperable systems to enable data discoverability via web services and catalogs. The vision of CeNCOOS data management is to be recognized in the ocean observation community as a trusted leader in data quality, interoperability and discoverability.

CeNCOOS partners with Axiom Data Science (referred to hereafter as Axiom) to provide a standards-based lifecycle data management framework that maximizes the discoverability, accessibility, and usability of data and information products and ensures their sustained use.

CeNCOOS leverages Axiom's data systems that also support AOOS and SECOORA to use common infrastructure which enables the dedication of more funds to system advancements and innovation than would otherwise be possible. The relationship between CeNCOOS and Axiom is a collaborative partnership designed not only to serve the needs of CeNCOOS, but also to allow for greater contributions to the larger IOOS community. CeNCOOS works closely with Axiom to develop and update data management plans, statements of work, facilitate the flow of data, and ensure a coordinated end to end system. The standards and protocols, and annual work plans of the CeNCOOS DMAC Sub-system are revised annually by the CeNCOOS DMAC Sub-system Management Team and reviewed by the CeNCOOS Data Management Committee. Axiom implements recommended and standard practices as defined by the IOOS Data Management and Communications (DMAC) committee and more specifically those in the Guide for IOOS Data Providers (https://ioos.noaa.gov/data/data-standards/). These practices apply to data archive, data discovery, data serving (web-based browsing), data transport (binary access to data), metadata, information technology (IT) security and data quality assurance and quality control (QA/QC). In addition, the system meets CeNCOOS data management practices that are guided by regional needs with a focus on quality, interoperability, trust, and discovery.

CeNCOOS provides access to catalog-level information and to data sets via web services and the CeNCOOS Data Portal. Both web services and the portal are administered by Axiom. The main web service uses Open-source Project for a Network Data Access Protocol (OPeNDAP), built around the Thematic Real-time Environmental Distributed Data Services (THREDDS) Server (TDS). Axiom maintains an Environmental Research Division's Data Access Program (ERDDAP) server that also provides DAP and web map service (WMS) capabilities. Axiom maintains a Geoserver for geospatial data. Geospatial datasets are accessible via Open Geospatial Consortium (OGC) web feature service (WFS) and WMS services. Some data streams are provided to the Global Telecommunications System (GTS) by providing the data via an intermediary such as National Data Buoy Center (NDBC), which posts the data to the GTS.

CeNCOOS maintains a website (https://www.cencoos.org) which acts as a higher level interface for data discovery and access (i.e. linking to the Data Portal). The site provides digestible project descriptions and documentation and serves as a platform to communicate current events and stories of interest for the region. This space allows CeNCOOS to bridge science communication with data discovery, by linking directly to the Data Portal within articles. The CeNCOOS website is built using the content management system Drupal and is administered by CeNCOOS program staff.

The CeNCOOS web and social media provide our broadest outreach, serving data and information products to a region-wide audience. Both the web and social media provide

two-way communication. Users can comment on the information and data products, provide feedback and suggestions, and thus shape the flow of information so that it meets their needs. Our help-desk ticket system tracks and organizes our response to feedback. CeNCOOS maintains existing product development and communications activities, internet-based education and outreach, and continues the support for investigators at universities and their associated outreach activities. CeNCOOS produces and implements a communications plan that identifies education and outreach materials to be produced.

For each data stream, dissemination (access nodes) is described by one or more of the following terms: GTS, THREDDS, ERDDAP, WMS, CeNCOOS Data Portal, CeNCOOS website.

CeNCOOS is implementing recommended and standard practices as defined by the U.S. Integrated Ocean Observing System (IOOS) Data Management and Communications (DMAC) committee, with specific consideration to meet the core capacity requirements outlined in Contributing Data to IOOS. This will ensure data collected by CeNCOOS and member entities is distributed on the CeNCOOS web site and are managed according to best practices identified by NOAA/US IOOS. This also ensures that appropriate metadata and QA/QC practices are followed and that the data are of a known quality to the end user. These practices apply to data standards, metadata and data, transport and access, archival, information technology (IT) security, quality control and quality assurance, described in the NOAA IOOS Program Office DMAC White Paper (v1.0), and data management and communications DMAC requirements for IOOS Regional Associations and other IOOS grant recipients who are providing data to IOOS.

The CeNCOOS Data Management and Communications System (referred to hereafter as the CeNCOOS DMAC Sub-system) must adhere to these practices, and this CeNCOOS DMAC Sub-system Plan provides the approach to the necessary implementation, describing how data are ingested, managed and distributed from the source to public dissemination. The CeNCOOS DMAC Sub-system Plan is organized as follows:

- Section 2 provides an overview of the CeNCOOS Data Management and Communications, describing: the function, goals, and objectives of the CeNCOOS DMAC Sub-system and, details related to the CeNCOOS DMAC Sub-system Management Team roles.

- Section 3 briefly describes the CeNCOOS data resources, defines data categories and asset types, and describes how the data categories are handled in the plan.

- Section 4 presents the CeNCOOS DMAC Sub-system statement of work and includes descriptions of the system computing infrastructure including details about the

processes related to data ingestion, standards for format and content, metadata and data discovery, quality control procedures (including procedures for data that cannot undergo quality control) and flagging protocols. Additionally, this section covers policies for stewardship, public access and dissemination, data archival and preservation, and data system performance and security measures.

This document, unless superseded, pertains to a period of five years from December 1, 2023 through December 1, 2027.

2. CeNCOOS DMAC SUBSYSTEM

The mission of the CeNCOOS DMAC Sub-system, is to acquire, archive, and share coastal and marine data and information products to meet the needs of CeNCOOS stakeholders and the national US IOOS program. CeNCOOS uses a data management system that allows a complex array of oceanographic and environmental data types to be well organized, discoverable, accessible, and understandable. The CeNCOOS DMAC Sub-system employs a distributed data management approach, which allows data to seamlessly interchange between participating agencies. The system is composed of an internal master node coupled with external data provider nodes. External data providers include stakeholders, partners, and CeNCOOS funded projects who produce, and manage and share data. This distributed configuration increases capacity and technical knowledge within individual groups, allowing them to better meet their own internal data management goals. The distributed architecture leverages hardware, bandwidth, and staff resources across multiple systems and groups. Utilization of currently available external data feeds for sensor, remote sensing, and other data sources improves access to data for CeNCOOS users with minimal effort.

Integrating available interoperable data feeds into data access applications and data management systems adds a variety of resources at a low cost. Large quantities of real-time and historical sensor information, remote sensing satellite information, and marine habitat and biological data for the CeNCOOS region are openly available for use through interoperability protocols. For example, National Aeronautics and Space Administration (NASA) Earth Observations (NEO) provides an expansive array of long term oceanographic, climate, and atmospheric remote sensing datasets. Real-time and historical sensor data feeds for the CeNCOOS region are available for hundreds of sensors via the NDBC, the Center for Operational Oceanographic Products and Services (CO-OPS), National Estuarine Research Reserve System (NERRS) and other NOAA programs. Additional sources of interoperable data include those hosted at NASA's Jet Propulsion Laboratory (JPL), U.S. Geological Survey (USGS) TerraServer, and

other research organizations. CeNCOOS integrates all of these data and make them available on the CeNCOOS Data Portal.

