OOI CYBERINFRASTRUCTURE
NSF OCEAN OBSERVATORIES INITIATIVE

OOI – CyberInfrastructure

Requirements Workshop
Rutgers University, New Brunswick, NJ
July 23-24, 2007

Workshop Report

FINAL

November 2007

NSF JOI Joint Oceanographic Institutions UCSD

Document information

Project	Network for Ocean Research, Interaction and Application (NORIA)
Document Custodian	Oscar Schofield/Scott Glenn (Rutgers University, NJ) OOI CIIO Project Scientists
Approval	Frank Vernon
Created on	September 26, 2007
Last Changed on	November 8, 2007
Document Status	FINAL

Document History

Date	Version	By	Description of Changes
2007-09-26	0.1	M. Meisinger	Initial workshop report from notes
2007-10-03	0.2	M. Meisinger	Comments from M. Arrott, A. Chave addressed; minor format changes;
2007-10-11	0.3	M. Meisinger	Edits by I. Krueger and O. Schofield included; comments addressed
2007-10-18	0.4	M. Meisinger	Minor typo fixes, added scientific figures and group picture; figure captions by O. Schofield
2007-11-08	1.0	M. Meisinger	State changed to FINAL after post-workshop validation by participants

Table of Contents

OOI - CyberInfrastructure

Requirements Workshop at Rutgers University
Outcome and Summary

1 Executive Summary

The OOI CyberInfrastructure IO organized a requirements workshop July 23-24, 2007 at Rutgers University (NJ), inviting oceanographic science users of the mid-Atlantic ocean modeling and coastal ocean observing community for the purpose of CyberInfrastructure user requirements identification and elicitation, as well as an early outreach measure to immediate CI user communities – in this case the ocean modelers. The Rutgers University's Coastal Ocean Observation Lab (COOL) provided the scientific environment for a 1.5 day workshop that covered introductions to the planned CI, current oceanographic science presentations, CI requirements discussions, domain modeling sessions and feedback opportunities.

Workshop outcome and results include

- User requirements for the CI out of the modeling community
- Validated existing conceptual architecture system requirements
- Instrument/resource domain model developed as workshop session result
- Collected workshop materials including introductory presentations (OOI, CI, science) on CD
- Preparatory domain models covering parts of the ocean science

2 Introduction

2.1 Goals and Background

In order to provide the U.S. ocean sciences research community with access to the basic infrastructure required to make sustained, long-term and adaptive measurements in the oceans, the National Science Foundation (NSF) Ocean Sciences Division has initiated the Ocean Observatories Initiative (OOI). The OOI is the outgrowth of many years of national and international scientific planning efforts. As these efforts mature, the research-focused observatories enabled by the OOI will be networked, becoming an integral partner to the proposed Integrated and Sustained Ocean Observing System (IOOS; www.ocean.us). IOOS is an operationally-focused national system, and in turn will be the enabling U.S. contribution to the international Global Ocean Observing System (GOOS; http://www.ioc-goos.org) and the Global Earth Observing System of Systems (GEOSS; www.earthobservations.org). Additionally the OOI will provide an ocean technology development pathway for other proposed net-centric ocean observing networks such as the proposed Navy's Littoral Battlespace Integration and Fusion program.

The OOI comprises three distributed yet interconnected observatories spanning global, regional and coastal scales that when the data is combined will provide will allow scientists study a range of high priority processes highlighted by the community. The OOI CyberInfrastructure (CI) constitutes the integrating element that links and binds the three types of marine observatories and associated sensors into a coherent system-of-systems. The objective of the OOI CI is provision of a comprehensive federated system of observatories, laboratories, classrooms, and facilities that realizes the OOI mission. The infrastructure provided to research scientists through the OOI will include the sea floor cables combined with water column fixed and mobile systems. Central to these observational platforms are junction boxes which provide power and two-way data communication to a wide variety of sensors at the sea surface, in the water column, and at or beneath the seafloor. The initiative also includes components such as unified project management, data dissemination and archiving, and education and outreach activities essential to the long-term success of ocean observatory science. The vision of the OOI CI is to provide the OOI user, beginning at the science community, with a system that enables simple and direct use of OOI resources to accomplish their scientific objectives. This vision includes direct access to instrument data, control, and operational activities described above, and the opportunity to seamlessly collaborate with other scientists, institutions, projects, and disciplines.

A conceptual architecture for the OOI CyberInfrastructure was developed and published by a NSF committee in 2006 (see http://www.orionprogram.org/organization/committees/ciarch) [CI-CONARCH]. It describes the core capabilities of such a system. Initial requirements were derived from similar CyberInfrastructure and oceanographic projects. Please refer to Appendix C for further OOI background information.

In May 2007, the consortium lead by SIO/UCSD including JPL/NASA, MIT, MBARI, NCSA, NCSU, Rutgers, Univ Chicago, USC/ISI and WHOI was awarded the development of the CI as Implementing Organization (IO), based on the NORIA (Network for Ocean Research, Interaction and Application) proposal [NORIA]. Ongoing efforts include architecture and design refinement and consolidation towards a preliminary design level as well as validation and refinement of the existing CI requirements by direct involvement of prospective CI user communities. The requirements elicitation and management process is planned to be an ongoing activity in close collaboration with the user communities involved.

The first direct science user involvement occurred during the first CI requirements workshop, July 23-24, 2007 at Rutgers University. Summary and outcome of this workshop are the content of this report.

This workshop is one of three in a series of CI architecture and design team organized workshops to identify and elicit requirements from domain users. All workshops are targeted mainly at the ocean modeling communities.

- Workshop one covered the Mid-Atlantic community with a focus on the range of data assimilative numerical continental shelf models..
- Workshop two will cover the Pacific Northwest communities, which includes the Regional Scale Node (RSN) area and Neptune Canada. Location possibilities are UW at Seattle and Oregon State University. Target date is after the New Year early in 2008.
- Workshop three will cover the Global-Climate modeling communities and include the SoCal institutions. The probable locations will be at University Consortium of Atmospheric Research (UCAR) in Boulder Colorado also early in 2008.

Goals of the first workshop described in this report were:

- Establish an initial direct contact to interested CI user communities, here the mid-Atlantic ocean modeling community
- Provide the CI engineering team insight into the systems and issues faced by the coastal ocean modeling community.
- Identify and elicit user requirements for the CI from the view of this specific community
- Develop a thorough domain understanding by direct collaboration with domain scientists in order to increase language tangibility
- Provide a basis for further requirements elicitation and domain modeling in subsequent installations of this workshop and in ongoing requirements and architecture design work
- Validate existing previously collected CI system requirements

2.2 Preparation

In preparation for this workshop, the architecture & design team analyzed materials from different sources. The purpose of this work was the following:

- Familiarization with the specific concepts and language of the targeted scientific users
- Familiarization with general oceanographic science concepts
- Display systems engineering notations and methodology using a well-understood example
- Provide a system model baseline and basis for refinements of these models during the workshop
- Establish communications basis between architecture design team and science user participants

The first topic analyzed and the resulting model covers the initial version of the science question traceability matrix documents [SQTM]. Resulting models were a scientific observation domain model (see Figure 7), a scientific process activity diagram (see Figure 8), as well as instrument platform and instrument domain models (see Figure 9 and Figure 10).

The second analyzed document was the MARCOOS proposal [MARCOOS] that the Mid-Atlantic ocean observing and modeling community had recently been funded as a backbone component to the NorthEast IOOS network. It contains a current situational analysis of the research community and their tools, as well as a proposed plan to combine existing data sources and models to provide a coherent network to observe and forecast the oceans. Resulting system engineering domain models were the high level modeling process domain model (see Figure 11), an integrated MARCOOS tasks domain model (see Figure 12) as well as a MARCOOS proposed scenario domain model (see Figure 13).

All models can be found in the appendix as well as in the respective slide presentation materials. The state of the models is that of a working draft. They capture the modeled domain to a level of detail that was required to establish a common understanding and modeling baseline for this workshop. The created models will be further refined and integrated in CI requirements and design documents as needed.

2.3 Acknowledgements

This report was developed by the OOI CI architecture and design team; it contains input from many sources, such as the workshop presentations by the organizers and invited science users, the CI conceptual architecture, OOI science background information by the Project Scientists, and notes taken from various individuals. Workshop notes were taken by Michael Meisinger, Vina Ermagan and Jack Kleinert. Furthermore, this report contains summarizing and general statements by the organizers.

3 Presentations

3.1 OOI Project Overview

The presentation by Oscar Schofield, Rutgers University, covered the current status of OOI development: Implementing Organization (IO) selection is now complete with

- Regional Scale Node (RSN), developed by UW, comprising 5 regional sites
- Coastal/Global Scale Node (C-GSN), developed by WHOI-lead consortium, comprising 3 coastal and 6 global sites
- CyberInfrastructure (CI), developed by SIO- lead consortium

The OOI Pioneer Array will be one of the first assets to be created for the C-GSN in the Mid-Atlantic Bight (MAB) designed to study shelf slope exchange processes and to quantify the importance of episodic processes (storms, river plumes, etc.). The proposed Pioneer Array will provide a high resolution highly flexible array that can be nested within the more broadly distributed regional observing network that is being developed as part of the MARCOOS system. The combined nested capability is envisioned to enable a wide range of smaller-scale process studies and large dedicated science campaigns that will consist of multi-ship, multi-institution projects. All the expected activities will require integration.

The benefits for science through an integration of tools, processes, instruments, results, etc. are:

- Putting assets together allows for quantitative process studies
- Informs day-to-day decisions: Where to drive a boat, what instruments to use, where to place an instrument and how often must a location be sampled.
- Improvements in sampling strategies through better- informed decisions enabled by real-time data or ocean forecasts.
- Today, the input for proposals of what data streams to use is based on personal knowledge and communications with personally established contacts between researchers
- The mid-Atlantic modeling community is very small (15 institutions, 20 PIs for a given project). (This is probably small by CI standards, but this is likely the scale of a science campaign)
- Example for data integration: Inland rain, highly dynamic weather data sets (post analysis) combined with data collected by fleets of mobile platforms.

