OOI CYBERINFRASTRUCTURE
NSF OCEAN OBSERVATORIES INITIATIVE

OCEAN OBSERVATORIES INITIATIVE

OOI – CyberInfrastructure

Requirements Workshop
Integrated Observatory Management
Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA
May 28-29, 2008

Workshop Report

Elizabeth Rosenzweig, Alan Chave, Michael Meisinger

CANDIDATE

August 2008

CONSORTIUM FOR Ocean Leadership
DISCOVERY • UNDERSTANDING • ACTION

UCSD

Document information

Project	Ocean Observatories Infrastructure (OOI) CyberInfrastructure (CI)
Document Custodian	OOI CI Architecture & Design Team (ADT)
Approval	Report: Alan Chave (Woods Hole Oceanographic Institution), OOI CI IO System Engineer User Requirements: Frank Vernon (UCSD), OOI CI IO Deputy Director
Created on	May 15, 2008
Last Changed on	August 12, 2008
Document Status	CANDIDATE

Document History

Date	Version	By	Description of Changes
2008-05-15	0.1	M. Meisinger	Initial workshop report from notes and previous workshop report structure
2008-06-10	0.2	E. Rosenzweig	Updated to include information from IOM Workshop
2008-06-30	0.3	E. Rosenzweig	Updated with detail information for scenarios and requirements
2008-07-17	0.4	E. Rosenzweig	Updated to include feedback from A. Chave and Matt Arrott
2008-08-05	0.5	E. Rosenzweig	Final updates include final list of requirements and new comments from A. Chave
2008-08-08	0.6	A. Chave	Review
2008-08-08	0.7	M. Meisinger	Formatting for candidate publication, status changed to candidate

Table of Contents

OOI - CyberInfrastructure

Integrated Observatory Management
Requirements Workshop at SIO/UCSD, May 28-29 2008
Outcome and Summary

1 Executive Summary

This document summarizes the results of a requirements workshop held at UCSD on May 28-29, 2008. This workshop was held as part of an ongoing series whose goal was to further the understanding of user requirements and constraints for the planned Ocean Observatories Initiative (OOI) CyberInfrastructure (CI). The OOI CyberInfrastructure Implementing Organization (CI IO) Architecture Development Team (ADT) and System Engineer organized the workshop.

This workshop was the fifth in a series, succeeding four prior requirements meetings held between July 2007 and May 2008 (see [CI-RWS1], [CI-RWS2], [CI-OOP-WEB], [CI-DPG-WEB]). Invited participants included observatory operators, system engineers and scientists from the community as well as from the Regional, Coastal and Global Observatories of the OOI.

Goals of the Integrated Observatory Management workshop were:

- Capture knowledge from field operators about their systems, observatories and daily activities
- Elicit, identify and document science user requirements
- Validate, refine and prioritize existing science user requirements
- Provide an opportunity for information exchange between CI engineers and observatory operators
- Create and solidify outreach and communication with future CI user communities.

The workshop outcome and results include

- Additional CI user requirements provided by ocean observing community members
- Domain models elaborated during the workshop
- CI use case scenarios for ocean observing elaborated during the workshop
- Collected workshop presentation materials including introductory presentations (OOI, CI, science) on the OOI CI Confluence web site [OOP-WEB]
- Science user questionnaires for requirements elicitation (extended and short versions)
- Completed participant questionnaires

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 over a decade of national and international scientific planning. 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 Navy's proposed Littoral Battlespace and Fusion Integration program (LBSFI). Additionally, the global community spanning Canada, Asia, and Europe are also developing new ocean networks which all contribute to the GEOSS. Developing a robust capability to aggregate these distributed but highly linked efforts is key for their success.

The OOI comprises three distributed yet interconnected observatories spanning global, regional and coastal scales that, when their data are combined, will allow scientists to study a range of high priority processes. 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 realize the OOI mission. The infrastructure provided to research scientists through the OOI will include everything from seafloor cables to water column fixed and mobile systems. Junction boxes that 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 are central to these observational platforms. 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 of facility resources, and operational activities, along with 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 committee established by JOI in 2006 (see http://www.orionprogram.org/organization/committees/ciarch) [CI-CARCH]. It describes the core capabilities of such a system. Initial requirements were derived from similar cyber-infrastructure projects.

In May 2007, a consortium led by SIO/UCSD, including JPL/NASA, MIT, MBARI, NCSA, NCSU, Rutgers, Univ. Chicago, USC/ISI and WHOI, was awarded a contract to be the Implementing Organization (IO) for the development of the OOI CI. The first six months of the design phase has focused on architecture and design refinement and consolidation, and an initial science user requirements analysis and community involvement effort. In December 2007, the preliminary CI design [CI-PAD] was successfully reviewed in a PDR (Preliminary Design Review) by a panel of independent experts appointed by NSF, who provided very positive review comments.

Current activities are targeting the Final Design Review (FDR) in November 2008, where all requirements and design documentation, operations management plans together with cost estimates and feasibility analyses will be reviewed. Major activities towards FDR focus on completing a baseline set requirements at all levels of the OOI and CI, covering user, system and subsystem requirements, with clear tracing to OOI science objectives [SCIPROSP] and user community expectations. Further activities target advancing the CI design and that of its subsystems to the next level to be ready for the start of OOI MREF construction. During all OOI design and construction activities, the validation of any previously elicited and documented user and system requirements through the community will remain a primary concern. Direct involvement of prospective CI user communities is of paramount importance to the success of the program. The requirements elicitation and management process is planned to be an ongoing activity in close collaboration with the user communities involved throughout the design and construction phases.

Earlier science user involvement occurred during the first CI requirements workshop (RWS1), July 23-24, 2007 at Rutgers University and the second CI requirements workshop (RWS2), January 23-24, 2008 at UC San Diego. For each of these workshops, the outcomes were summarized in the form of publicly available reports [CI-RWS1, CI-RWS2].

This report covers the outcome of the fifth requirements workshop on Integrated Observatory Management (IOM). The workshop took place May 28-29, 2008, at Scripps in La Jolla, CA. Goals of this workshop were:

- Provide the CI engineering team with detailed insight into how existing ocean observatories are and how future ones will be managed and operated
- Identify and elicit new user requirements for the CI from the view of this community
- Validate, refine and prioritize existing user requirements
- Develop a thorough domain understanding through direct collaboration with domain scientists in order to increase language tangibility, and document this understanding in the form of domain models
- Refine and consolidate the basis for further requirements elicitation and domain modeling in subsequent instances of this workshop and in ongoing requirements and architecture design work
- Provide an opportunity for interchange between the CI and the OOI marine observatory IOs
- Advance the common understanding across the individual OOI teams

2.2 Science Background

The science motivating the OOI network is based on research community input collected over several decades. The community documents emphasize the need for simultaneous, interdisciplinary measurements to investigate a spectrum of phenomena, from episodic, short-lived events (tectonic, volcanic, biological, severe storms), to subtler, longer-term changes in ocean systems (circulation patterns, climate change, ecosystem trends). The introduction of in situ power and bandwidth will allow the transition from ship-based data collection to the management of interactive, adaptive sampling in response to remote recognition of an “event”-taking taking place. Sophisticated CI tools will enable individuals and communities of researchers to tackle specific research questions. This ambitious program will require an integrated management capability given the need to maintain a 24:7 distributed facility with a diverse suite of distributed users. Given this, a robust cyber-enabled suite of tools will be required to utilize and manage the network.

A mature network will enable the observatory operators by facilitating their interactions with many distributed users. This requires clear delineation between the domains of authority of the marine and cyberinfrastructure operators. It also requires clear policy and governance resources as well as the identification of evolving performance metrics. It requires the CI to be flexible and best serve the wide range of platforms to be deployed by the OOI. For example, for mobile assets such as autonomous gliders, limited and disparate connectivity via Iridium requires a series of pre-programmed management tools to be remotely deployed upon marine platforms or in the shore stations. These tools might include initiating different science behaviors based on user requests (flight patterns, sensor sampling frequency, or clustered fleet operations, etc.) or operator initiated management needs (response to battery power declines, poor connectivity, close to international shipping lanes, etc.). Many of these capabilities are distinctly different than the needs of the RSN where continuous real-time connectivity may not require as many automated behaviors, but will require significant management of a large suite of diverse sensors.