2.1 DMAC SUB-SYSTEM MANAGEMENT GOALS AND OBJECTIVES

The CeNCOOS Director, Data and Information Manager, and Product Developer, as well as Axiom comprise the DMAC Sub-system Management Team and are tasked with fulfilling the primary goals and objectives within the CeNCOOS Data Management Plan.

Goal 1: Provide Core Data Management Support to the CeNCOOS Program

1. Provide Technical Support for CeNCOOS cyberinfrastructure.
2. Develop and maintain web-based data portal.
3. Deliver real-time, delayed-mode and historical data for *in-situ* and remotely-sensed physical, chemical and biological observations.
4. Deliver model-generated outputs, including both nowcasts/forecasts and reanalysis, to CeNCOOS users.
5. Implement U.S. IOOS Quality Assurance of Real-Time Oceanographic Data (QARTOD) QA/QC checks for CeNCOOS real-time data feeds.
6. Archive data to the federal archives (e.g. National Centers for Environmental Information (NCEI)).
7. Provides system performance and security measures.

Goal 2: Provide DMAC support to the CeNCOOS program

1. Provide overall DMAC project management and oversight.
2. Participate in regional, state, national and international DMAC activities.
3. Engage with data providers to access, understand, and appropriately document data (metadata and QA/QC) that is ingested through the CeNCOOS infrastructure.
4. Facilitate communication between different data providers to leverage regional

and technical experience.

5. Participate in regional committees and task teams (including teams as determined by the Director, and the joint State-Federal Data Integration Initiative) in order to facilitate data integration and interoperability within the region.

6. Participate in national and cross-regional committees, workshops and task teams in order to further the development of a coordinated approach to IOOS data management.

7. Work closely with the CeNCOOS office, other data management awardees if selected, and appropriate advisory committees to implement identified user products, tools and web interfaces; develop product requirements; and beta test and refine products in order to increase their utility.

8. Provide reports as requested.

9. Develop detailed work plans with measurable timelines, deliverables, and performance metrics; and assist with proposal development.

Goal 3: Web Hosting and Support

1. Host and maintain the CeNCOOS Data Portal at https://data.caloos.org

2. Provide access to the user interface and visualization tools, data products, data query and access tools, decision-support tools, agency project tracking systems and databases, as well as IOOS Registry tools.

3. Work with CeNCOOS staff, CeNCOOS PIs, and member organizations to update the Data Portal periodically, in order to improve access to data, ingest new data, develop new tools, improve clarity and ease of use, and the overall “look and feel.”

2.2 DMAC Sub-system Management Team

The CeNCOOS DMAC Sub-system Management Team consists of the CeNCOOS Director, Data and Information Manager, and Product Developer and the Axiom Data Team (CeNCOOS PO staff members), the DMAC Sub-system Lead (Axiom Data Team), and the Data Management Technical Lead (Axiom Data Team). CeNCOOS Staff and subcontractor employee information and

CVs are included in Appendix A - DMAC Personnel Resumes. Additional information about management roles and responsibilities, including the advisory committee's role, is provided below.

The CeNCOOS DMAC Sub-system is advised by the DMAC committee, a panel of regional experts that advise CeNCOOS staff on data management practices and implementation on an ad hoc basis. The DMAC committee is composed of experts from partner institutions who serve on a voluntary basis. The co-investigators on the IOOS award and other grants received by the host institution have data management responsibilities that are described in the Axiom proposal sub-award statements of work.

2.2.1 Roles and Responsibilities

Director (Henry Ruhl, CeNCOOS). The CeNCOOS Director manages operations for RICE certification and standards compliance, of which the CeNCOOS DMAC Sub-system is a critical aspect and has direct oversight of the Axiom sub-award effort.

Data and Information Manager (Fred Bahr, CeNCOOS). The CeNCOOS Data and Information Manager is responsible for developing data management plans and statements of work with Axiom, serving as the project manager and as the CeNCOOS point of contact for the Axiom Data Team. The Data and Information Manager primary responsibilities are to facilitate data ingest, website integration, and general data management coordination. The Data and Information Manager works with the Axiom Data Team and data providers to ensure a coordinated system from data ingest through data visualization and dissemination, and integration with the CeNCOOS website. The role has direct responsibility for the acquisition, curation, and delivery of CeNCOOS data. The role is also responsible for coordinating data management planning and implementation with the IOOS program office and other regional associations, and in this role attends the annual DMAC meeting and other IOOS data management coordination activities (e.g., monthly calls, report and standards review and development). The Data and Information Manager liaises with CeNCOOS-supported data producers to ensure that appropriate guidelines are followed. The ultimate responsibility for the information management conducted by CeNCOOS-supported investigators rests with the CeNCOOS Director who implements the sub-awards.

Data Specialist (Marine Lebrec, CeNCOOS). The data specialist helps manage the integration of regional oceanographic datasets into CeNCOOS web products (data portals, websites) to meet stakeholder needs. The data specialist also coordinates with ocean observing networks (MBON,

iDOOS) to implement FAIR data best practices and comply with QA/QC standards. The data specialist works closely with the data and information manager and regional partners to shepard data sets into the DMAC.

DMAC Sub-system Lead (Rob Bochenek, Axiom Data Science). The role oversees the Axiom Data Team and the CeNCOOS DMAC Sub-system. The DMAC Sub-system Lead together with Axiom's Director of Programs, Stacey Buckelew, contribute to proposal development and general CeNCOOS data management reporting requirements. The DMAC Sub-system Lead is the main point of contact for all technical data-related questions and is an expert in managing large scale datasets related to the CeNCOOS mission. The DMAC Sub-system Lead submits quarterly reports to the CeNCOOS Director.

Data Management Technical Lead (Shane St Savage, Axiom Data Science). The role is responsible for implementing IOOS-recommended technologies for the collection, curation, delivery, and archive of CeNCOOS data. The technical lead advises CeNCOOS on the application of technologies that meet user and stakeholder needs and oversees their implementation.

Data Management Committee. The Data Management committee advises the CeNCOOS Director and provides guidance to CeNCOOS on both a strategic planning and operational level, regarding the acquisition, curation, and delivery of CeNCOOS data. The committee consists of data scientists and informatics experts who are familiar with coastal marine science data, standards and protocols for data stewardship (including IOOS recommended protocols) and familiar with existing and emerging technologies for the browsing and visualization and delivery of scientific data.

2.2.2 Communication

The Axiom Data Team meets with the Data and Information Manager weekly to provide updates on projects status and timelines and communicate any task related issues. Projects involve a kick-off meeting with CeNCOOS providing initial guidance, defining project scope and setting project milestones. Axiom personnel meet in person with CeNCOOS in Moss Landing at least twice during the year. The Axiom Data Team participates in at least one regional meeting (i.e. State, West Coast Ocean Partnership, or other conference) in addition to the IOOS/CeNCOOS meetings. Axiom maintains relationships with regional ocean data management groups (e.g. WCOP and State of California). The Data and Information Manager and the Axiom Data Team

both attend the annual IOOS DMAC meeting, in addition to maintaining communication with IOOS to keep up to date on national IOOS office expectations.