Challenges for the CI:

- Numeric modelers need to integrate many data streams, not only those of NSF origin
- The potential data streams are large
- There is a need to qualify data streams for a given study, for instance data streams coming from processing models or satellites that are beyond the control of the researchers and current community validation capabilities. Such steams can be highly contested and will likely be high scientific interest.
- In the numeric modeling domain, integration is lead by examples of success.
- Infrastructures need to be pertinent, relatively constant, easy to incorporate into a study and reliable, otherwise they won't be used by scientists

Gliders are an example class of instruments that are particularly interesting for systematic ocean research. Gliders will complement sampling from ships by autonomously following large specified sampling routes. These transects usually take about 4-6 weeks to complete and can span over 500 kilometers of the MAB. Gliders surface every 6 hours for data synchronization and instrument/mission adjustments. All instruments on board are controllable. After the mission period, the glider archive data is retrieved and a new route is defined. Other instruments that are of oceanographic use include drifters, which travel across the ocean at different depths, moorings providing high frequency time-series at a fixed location, remote sensing data collected from the international constellation of satellites and surface current measurements provided by HF shore-based/mooring networks.

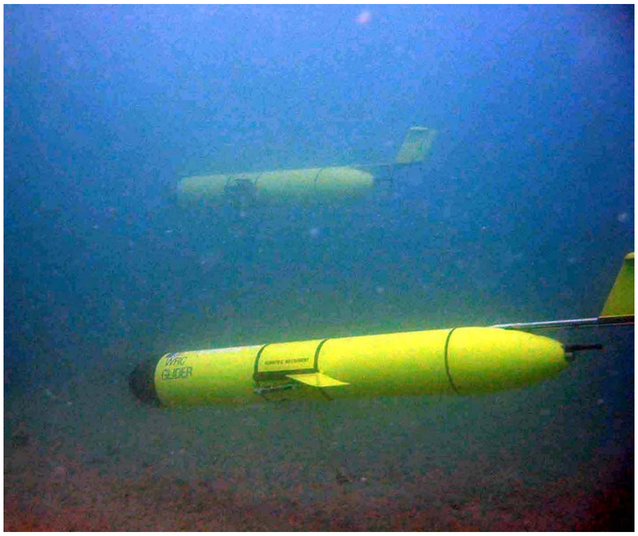

Figure 1: Two gliders flying in a coordinated mission in near-shore coastal waters,

Ideally, the ocean environment would be much better sampled so that better research questions can be asked and answered. The community, for instance, envisions nowcast models similar to weather models/forecasts describing the current status of the oceans in a continuous space-time representation. Such models do not yet exist for the oceans. For such models, it is important to know the uncertainty of the model itself and the uncertainty of the source instrument data. It is the current understanding that the uncertainty introduced by the models is an in order of magnitude larger than that of the observations.

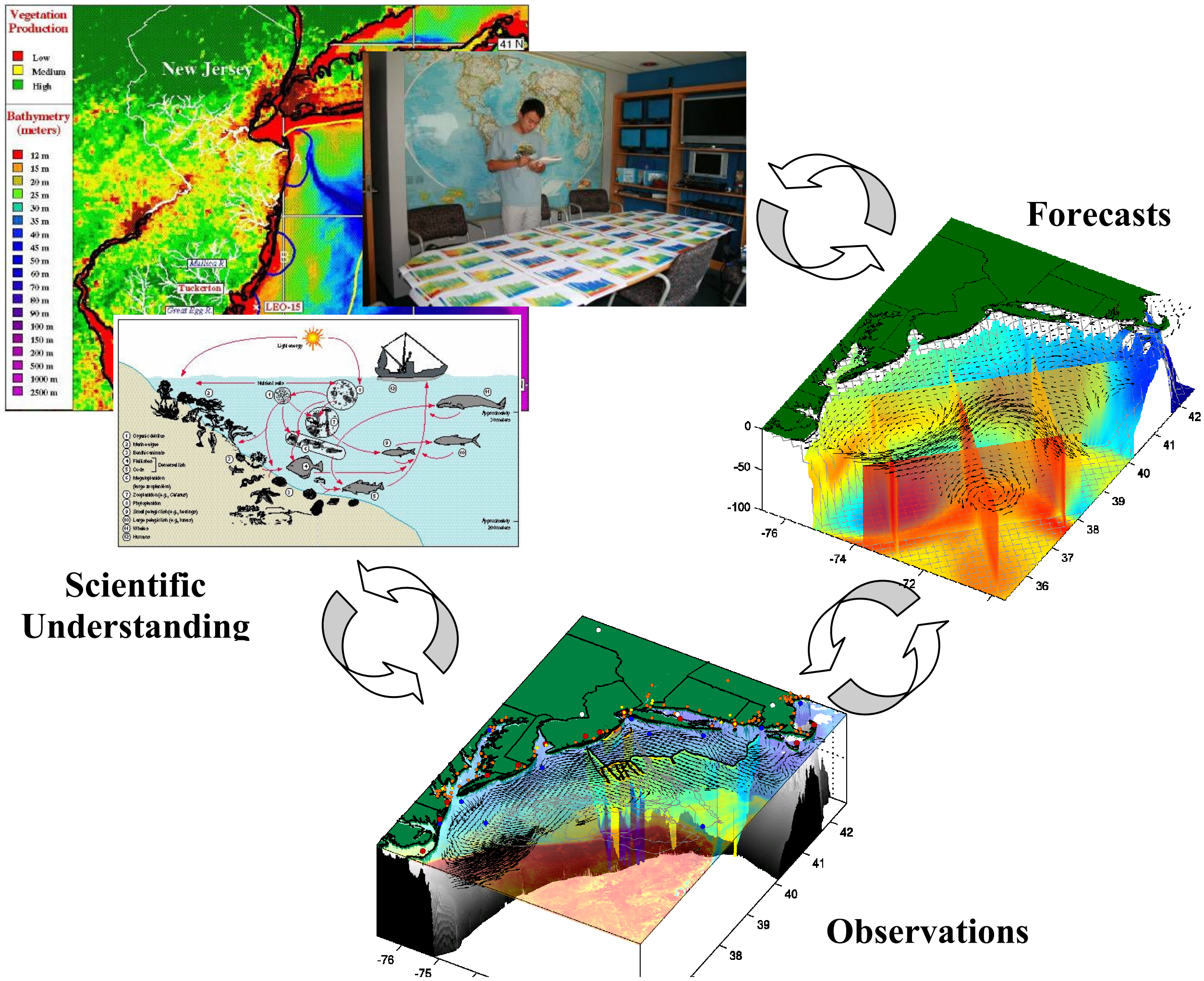

Figure 2: Three Commonalities to Ocean Observatories illustrating how the observations, models, and scientific understanding is linked in a mature ocean observatory. The CI provides the critical link between the three commonalities.

The presentation by Oscar Schofield emphasized the understanding that the ocean environment is chronically undersampled and improving the ability to sample the system is critical to enabling transformative new science. The concept of an ocean observatory differs from the more traditional approach of environmental monitoring using discrete instrument platforms in that an observatory bridges the gap between individual measurements in time and space and thus provides a means for mechanistic understanding. The time and space scale problems associated with ocean monitoring are numerous, and for the most part result from the inability to synoptically sample all of the desired biological, chemical, and physical measurements needed to resolve the state and trend of the coastal ocean. An alternative to measuring all things at all times is to couple an adaptive sampling strategy with a mechanistic modeling regime to resolve the processes that result in a particular ecotone state, as well as providing the information and tools to predict the direction the ecological system will take. By linking the sampling strategy to the modeling process, there is a refinement of both the data collection techniques and the modeling of the ecological processes.

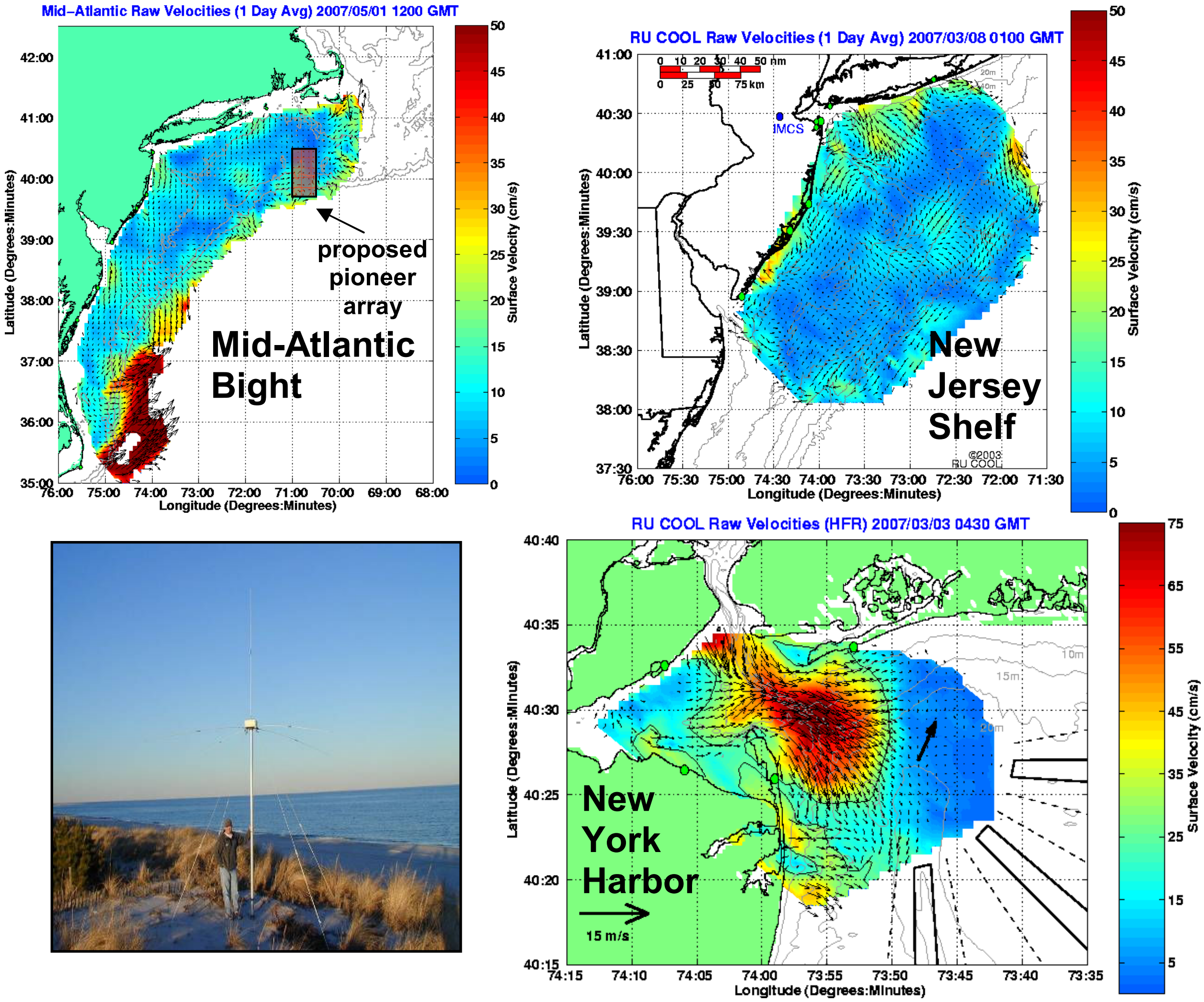

Figure 3: Nested Surface Current Maps measured by the NOAA CODAR network. Here is a highly valuable data source that will need to be aggregated with the OOI data collected by the Pioneer array (the grey box in upper left figure). Therefore a critical tool of the CI will be allow researchers to have access to all available data within their domain.