2.3 Outline

The remaining parts of this report are structured as follows: Section 3 summarizes the presentations given at the workshop and places them into the context of the scientific background. Section 4 documents the direct workshop outcomes, such as discussion summaries, domain models, elaborated scenarios and prioritized requirements. Section 5 lists all user requirements including new requirements identified in this workshop. Section 6 documents participant feedback to the organizers. The appendices contain further details about the workshop organization and background materials.

2.4 Preparation

The CI ADT has developed an extensive questionnaire with relevant questions for user requirements elicitation that was structured into selected categories. A shortened and tailored version of the questionnaire was sent to the workshop participants. The scientists were asked to provide answers to the questions prior to the workshop.

Each scientist was asked to prepare an overview presentation covering projects, research interests, and further relevant background information related to the OOI CI. The presentations were supposed to address the main topics covered by the questionnaire. The presentations covered approximately 15-20 minutes each, including questions.

During the workshop, the extended version of the questionnaire was used to structure the general requirements discussion session. Appendix B of this report documents the extended questionnaire.

2.5 Acknowledgements

This report was developed by the OOI CI ADT; it contains input from many sources, such as the workshop presentations by the organizers and invited science users, the completed participant questionnaires, the CI preliminary architecture and design, OOI science background information from the project scientists, and notes taken by Michael Meisinger, Elizabeth Rosenzweig and Emilia Farcas. Furthermore, this report contains summarizing and general statements by the organizers.

We profoundly thank the invited participants for their time and efforts during the workshop, and their valuable contributions to the OOI CI requirements elicitation process. Furthermore, we would like to thank them for their efforts in filling out the participant questionnaire and providing further materials after the workshop, and for reviewing and validating this report.

2.6 Disclaimer

The contents of this report reflect the understanding and analyses of the CI ADT based on written workshop notes and general background. Errors in transforming them into this report are the responsibility of the CI ADT. No statements in this report are verbatim quotations of participants; there were no audio recordings of the discussions taken during the workshop.

3 Presentations

3.1 OOI CI User Requirements Elicitation Process

Frank Vernon from Scripps welcomed the workshop participants and described the process for science user requirements elicitation. He presented the requirements elicitation process and described the purpose of systematic and iterative requirements elicitation efforts involving multiple user communities over the course of the OOI CI project. The OOI project is preparing for Final Design Review in November 2008. A set of important activities covered in this workshop is refining and expanding user requirements for the OOI integrated observatory as viewed and accessed by the observatory operators.

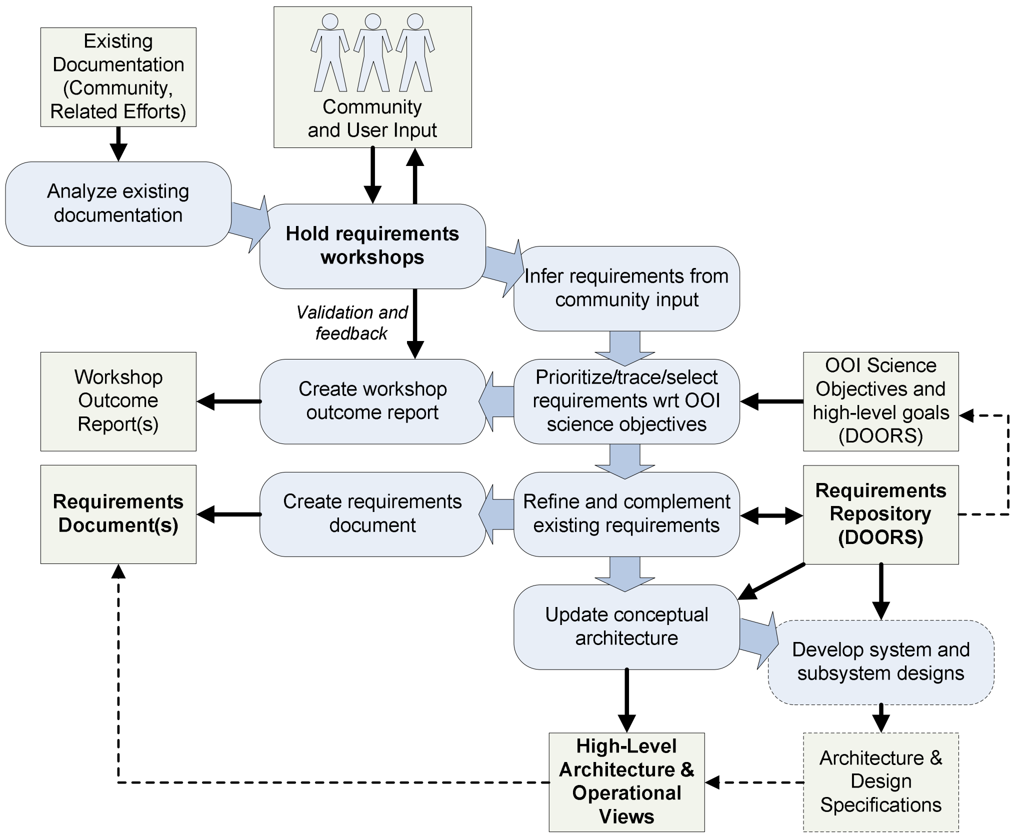


Figure 1: CI user requirements elicitation process

3.2 CI Overview, Requirements, Architecture

Matthew Arrott (UCSD/Calit2), OOI CI Project Manager, provided an overview of the OOI CyberInfrastructure project and the CI project organization. The main goal of the CI is supporting the three main research activities of observing, modeling, and exploiting knowledge through a set of well-rounded resources and services. The CI infrastructure will be distributed across the country and will have points of presence at the sites of the main OOI observatory components on the east and west coasts.

The design process involves several iterations that advance the understanding of requirements and design. Previous design cycles led to the conceptual architecture, the proposal for the OOI CI and design refinements for Preliminary Design Review (PDR) in December 2007. The current iteration emphasizes refining requirements and design for FDR.

One goal of this workshop is detailing out the requirements for (1) the direct access of instruments through commands sent to them, (2) interaction with instrument resource agents through observation plans and (3) interactions and resource allocations between the infrastructure and distributed physical resources. Communities of interest include the OOI marine observatory operators that manage the observatories and their infrastructure resources.

3.3 Project and Research Overview: SCCOOS

Eric Terrill's presentation described the Southern California Coastal Observing System (SCCOOS) network that includes buoys, moorings and gliders, as well as six automated shore stations. This observatory runs 24 hours a day, 7 days a week. Since it is a 24/7 station, it means that staff has to be on call at all hours.

The network seeks to extend climate-relevant physical and biological sampling to near-coast state waters. The stations complement similar data from expeditionary sampling closer to the coast. An integrated climatology based on the two will be developed in the future.

This ocean observatory was developed in a build-test-build environment. This iterative process facilitates quick integration of new learning and experience. As a result, data management needs to be tightly integrated with observers so that they become part of the system.

The project team feels that a vested user community for OOI should be identified and kept engaged. The SCCOOS team was concerned that the OOI team would be voluntary and that would marginalize some parts of the community who didn't have the time or resources to volunteer to be involved.

3.4 Project and Research Overview: Glider Network Management

Mark Moline (CalPoly, San Luis Obispo) presented a high level overview about managing a network of gliders on the southern California coast that includes HF radar; water quality sensors, AUVs, and gliders that perform integrated ocean sampling. The HF radar includes long and standard range instruments. The glider network covers the mid-California coastline and is considered part of the mid and northern California coastal group

Mark and his team track issues and problems with the gliders, as well as providing education and outreach in the bay near San Luis Obispo. The system is developed to meet user needs, and educate communities to understand issues such as pollution and its impact on sea life.

The computer system has a visually-appealing dashboard to view different parameters on water quality with live data about the system. One focus of the system is educating and assisting people interested in purchasing AUVs. These operators report the system to be useful as well as having good organization and utility. In addition, he recommends everyone visit the website.

The vehicle network tracks several types of mobile platforms; among the main types are AUVs that are typically out for 3-4 days. The process of uploading the data for AUVs is similar to the process for gliders. Once they come back, the mission data are uploaded and put on the website. The next step is filtering them with location, objective, distance and other important data.