2.2.3 Axiom Data Team

The Axiom Data Team is composed of staff from Axiom and they are involved with all aspects of the CeNCOOS data flow, including data ingestion, creation of metadata, conversion, discovery, maintenance of data feeds, storage, and any necessary archival services. Its primary goal is to gather and serve data important to CeNCOOS end users via standard services as recommended by the IOOC and the IOOS Program Office (e.g., OPeNDAP, ERDDAP, THREDDS, etc). It is also tasked with managing and archiving any CeNCOOS-funded and non-funded data generated by oceanographic models, buoys, or other devices to enable generation of data products.

The Axiom Data Team is responsible for the design and deployment of a DMAC Sub-system to meet the needs of the CeNCOOS user-base. This system must provide the functional components required by IOOS RCOS as described in this plan. The team offers comprehensive technical solutions to data management needs, underpinned by a scalable, open source system that uses existing and emerging software, high performance computer clusters, and interoperability services. This data management system provides an environment that increases the access and use of data by all user groups and allows data management staff to rapidly develop new capabilities and tools to meet emerging user needs.

The Axiom Data Team is dedicated to providing data management and informatics support for CeNCOOS, and provides development capabilities for map-based data portals, spatial planning tools, and data management frameworks which transfer and ingest data from external systems via interoperability protocols. The team ensures transparency and communication between client and contractor about design requirements and development progress, and continually researches and employs new technologies to extend the capabilities of digital information and computer analysis systems.

CeNCOOS and Axiom personnel maintain regular communication with the U.S. IOOS Program Office through a variety of mechanisms including in-person meetings, phone calls & webinars, email conversations, and GitHub repositories. The continuous communication ensures that the DMAC team and Axiom is aware of all new practices and protocols, as promulgated by the Interagency Ocean Observation Committee (IOOC) and the IOOS Program Office, and understands how and when to implement them.

2.3 CeNCOOS Procedures for Evaluating the DMAC Sub-system

The CeNCOOS DMAC Sub-system Management Team contracts for professional data management services from Axiom. The Axiom Data Team selection followed a process of broadly soliciting competitive proposals to provide web portal, data management, communication and user product services for CeNCOOS for up to 5 years. The Axiom employee CVs have been evaluated and new CenCOOS DMAC Sub-system contributors will be evaluated before starting in such a role. CeNCOOS procedures followed during the solicitation, evaluation, and selection of contractual data management support are fully described in CeNCOOS Framework for Decision Making. CeNCOOS solicited proposals for two sets of services, and proposers were encouraged to bid on one or both of these components, separately or combined.

I. Data Management and Communication Services: Services include providing data management support (data ingestion, metadata, relational database development and maintenance) and communication services (web portal, data clearinghouse, coordination, and communication), building upon the hardware, software, query tools and products developed over the previous years, and following the national IOOS Program guidance.

II. User-Driven Product Development Services: Develop user-driven products and associated interface and visualization tools that will be maintained by and interact with the data system developed under component I above.

The Axiom contract is administered by the CeNCOOS Director and implemented by the CeNCOOS Data and Information Manager. The Axiom contract scope of work lists the services which the contractor must be able to provide and this in turn is used to evaluate the quality of work provided by the contractor. The CeNCOOS Director coordinates quarterly evaluations of the DMAC services contract including discussion of Axiom employee performance. The Axiom subcontract is also reviewed annually as part of the Governing Council program review. Annual reviews take place directly following CeNCOOS Fall Science Impact meetings. The CeNCOOS Director will use recommendations from the program review to inform future subcontracting agreements. See Framework for Decision-making Section 13 (page 8). In addition the Data and Information Manager has weekly meetings with and progress evaluation discussions

The Data and Information Manager is hired by the CeNCOOS Director and reports directly to him/her with annual evaluations. Continued employment of all CeNCOOS Program Office staff is dependent upon responsible execution of the duties incumbent to the position they hold and in accordance with MBARI's personnel manual (not publically available).

3. CeNCOOS DATA RESOURCES AND ASSET TYPES

The CeNCOOS DMAC Sub-system provides data to the public from multiple sources including CeNCOOS-funded projects and data from numerous and diverse external federal and non-federal organizations.

3.1 Observational Data Types

The CeNCOOS data inventories include multiple types of data, including real-time data, near real-time data, and historical data. CeNCOOS defines each data type in a consistent manner with IOOS Guidelines as follows:

- *Real-time data* are ingested, served, and displayed by the CeNCOOS DMAC Sub-system at the same frequency the data are collected (and sometimes reported) by the originator with little to no delay. Real-time assets primarily include shore stations, HF Radar, gliders, oceanographic buoys and numerical model data.
- *Near real-time data* are ingested by the CeNCOOS DMAC Sub-system at the same frequency that the data are made available; however, there is some delay (hours to days) between data collection and when the data provider makes it available. Examples of near real-time assets include satellite images and derived satellite products.
- *Historical data* are data that are one month old or older. Historical data are sometimes collected in real-time and then archived, and sometimes ingested from local or national archives on request.

3.2 Data Categories

The CeNCOOS data types are divided into five major categories that determine the level of documentation and quality control (QC) that is required for the data assets within each category:

1. Federally Sourced Data
2. Model Products
3. Static Data Products
4. Funded Data Streams
5. Regional Partners Data Streams (not CeNCOOS funded)

3.2.1 Federally Sourced Data

Federally sourced data incorporated into the CeNCOOS Data Portal are quality controlled following rigid data management and archival processes by the federal agency collecting the data. These data only require generic documentation by CeNCOOS on how these data are ingested and made available to the public (Section 4). As of the writing of this plan, federally sourced data served by the CeNCOOS DMAC Sub-system are all exempt from detailed data stream documentation. Federal sources include the National Oceanographic and Atmospheric Administration (NOAA), the U.S. Geological Survey (USGS), the National Weather Service (NWS), and the National Estuarine Research Reserve System (NERRS). See the Appendix B - Federal Data Product Inventory for a complete list.

3.2.2 Model Data Products

Model outputs and products served by the CeNCOOS DMAC Sub-system may incorporate or assimilate observational data (e.g., all bathymetric charts served by CeNCOOS are from gridded models derived using “true” observations). These models are considered a product that falls outside the realm of “true” observations therefore they are exempt from detailed data stream documentation. See the Appendix C - Model Data Product Inventory for a complete list.

3.2.3 Static Data Products

CeNCOOS static data products are typically derived from observed data, but are displayed in a way that the original data are no longer reproducible and cannot be used to assemble a numerical observational dataset in time or space. Other types of static data products are merely representations of fixed political or legal boundary information. These products fall outside the realm of CeNCOOS observations and are exempt from detailed data stream documentation. See the Appendix D - Static Data Product Inventory for a complete list.