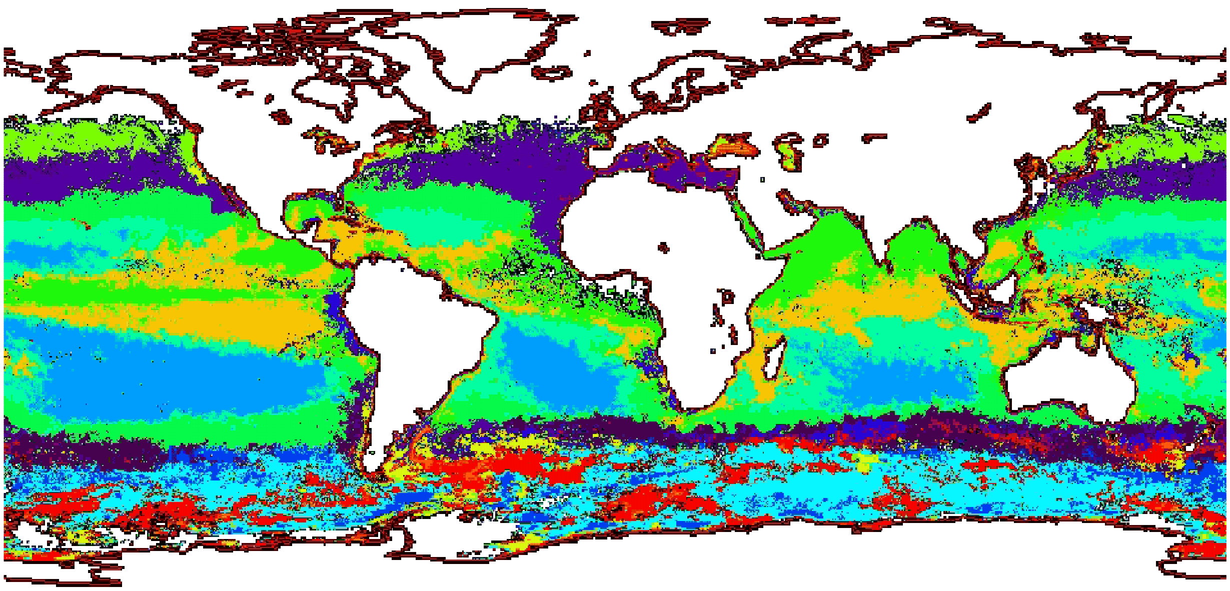

Figure 4: CI will allow new products to be developed. Here aggregated data from multiple satellites are combined to generate new products. Here Sea surface temperature and ocean color data were combined using bioinformatics algorithms to map objectively the biotic provinces in the world's oceans.

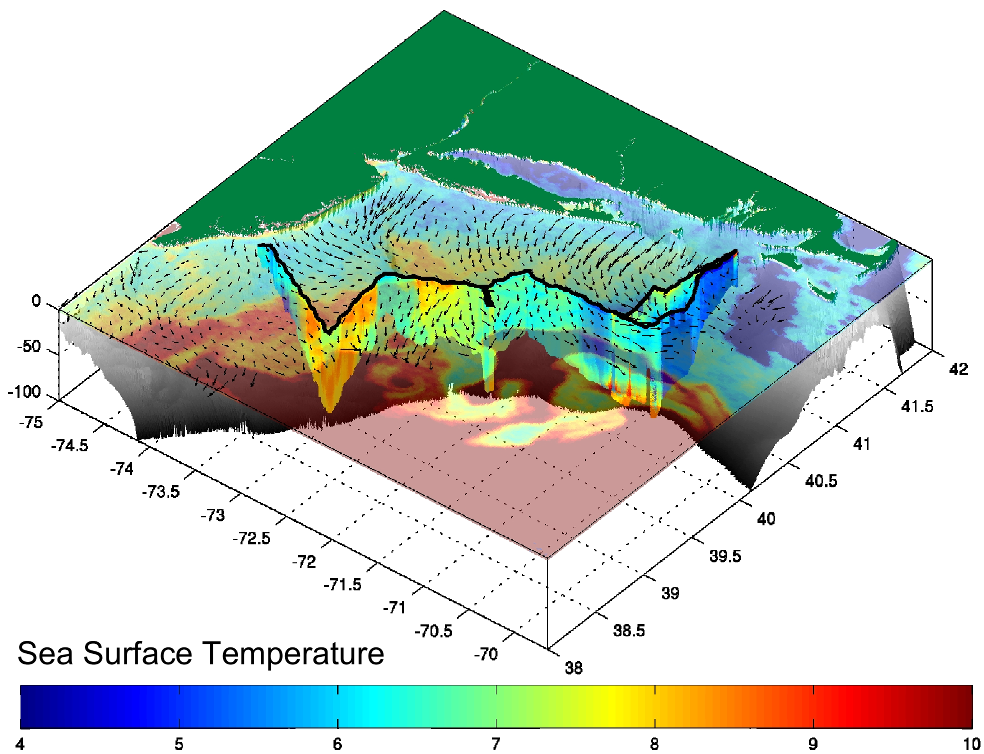

Figure 5: CI should provide the ability aggregate many disparate data sets that might all have different frames of reference.

Traditionally climatologies have been used to provide the conditions on continental shelves; however, these approaches ignore episodic events that are believed to play a disproportionately important role for many physical, chemical, and biological processes. Given this, in a typically under-sampled ocean, ocean forecast errors are dominated by uncertainties in the model initialization; therefore ensemble forecasts with differing initial conditions are often used to identify regions in which additional data are required. Hence the models provide insight into what has not been sampled and guidance for field efforts. With potential advances associated with OOI and IOOS, scientists are hoping to no longer consigned to operating in an under-sampled environment. The OOI network will provide spatially extensive observational updates on time scales of an hour or better which will alter the entire paradigm for adaptive sampling and nowcast/forecast modeling. In the well-sampled ocean, forecast errors may now be dominated by uncertainties in the model formulations or boundary conditions, and ensemble forecasts with differing model parameterizations can be used to identify regions in which additional data can be used to keep a model on track. In the time it takes to prepare the ensemble of forecasts for the well-sampled ocean, additional data has arrived, and on-the-fly model-data metrics can be used to quantify which forecast in the ensemble is most accurate. Thus in the well-sampled ocean, the observations are used to improve our understanding of errors associated with fundamental model assumptions. Please refer to Appendix C for a more detailed scientific background for the OOI.

3.2 CI Overview, Requirements, Architecture

Matthew Arrott (UCSD/Calit2) provided an overview of the OOI CyberInfrastructure conceptual architecture, current stage of architecture and design and the CI project organization. During the presentation, the complementary nature of OOI and IOOS was discussed and the importance of IOOS was acknowledged; it already has a national agenda focused on operationalization. IOOS is likely to be among the earliest beneficiaries of the OOI system.

Alan Chave (WHOI) presented requirements elicitation and management methodology and the purpose of systematic and iterative requirements elicitation efforts involving multiple user communities over the course of the OOI/CI project.

Ingolf Krueger (UCSD/Calit2) presented domain modeling notation and techniques, which is required as a structured way to capture domain knowledge for requirements purposes but also as foundation for system architecture and design. As discussion basis and case examples, he presented domain models of ocean science observation processes and of the MARCOOS proposal that was provided before the workshop as background information about the mid-Atlantic modeling community's research topics.

3.3 Science Presentation: Regional Circulation modeling for the mid-Atlantic Bight & Gulf of Maine (MABGOM)

Rouying He of NCSU presented scientific background and current challenges around his circulation model of the mid-Atlantic and Gulf of Maine area. Objectives are to develop realistic simulation models that provide data on circulation with continuous time and space coverage. The modeling approaches used include HYCOM and ROMS models. Several input data streams are required for lateral and surface boundary conditions and to provide model validation.

Specific statements:

- Acquiring access to data sources and getting data streams in the right format and quality is a time consuming task in the everyday research work.
- Input Data for forecast models is retrieved from the source location by a pull mechanism; such data can, for instance, be in a format that aggregates source data by the hour. Current models do

not require real-time streaming or data-push from the source; this, however, would enable next generation models.

- Ability to aggregate data from different sources would significantly improve current scientists' daily work.
- Different organizations produce similar or same output (same variables). The ability to select among available/suitable data in an automated way would be a good improvement
- Individual researchers invest significant time in specific data conversion, and transformation solutions. More efficient solutions to these tasks will leave more time for research. Existing tools that provide support for these tasks are difficult to adopt, however.

3.4 Science Presentation: Prediction and Analysis for Mid-Atlantic Bight Ocean Observatories

John Wilkin of Rutgers University presented scientific background on ocean nowcast and forecast models for the mid-Atlantic bight. The approach uses current measurements using a model-observation comparison to tune model parameters to provide more accurate simulations. This leads to "relative future observations" and 4D-VAR models. The models are used for the coast guard to predict the path of the drifters, searching operations, seashell toxic substances path for seafood industry, tide heights for ship cargo delivery at ports. Further analysis includes a long time mean condition derivation from models and comparisons between coastal temperatures of different parts of the coast at different times. Example input data are from glider runs, ship runs, moorings, CODAR, Satellite, etc.

Specific statements:

- OPeNDAP has revolutionized the handling of data. As a disadvantage, however, researchers can make their entire data set available to the outside world but the outside world has to sort and select their subset of interest.
- NetCDF files and file headers that define the data format are widespread in the community. NetCDF provides capabilities for mapping data table columns for conversion purposes. Such mappings are problematic because they are hard-coded in hard-to-read files, making use of constants that are defined based on proprietary interpretations in order to define concepts such as namespaces etc.
- Currently, domain knowledge informs many model parameters, such as the forecast/modeling intervals used in models
- Terminology
  - Data assimilation: Bringing observed data together with forecasts, interpolations and short-term predictions in balance with known scientific processes.
  - Data analysis: Modeling historic observations as close as possible
  - Data forecast: Predicting future observations
- The CI team should evaluate the ESMF (Earth Systems Modeling Framework), which enables interoperability between models.
- Should evaluate ECMWF (European medium term weather forecast). They have good medium scale models, accessed by local organizations for precise local scale models plus infrastructure.
- Beneficial would be to follow the Climate Forecast Metadata standards.