One of the system's goals is to organize the database for the user. However, the team also feels that they have to continually balance their priorities between user and development needs. They want to make their system usable, but they also need to make sure it is functional and often the resources are limited. In addition, they have become a hub for other projects with organizations such as:

- JPL

- Wetlabs
- Raytheon
- Teledyne RDI
- Reason
- SeaBird

Challenges:

- How does one integrate data from two ships if the data are disparate?
- The system takes into account time and space data. It looks at data taken in the same location over time. This creates some overlap that is useful in analyzing what happens in a specific location over time. This also creates issues when there are multiple investigators, (12 scientists, more then 1 PI) all analyzing data simultaneously.
- The problems are complex because even if you can incorporate the information, how do the process crews integrate the glider and other data?

3.5 Project and Research Overview: VENUS Observatory Management

Paul Macoun (Univ. of Victoria) presented an overview of the VENUS Project (www.venus.uvic.ca), which is a Canadian scientific coastal ocean observatory facility. This observatory has been operational for the past two years; instruments have been installed since 2006, and will continue to be installed in fall 2008. The VENUS infrastructure works with COTS; equipment manufacturers are engaged to test their instruments on the platform. There is an intention to move to Ethernet-based instruments in the future. Ethernet equipment is from Cisco, Moxa, Perle, and custom power equipment

The Data Management and Archive System (DMAS) allows data from all new instruments to be available to everybody. The data are raw at the initial stage, but since the user is the PI it is not an issue since they can make sense of that data. The DMAS uses a virtual private network (VPN) from shore stations to transmit data from the instruments. There are discussions for creating processes for more flexible operations by setting up services that would require payment/funding for operations and maintenance of the instruments.

Paul described the big challenges as: power, access and troubleshooting. Instruments can produce acoustic interference with each other. Avoidance requires smarter software drivers enabling acoustic communications only when other instruments are silent. Ship time and deck space is at a premium. Everything is designed with ROV access in mind. If the ROV capability is not available, the instrument platform is hardly of use. There are a lot of data flowing – CTD sampling at 1/min is not an issue, but other instruments produce up to 50-100GB/day, and storage is becoming a problem. This requires a minimum bandwidth through the system of about 100 Mbit/s. Most operations and maintenance is done through manual intervention via VPN. Data are stored into an Oracle database and made available via a webpage (plots, time series, acoustic data, other images and video streams). QA/QC is not fully implemented yet. Any new instrument is pre-tested prior to deployment in water. DMAS drivers are developed at that time. Security issues with the US and Canadian Navies exist regarding hydrophone sensors.

The system has one primary database that feeds to end-users through the project website, and one archival database at a different facility. Currently, there are no requirements for data streaming, although it would be possible from the main database. The delay of the data stream would be on the order of seconds to minutes, depending on the instrument. Some instruments provide data in hourly or daily chunks; hence the delay could be larger, but is well known in advance. At present, there is no data fusion in place, but there are plans to provide an integrated view of multiple sensor data over a specified period. The system allows the end-user to download the data in multiple formats. Data from the instruments comes in multiple formats, requiring a variety of drivers that convert them into the format accepted by the storage database. The equipment manufacturers have been approached to add custom XML data formats to the instruments and some have provided such capability. Video data are not available yet; only some still cameras (VGA resolution) are deployed and viewable by the end-users when the PIs turn them on.

3.6 Project and Research Overview: MBARI Operations

Craig Dawe (MBARI) presented the MARS project, which was installed a few months ago but suffered a catastrophic failure, and has never been operational. The installation is based at the MBARI facility in Moss Landing, California. MBARI has research ships and remotely operated vehicles that they use for their work.

The MARS project will allow scientists to perform experiments 900 meters below the surface of Monterey Bay. It is a cabled observatory that serves as a test bed for engineering, education, and science projects.

One goal of MARS is maintaining quality by duplicating instruments, but this has not been accomplished to date.

- On the shore side there are instrument proxies.
- Most of the metadata for the system are entered by hand.

MBARI's plan is to perform updates to the configurations regularly and check it on a monthly basis, but perform regular upgrades about once a year, or in general as infrequently as possible. MARS is different from other observatories with regard to access to the assets for maintenance because it is harder to get to due to ts depth and the distances involved. The system includes a 51 km long power/fiber optic cable.

<u>Specific Statements:</u>

- There are many steps in adding an instrument that slows the process of installation down.
- The approval process for deploying new instruments was long, but it has been streamlined to be more efficient. MBARI is continually working to improve their processes.

3.7 Project and Research Overview: Earthscope ANF

Jennifer Eakins (Scripps) presented the Earthscope Array Network Facility, which is an element of a NSF Major Research Equipment (MRE) program to provide a set of observational facilities to promote the science of solid earth geophysics. The presentation focused on the transportable array (TA), one component of the Earthscope project. This is a land based sensor network, currently deployed in the western portion of the United States. Its goal is understanding the structure and evolution of the North American Continent – structure, deformation and properties that control earthquakes and volcanic eruptions. In addition, the Plate Boundary Observatory, another component of Earthscope, was described.

The TA project science objectives include:

- Earth structure determination over the broadest possible scales, looking at data coming from the network in order to understand events in the environment
- Focus studies of teleseismic events
- Consistent and uniform determination for earthquake characteristics in the US

The TA project design objectives include:

- Quasi rectilinear grid of receiver points over the entire USA
- Standardized instrument and acquisition characteristics
- Continuous 24/7 acquisition of data over a time range of at least several years

TA Data processing requirements include:

- "Streaming" data flow model throughout the system – data driven as opposed to event or schedule driven
- Acquisition and processing software must operate properly in this environment
- 24/7 operation requires software to run reliably over long periods of time and be highly fault tolerant (e.g., the need exists to adjust to temporary loss of communications)
- 24/7 operation requires high levels of automation which is challenging since there is currently no 24/7 staffing
- Processing system must support rapid problem determination and resolution

TA sensor calibration involves:

- The system includes an automated process to command, capture and analyze calibration signals applied in situ using Antelope.
- System interprets calibration analyses to verify amplitude and phase response
- Calibration process is applied to all stations at beginning and end of deployment.
- Data are archived as a data product, as opposed to being stored in the raw data format

TA challenges:

- The sensor network is expanding, and as a result the system must be scalable. If it is not, then there are gaps created in data collection, causing errors.
- It is hard to predict all possible problem scenarios. Even though they ran tests, they didn't account for all situations, and so didn't pick up some important gaps
- Quality assurance is a major concern, so analysts look at data everyday to insure that the data are good
- Different types of scientists are using data and the project has more exposure than before
    - Science community is expanding as the TA provides fantastic data, so it has to be readily available and usable.

Another component of Earthscope is the Plate Boundary Observatory (PBO) whose mission is to install, reconfigure and maintain 852 sensor sites, 103 borehole strain meters (BSM) and 5 laser strain meters (LSM) on-schedule, on-budget and to specification within five years. Taken together, these instrument types span the broad temporal and spatial spectrum of plate boundary deformation anticipated in the project area. GPS measures millimeter-scale ground movement on time scales of days to decades and over large spatial scales

The geodetic focus of the Plate Boundary Observatory addresses the following scientific questions:

- What are the forces that drive plate-boundary deformation?
- What determines the spatial distribution of plate-boundary deformation?
- How has plate-boundary deformation evolved?
- What controls the space-time pattern of earthquake occurrence?
- How do earthquakes nucleate?
- What are the dynamics of magma rise, intrusion, and eruption?
- How can we reduce the hazards of earthquakes and volcanic eruptions

PBO challenge- sensor calibration

- Sensors get re-used so calibration can't changing when they are moved
- How have data been affected by moving the sensor? This can be analyzed by looking at data over successive sensor deployments. Did data change over the life of a sensor?

3.8 Project and Research Overview: Regional Scale Nodes Infrastructure

Mike Kelly (UW APL) presented the deep water Regional Scale Nodes (RSN), a cabled underwater research facility that is planned to be installed off the coasts of Washington and Oregon. The RSN will include 1500 km of fiber optic cable providing high power and bandwidth to the seafloor and throughout the water column. This will provide a constant stream of data in real time.

The RSN is a Greenfield environment, which means it is a brand new, to be installed environment. The RSN is made up of a cross-functional team engaged in process mapping of operational procedures.

The team used process maps to define operations and support a system template for process mapping called Telecom Operations Process Map (called TOM). The value of using templates is high:

- They are strong
- Well tested
- Provide a generic process map
- The framework is readily available
- The structure is flexible enough to support either high volume retail or low volume wholesale

Specific Statements:
What does RSN expect from CI?