3.2.4 Funded Data Streams

Data funded by CeNCOOS fall into their own category. The primary processes involved with data management include data ingestion, standards and format, metadata and discovery, quality

control, stewardship and preservation, access and dissemination, archival and security. Descriptions of the processes that consistently apply to all data streams are provided in Section 4. Additional data management documentation unique to individual data streams are provided through a systematic Data Stream Plan template that follows the RCOS Certification Guidance DMAC requirements (section 1-6) and the NOAA Data Sharing Template. Use of a custom Data Stream Plan template facilitates consistent documentation, and streamlines future additions and edits to existing data stream protocol.

The Data Stream Plans use a consistent and comprehensive set of questions designed to describe how data streams are handled and managed end to end. Grouped variables may originate from a single platform type (e.g., a mooring that provide temperature, salinity, and dissolved oxygen data, all of which are treated in a standard way); a data type that is handled similarly across all platforms (e.g., webcam imagery); or originate from a single data source (e.g., Moss Landing Marine Labs or Monterey Bay Aquarium Research Institute).

Quality Control descriptions included in the Data Stream Plans may follow one of four paths for a given data stream:

1. Follows prescribed QARTOD guidelines (required for real-time data only if a QARTOD Manual exists for the variables in the data stream).

2. When QARTOD guidelines do not exist, some other suitable form of QC implementation is conducted and described;

3. A description of the QC completed by the data provider (e.g., brief description or link to QC protocols performed at the source).

4. Data are considered exempt from QC documentation or requirements if federally sourced.

See the Appendix E - Funded Data Stream Inventory for a complete list including links to individual Data Stream Plans.

3.2.5 Regional Data Streams

Regional Data Streams are defined here as any data resource that does not fit into the exempt categories already discussed: federally sourced data, model product, and static data product and are not funded by CeNCOOS. These include, for example, regional data provided by local or

state agencies, private companies supporting maritime activities in coastal waters, university projects, and research studies funded and conducted by local entities. None of these data streams served by the CeNCOOS DMAC Sub-system originate directly from CeNCOOS funding. Most regional data originate from sole source providers affiliated with other entities (research, private, NGO, etc.). Occasionally, a federally sourced data asset is manipulated in some fashion prior to display and, therefore, requires documentation (e.g., federal satellite data that is transformed from a NSIDC-binary format into netCDF). Data streams may be of any data type: real-time, near real-time, historical, or citizen science. Leveraged projects in which CeNCOOS helps support but does not fund may also fit into this category. Due to the external nature of these data streams they are usually exempt from detailed data stream documentation. On occasion, however, a data stream that would normally be considered exempt will require documentation in a Data Stream Plan:

- Data products that include representations that can be used to reproduce numerical data in time or space are considered observing data, are treated as a Regional Data Stream and are further documented in a Data Stream Plan.
- A federal data source that is translated or transformed in some way between the source at ingestion to the CeNCOOS access point of delivery (e.g., smoothing, block averaging).

See Appendix F - Regional Data Stream Inventory for a complete list including links to individual Data Stream Plans, as applicable.

4. CeNCOOS DMAC SUB-SYSTEM ARCHITECTURE AND WORK PLAN

4.1 Computing Cyberinfrastructure

CeNCOOS DMAC Sub-system employs a framework for managing a variety of ocean data types (*in-situ* and remotely sensed data streams, multidimensional grids, geographic information system (GIS), and other structured formats). This framework, developed by Axiom, exposes managed data through interoperability systems and uses several user interface tools that allow the data to be discovered and explored by the broader community. Use of this framework to power the CeNCOOS DMAC Sub-system enables CeNCOOS to rapidly ingest or connect to data sources relevant to CeNCOOS and develop advanced user tools and data products efficiently.

The system is divided into four tiers, which separate the suite of technologies composing the

system. See Figure 2 for a diagram of the system.

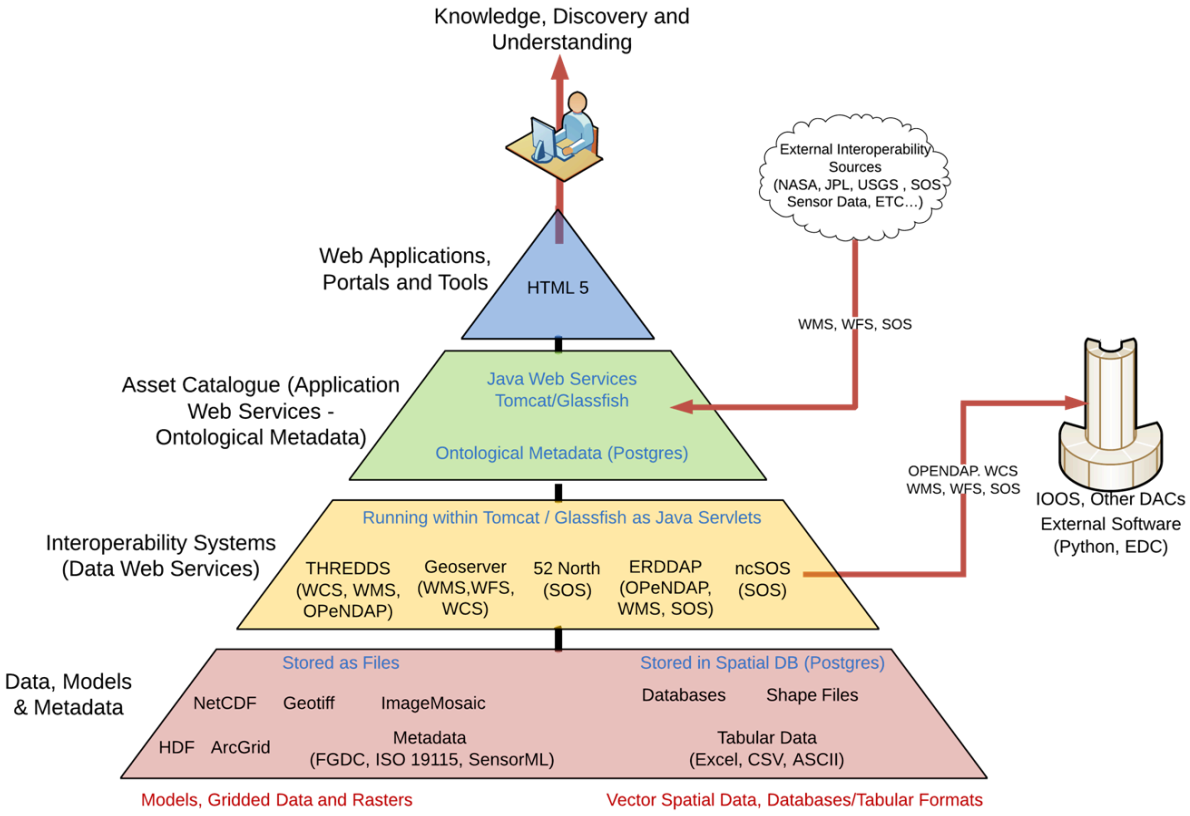

Figure 2. CeNCOOS DMAC Sub-system framework showing the flow of data through logical technology tiers, enabling discovery of data that enables understanding the ocean and coastal environments.