3.5 Science Presentation: Advanced Fisheries Management Information System (AFMIS) for the Northwest Atlantic Ocean

Wendell Brown of the University of Massachusetts, Dartmouth, presented a design for a fisheries management information system that was developed some time ago in an effort to provide an advanced integrated model, nowcast and forecast system including data from many environmental and fisheries data bases. Models include a physical ocean model, an ecosystem model and a fish model. The models apply

Principal Components Analysis (PCA) to analyze different model data, and to find features in data. Examples for such oceanic "features" are eddies, upwellings, etc. Such features can for instance be identified by specifying the physical boundaries of a feature, which allows for the retrieval of data about the feature. Characteristics about various types of features can then be abstracted out and studied such as salinity or seasonality.

4 Discussion Sessions

4.1 CI Questions Discussion

This first discussion session takes the shelf-slope exchange process and the Pioneer Array as scientific topic to elicit requirements and CI related questions. All workshop participants had been tasked on the previous day to come up with specific questions related to the CI reflecting this scientific topic. The Pioneer Array is intended to provide a persistent set of sensing assets to which researchers can add temporary assets for specific experiments. It provides sampling assets (buoys) that can be controlled (i.e., turned on and off, etc.)

The following statements and questions were identified during the session. They represent raw information that directly feeds into the requirements elicitation and architecture design efforts:

- Will the commands for controlling the assets be passed through the CI?
- There are different quality levels of data streams (reliability depends on agency, confidence intervals, pre-processing). For instance: individual researchers apply cloud removing algorithms to data streams to obtain data for coastal areas; NASA, for instance, entirely removes the coastal areas form the data to ensure high data accuracy. Researchers will need to make a choice considering this trade-off. Providing the flexibility to allow individual researchers to choose is extremely important.
- Existing data is scattered (hard to find), different formats, different QA/QC, hard to take advantage of unrelated experiments.
- It is important to scientists to have opportunistic experiments scheduled within long-term planned experiments
- What experiments are happening and when they are happening should be exposed to the community to facilitate coordination between PIs.
- Many CI issues are discussed extensively in DMAC, but not with the same focus. OOI CI covers effectively DMAC topics. OOI CI has a more comprehensive view on these topics.
- There can be reluctance among scientists to make new data freely available right off the source. Vision: this protective model will fade out in favor of a community/shared data model.
- Example: Data created on ship. Made available real-time to other ships. At the end of the field season, data is transferred to system, processed and archived.
- An oceanographic model has a code base that requires a compile environment. If there are enough traces kept, reproducing a model can be a manual process if needed. Meta-data is needed.
- A model can be bundled and delivered standalone as a file archive, installable with make. A model has source code, input transformation scripts, configuration files etc. The model size is small, a few 100K LOC source code. More complicated is solution to distribute it onto computational nodes.
- New model versions are created frequently in the active phase, for instance every 12 hours. An active model had many people working on it, continuously improving it.
- For such numeric models, SVN (subversion) is used, for instance, as file repository. This enables a later check out of particular versions of the model. However, model version does not describe the execution environment it was targeted for.
- CI needs individual- PI-driven, versioning of data and modeling resources.

- Large experiment model output can be up to 1TB result data (grid size 300 x 200 x 30, 1-10 million grid points, every 3 days for 2 years, 12 variables). That's one grid within a set of nested grids. When having a high-end biochemical model: more than 65 variables.
- For instance a single ROMS model run will cover 10,000,000 data points.
- As infrastructure visualization, a light-weight, easy to understand and use interface should be considered. DMAC, for instance, implemented an interactive display, which can take more time, because it offers significant interactivity. There was a need for interaction and scripting.
- Meta data: Contains sources, used, algorithm, provenance (origin), citation.
- There might be multiple versions of one data stream: for instance after recalibration of sensors. Are these considered derived products or as different versions of the data? There is a need for an ability to update/withdraw data with notifications to users.
- Typical data producers in OOI include: OOI, NOAA (IOOS), EPA, ARS, DoD, DHS, different PIs, Intl. partners
- Typical data consumers in OOI: Scientists, E&O, Agencies, Local and Regional companies, State environmental managers, fishermen, International scientists etc.
- The Raytheon IOOS study lists the top data products for the IOOS community and origin. IOOS, however, excludes many topics interesting for scientists (NSF concerns). OOI enables and focuses on the NSF contribution. IOOS products have a much higher degree of contract (liability, availability), not applying to OOI
- The CI Design Team requires getting scientific workflows from domain scientists, broken up by steps to identify the possibly enabling services of the infrastructure.

4.2 Deployment Scenario Discussion

The second discussion session covered a hypothetical first CI deployment scenario within the modeling community. The following list contains statements and questions from this session:

- The community should agree on 1-2 numeric modeling systems for reanalysis and once developed transfer the responsibility to an operational group. At this point, however, the research community still needs access to the model source code and development of the models should continue even after they have been operationalized on a common infrastructure.
- It requires significant effort to get community buy-in for general community models. It would be difficult to develop an NSF- sanctioned model with community buy-in in a short time frame (less than 5 years).
- The concept of OOI-CI running and governing models seems a viable vision, but for more distant future. Current perception of scientists often is that releasing a model to the infrastructure is "throwing it over the fence"
- The critical resource of an operational model is the operator and the effort needed
- Example existing community model: Climate model, downloadable from public source. Scientists can fully modify it, run it locally. Only the public version is maintained by the governing organization.
- The OOI will not provide any computational power for model computation. Instead OOI-CI will do allocations on national grid (such as Teragrid). Only hardware deployed will be for real-time processing.

5 Workshop Outcome

5.1 Jointly Developed Domain Model

The workshop participants jointly developed a domain model in UML class diagram notation as a whiteboard exercise covering the relationships and entities surrounding oceanographic science instruments and their control.

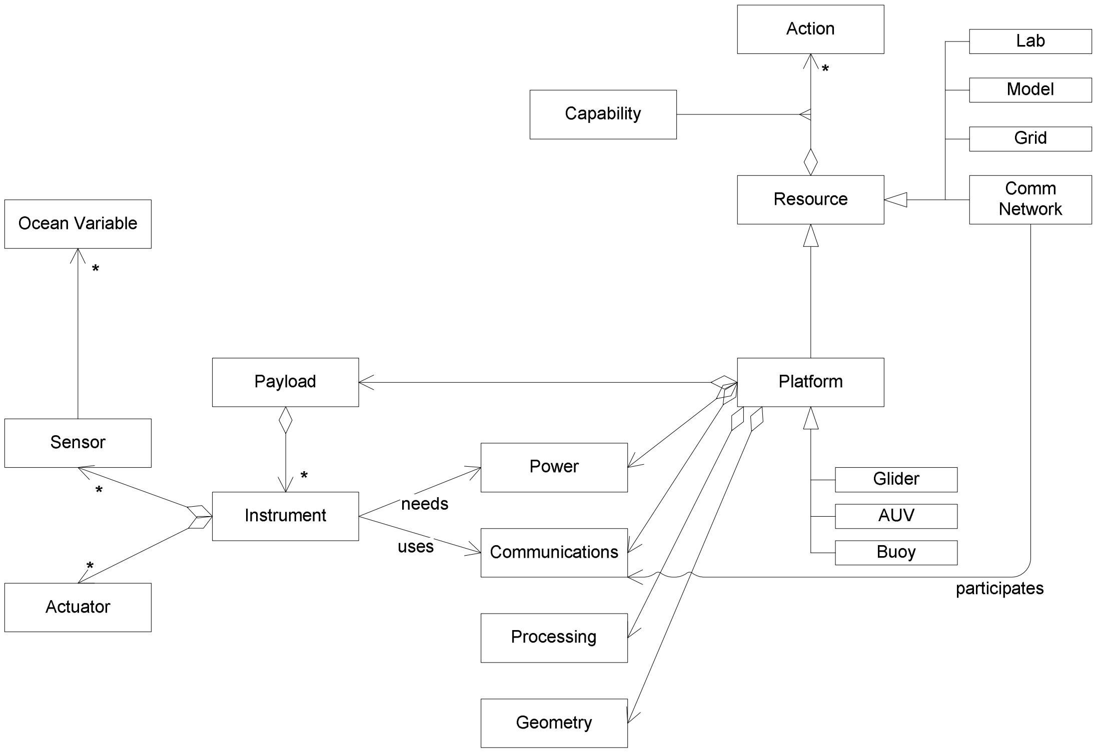

Figure 6: Instrument Domain Model

5.2 Requirements Elicitation Process

The requirements listed in the next section were identified, by

- direct identification during the science presentations and related discussions
- joint elaboration during the working sessions
- inference from the presentation contents and user statements by the ADT
- abstraction and generalization from requirements and statements during the workshop

All listed requirements have been vetted through the participating user community in a post-workshop agreement process.

In order to uniquely identify the elicited requirements, each requirement in this report follows a standard template. Each requirement contains a unique identifier (RWS1-Rn), a requirement label and optionally marked explanation, scoping, example, and detailing paragraphs. Requirement labels are constructed in a schematic way.

The listed requirements are atomic, i.e. they cover one requirement statement only and do not contain sub-requirements. However, requirements might be related and one requirement might be influenced by another requirement.

Requirements are grouped into the following requirements categories:

- Integration and Interfacing
- Resource Management, Data Storage and Data Management
- Data Analysis, Modeling and Dissemination
- Presentation and User Interfaces

- Documentation and Development Process
- Security, Safety and Privacy Properties
- Quality Properties

5.3 Identified CI User Requirements

This section contains the elicited workshop requirements, grouped into categories and phrased according to a template as described above. Requirements marked with an asterisk [*] can be directly or indirectly traced to statements made during the workshop. The remaining requirements are inferred from the context of the application domain or from common domain knowledge.

5.3.1 Integration and Interfacing

This category contains requirements related to the external data and application interfaces of the CI and their integration into a coherent CI proper.

RWS1-R1: The CI shall support distributed resources and actors [*]

Explanation: A resource is any entity associated with an observatory that provides capability, and includes instruments, data, workflows, networks and more. An actor is an entity external to an observatory that interacts with it, and may be a human or a machine.