- Software monitoring/surveillance performance and provision of the service-sensors layer
- Programmatic trouble ticket system,
- Monthly performance reports of individual services (sensors)
- Biannual power and bandwidth forecasts
- Technically literate, 24/7 network management center
- Future expectations: service driven organizations, strong process in place to use TOM,
- RSN observatory management system for control and monitoring of the network layer,
- Collaborative ocean visualization environment (COVE), RSN observatory visualization tool
- Access to CI service (sensor) layout tools,
- Build and maintain-mooring log, architect test data and, maintenance log for every series of instrument assemblies including sensors, power and communication systems.

3.9 Project and Research Overview: Coastal and Global Observatory Infrastructure

The goal of the current system is to connect sensors within the ocean utilizing 27 moorings and 29 vehicles. The current system includes surface moorings, sub-surface moorings, global gliders, and coastal gliders. The system is planning the following changes in the near future:

- Moored profilers
- Adding acoustic modem to subsurface moorings and gliders

Current analog systems perform benthic experiments using mixed cabled and uncabled platforms that perform data archiving. Ancillary data are essential and they need a management system for calibration and checkout data.

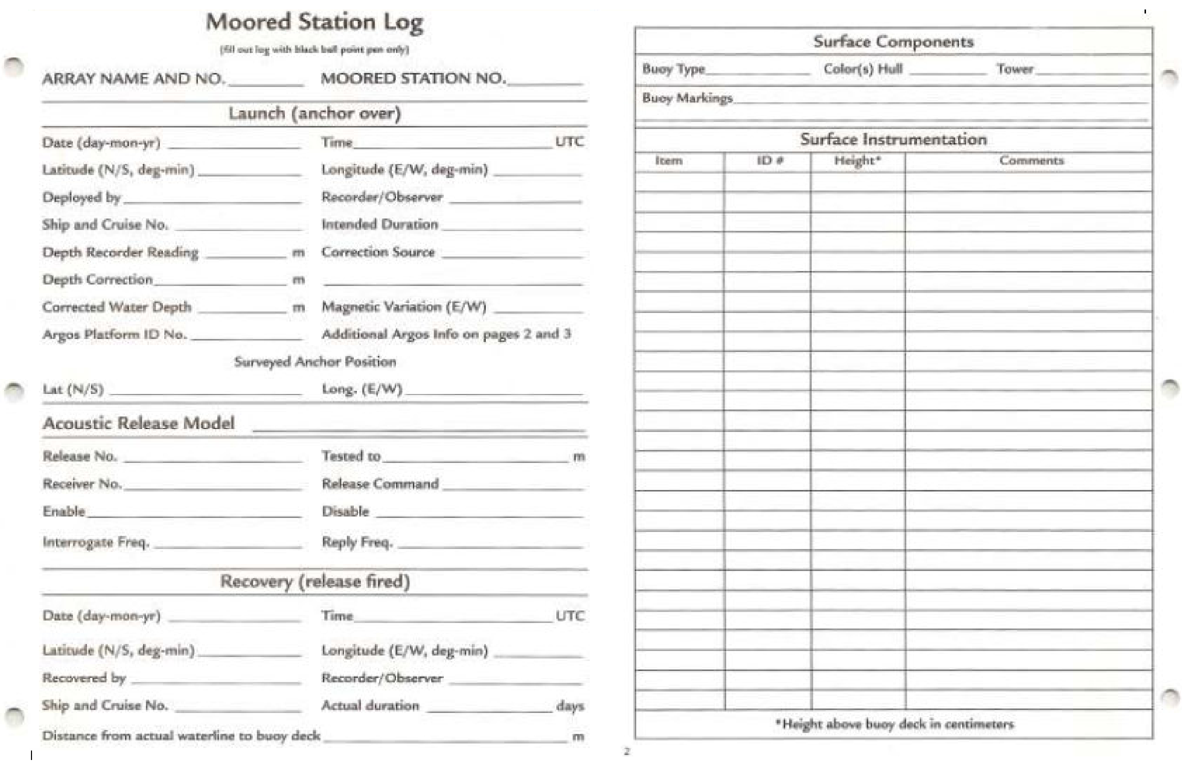
The team identified the following system needs:

- Data archiving and quality control
- Easy to access interface to allow user to interact with the data
- Secure environment to allow science operators to control and change operation of systems- who is allowed to do what, require models of systems

- Database with inventory and operational status of all assets- who has what sensor available and what is its status?

The team also expressed the desire to have a live person involved during deployment, since there is a need for a human to provide information for deployment of mooring. The reason is that a person can see what is happening and know what is important to report, and can make decisions in situ and on the fly.

The team currently uses a “Moored Station Log” to keep important information. The team provides the information via a form they fill out. They felt strongly that this type of form is essential because the format is institutional. The scientist or engineer fills out the form before they go out to sea and includes details such as time in water, and pre and post cruise calibration data. Other data that are stored in the logs system include contact information trouble reports and status reports



Moored Station Log

(fill out log with black ball point pen only)

ARRAY NAME AND NO. ______ MOORED STATION NO. ______

Launch (anchor over)

Date (day-mon-yr) ______ Time ______ UTC
Latitude (N/S, deg-min) ______ Longitude (E/W, deg-min) ______
Deployed by ______ Recorder/Observer ______
Ship and Cruise No. ______ Intended Duration ______
Depth Recorder Reading ______ m Correction Source ______
Depth Correction ______ m
Corrected Water Depth ______ m Magnetic Variation (E/W) ______
Argos Platform ID No. ______ Additional Argos Info on pages 2 and 3

Surveyed Anchor Position

Lat (N/S) ______ Long. (E/W) ______

Acoustic Release Model ______

Release No. ______ Tested to ______ m
Receiver No. ______ Release Command ______
Enable ______ Disable ______
Interrogate Freq. ______ Reply Freq. ______

Recovery (release fired)

Date (day-mon-yr) ______ Time ______ UTC
Latitude (N/S, deg-min) ______ Longitude (E/W, deg-min) ______
Recovered by ______ Recorder/Observer ______
Ship and Cruise No. ______ Actual duration ______ days
Distance from actual waterline to buoy deck ______ m

1

Surface Components

Buoy Type ______ Color(s) Hull ______ Tower ______
Buoy Markings ______

Surface Instrumentation

Item	ID #	Height*	Comments

*Height above buoy deck in centimeters

2

Moored Station Number ____________________

Item No.	Length (m)	Item	Inst No.	Time Over	Notes	Data No.	Depth (m)	Time Back	Notes
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									
18									
19									
20									
21									
22									

The team wondered if some of this could be automated, since many of the data about the ship are already known. Information such as time and location could be automated through the simple use of clock and GPS systems.

The system includes an operations management center which performs the following functions:

- Monitor health and status
- Plan observatory activities
- Operations support tasks
- System life cycle, record, retrieve
- System contact log-record, store, retrieve, report system trouble reports

4 Workshop Outcome

4.1 Questionnaire Response Analysis

The CI ADT received substantial input from the participating scientists through the questionnaires that were handed to them prior to the workshop with the request to provide answers to as many questions as possible. The input from the questionnaires went directly into refining and validating the science user requirements. Selected statements are listed in various sections throughout this report, in particular the individual participant presentation sections and the general requirements discussion section.

4.2 User High Level Issues

The workshop brought several high level issues to the surface and provided a venue for users to discuss their concerns. The comments shared by most participants are listed below.

Specific statements

- The CI must have streaming real-time data, processing and modeling. This means the system is data driven, not event or schedule driven. The system examines the data and can detect if something is wrong with a data stream. If that does occur, an operation will be triggered to solve the problem. This focus on data includes:
  - Seamless integration with archived data
  - Export of data in useful formats
  - Mechanism to evaluate the quality of data, both automatically and by humans
  - Connection of sensor data into real-time models with feedback to instrumentation
- Acquisition and processing software must operate properly in the CI environment
- Our operation is collecting data 24/7. This type of operation requires software to run reliably over long periods of time and be highly fault tolerant (e.g. immune to temporary loss of communication)
- 24/7 operation requires high levels of automation (no 24/7 staffing available so system has to be able to correct problems as they see them.)
- Processing system must support rapid problem determination and resolution
- Connection of sensor data into real-time models with feedback to instrumentation
- User interactions must be allowed with individually and system owned instruments.
- Users need to have access to all data in near real time
- It is important that there is a scalable architecture from instrument to end-user
- Many users want the CI to improve their current situation. They posed the following questions and issues:
  - How do I use the data that are available to me to better study the ocean? I think the CI is key to this.
  - How do I document my ocean exploration? Will the CI allow us to record numerically and verbally the day to day decisions, discoveries, etc?
  - I need a tool to overcome information overload but also allow me or very advanced graduate students to drill down to the individual bytes when the science or engineering requires it.
  - I am time limited; the CI should provide the tools to improve my management and science productivity

4.3 User Requirements Discussion

The workshop participants discussed the purpose and need for gathering requirements from the user's point of view. The fact that the cyber-infrastructure has such a large and diverse user base underlines the need to have clear requirements that can be traced directly to user needs. This process for the workshop scenarios and requirements gathering started by first listening to scientist and engineers describe their work.