CeNCOOS partners' data, models, and metadata are ingested autonomously into the back-end data system through a series of harvesting mechanisms written in Java, Scala, and Python that make use of lower-level interfaces (e.g., file transfer protocol [FTP], hyper text markup language [HTML] and ad hoc service application program interfaces [APIs]). Data files are processed during the ingestion process and loaded into a clustered file storage and database system (GlusterFS and Postgres). A suite of interoperable systems connect to the data storage, including GeoServer, THREDDS, ERDDAP and ncWMS , and they expose data feeds through WFS, WCS, WMS and OPeNDAP protocols. The CeNCOOS asset catalog is a database containing ontological information describing the dimensional characteristics (space, time, unit, measured parameter and taxonomy) of each known data resource and how these characteristics relate to each other across data sets. References to both internally- and externally-hosted data feeds are stored in the ontological database and provide the user with a harmonized set of interfaces for consistent

access to data and visualizations. Sensors, numerical model output, and remotely sensed observational grids are mapped to common characteristics (space, time, and climate forecast parameter) for comparison across sources. Data sets are further mapped across keywords and, if applicable, Integrated Taxonomic Information System (ITIS) records. The asset catalog also exposes web services providing external access to metadata in the database and provides a method for indexing metadata across multiple formats and types using ElasticSearch, a scalable, Apache Lucene based, clustered search engine. The underlying system architecture works together to allow users to rapidly discover, access and use data through web-based applications and tools developed using modern web development languages and libraries.

4.2 Data Ingestion

Observations and information are ingested into the CeNCOOS DMAC Sub-system from a variety of sources, including both historical and real-time observations, forecast, nowcast, and hindcast model outputs, GIS information, and synthesized products that can be useful for layering with other data in the CeNCOOS DMAC Sub-system. Each data asset ingested into the system has its own level of data processing maturity and quality with respect to the metadata available.

Data has the ability to be ingested into the system using one of several pathways:

1. Contribution by the originator
2. Direct access or harvest from the originator website (real-time sensors)
3. Auto submission pathway from the Research Workspace

CeNCOOS -funded partners provide data to CeNCOOS in a timely manner, stipulated in the US IOOS descope proposal. When possible, data are served in real-time. In cases where projects do not produce real-time data, the project PIs are responsible for making sure data becomes accessible by CeNCOOS as soon as possible.

4.3 Standards for Format and Content

4.3.1 Shared Data File Formats

CeNCOOS provides nearly all data in four open and standardized forms:

1. *Network Common Data Form (NetCDF*) - a self-describing, machine-independent data

format that CeNCOOS uses primarily for raster (gridded) data. Some data stored as unstructured grids use this format as well.

2. *Comma Separated Values (CSV)* - a human-readable ASCII format that is nearly universally accepted by spreadsheet and programming languages. CeNCOOS uses CSV formats to allow users to download: (1) time-series extractions from raster data, and (2) GIS vector and polygon information (e.g., boundaries).

3. *Shapefile* - an open geographic information system format for point, vector, and polygon data. CeNCOOS allows users to download shapefiles of static GIS layers such as boundaries, biologic distributions, etc.

4. *Portable Network Graphics (PNG*) - PNG is a lossless image format provided as an alternative to shapefiles in the CeNCOOS catalog. PNGs are limited in use as they are pre-projected, pre-scaled, and pre-sized images of data layers. CeNCOOS provides PNG files as example WMS requests, which are useful to users who cannot access GIS services and who do not understand how to manipulate WMS requests.

4.3.2 Data Access Points

Access points provide standardized, documented services that allow users to download data from CeNCOOS without having to make person-to-person data requests. These standard services are provided through four main platforms:

1. Thematic Realtime Environmental Distributed Data Services (THREDDS) - THREDDS is a set of services that allows for machine and human access to raster data stored in NetCDF formats. THREDDS provides spatial, vertical, and temporal subsetting as well as the ability to select individual dimension or data variables to reduce file transfer sizes. CeNCOOS provides THREDDS access points for raster (gridded) data and discrete time-series observations stored in NetCDF format.

2. Environmental Research Division's Data Access Program (ERDDAP) - ERDDAP is a common data server that provides access to subsetting and downloading data in a variety of formats. CeNCOOS provides ERDDAP access to all time-series data in the region, a subset of gridded data, and some GIS-data based products.

3. GeoServer - GeoServer is used to serve image tiles and provide download formats for tabular GIS data.

4. ncWMS - ncWMS is used to serve image tiles for gridded datasets (NetCDF).

Service protocols provided by these platforms include:

1. Open-source Project for a Network Data Access Protocol (OPeNDAP) - OPeNDAP is a protocol that can transfer binary or ASCII data over the web. Like THREDDS, it provides spatial, vertical, and temporal subsetting and the ability to select individual variables to reduce file transfer sizes. Unlike THREDDS, requested data are provided as non-NetCDF, structured output. OPeNDAP output can be imported directly into graphical programs such as GrADS, Ferret, or R. CeNCOOS provides OPeNDAP access points for raster and time-series data.

2. Environmental Research Division's Data Access Program (ERDDAP) is a common data server that provides spatial, vertical, and temporal subsetting and the ability to select individual variables to reduce file transfer sizes. ERDDAP provides the user with multiple options for output of data served by CeNCOOS. These outputs allow the user to obtain the data in the format that they prefer. This server provides access to all but gridded and raster data.

3. Web Map Service (WMS) - WMS provides machine access to images, which can be used by individuals or programs (e.g., tiling services). Accessing programs use GetCapabilities requests to ask for image data in whatever format they require, which allows them to gather image tiles over specific areas with the projections, styles, scales and formats (PNG, JPG, etc.) that fits their needs. CeNCOOS provides WMS access points for point, vector, and polygon information, as well as raster data.

4. Web Feature Service (WFS) - this service provides machine access to the vector elements of static layers. CeNCOOS provides WFS access points for point, vector, and polygon information, as well as time-series and raster data.

CeNCOOS also provides data as downloadable files including NetCDF, CSV, and JSON formats, usually by leveraging services of the platforms described above. Project-specific data are served in their native file formats.

The flow of data from the source to CeNCOOS Data Portal follows the same general path for all sources as illustrated in the following flow diagram (Figure 3). For cases where the data are transformed or modified in any way, an explanation is provided in the individual Regional Data Stream Plans for that particular instance. This includes format translations or aggregations of component data streams into an integrated product.

Though CeNCOOS relies on local investigators to provide best practices for QA on their activities related to data submitted to CeNCOOS, part of the data ingestion process is to establish adequate metadata and provide metadata links that provide the necessary background information to establish the purpose of the data and expected quality.