RWS1-R2: The CI shall facilitate user offline operation [*]

Explanation: Users should be able to interact with the CI on a regular basis without the obligation to be always connected and online. This, for instance, includes batch data stream update and download for subscribed data resources. This also includes data caching and buffering while either a resource or the connected user/application is offline.

RWS1-R3: The CI shall facilitate adding new resources and applications [*]

Explanation: New proposals and grants lead to new/updated hardware in existing observatories as well as to new observatories. The CI shall be flexibly extensible with such resources.

RWS1-R4: The CI shall facilitate the translation between specified data and message formats [*]

Explanation: OPeNDAP and NetCDF as protocols and message formats shall be supported

Scoping: specificity needed

RWS1-R5: The CI shall facilitate the translation between user-specified message formats [*]

Explanation: Encompass the provision of documented connectors to allow the user to provide transformer code.

Scoping: specificity needed

RWS1-R6: The CI shall provide application program interfaces (APIs) to CI services [*]

Explanation: APIs are required by external user applications that interact with CI services and provide user/organization-specific extensions to common CI services and functionality.

RWS1-R7: The CI shall provide synoptic time throughout the OOI observatories [*]

Scoping: The accuracy needs to be specified.

RWS1-R8: The CI shall utilize open standards and software to the maximum possible extent [*]

Explanation: Open standards and software facilitate easy integration of heterogeneous resources and applications, increasing maintainability and extensibility. Open software also permits user specific extensions and modifications. Often, source code and documentation are publicly available, which facilitates user understanding and proposal of changes to the CI. In some instances, proprietary software packages (such as Matlab) may be used where no open source substitute exists.

5.3.2 Resource Management, Data Storage and Data Management

This category contains requirements related to the storage, archival, meta-data annotation, aggregation and harmonization of data resources, and data streams, as well as generic kinds of CI resources, such as instruments and models.

RWS1-R9: The CI shall provide a catalog for all resources under CI governance [*]

RWS1-R10: The CI shall provide the capability to discover all resources based on provided selection criteria [*]

Explanation: Selection criteria apply to resource descriptions, meta-data, parameters, location, observatory, etc. Discovery includes resources connected to the OOI observatories, as well as user-provided electronic and data resources.

RWS1-R11: The resource catalog shall include information about physical samples

Explanation: Physical samples refers to biological, chemical or geological samples retrieved from the seafloor or water column. This implicitly contains the requirement that metadata be associated with physical samples.

RWS1-R12: The CI shall support links to non-OOI resource catalogs and metadata

RWS1-R13: The CI shall provide unique identification for resources, including data streams and data sets [*]

Explanation: Unique identification applies to both human and machine interactions.

RWS1-R14: The CI shall provide pointers from entries in the resource catalog to the resource subject [*]

RWS1-R15: The CI shall provide pointers from entries in the resource catalog to their associated metadata [*]

RWS1-R16: The CI shall bind metadata to all resources connected to an OOI observatory from inception to removal [*]

Explanation: This is a cradle-to-grave requirement to associate metadata with resources, but does not specify the metadata format or content.

RWS1-R17: The CI shall incorporate information on citation and correspondence of resources into the bound resource metadata [*]

Explanation: The term citation refers to statements about the use (including its outcome) of a given resource by another resource or actor. Correspondence refers to a statement of association between two or more resources.

RWS1-R18: OOI-standard metadata shall include, but not be limited to, a complete description of behaviors, content, syntax, semantics, provenance, quality, context and lineage [*]

Explanation: Metadata standards will be externally imposed since the OOI is federally funded, but the OOI standard will probably need to go beyond them. The term behaviors refers to the inherent characteristics of a resource (such as the range of sample rates that an instrument is capable of). The term content refers to the characteristics of any externally-presented information provided by a resource (for instance what an instrument measures, including calibration information). The term syntax refers to a model for the resource content based on structure. The term semantic refers to a model for the resource content based on meaning. The term provenance refers to the resource origin. The term quality refers to information on the QA/QC status of a resource. The term context refers to information about resource usage (such as the geographic location of an instrument). The term lineage refers to information about the evolution of a resource (such as versioning of data due to QA/QC.

RWS1-R19: The CI shall relate different data streams that are based on the same source data [*]

Explanation: It shall be possible to identify different data streams that are based on the same instrument source data with possible but not necessary differences in sampling rate, quality of service parameters, meta-data annotations or post-processing algorithms applied.

RWS1-R20: The CI shall offer data stream subscribers fallback options with similar data in case of original data stream unavailability [*]

Explanation: In case of temporary or permanent unavailability of the subscribed original data streams, the CI shall offer data alternatives which are based on the same source data. If desired by the user, this fallback could be transparent.

RWS1-R21: All data or data products associated with an OOI observatory shall be archivable

Explanation: Whether or not to archive a given data set is a policy decision of the OOI operators that may be driven by economics.

RWS1-R22: The CI shall facilitate the archival of versioned data

Explanation: Versioning may occur due to changes in QA/QC state.

RWS1-R23: The CI shall verify the accuracy of archived data throughout the OOI life cycle

Explanation: This encompasses the requirements to verify that archived data accurately reflect the original, and that archived data are protected from loss due to media degradation or technology change.

RWS1-R24: The CI shall ensure that archived data are up to date

Explanation: This encompasses the requirement to make sure that the latest version of data are archived.

5.3.3 Data Analysis, Modeling and Dissemination

This category contains requirements related to the manipulation and dissemination of data resources, data streams etc.

RWS1-R25: The CI shall facilitate the integration of multiple data streams or data sets into a single stream or set, including elimination of redundant entries [*]

RWS1-R26: The CI shall support notification of changes in resource state [*]

Explanation: The term state refers to behaviors or characteristics that persist (for instance whether an instrument is on- or off-line, or changes in QA/QC state for archived data). Notification applies to subscribers of data streams.

RWS1-R27: The CI shall provide a standard set of tools to compose and execute processes [*]

Explanation: A process is a sequence of human or machine executed steps that are applied to a data stream or data set. This included data cleaning, filtering, aggregation as well as modeling algorithms.

RWS1-R28: The CI shall facilitate data manipulation such as re-projection, re-gridding, subsetting, averaging, filtering and scaling [*]

Explanation: These data manipulations shall facilitate model integration and combination.

RWS1-R29: The CI shall facilitate alignment of data gridlines based on resource meta-data when combining multiple models [*]

RWS1-R30: The CI shall facilitate publication of processed data streams as new data streams [*]

Explanation: This includes data streams containing filtered, processed, aggregated data as well as model computation and simulation output. Such computed data streams shall be treated similar to source data streams and have similar properties, including unique identification, catalog entry, meta-data etc.

RWS1-R31: The CI shall provide subscription facilities to data streams [*]

Explanation: This includes data streams containing unprocessed source data as well as processed and aggregated data

RWS1-R32: The CI shall provide time zone conversion capabilities for subscribed data resources [*]

Explanation: Timed data in data streams shall be automatically converted to users' preferred time zone if so desired.

RWS1-R33: The CI shall provide resource access statistics

Explanation: The CI shall keep track of resource access and usage and provide statistics based on this collected data to interested parties as data stream or on request.

5.3.4 Presentation and User Interfaces

This category contains requirements related to the management and user interfaces of the CI as presented to human and machine users.

RWS1-R34: The CI shall provide web-based user interfaces

Explanation: This shall include a portal into the CI, and hence into the OOI observatories.

RWS1-R35: The CI shall provide the capability to make OOI-standard metadata human readable [*]

RWS1-R36: The CI shall facilitate the integration of user-friendly 4D data visualization tools [*]

RWS1-R37: The CI shall facilitate resource listing based on user selected sort criteria [*]

> Explanation: Different sorting options depending on the type of resource and meta-data shall be provided to the user to select from.

RWS1-R38: The CI shall provide real time tailorable data plotting capabilities [*]

5.3.5 Documentation and Development Process

This category contains requirements related to the development process of the CI, user involvement, availability of CI development materials and artifacts, documentation etc.

RWS1-R39: Web-based documentation for all components of the CI shall be available

RWS1-R40: A mechanism to incorporate user-suggested modifications to the CI shall be provided

> Explanation: The mechanism must include checks to ensure compatibility and consistency with observatory policies (such as security).

RWS1-R41: CI source code developed by the CIIO shall be publicly available

RWS1-R42: The CI shall provide documentation for any application program interfaces (APIs) to CI services

5.3.6 Security, Safety and Privacy Properties

This category contains requirements related to security, safety and privacy aspects of the CI.

RWS1-R43: The CI shall provide mechanisms to enforce user privacy policies [*]

> Explanation: The policies will presumably be defined by the OOI contractors and NSF in consultation with representatives of the user community.

RWS1-R44: The CI shall provide for the sharing of resources subject to specified policies [*]

> Explanation: The policies will be set by the OOI operators, and are constrained by resource providers and external entities such as the Navy.

RWS1-R45: The CI shall provide access to resources subject to use policy [*]

> Explanation: The extent of access depends on explicit resource policies set by the OOI operators and resource provider. This particularly applies to accessing resources discovered in the resource catalog and affects the extent of linkage provided from the catalog to the resource.

5.3.7 Quality Properties

This category contains requirements related to non functional properties of the CI including quality properties, scalability and maintainability requirements.

RWS1-R46: The CI shall deliver messages with accuracy comparable to that of the Internet

> Explanation: to be specified

RWS1-R47: The CI shall support real time, guaranteed delivery, pull mode, streaming and register to receive communication capabilities

Explanation: Real time, in this context, refers to minimum delay commensurate with latency on the channel. Guaranteed delivery refers to storage of a message until an acknowledgement receipt is received. Pull mode refers to storage of a message pending receipt of an explicit request for it. The term streaming refers to an asynchronous, continuous format. The expression register to receive refers to the statement that a given actor or resource must explicitly ask to receive a given stream, at which point it is sent directly.

6 Workshop Conclusions

6.1 In- Workshop Statements

The following sections list feedback statements from the workshop participants that were given during and at the end of the workshop in specific feedback sessions. The statements are listed anonymously and in no given order. Statements from each different person are grouped together in order to ease understandability. Statements might be redundant, overlapping and contradictory due to the fact that they originate from different individuals.