The second stage of the workshop included specific user case studies that described the different scenarios experienced by the users. The scenario discussion yielded several use cases that act as models that ensuing requirements can be built from.

The team running the meeting agreed to use the time in the workshop to generate scenarios and requirement lists. This was done in a brainstorming session, with no prioritization. It was clear that the requirements would be compiled, organized, and prioritized after all of the data from the workshop were aggregated.

4.4 CI Use Case Scenarios

In this session, the charge for the workshop participants was brainstorming and discussing a hypothetical use scenario for a transformative community cyber-infrastructure. The following list documents this use scenario:

4.4.1 Use Case Scenario 1: A day in the life of a Test Pier Operator

This scenario focused on the environment of a test pier. The group walked through a day in the life of a Test Pier Operator. The first important issue was that there are so many different projects on a test pier that there is no such thing as an average day. Using that knowledge as a point of departure, the group mapped many different scenarios that could occur.

Most users thought there were great advantages in putting an instrument out on the pier since it provides an environment to test the instrument. Some experiments are done right at the pier and have the advantage of getting data close to shore thus easily accessing the work. Some of the drivers for putting instruments on the pier are looking at water quality and fishery habitat observations in the location of the pier.

Different user types, environments, and activities are listed below:

<u>WHO: User</u>

- Industry people testing new equipment
- Students with projects from campus (built their own instrument or bought one to deploy, also other university institutions)
- JPL, satellite prototype validation
- Operator; has real world experience, past their science project at school, masters level at least, has only seen 2 out of 15 operators so far who have this kind of experience.

<u>WHERE: Environment</u> (on pier or on ship)

- Observation network on the estuary
- Observatory team worked with the users to find parameters of interest in the environment
- Other environmental issues to understand include water quality at creek input, including septic sewage system, etc.

<u>WHAT: Situation</u>

Each experiment has different data needs. Some need to view its outcome as providing a final report; others want to see real-time data collection. Other parameters include:

- Certain data thresholds
- Data stream, data breaks

- Time requirements:
    - Days
    - Time segments, e.g.: 15 minutes with 2-minute bursts.
    - Other time periods

Pier Operations:
Two key teams work on the pier, operations, and management. Each team has a different focus and is concerned with different issues.

The operations team is focused on the instruments. This team is responsible for installing sensors, and they estimate installation of new sensors on the pier every two months, mostly as short-term deployments. The concern from this perspective is control of the resource. This includes insuring the resources have power and are online

The IT management perspective looks at the core instrumentation and how it interacts with the system. They are concerned with data flow in and out of the instruments as well as communication between nodes on the network.

<u>Focus:</u> The focus for this scenario is on IT systems that are used to operate the processes and steps. This system supports observational studies that have specific goals of their own. Each study has its own focus so the IT system has to be ready to support any study.

IT infrastructure issues and needs:
- In order to have complete results, they need regional integration of data
- Data collection is a separate function from data processing
- Need simpler metadata capture
- Need to provide schedule and bookkeeping information
- Need an awareness of redundant assets such as maps, etc.
- System must be able to ingest and process information from other systems, locations
- System does not rely on specific locations
- System provides metadata for all instruments online, accessible by prospective users

Metadata issues
- Core instrumentation is reconfigured when you add new data stream, creating new metadata
- Team has limited time, resources, and other priorities, so metadata work does not always get done. Suggestion is to have a person who is assigned to this issue.

<u>Process for setting up new instrument on the pier</u>
1. Set up data portal
    a. Create connection to the instrument
    b. Determine the parameters of the data set that is anticipated to come from the instrument.
2. Schedule for installation
3. Test instrument
4. Test on shore (one day)
5. Test on infrastructure with cable/telemetry (one day)
6. Install instrument
7. Work on upgrades, change of data streams

<u>Specific Statements:</u>

General Challenges

- Getting sufficient funding, since northern and southern coastal observatory systems might be going after the same money
- HF radar program is a good example of a cross regional program. There was an expressed need for more of these types of programs.
- No predictable schedule for projects or classes on the pier, and this makes it hard to plan, thus creating problems.
- Developing user requirements and stakeholder ideas. Teams don't know how to find the requirements, they feel it slows the program and they are forced to interact with federal folks. Most teams don't know how to elicit and develop requirements, and would appreciate a system for that. Both the IT and observatory operators don't have training in writing requirements for experiments they run on the pier.
- Changing situations create changing requirements: Sometimes scientists don't know what they are looking for and don't have requirements for their instruments. Other times, the scientist wants to use the instruments for a bit to see what they can do with the instruments. Scientists might configure the instrument for the highest frequency sampling at the maximum rate, then have too much data and they do not always have a good mechanism to deal with them efficiently. Sometimes, they don't even want the detailed level of data.
- HARD PART, reviewing and analyzing data and coming up with resulting conclusions that result in a published paper.
    - Issue of having enough people to do the job
- Collaboration is an issue, right now they use email and the phone, but are considering other systems to help do this, time commitment to use the collaboration system; is it better to do it live- too much delay, asynchronous, maybe a Web 2.0 solution?
- Data collection shouldn't be 90% of effort- should be 20% and 20% using it- priority should be focused on the science- come up with results
- Instruments staying workable- insure it is used repeatedly which keeps it working
- Real estate issue on pier, who gets to put their experiment where becomes important to the users
- Scheduling conflicts
    - Timing when the experiment comes in and coordinating with anything else that is going on such as classes
    - Staff can't be in two places at one time, so how do they manage their time and schedules?

<u>Operator challenges: specific statements</u>

- Producing data suitable for many different types of specific users
- How to configure an installed instrument
    - What is the ideal sampling rate
- Address data load/delivery limitations
- Push regulations that are not effective
- Insure data analysis on a regular basis by qualified people
- Insure QC/QA operations on data streams
- Prioritizing workloads for staff and project, data side often ends up on the bottom
- Scale existing observing systems
    - Stay manageable
    - Manage complexity
    - Manage multi-stakeholder needs
- Key is collaboration, need other people to help do the job, current tools work and they are looking at other ideas: ocean portal, but high time commitment, high latency.
- Tools combined with voice could really help-experiment planning and adaptive sampling

4.4.2 Use Case Scenario 2: A day in the life of a complex sensor network manager

What: Land sensor array that includes 400 stations currently in the western portion of the central United State of America.

Focus: Seismometers to study earth's velocity structure

WHO: User: Jennifer is lead, in charge of data flow- real time system has a team of 4 or 5 other people

- Operators
- Data analysts- QC
- Scientist who uses data
- Deployment teams, field service teams
  - Weekly meetings for team to prioritize which one to look at
  - Scheduling happens between the teams
  - New Mexico has a shop to fix the instruments

TEAM

- Jennifer is both the real time operations systems and metadata person
- First person on the system in the morning deals with problems
- Communication between people is carried out through phone or email
- Skill set includes:
  - UNIX
  - Trouble shoot /dig in to details
  - Multitask
  - People skills
  - Personal interest, but with a boundary

WHERE: Environment: Operators works from a home office

- Don't interact on a personal basis with other team members
- Most work from a home office and they find that works well because you get more done

Actions, tasks:

- Communication through emails:
  - Starts with reviewing emails, looking for alarms, looking for data
  - Check the running scripts that parse the following information:
    - Stations that report no data
    - Stations that report gaps in data
    - If system has found a problem and activates a problem script, operator checks on results to insure the correct action was taken. Examples of problems that can activate a script include:
      - The position of the instrument is monitored and has to be in the proper location. If it is off position, then the system informs the operator.
      - Any data out of range
- Check daily station reports
  - Status power issues
  - Condition of sensor- is there water in it (these are dry land sensors, so there shouldn't be water)
- Read standard digests of system information
- React to the problems, sometimes there are issues

- Don't go out to fix problems, but delegate problems to service people. Then operator must track and file reports that get sent back from the service team in the field.