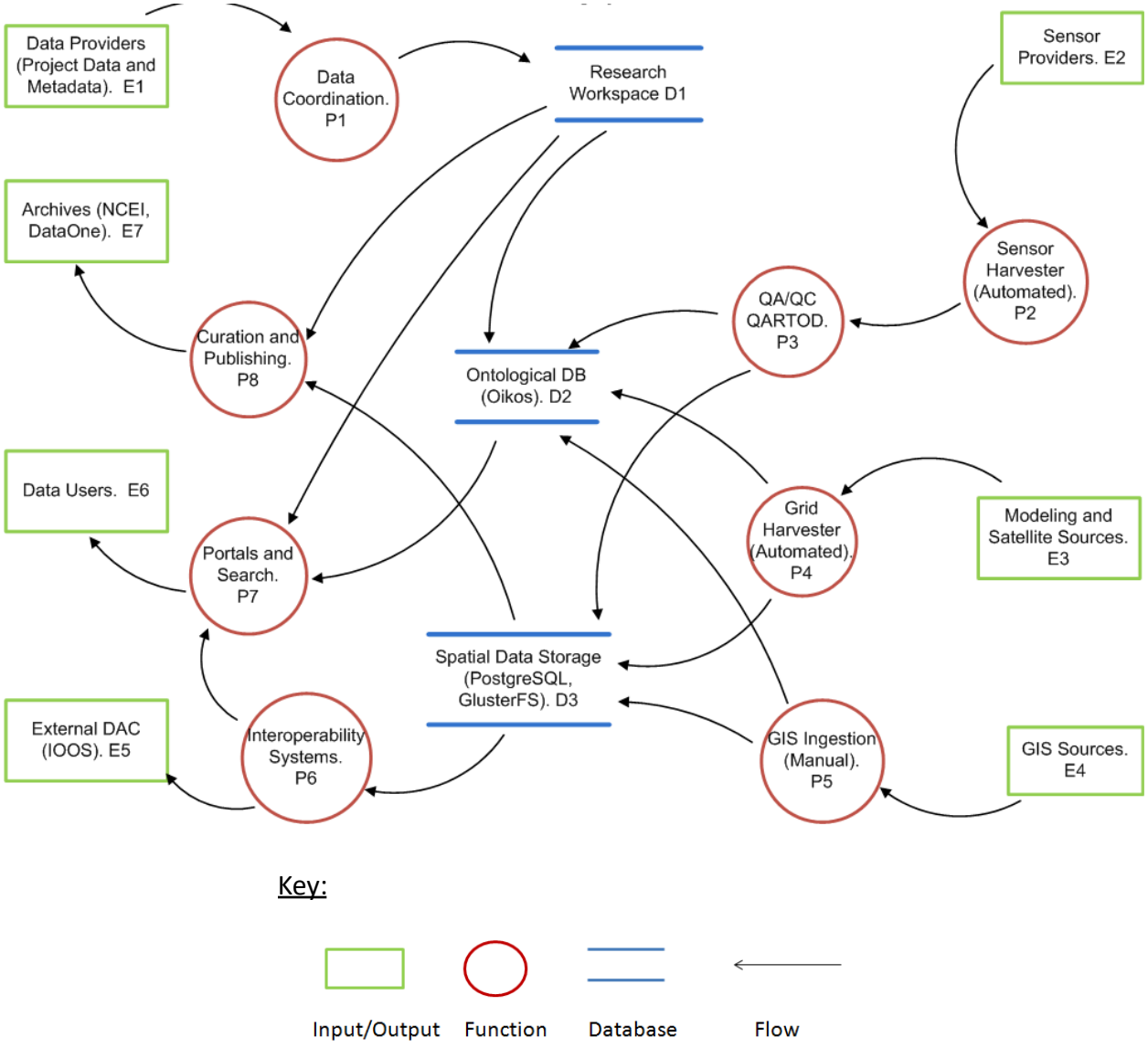

Figure 3. The data flow diagram is a graphical representation of the "flow" of data through the CeNCOOS DMAC Sub-system, modeling its process aspects. The diagram makes use of the Yourdon/DeMarco notation. From right to left: the data sources are inputted to the system (green squares), advance through the system using various processing functions (red circles), are stored within internal databases (blue lines), and are processed and transformed in a standardized manner (red circle) to serve publicly-accessible information products to various

outputs (i.e. end users; green squares).

The Data Portal and underpinning infrastructure and service platforms are managed by Axiom Data Science; resources powering these services are leveraged among other regional associations (AOOS and SECOORA).

4.4 Metadata and Data Discovery

CeNCOOS requires standards-compliant metadata for project-level data (CeNCOOS or IOOS-funded projects). Though CeNCOOS does not require specific metadata standards for ingesting other types of data, most modern data submittals are accompanied by standard ISO/FGDC metadata records. However, many older data sets come with informal metadata documentation that is variable in terms of completion and detail required by modern standards, and some are only accompanied with narrative information. In these cases, CeNCOOS works towards making the source information easily accessible to the end-user by providing links to source data or data providers, and making all available metadata information that came with the data available in the CeNCOOS data catalog. Details and availability of metadata are discussed in individual Data Stream Plans.

4.5 Quality Control Procedures

A primary mission of CeNCOOS is to serve as a regional data assembly center (DAC), aggregating data from local and federal sources and making them available, accessible, and understandable to the public. Quality assurance (QA) relates to procedures undertaken during the experiment and/or instrument design phases of data collection, ensuring that all the data collected are as accurate and precise as possible. Providing very few data collection devices itself, CeNCOOS is reliant on individual data providers to provide adequate QA procedures, and they will not be discussed in this document. CeNCOOS Principal Investigators maintain equipment inventories, shipping logs and instrument history logs for equipment owned and/or operated by CeNCOOS. All CeNCOOS instruments are calibrated, validated, operated, and maintained in accordance with manufacturer's guidance and as recommended by the principal investigators responsible for the equipment.

Quality control (QC) processes implemented by CeNCOOS are used to identify and flag or remove bad or suspect data after data collection. Sharing these protocols and quality flags are an important component of publicly serving data.

4.5.1 CeNCOOS Implemented QC Protocols

CeNCOOS does not receive or serve any raw data transmitted directly from stations, so any applied QC procedures administered by CeNCOOS are in addition to those applied by the data provider. For many historical datasets, CeNCOOS provides the same data (though sometimes in converted formats) that are available from the source provider, which in many cases is a requirement stipulated by local providers. Any QC procedures that are documented and made available to CeNCOOS by the providers are included in the QC section of the individual Data Stream Plans.

CeNCOOS currently uses the IOOS QARTOD library (ioos.github.io/ioos_qc) for the station data. If a shore station operator has QC procedures those take precedence of the use of QARTOD as they are understood to be the experts concerning their data. CeNCOOS currently applies the QC procedures listed below to real-time and historical observation data before it is served by the CeNCOOS DMAC Sub-system. These tests include the following:

1. *Syntax Test:* regional data sources use various techniques and technologies to make data available. Some have standardized data storage protocols and provide services (e.g. using ERDDAP or custom CSV endpoints), while others simply present data on an html web page or hosted text file. Therefore, each regional source requires a custom syntax test, which merely checks for parity errors by testing if data can be extracted from the downloaded or scraped data. If no data can be extracted, the test fails.