6.1.1 Feedback from science users

- Discussions during the first presentations outside of the agenda were helpful for an initial understanding of the current project status and workshop goals
- An iterative approach to system design is very valuable. Interactions with domain users should continue along the system design and development
- Language tangibility increased by discussions in this workshop
- The presented domain modeling examples were very helpful. Especially the MARCOOS diagram, where people already knew the content in details.
- The provided previously circulated example use case scenario and the ConOps document were helpful.
- The CI overview and context setting presentations were important to understand the context of the CI development and the workshop.
- Helpful about the requirements methodology presentation were in particular the given good/bad examples.
- Very helpful in the domain model presentation were the example models of the topics covering the scientists' own work.
- Should get workshop into details earlier. Context setting could be accelerated
- Possible topics in an initial workshop session: Take questions initially in order to not to loose people
- The participants agree on success of the workshop and on deepened common understanding

6.1.2 Feedback from the organizers

- The planned agenda and discussions worked mostly very well. The extensive morning discussions along the initial presentations established the successful context for the workshop. The context setting part should be streamlined, to not overwhelm participants. The first discussion session should start in the second half of the first morning.
- The science user participants worked together very well and their presentations were very informative and valuable despite the fact that the presentations were mostly monologs.
- For future workshops, a time limit should be set and communicated for science presentations. A questionnaire should be sent to the science users ahead of the next workshop asking to provide information in a number of categories. This could create a preceding exposure and required mental effort.
- The modeling session at the end of the first day could help streamline the discussion.

- It is important to take the breaks at designated times.
- The context setting and mutual getting to know part during the first half day should not be rushed. Rushing could hurt the process. The first half day is required for establishing trust relationships, so that second half can be productive.
- Should extend workshop to two days not to cut valuable discussions off.
- Use workshop location without many distractions
- Target audience in future workshops might be different. Not all audiences might be as computer/software modeling experienced; therefore, different strategies are required
- One integrated project- context setting presentation with overview, requirements, architecture would help to avoid unplugging and unrest.
- Having an explanation of the domain modeling notation upfront and maybe the example model itself can further understanding.
- There is a need for more explicit requirements extraction and validation work in the next workshops

6.2 Conclusions from the Organizers

The first OOI CyberInfrastructure requirements workshop hosted by the Rutgers University's Coastal Ocean Observation Lab (COOL) was very successful in providing valuable outcome for the further CI architecture and design efforts, and in fostering the mutual understanding of prospective CI user communities and the CI design team. Direct outcomes, such as the list of identified requirements, the jointly developed instrument domain model, and the domain models developed by the CI design team in preparation for the workshop – all contained in this report – will be valuable assets in the subsequent CI design efforts leading to Preliminary Design stage and beyond. Further results include a validation of the requirements previously collected for the Conceptual Architecture, and first outreach measures to future CI user communities.

The subsequent two planned requirement workshops are expected to deepen the mutual understanding, and further refine, complement and validate OOI CI requirements and design, based on the previously elaborated results through a complementation with input from a different user community.

All presentation materials can be found on the workshop CD.

Appendices

A OOI Project Background

Almost two centuries of ship-based expeditionary research and three decades of satellite observations have successfully revealed basic descriptions of the oceans and their interactions with terrestrial and atmospheric systems. As understanding has advanced, it has become increasingly apparent that many critical processes occur at temporal and spatial scales that cannot be effectively sampled or studied using these traditional tools. Ship-based studies, for example, are limited in their ability to investigate episodic events or to study dynamic systems that change over time periods longer than a few months. The signal levels associated with some phenomena (such ase.g., secular changes in temperature of the ocean or the motions of tectonic plates) are small compared to background fluctuations, and require measurements over months to years just to resolve a signal. Some processes occur only during certain times of the year or during energetic events, such as storms, when measurements from shipboard platforms are simply not possible. Therefore, the National Science Foundation has begun to implement the Ocean Observatories Initiative (OOI) to advance science, technology, education, and public awareness by building an interactive suite of ocean observatories.

The OOI will provide users with the means to explore the oceans for decades through three interconnected systems spanning global, regional and coastal scales. The global component addresses planetary-scale problems via a network of moored buoys linked to shore via satellite. A regional cabled observatory will 'wire' the Juan de Fuca plate in the Northeast Pacific Ocean with a high speed optical and power grid serving a wide range of tectonic, physical oceanography, hydrothermal vent and biogeochemistry issues. The coastal component of the OOI will expand the growing set of national coastal observing assets and provide extended opportunities to adaptively study high frequency forcing in coastal environments. These systems are to be bound together by a cyberinfrastructure (CI) backbone that will integrate the OOI observatories, and associated sensors and models, into a coherent system-of-systems with full two-way communication capabilities to facilitate direct interaction with the oceans. Additionally, the CI allows the existing network to also access a diversity of other (i.e., non-OOI) data streams and provide a coherent 4-D view of the ocean for any researcher. The completed OOI will allow anyone to interactively view any element of the global system, and mirrors ongoing international efforts by Canada, Japan and Europe.

The science to be addressed by the proposed ocean observatories will span many themes including climate change and biogeochemical cycling; ecosystem dynamics, turbulent mixing and biophysical interactions; and fluids & life in continental margin sediments. The broad range of science questions will be addressed with a distributed network of deep ocean moorings, a plate-scale Regional Scale Nodes (RSN) element, a long term coastal time series Endurance array, and a coastal adaptive sampling Pioneer array. The distributed network of hardware will enable science that spans time and space scales beyond any of the individual components of the OOI through an advanced CI capability that allows any actor two-way interactivity, command & control, and data discovery through both real-time, and historical data archives. The user requirements of the CI are diverse, and vary with specific science needs that have been identified in a series of NRC and community science documents. Specific science goals that will require CI capabilities include:

● *Data discovery and acquisition from real-time data streams and historical distributed data archives.* The ability to utilize real-time and historical data will be central to enabling hindcast and data assimilative nowcast and forecast modeling efforts focused on a range of science problems such as: What are the temporal dynamics in ocean mixing processes? What is the dissolved and particulate distribution and speciation of trace elements, nutrients, and carbon species in the water column? What is the sequestration of carbon from the surface ocean to the deep sea? These questions will require aggregation of all archived data distributed between OOI components and national/international science agencies to provide sufficient spatial data for hindcast studies of the coupling between the physics, chemistry, and biology of the water column. Real-time data required by data-assimilation numerical models will be used to conduct predictive skill experiments utilizing the well-sampled ocean of the OOI environment to assess which models best predict ocean conditions. At the heart of these questions is the requirement for spatial time series to statistically characterize ocean processes over a range of spatial (millimeters to tens or hundreds of kilometers) and temporal scales (seconds to decades) to resolve everything from turbulence to seasonal and interannual dynamics. Examples of non-OOI data streams might include ocean data collected by NOAA's Integrated Ocean Observing System (IOOS) or ARGOs program, satellite data products collected by NOAA-NESDIS, NASA, and/or international science partners.

● *Adaptive sampling:* One of the powerful capabilities of the OOI is the ability to adaptively sample the ocean based either on model forecasts or on real-time data. This will be a new capability for oceanography that will allow scientists to study episodic events and/or explore unexpected processes in depth. Questions that this OOI capability will be absolutely key for include: What are the air-sea exchange processes that occur during severe weather events? What processes determine the transport of carbon, nutrients, and planktonic organisms in coastal ocean ecosystems? What are the key processes that drive the sequestration of material from the coastal ocean to the deep sea? What is the shelf-slope exchange of heat, salt, and nutrients, and what is the corresponding impact on the marine food web? For example for shelf slope exchange studies the decorrelation length scales are equal in both the alongshore and cross-shore directions, traditional sampling cannot resolve the appropriate scales, and the available data hinder numerical studies of the potential key processes regulating the exchanges. Therefore, the array needs to adaptively sample the shelf/slope exchange of heat, salt, and biological material in order to identify and quantify the dominant driving mechanisms. Studies on the shelf-slope exchange processes will be a focus of the Pioneer Array consisting of moorings, autonomous underwater vehicles (AUVs) with docking stations, and gliders. Another example is use of the RSN to study earthquake and vent phenomena where mobile assets might be dispatched to collect spatial maps over vents fields triggered by tectonic activity. The CI network will allow researchers to adaptively sample the environment using a variety of mooring and AUV systems that will need to be operated collectively to provide the real-time ability to access data from the ocean, and adjust either the sampling from individual sensors on a mooring or to initiate the coordinated movement of multiple mobile assets. The CI provides an easy-to-use, real-time framework for scientist actors to adjust their initial sampling strategy, which is typically based on preconceived notions, when unexpected events or processes are encountered. These undiscovered processes are likely to be the norm during the OOI field campaigns, as the oceans remain chronically under sampled.

● *Automated ocean surveillance for event detection and classification procedures.* Traditional attempts at adaptive sampling are driven with scientist actors initiating and guiding changes in the sampling. The CI will expand on this by providing a capability for the network to initiate sampling in an automated fashion based on real-time data. The CI will allow scientists to program behaviors that then drive the network responses to events. This requires a robust command and control capability. The network behaviors will be initialized by data or derived data products. For example, recent work by NASA is developing bioinformatics algorithms that combine all available data streams to objectively classify unique water masses that could conceivably initiate increased chemical and biological sampling by the OOI assets as new water masses pass into the study area. Science questions that will rely heavily on this capability include: What are the air-sea exchange processes that occur during severe weather events? How do storms drive benthic resuspension processes, and do these resuspension events play a critical role in initiating phytoplankton blooms? What physical and biological processes lead to hypoxic conditions in the coastal ocean? What regulates the transport and transformation of nutrients on continental shelves? For example, wave and optical data from OOI moored sensors might be used to detect the onset of storm-induced mixing and resuspension. This would automatically trigger both vertical profilers and mobile assets to study the time-dependent spatial evolution of the resuspension and corresponding transport in the water column.

The automated activity may include single sensors or swarms of mobile assets, all of which have been provided appropriate behaviors for the mission at hand.

● *Facilitating collaborative investigation that coordinates science campaigns across teams of researchers*. Ocean observatories have demonstrated the ability to act as magnets for large science campaigns. For example, field campaigns planned around the existing ocean observatory in the Mid-Atlantic Bight have repeatedly involved hundreds of scientists, multiple ships, 7 aircraft and upwards of 50 moorings. Many of these assets were in addition to the existing observatory which provided the environmental context. This allowed the non-observatory science assets to focused on the specific process of interest. The OOI will provide the infrastructure thus provides interactive framework and data that allows a single PI effort to become a component of large interdisciplinary study.

● *Coupling analytical capabilities to observational data stream.* Models and observations rarely operate in a two way interactive mode in real-time. The CI provides the opportunity to enable real-time data analysis and models for the synthesis across the multiple dimensions of a process of interest. Here suites of models can made available to an OOI researcher as an analysis tool to understand their observational data in order to allow for hypothesis testing. For example biological oceanographers are on the cusp of conducting predictive skill experiments given the newly available data from the ocean. Experiments in this "well-sampled" ocean may be allowed to run ensembles of biological models and then determining which models perform well or poorly. This information will allow then allow the scientist to ask "why" does one model perform well.