Metadata activities:

- Monitor 24/7
- If system is good then go on to metadata work- 2 days a week
    - Summary of service reports
    - New stations installed
    - Stations removed
- Review all email to see if there are new installations, equipment changes,
    - There are 10-14 changes a week
- Interrupt flow to database to update and input metadata
    - Enough buffering to keep data coming in from other sources,
    - Upsets people who have different use for data than it was designed for
    - Adding tools to send out emails to let users know that the system is coming down
- Update local database
- Ship metadata to users
    - Sends out email- finds it is inefficient
- Use Antelope software package
    - Transfer files into ORB server
    - Once it is up, other people can use it
- Programming on other days
    - Develop tools, displays, add value

Other tasks:

- Support- respond to queries from users about data
    - Can be from downstream scientific users, from service folks, etc.
    - Conference calls for meetings
    - Travel to Scripps: Every 6 weeks
    - Meet with users of data, phone, sometimes meet with these users in person
    - Attend scientific meetings
        - AGU
        - IRIS
        - SSA conferences

Specific Statements:

- Service reports helps understand what is happening with data- not always submitted in a timely fashion
    - Only get paid if they get the reports submitted
- Installation reports become part of the metadata ("site is in a field with cows")
- Tracking systems are a bit clumsy and could be streamlined, perhaps automated
- People have sensors on their land, most people love it and are accessible, but scheduling has to be done in advance
- Have test vault, swap components and test them
- Trouble happens sometimes if the project lead is out of the office, how does the information get stored, how are problems taken care, need chain of command more clearly defined
    - Not much cross training and if someone is away then there is a problem
- Communication needs to be well established; since the groups are far apart, it took two years to meet some of the people on the team
- Service teams don't always communicate about their work

- Short on sensors so sometimes they are used even if they are not perfect
- If they don't get data from a sensor, they monitor the state of health
- Could be environmentalist nearby or other field issues, so sensors are not consistent
- Track updates, and if they were shipped downstream
- CI tool to track metadata updates would be great
- Need people on staff who can multitask and get the job done; start something and be able to follow through
- Multitasking and prioritization issues
- Keeping up with metadata updates, current system seems to work, but since there are 400 stations and the ones in existence before them, it is hard to track
- Improve data visualization of images on website- sometimes it takes two years to get these up on website, now only raw data gets up
- Data upload from field teams a problem (need an app for this)
    - No automated discovery
    - Field folks don't want to be using web services since they are in the field. These field technicians often work in extreme weather and can't get on the Internet until they are back in the office.
    - Email form has to be fed into database
    - Even sending email is a problem
    - Too time consuming
    - Right now it is a form on a computer, and they email it to her, but they won't go on a website
        - Scalability
        - Form was simple and it worked
        - Maybe the website was too complicated
- Security
- Scalability with the computer systems, great at 100 sensors, then when they went up to 400 it brought system to its knees.

4.4.3 Use Case Scenario 3: Processes for Operation and Management: Group Activity

How do you get something in the water? What are the processes involved in doing the science? How does the CI get involved in that process?

This scenario starts with how you get something in the water, plan and propose, what are the processes involved in doing the science and what is needed from CI?

Process/Tasks for O&M	Who	CI Involvement	Assumptions, Policies, Steps
Proposal for new resources on infrastructure • Science • Logistics	PI- agency (NSF)	Portal • Manual guide for users • Form + PDF	Agency (ex.: RIDGE, NCAR) checks with infrastructure Feasibility
Instrument preparation • Labeling/compatibility • Instrument driver create/test + database setup • Testing phase • Integration	Observatory Ops Data mgmt. Group – interfaces/drivers, database; Sensor mng. – Integration, configuration,	Driver exists?	Money + instrument + configuration comes from PI Presentation/ Visualize – with std. front-end, provided by PI

	state/monitoring		
Define and handle trouble tickets	CI User > Ops > role / individuals Issues generated by user, ops, network, thresholds, alarms, errors Order process items, change requests, orders Maintenance, performance, networking activity	Portal • Reporting mechanisms • Progress tracking • Filter, prioritize, categorize • Metrics, types • Escalation • Queues based on job functions • Notifications based on type • Audit trail, tracing • Policies	
Provide a portal/face of the network "Integrated Observatory" • Phone, Email • Electronic	CI as service/user organization	• Frameworks to respond to IOs • Track conversations • Trouble tickets system • User guides • Visualization of resources on map, with details and status • Data access • Per resource/sensor info - Status - Ancillary info – history, manual, maintenance - Data products available - Contact info • General contact form with dispatch • Control instruments - Set alarms, sample rates, basic functionality (for authorized users) • Notice of danger areas/events (by region) • News, changes, public relations, information • Views, plans of network activities, outages • Bibliography, OOI descriptions, products, global library, knowledge compilations • Public outreach/education info • User specific dashboard/space • Subscribe to data streams, news	

		• Show credit for funding, sources, user agreements • View user specific policies, quota, access control • QA/QC policies • Find other users, form communities, social networking, discussion forums • F.A.Q. site, Knowledge base, • Press releases • List of experiments nested, uses, hypotheses, project spaces, models that use the data or that operate over the same region • Data analysis and visualization tools • Links to other experiments/data/projects inside and outside • Discover data • Data mining & discovery - Explore the database - Limits by variables (search by scope) • Publish information – about papers and data • Data entry forms, cruise logs, mission logs • Availability of instruments and sensors • Public	
Monitoring the state of health and notifications for events, alarms • Online, with CI • Power failures • Fire, etc.			
Scheduling and planning of resources: • Ships • Network • Phased integration		View schedules	
Testing qualification, approval, certification of instruments according to system model • Qualification procedure • Input analysis • Network testbed • Simulation/staging			
Acquire the data • Instrument			

• Network • QA/QC			
Commission instrument after deployment • Resource deployment • Feedback of written notes, … into system			
Resource management • Manage and view blackout dates • Asset management for govt. reporting • Control instrument • Collect and present metrics • Inventory management – location, schedule, condition • Manage access security/governance policies • Manage dynamic allocation constraints – request, monitor, enforce, view • Scheduling of enabling resources • Maintenance and configuration mng. • Expanding data storage • Forecasting capacity and needs • Physical sample archiving & tracking • Reporting mechanism • Integration, tracing with trouble tickets, service mng. Tools		Metadata for sensors/instruments --- • Make, model, serial no. • Calibration information, period • Unique name • Transfer function (e.g., linear filter) • Location, orientation • Performance specs • Hosting instrument (piggy back) • Relation to other instruments • Primary site: location designator • Time-scoped information • Reference designators (the role of that site/sensor) • Physical characteristics (battery, radio, mass,) • Sensor settings, config • NSF imposed: inventory, acquisition, service, cost estimate, disposition • History: deployment, recovery, repair,	Action --- • Change comm. parameters (speed) • Change state (on/off/pause) • Change mode (self test, debug mode, self caliber. mode, diag. mode) • Set limit for faults, monitoring range • Set ground fault monitor on/off • Voltage, current, ground fault • Configure port isolation • Retrieve on-board archive data • Enable/disable fault reporting • Switch on direct access mode • Query information about device • Remote power control
Adaptive sampling • Reconfirmation • Conflict resolution			
IT System Management			

Trouble ticket – any user/alerts/operator/network itself may submit a ticket via phone or online. The ticket goes to the op. center, where it is analyzed then dispatched.
VENUS – tickets are associated with the data and the related tools. It is not currently applied to data delivery issues, power, etc.