2. *Gross Range Test:* this test checks data values against minimum and maximum values defined for each parameter. In addition to variables outlined in QARTOD manuals, CeNCOOS includes gross range tests for variables not covered by QARTOD. Values outside of the prescribed parameter ranges are rejected and replaced with missing value flags in data storage connected to access points and the graphic displays. See Appendix G - Gross Range Test Table for the current values used in Gross Range tests.

3. *Timing Gap Test:* CeNCOOS also implements a type of “time-gap check" that informs observational assets (e.g., weather stations) displayed on its "Real-Time Sensor Map". If no data are received from an existing observational station for four hours, the icon on the map changes from a scaled color to a small grey dot. If no data are received from an existing observational station for one week, the asset is automatically removed from the map, but assets are still made available on a historical sensor map. This “time-gap check” does not flag data or gaps in the underlying CeNCOOS data storage, thus is not considered adequate to satisfy the time-gap test proposed by QARTOD. The above is done instead of inserting a time with a missing value and a fail flag as the QARTOD

manual appears to suggest for the time-gap test.

4. *Flat Line Test:* this test checks for consecutively repeated values within a tolerance. The tolerance will depend upon the variable. A point is considered suspect after 6 hours. If the value has not changed after 12 hours the variable fails the test.

5. *Spike Test:* this test tests how much difference there is between a data point and its neighbors. The suspect and fail thresholds will depend upon the variables. The implementation follows the QARTOD manual.

6. *Rate of Change Test:* this test checks if a value exceeds a threshold between consecutive points. If the data fails this check it is marked as suspect. There is no fail condition.

7. *Location Test:* this test is only applied to moving platforms. It is not required for fixed stations. The test checks for impossible locations (a fail) or unlikely displacement (suspect). The displacement test threshold will depend upon the type of platform and what is considered an abnormal range of displacement.

8. *Attenuated Signal Test:* This test checks whether a signal has been affected by fouling or some other effect that reduces the amplitude of the signal. The implementation follows the QARTOD manual.

9. *Climatology Test:* This tests the data relative to the historic time-series values from the station. This was also implemented using the IOOS instance of the QARTOD code. It is applied to signals with more than 2 years of data.

4.5.2 CeNCOOS Planned QC Protocols

CeNCOOS is in the process of implementing policies outlined in the US IOOS Quality Assurance of Real-Time Oceanographic Data (QARTOD) manuals. As new data are ingested into the CeNCOOS DMAC Sub-system they will be assessed and classified accordingly to allow for full documentation as described in this plan, including Data Stream Plans for new assets that do not come from federal sources and that will be archived by CeNCOOS. As new QARTOD protocols are updated and new parameter manuals developed over time, Data Stream Plans will be updated accordingly to include newly required QARTOD implementations. When QARTOD guidelines do not exist for a variable, other suitable forms of QC implementation are conducted and described.

QARTOD procedures have been implemented using the IOOS QARTOD library. Q Any existing quality flags passed from the data source provider will be stored alongside the data and

available for visual discovery and download via the public-facing CeNCOOS Data Portal. QARTOD tests are run within the CeNCOOS DMAC Sub-system for all the real-time *in situ* observing streams for sources in the Funded Data Streams and Regional Data Streams categories that do not already provide quality flags.

QARTOD calculations vary considerably in terms of complexity: for example, the Flat Line Test can be applied to any variable regardless of sensor type or source, the Gross Range Test requires specified boundaries that would vary by parameter, and the Climatology Test requires at least two year's worth of data and boundaries that vary by parameter type and station. These have been implemented by Axiom using the IOOS instance of the QARTOD python library. The Climatology test has been implemented for variables with a long enough time-series for that to make sense. The implementation of QARTOD tests by CeNCOOS will have different processes depending on the data type, which are real-time data, historical data, citizen data, and federal data. These are further defined below.

1. Real-time Data

The CeNCOOS DMAC Sub-system will ensure that quality control (QC) standards are implemented and QC flags made available for all real-time data that are not received from a federal source. CeNCOOS currently serves various non-federal data streams that require QARTOD QC tests implementations. Currently, real-time observations ingested by CeNCOOS DMAC Sub-system have a minimal CeNCOOS QC tests applied (syntax test, gross range test, time-gap check are applied at ingest. The other tests are applied to the time-series during a post processing step by Axiom at periodic intervals). No QC flags are associated with these data. As some data providers document existing QC performed at the source by the data originator (e.g., MBARI M1 buoy), links to these procedures or a brief summary of the QC performed are provided in the applicable Data Stream Plans (see Appendix E - Funded Datastream Inventory). The NWS, NERRS, the IOOS HFR DAC, and the IOOS Glider DAC all ingest and perform extensive QC on the raw data collected by these platforms prior to making them available to the public. CeNCOOS ingests these data from these programs for display in the DMAC Sub-system, and are not required to perform additional QC on these assets. Required QARTOD tests on non-federal real-time data (either by CeNCOOS data providers or at the CeNCOOS DAC) have been implemented as described above.

2. Historical Data

CeNCOOS currently applies two standard QC procedures to historical time series sensor data

ingested through the DMAC Sub-system (syntax test, gross range test). When possible, CeNCOOS provides documentation of QC on historical data, including QC procedures that are implemented by the data provider, in addition to any additional processing or QC that occurs after data ingestion but prior to data dissemination via the CeNCOOS DMAC Sub-system. When QARTOD applies, data assets that were previously reporting real-time data and that have had their historical data stored and made available in the CeNCOOS DMAC Sub-system will follow the same QC protocols as the real-time data (e.g., non-federal weather data). Data with QC flags will become available in the Data Portal as QARTOD and other QC protocols are fully implemented (as described above).

4.6 Stewardship and Preservation

CeNCOOS stores ingested data in a secure, professionally managed external facility. Data are stored on storage volume with multiple levels of fault tolerance and can survive multiple server and drive failures. In addition, data are backed up to Backblaze, an offsite backup service provider. CeNCOOS currently has total storage space for over 1.8 petabytes of data. CeNCOOS stores all aggregated data, be it real-time sensors, forecasts results, static GIS layers, etc., indefinitely beyond the life of each individual project. This means that real-time sensor feeds will become historical sensor feeds one-month after collection, and it allows CeNCOOS to grant users rapid web-based access to all sensor data (federal and nonfederal) since CeNCOOS began aggregating feeds. The only assets that are not kept indefinitely in storage are webcam images, NERRS data (as it is strictly prohibited in their terms of service), and forecast products that have been replaced with a more accurate forecast.

4.7 Providing Public Access and Dissemination

All data served on the CeNCOOS DMAC Sub-system are fully available to the public and have no data restrictions or embargo periods placed on them. New datasets from either new or current data providers received by CeNCOOS are immediately available to the public after data ingestion and documentation is complete; however, they are not added to the searchable data catalog. Datasets are added to the searchable, public catalog only after the data provider is brought into a feedback loop to comment on the metadata, usage notes and citation information regarding their dataset. Once published in the catalog, products are promoted via the CeNCOOS website, social media accounts, outreach campaigns, and an email newsletter.