B Workshop Preparation Materials

B.1 Research Questions Traceability Matrix Analysis

Figure 7 shows a simplified ocean observation and modeling workflow as UML activity diagram.

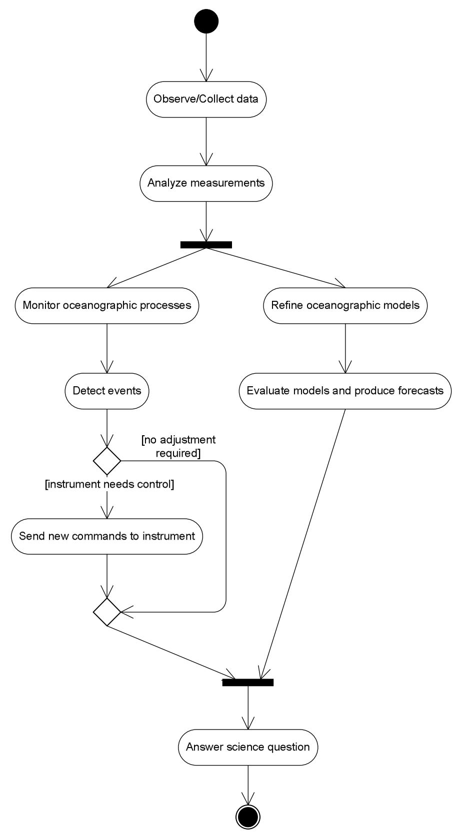

Figure 7: Scientific Observation Activity Diagram

Figure 8 provides a domain model for ocean observation and modeling as UML class diagram

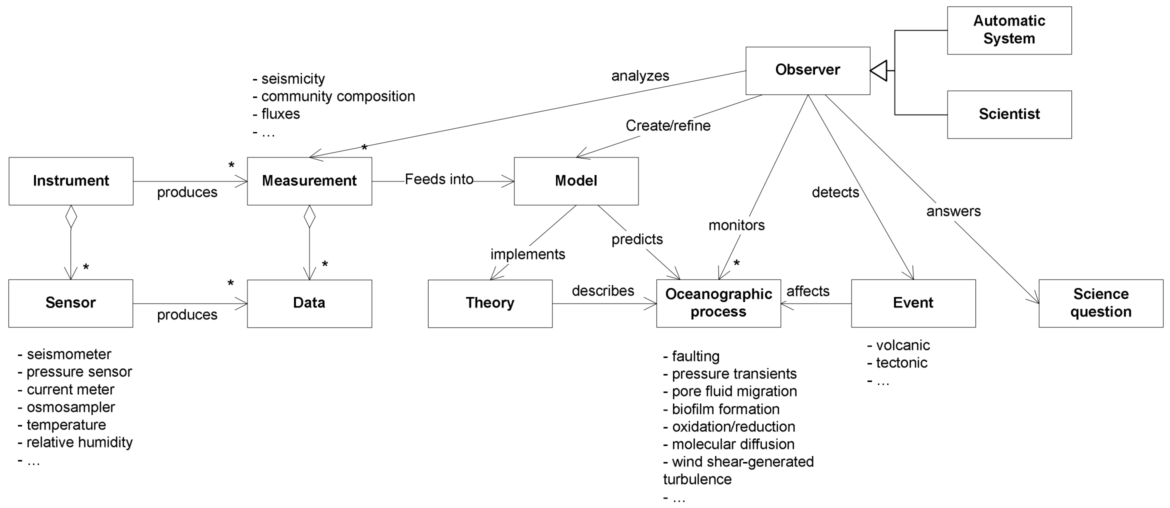

Figure 8: Scientific Ocean Observation Domain Model

Figure 9 provides a domain model for instrument platforms, extracted from the science question traceability matrix.

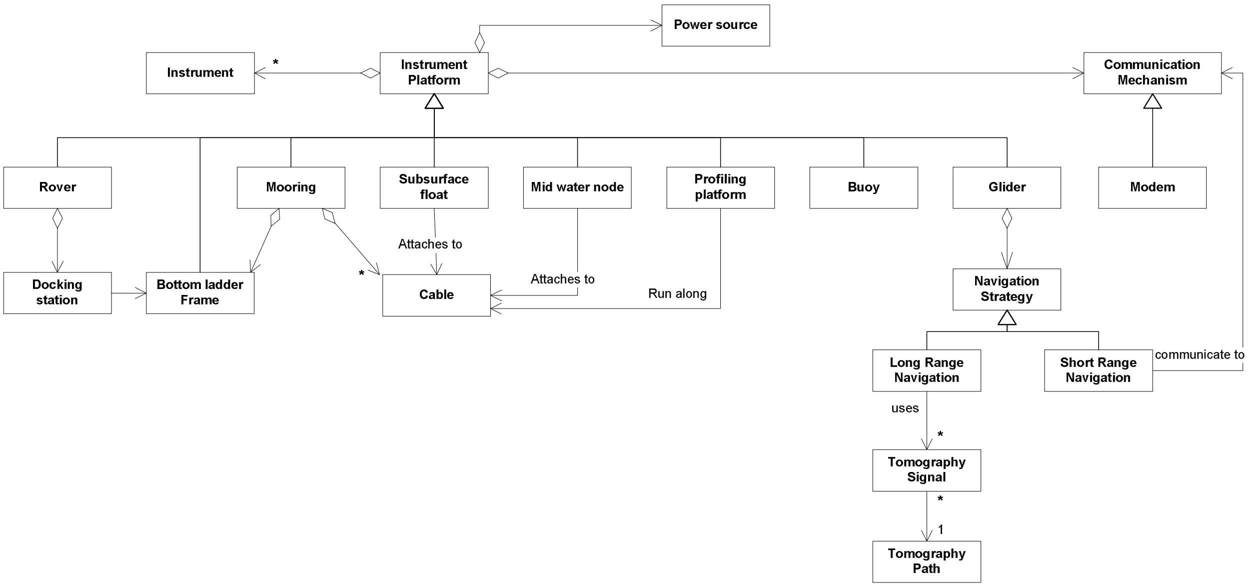

Figure 9: Instrument Platform Domain Model

Figure 10 provides a domain model for instruments, extracted from the science question traceability matrix.

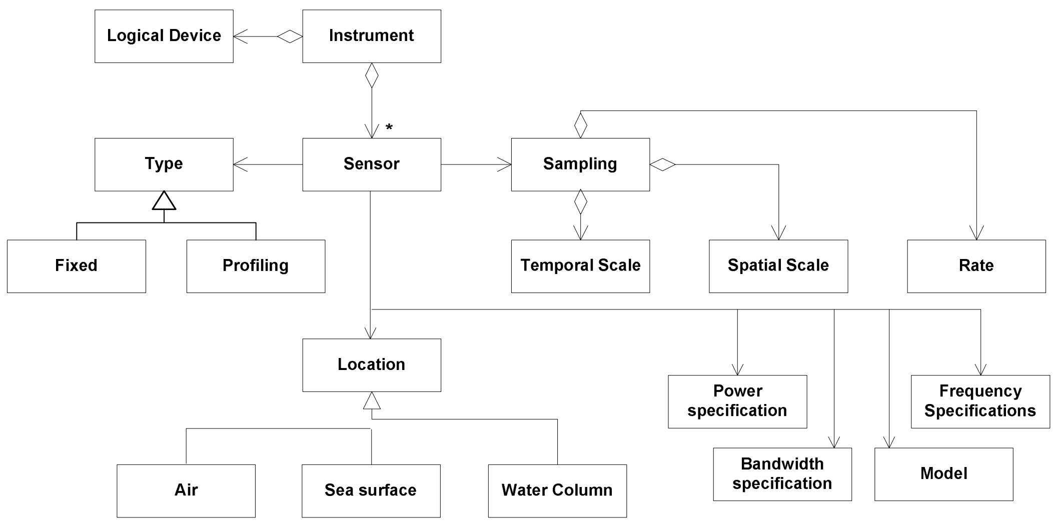

Figure 10: Instrument Domain Model

B.2 MARCOOS Proposal Domain Analysis

Figure 11 provides a domain model for the mid-Atlantic ocean observation and modeling data flow as proposed in the MARCOOS proposal.

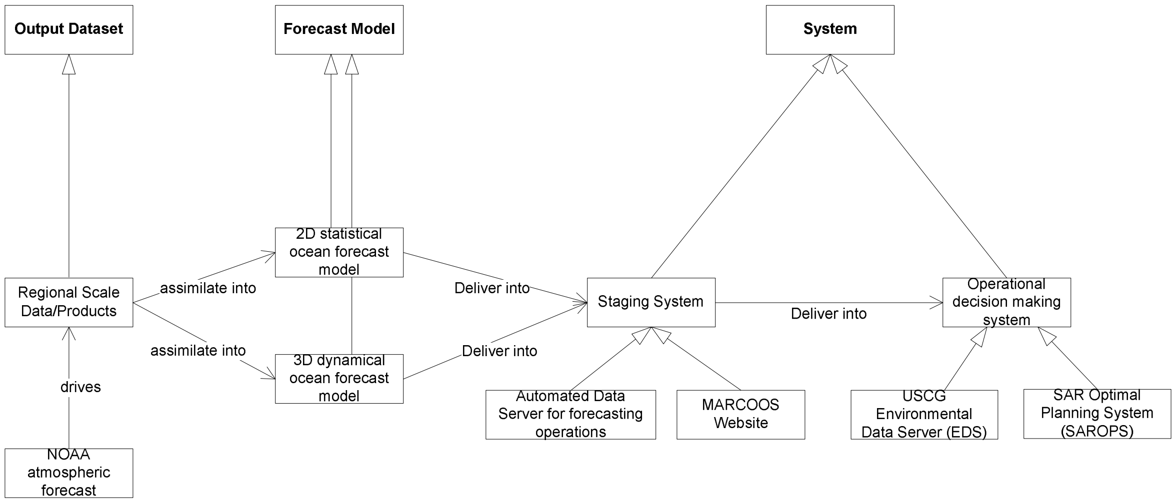

Figure 11: MARCOOS Data Flow Domain Model

Figure 12 provides a domain model for the mid-Atlantic ocean observation and modeling tasks and responsibilities as well as goals as proposed in the MARCOOS proposal.