Portal

- Contact info
- Status reporting
- Visualization capabilities

Resource management

- Issues to investigate: Ownership of data, metadata; Data access policies; Regulations; Standards compliance; Capabilities compliance

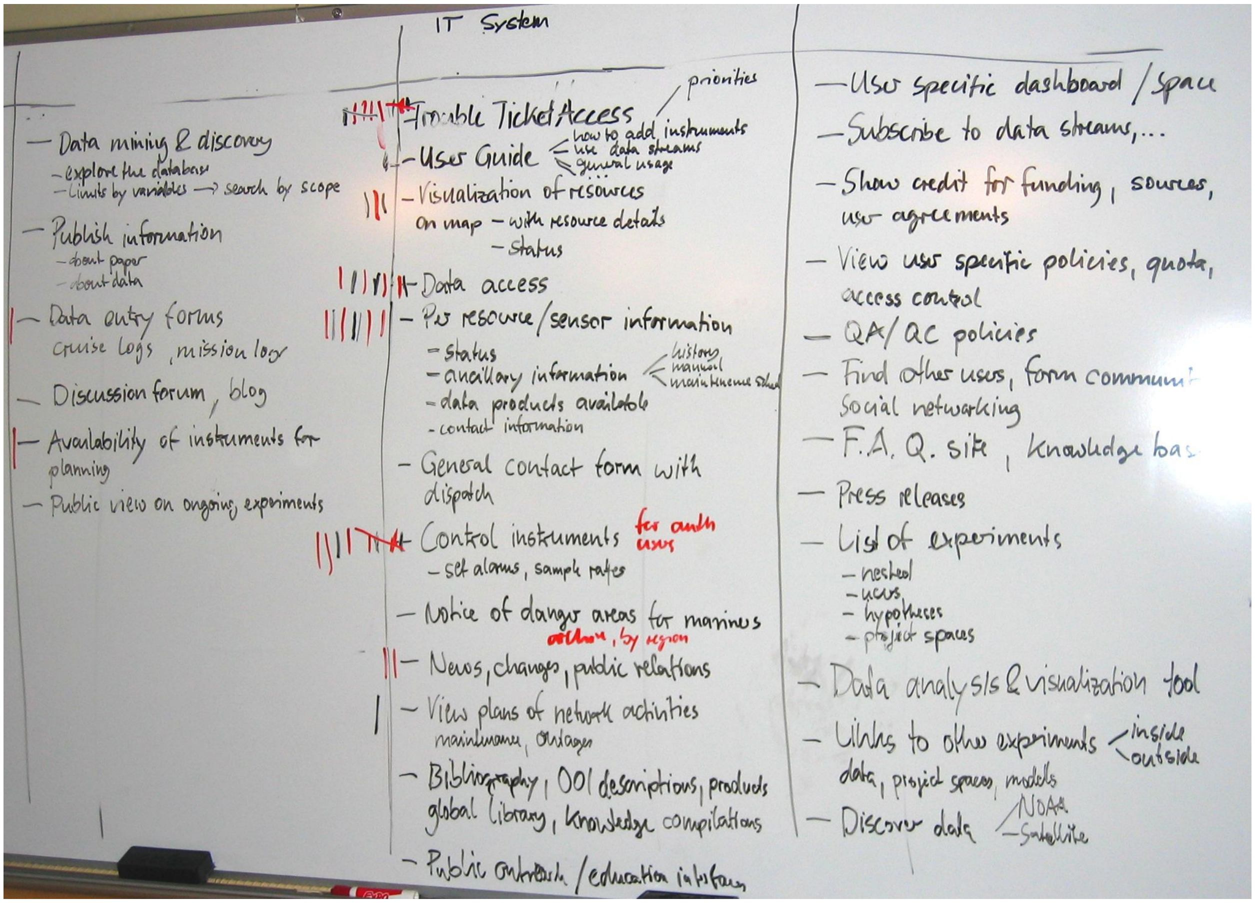


Figure 2: Group exercise voting on important features.

Vote what is more important. Listed below with more important item first.

- Trouble ticket access
- Per user, sensor information
- Control Instruments
- Visualization of resources
- Data access
- News, changes, public relations
- View plans of networks
- Data entry forms
- Availability of instruments for planning

<u>Specific Statements:</u>

- A question was “What are the expectations from an instrument, what are its contracts with the CI or its basic set of specifications from the ops point of view (assuming it is developed, tested, deployed and part of the running system)?”
- Policy regarding power – ops is allowed to turn on/off

- One type of interface for the infrastructure and another set of various low level ones to get to the instrument.
- The maker of the driver has the knowledge about the control of the instrument as is expected to create the "management interface." All instruments are expected to provide a minimum set of control capabilities, although an overall encompassing set is probably infeasible.

5 Science User Requirements

5.1 Requirements Elicitation Process

The requirements listed in the next section represent the current collection of science user requirements for the OOI CI. Some of the requirements were identified in prior requirements workshops and partially validated by the participants. Further requirements originate from the analysis of related cyber-infrastructure efforts. The remaining requirements were identified through a thorough post-workshop analysis process. Requirements were either directly stated by the participants during the workshop discussions, called out in the participant questionnaires or inferred through a requirements analysis process by the CI architecture and design team. Requirements are grouped into categories and formatted according to a template as described below.

In order to uniquely identify the elicited requirements, each requirement in this report follows a standard template. Each requirement contains a unique identifier issued by the DOORS requirements management system. Furthermore, each requirement contains a label and an explanation. Requirement labels are constructed in a schematic way.

The listed requirements strive to be atomic (i.e., they express one idea only and do not contain sub-requirements). However, requirements might be related and one requirement might be influenced by another requirement. Further, the explanation might contain additional details about the requirement.

5.2 OOI Cyber User Requirements

This section contains a list of science user requirements as exported from the OOI cyber user requirements DOORS module on 7/31/08. It shows the identifiers and requirements labels and omits explanations and further attributes, such as priority. Please refer to [OOI-CU-REQ] for a full generated view containing all attributes. Requirements are grouped into categories, as indicated by the bold labels in the table. The numbering reflects the structure of the DOORS module. The requirements list contains all CI user requirements to date. Requirements that are traceable to the Integrated Observatory Management requirements workshop are marked in italics.

ID	Requirement / Category Heading
	4.1 Resource Management
L2-CU-RQ-50	The CI shall support distributed resources, applications and actors
L2-CU-RQ-51	The CI shall provide the capability for a given resource to initiate change in another resource
L2-CU-RQ-52	All resources under CI governance shall be identifiable
L2-CU-RQ-53	All resources under CI governance shall be authenticatable
L2-CU-RQ-54	All resources under CI governance shall be authorizable
L2-CU-RQ-55	All resources under CI governance shall be auditable
L2-CU-RQ-56	The CI shall incorporate a policy-based decision system for the management of CI-governed resources
L2-CU-RQ-57	The CI shall ensure that resource utilization is governed by the rights and allocations of the initiating actor
L2-CU-RQ-58	The CI shall enable non-persistent connection of resources, users and applications
L2-CU-RQ-59	The CI shall act as the facilitator and broker for resource usage
L2-CU-RQ-60	*The CI shall schedule resource usage based on capacity, capability and availability*
L2-CU-RQ-61	The CI shall support the evolution of resources under CI governance
L2-CU-RQ-62	The CI shall support the resource life cycle, providing notification to resource providers and