The CeNCOOS DMAC Sub-system provides a variety of environmental and socioeconomic data resources in a one stop data portal, free to the public, with data originating from CeNCOOS

funded data providers, federal and state agencies, local municipalities, academic institutions, research organizations, private companies, non-profit organizations, and community observers. Data that is collected with Federal funds are freely available without license or rights retained by providing institutions, excepting cases where rights can be retained within Federal guidance such as with Tribes. Data access is defined here as being permitted to download data through the CeNCOOS Data Portal. Occasionally, a data sharing agreement between CeNCOOS and a data provider will identify the existence of intellectual property rights (IPR) to the data and this is noted in the applicable Data Stream plan. However, IPRs do not restrict access to any of the data that is freely served through CeNCOOS Data Portal. IPR information is provided to show its provenance . It is a best practice to always clearly give credit to the data source (the originator) and data provider (in this case CeNCOOS) in any work or publications that emanate from using data accessed via the CeNCOOS Data Portal and to similarly provide clear data attribution in the CeNCOOS Data Portal itself.

4.8 Data Archival

As a federally-funded program, CeNCOOS is required to submit data it generates to a national archive center. CeNCOOS has an active Submission agreement to archive data with the National Center for Environmental Information (NCEI) (see Appendix H - NCEI Archival Agreement). The agreement framework is being renegotiated with NCEI and is in the process of transferring that agreement to Axiom Data Team. The CeNCOOS DMAC Sub-system Management Team are working with the NCEI to assist with the transfer and preservation of appropriate data types. CeNCOOS has worked and consulted with several NCEI staff members, including Matthew Biddle (former), John Relph, and James Partain, on automating the submission of CeNCOOS-funded data assets to the NCEI. Julie Bosch will advise the CeNCOOS staff on the data submission forms and all necessary procedures. Automatic CeNCOOS submissions for long-term archive at NCEI are currently limited to *in-situ* data from a handful of regional observing networks (listed in Table 1).

CeNCOOS also archives data through two national IOOS Data Assembly Centers (DAC): the IOOS HF Radar DAC and the IOOS Glider DAC. Both national DACs are responsible for archiving all of their data, including that submitted by CeNCOOS, with NCEI and is beyond the scope of this document. For more information see the National Glider DAC and the National HFR DAC.

Source	Station Name	External Identifier	NCEI Accession	Status

California Polytechnic State University	Morro Bay - T Pier	CeNCOOS *in situ* water quality monitoring at Morro bay	171314	active monthly archive through June 2022
	San Luis Bay - Cal Poly Pier MET	CeNCOOS *in situ* meteorological monitoring at the Cal Poly Pier San Luis Obispo.	171315	active monthly archive through May 2022
	San Luis Bay - Cal Poly Pier Shore	CeNCOOS *in situ* water quality monitoring at the Cal Poly Pier San Luis Obispo.	171316	active monthly archive through June 2022
Humboldt State University	Trinidad (tdp)	CeNCOOS *in situ* water monitoring data at Trinidad Head, California	171318	active monthly archive through May 2022
	Humboldt (Formerly Dock B Shore Station) (HBC)	Humboldt Dock B Shore Station	171317	historical station; archive through March 2013
	Humboldt	CeNCOOS *in situ* water quality monitoring at Humboldt Bay Pier.	172588	active monthly archive through October 2021
Monterey Bay Aquarium	Monterey Bay Aquarium Seawater Intake	*pending*	-	new station effective April 2018, archive pending
Moss Landing Marine Lab	(MLSC1) Moss Landing Marine Laboratories Sea Water Station	CeNCOOS *in situ* water quality monitoring using the seawater input for Moss Landing Marine Laboratory.	171324	active monthly archive through October 2021
	Monterey Shore Station	CeNCOOS *in situ* water quality monitoring at Monterey Bay Commercial Wharf.	171312	active monthly archive through July 2018
San Francisco State University	Carquinez	CeNCOOS *in situ* water quality monitoring at Carquinez at the California Maritime campus.	171319	active monthly archive through August 2021
	Tiburon Pier, San Francisco Bay, CA (TIBC1)	CeNCOOS *in situ* water monitoring data at Romberg Tiburon Center Pier	171313	active monthly archive through December 2019
University of California Davis, Bodega Marine Laboratory	Bodega Bay (BML_WTS)	CeNCOOS *in situ* water quality monitoring at Bodega Marine Laboratory seawater intake.	171320	active monthly archive through November 2023
	Fort Point	CeNCOOS *in situ* water quality monitoring at Fort Point Pier.	171321	active monthly archive through November 2023
	Hog Island Oyster	CeNCOOS *in situ* water quality monitoring using a SBE_TSG and SeaFET at Hog Island Oyster Company in Tomales Bay.	171322	active monthly archive through April 2019
University of	Santa Cruz	CeNCOOS *in situ* Water monitoring	171323	active monthly

California, Santa Cruz		data at the Santa Cruz municipal wharf		archive through December 2019
Exploratorium	Exploratorium	*pending*	-	new station effective April 2018, archive pending
Wiyot Tribe	Indian Island	CeNCOOS *in situ* water quality monitoring at Indian Island	171311	active monthly archive through March 2017

Table 1. Funded Data Streams Inventory for Archive of *In-situ* Observations Data to NCEI as of 2018-07-30

4.9 Performance and Security

All physical infrastructure required to support the CeNCOOS Data Portal and its underpinning systems is located in a data center designed and maintained by Axiom staff. Resources in the data center include more than 2,500 processing cores arranged in a series of interconnected blade arrays, as well as slightly more than 1 petabyte of usable storage that includes multiple redundant backups. Compute nodes and storage nodes are connected over a low latency, converging network fabric (40 Gb/s Infiniband). GlusterFS is employed as a storage software abstraction layer that enables clients and storage servers to exploit data transfer over Remote Direct Memory Access (RDMA) protocols. This configuration enables data throughput from the storage clusters to the compute clusters to reach speeds greater than 160 Gb/s in high-concurrency situations. The CeNCOOS DMAC Sub-system provides the following enterprise-level infrastructure capabilities:

- Security and Redundancy: The CeNCOOS DMAC Sub-system is maintained by Axiom at its data center, collocated with the Pittock Internet Exchange in Portland, OR, part of the West Coast US internet backbone. There, the data center benefits from the low-latency, high-bandwidth internet connection, and network and power reliability. All data center resources are protected by several levels of onsite redundancy and backup, with offsite backup using Backblaze B2. This design ensures that multiple redundant copies of data exist in addition to web application servers. Several layers of physical hardware (enterprise-level firewalls) and system monitoring software (Nagios) are also in place to provide hardened cyber security.

- Capacity and Performance : High Performance Computing (HPC) has been a component of the CeNCOOS DMAC Sub-system technical strategy since early 2011. Axiom operates

its own private cloud of compute and storage resources that data managers can provision to specific tasks and roles. The current numbers of processing cores and storage volumes are scalable to allow additional resources to be added as necessary. The large GIS, model, and remote sensing datasets within the system require HPC environments to be visualized and queried over web-based interfaces. Because HPC is achieved through load balancing and parallelization, these types of systems also provide the added bonus of high availability and redundancy.