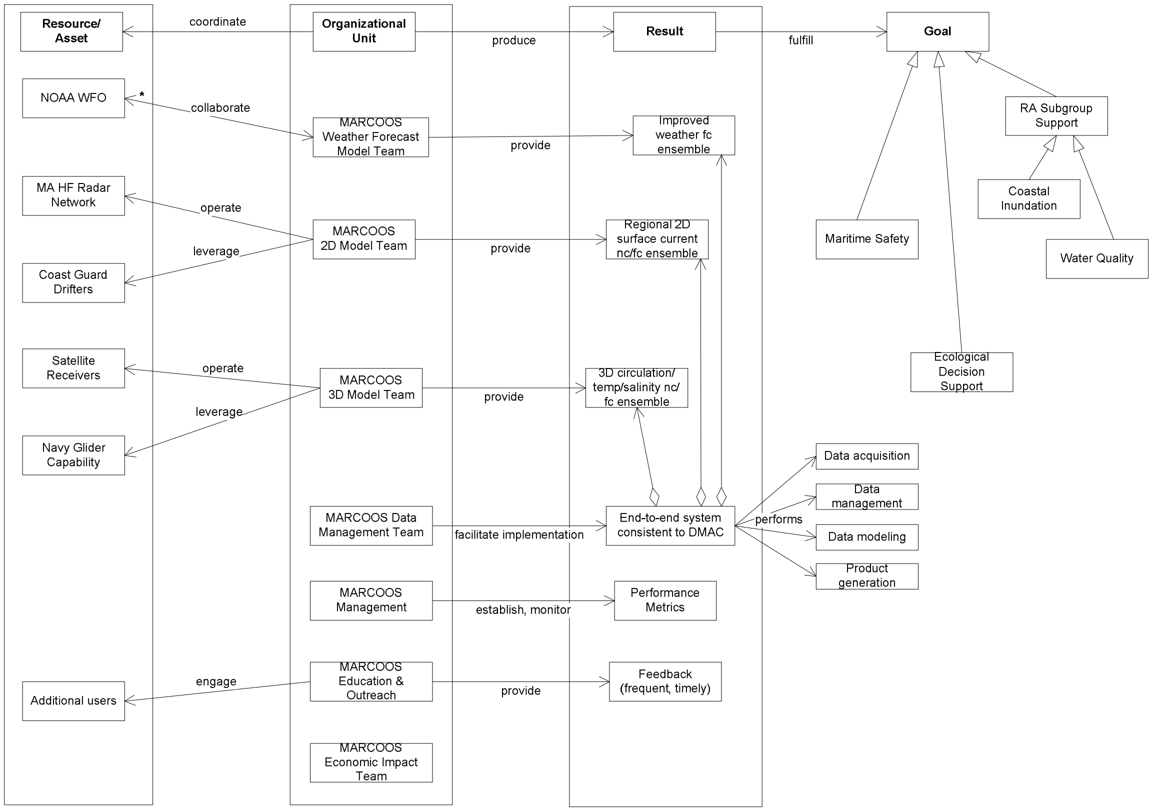

Figure 12: MARCOOS Project Tasks and Results Domain Model

Figure 13 provides a domain model for the mid-Atlantic ocean observation and modeling project domain as proposed in the MARCOOS proposal.

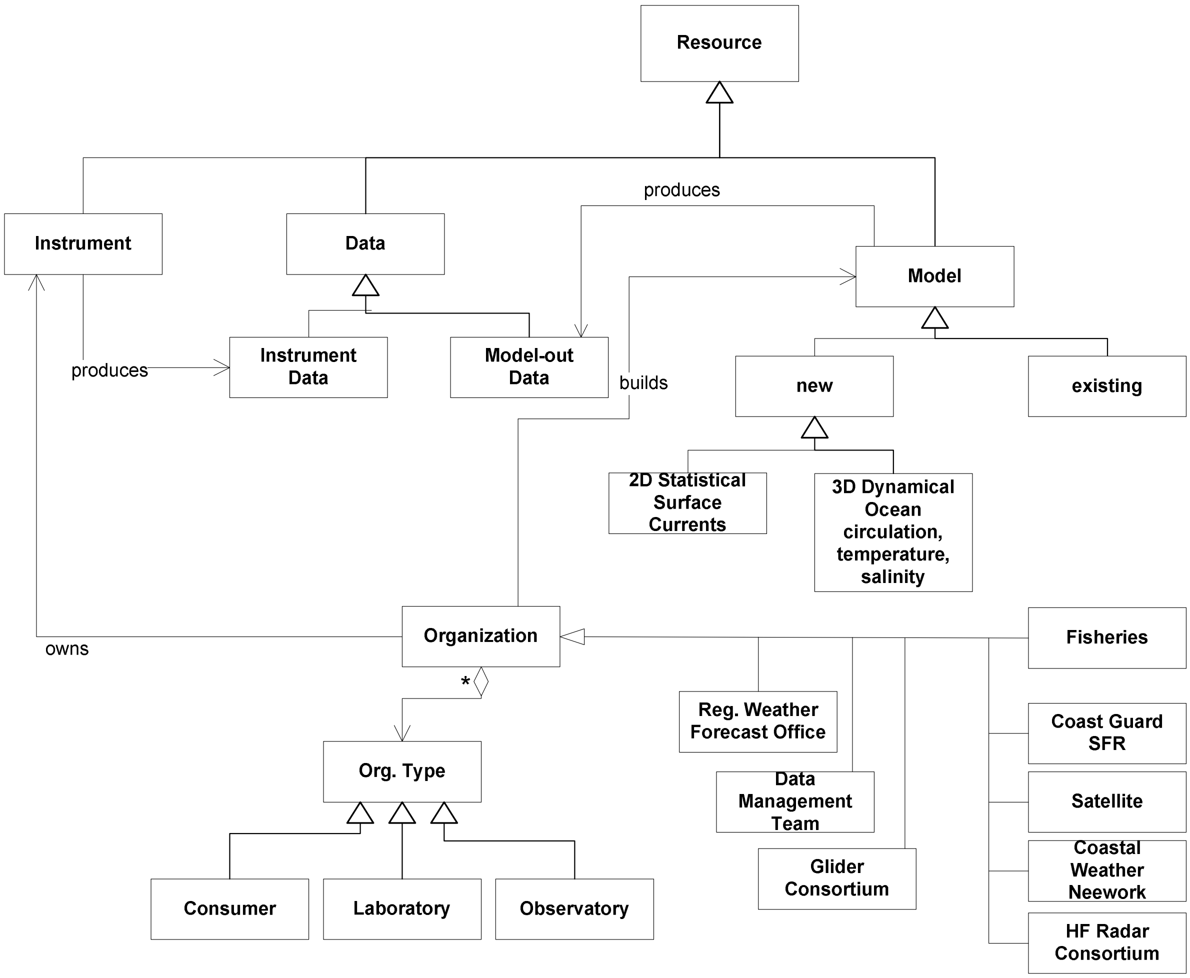

Figure 13: MARCOOS Source and Result Data Domain Model

C Workshop Agenda

Day 1, July 23, 2007 (Monday)

Time	Presenter(s)	Topics
08:15 AM	O. Schofield, S. Glenn	Welcome & Introductions
08:30 AM	O. Schofield	Meeting Purpose & Agenda Status of OOI & 3 Implementing Organizations Scientific User Requirements Workshops Background & High Level Science Questions
09:30 AM	M. Arrott	CI Project Overview, Conceptual Architecture/Requirements
10:30 AM	A. Chave	Concept of Operations and Role of Requirements
11:00 AM	I. Krueger	Overview of Requirements Gathering Process/Methodology/Notations
11:45 AM	O. Schofield	Preliminary Feedback Session
01:30 PM	R. He	Regional Circulation modeling for the mid-Atlantic Bight & Gulf of Maine (MABGOM)
02:30 PM	J. Wilkin	Prediction and Analysis for Mid-Atlantic Bight Ocean Observatories
04:00 PM	W. Brown	Advanced Fisheries Management Information System (AFMIS) for the Northwest Atlantic Ocean
05:00 PM	O. Schofield	Day 1 wrap-up and feedback session

Day 2, July 24, 2007 (Tuesday)

Time	Presenter(s)	Topics
08:30 AM	O. Schofield	Overview of Results & Revisions Charge for Final Discussion
09:00 AM	O. Schofield, I. Krueger	Discussion Session: CI Science Questions and Deployment Scenario
12:00 PM	O. Schofield	Final feedback session Workshop Adjourns

D List of Participants

Name	Organization	Project Role
Arrott, Matthew	UCSD/Calit2	Project Manager
Brown, Wendell	UMass Dartmouth	Domain Scientist
Chave, Alan	WHOI	System Engineer
Ermagan, Vina	UCSD/Calit2	System Architect Support
Glenn, Scott	Rutgers	Project Scientist
He, Ruoying	NCSU	Domain Scientist
Klacansky, Igor	UCSD/Calit2	System Architect Support
Kleinert, Jack	Raytheon	System Engineer Support
Krueger, Ingolf	UCSD/Calit2	System Architect
Meisinger, Michael	UCSD/Calit2	System Architect Support
Schofield, Oscar	Rutgers	Project Scientist
Wilkin, John	Rutgers	Domain Scientist

Figure 14: Group Picture at Rutgers University's COOL Operations Center

E Abbreviations

Abbreviation	Meaning
AUV	autonomous underwater vehicle
CODAR	Coastal Radar
ECMWF	European Centre for Medium-Range Weather Forecasts
ESMF	Earth System Modeling Framework
GTS	Global Telecommunications Network
IOOS	Integrated Ocean Observing System
MPI	Message Passing Interface (used in parallel computation)
NCAR	National Center for Atmospheric Research
NetCDF	Network Common Data Form
NOAA	National Oceanic & Atmospheric Administration
ONR	Office of Naval Research
PCA	Principal Components Analysis
ROMS	Regional Ocean Modeling System
RSN	Regional Scale Node
SST	Sea Surface Temperature
THREDDS	Thematic Realtime Environmental Distributed Data Services

F References

Reference	Citation
[CI-CONARCH]	CI conceptual architecture and initial requirements, available at http://www.orionprogram.org/organization/committees/ciarch
[MARCOOS]	MARCOOS Proposal
[NORIA]	NORIA Proposal, 22-Dec-2006
[SQTM]	Science Question Traceability Matrix documents, working draft version of 6/18/07. Contains • Fluid rock traceability matrix of 6/18/07 • Carbon Cycle traceability matrix of 6/18/07 • Endurance Array Science Matrix of 6/18/07 • Example Science Use Case 2 XLS of 6/18/07 • Global Science Questions Draft of 6/18/07