ID	Requirement / Category Heading
	consumers when manual intervention is required
L2-CU-RQ-63	*The CI shall provide a catalog listing all resources under CI governance*
L2-CU-RQ-64	*The CI catalog shall provide status information for all resources under CI governance*
L2-CU-RQ-65	All resources under CI governance shall be discoverable, either directly, by content or through their associated metadata
L2-CU-RQ-66	Multiple actors shall be able to simultaneously discover the same resource
L2-CU-RQ-67	The CI shall integrate resource discovery with resource access subject to policy
L2-CU-RQ-68	The resource catalog shall link entries to the associated metadata
L2-CU-RQ-69	The resource catalog shall incorporate information about physical samples
L2-CU-RQ-70	The CI shall cross-reference CI-governed resource catalogs and external resource catalogs
L2-CU-RQ-71	The CI shall enable discovery of all information resources that are derived from a given original information resource
L2-CU-RQ-72	The CI shall provide resource subscribers automatic and manual fallback options with similar characteristics in case the original resource becomes unavailable
L2-CU-RQ-73	The CI shall provide services to group resources
L2-CU-RQ-74	The CI shall provide registration services for resource notification
L2-CU-RQ-75	The CI shall automatically register resources for notification to the observatory operator
L2-CU-RQ-76	*The CI shall provide notification of resource state change to all resource subscribers*
L2-CU-RQ-77	The CI shall bind metadata to all resources under CI governance throughout the resource life cycle
L2-CU-RQ-78	The CI shall support standard OOI-standard metadata content that includes, but is not limited to, a complete description of resource behavior, content, syntax, semantics, provenance, quality, context, citation, correspondence and lineage
L2-CU-RQ-79	The CI shall specify and utilize a standard vocabulary
L2-CU-RQ-80	The CI shall maintain the relationship between OOI standard metadata and the vocabulary
L2-CU-RQ-81	The CI shall allow resource discovery utilizing the standard vocabulary
L2-CU-RQ-82	The standard vocabulary shall accommodate information on physical samples
L2-CU-RQ-83	The CI shall provide data generating resources using proprietary metadata formats with a means to transform them to OOI standard metadata
L2-CU-RQ-84	*The CI shall support the provisioning of OOI standard metadata*
L2-CU-RQ-85	The CI shall verify compliance of metadata with the OOI standard
L2-CU-RQ-86	*The CI shall update resource metadata within 5 seconds of resource reconfiguration*
L2-CU-RQ-87	The CI shall provide services for control and monitoring of observatory infrastructure resources
L2-CU-RQ-88	*The CI shall provide services for pervasive resource monitoring and control*
	4.2 Data Management
L2-CU-RQ-90	The CI shall be capable of archiving all data and data products associated with an OOI observatory
L2-CU-RQ-91	The CI shall act as a broker for CI-managed data products
L2-CU-RQ-92	The CI shall ingest data with variable delivery order
L2-CU-RQ-93	The CI shall support the delayed distribution of temporarily sequestered data
L2-CU-RQ-94	*The CI shall ensure the integrity and completeness of all archived data products throughout the OOI life cycle*
L2-CU-RQ-95	The CI shall ensure that all archived data products can be restored to their most recent state
L2-CU-RQ-96	The CI shall provide a topic-based (publish/subscribe) data distribution infrastructure
L2-CU-RQ-97	The CI shall provide registration services for data subscriptions

ID	Requirement / Category Heading
L2-CU-RQ-98	The CI shall publish unprocessed raw sensor data
L2-CU-RQ-99	The CI shall archive unprocessed raw sensor data
L2-CU-RQ-100	*The CI shall support the publication, distribution and archiving of different versions of the same data product or stream*
L2-CU-RQ-101	The CI shall support real-time data delivery
L2-CU-RQ-102	The CI shall support guaranteed data delivery
L2-CU-RQ-103	The CI shall support store until requested (pull mode) data delivery
L2-CU-RQ-104	The CI shall support streaming data delivery
L2-CU-RQ-105	*The CI shall integrate multiple data streams or data sets into a single stream or set, eliminating redundant entries*
L2-CU-RQ-106	The CI shall support peer-to-peer communication between discoverable resources
L2-CU-RQ-107	The CI shall support secure data delivery
L2-CU-RQ-108	The CI shall adapt data delivery in the presence of limited available bandwidth according to policy
L2-CU-RQ-109	The CI shall notify registered resource users when data delivery cannot be achieved due to low available bandwidth
L2-CU-RQ-110	The CI shall adapt data delivery in the presence of high channel latency according to policy
L2-CU-RQ-111	The CI shall notify registered resource users when data delivery cannot be achieved due to high channel latency
L2-CU-RQ-112	The CI shall publish data from external data sources, data bases, and data distribution networks from related scientific domains.
L2-CU-RQ-113	The CI shall provide support for large volumes of data
L2-CU-RQ-114	*The CI shall archive and catalog text, images, pdf, .doc files and spreadsheets*
L2-CU-RQ-115	The CI shall flag and notify data stream and data set state change
L2-CU-RQ-116	*The CI shall flag and notify redundant data and metadata*
L2-CU-RQ-117	*The CI shall acknowledge requests for data with an estimate of delivery latency*
L2-CU-RQ-118	The CI shall credit data publishers when data products are accessed
L2-CU-RQ-119	The CI shall provide services and interfaces for the acquisition of bulk data
L2-CU-RQ-120	The CI shall associate bulk data with their metadata and related data products
	4.2.1 Data Transformation
L2-CU-RQ-122	The CI shall support the moderation and auditing of published data
L2-CU-RQ-123	The CI shall provide services for interactive and automated data quality control (QC)
L2-CU-RQ-124	The CI shall perform automated quality control of observational data products in near real-time
L2-CU-RQ-125	The CI shall provide standard and user-defined methods to assess the quality of data
L2-CU-RQ-126	*The CI shall specify data models for resources based on characterization of structure (syntax)*
L2-CU-RQ-127	*The CI shall translate between standard syntactic data models without loss of information*
L2-CU-RQ-128	The CI shall support translation between user-specified syntactic data models
L2-CU-RQ-129	The CI shall specify data models for resources based on characterization of meaning (semantics)
L2-CU-RQ-130	The CI shall support mapping between senders and receivers using the standard vocabulary without loss of information
L2-CU-RQ-131	The CI shall provide capabilities to define event detectors
L2-CU-RQ-132	The CI shall provide event detection services
L2-CU-RQ-133	The CI shall provide registration services for event notification

ID	Requirement / Category Heading
L2-CU-RQ-134	*The CI shall provide notification of detected events*
L2-CU-RQ-135	*The CI shall provide versioning for detected events*
L2-CU-RQ-136	*The CI shall update data sets as sensor calibrations become available*
L2-CU-RQ-137	The CI shall be able to accumulate knowledge about the scientific interpretation of observational data from manual mapping and linking of variables between different data sets
L2-CU-RQ-138	The CI shall be capable of co-registering data from different instruments in space and time
	4.3 Research and Analysis
L2-CU-RQ-140	The CI shall suggest suitable data products, observation resources, analysis tools, visualization tools and other OOI resources based on user-specified research questions using the standard vocabulary
L2-CU-RQ-141	The CI shall support interactive data analysis and visualization through tools and user interfaces
L2-CU-RQ-142	The CI shall provide a standard, extensible set of data processing elements that provide data assimilation, alignment, consolidation, aggregation, transformation, filtering, subsetting, averaging and scaling
L2-CU-RQ-143	The CI shall provide capabilities for analysis and presentation of environmental data at specified sites
L2-CU-RQ-144	*The CI shall support the integration of external analysis tools*
L2-CU-RQ-145	The CI shall provide capabilities to transform between coordinate systems
L2-CU-RQ-146	The CI shall provide capabilities to transform between map projections
	4.4 Ocean Modeling
L2-CU-RQ-148	The CI shall enable the efficient configuration, execution, and debugging of numerical ocean models
L2-CU-RQ-149	The CI shall support the interaction of model developers and non-expert model users
L2-CU-RQ-150	The CI shall provide capabilities to tune numerical models
L2-CU-RQ-151	The CI shall provide a virtual model environment and simulator to determine optimal model inputs, parameterizations and outcome qualities
L2-CU-RQ-152	The CI shall enable the sharing of ocean modeling, data assimilation and visualization components, including the extension of models with new model components
L2-CU-RQ-153	The CI shall provide a repository and sharing capabilities for numerical model algorithms, model configurations, data processing tools and documentation
L2-CU-RQ-154	The CI shall archive numerical model workflows under configuration control
L2-CU-RQ-155	The CI shall recompute model data products using archived workflows
L2-CU-RQ-156	The CI shall enable the modification of archived workflows
L2-CU-RQ-157	The CI shall provide an environment for the development of community numerical models under community process support
L2-CU-RQ-158	The CI shall provide a non-restricted environment for the development of independent numerical models
L2-CU-RQ-159	The CI shall support the nesting of ocean models at different geographical scales
L2-CU-RQ-160	The CI shall provide a framework for the adaptation of model resolution to the available resources
L2-CU-RQ-161	The CI shall support model ensemble definition, execution and analysis
L2-CU-RQ-162	The CI shall publish both elements of and aggregated ensemble data products from ocean models
L2-CU-RQ-163	The CI shall support flexible high performance model execution
	4.5 Visualization
L2-CU-RQ-165	The CI shall provide interactive 2D, 3D and 4D visualization tools

ID	Requirement / Category Heading
L2-CU-RQ-166	The CI shall provide 3D visualization of sensor locations and their environment
L2-CU-RQ-167	The CI shall support the integration of external visualization tools
L2-CU-RQ-168	The CI shall provide extensible, configurable visualization capabilities for data streams
L2-CU-RQ-169	*The CI shall provide a zooming interface for all visualizations with at least three levels of detail*
L2-CU-RQ-170	*The CI shall provide a user interface system that includes at least two different views of the data*
	4.6 Computation and Process Execution
L2-CU-RQ-172	The CI shall support the execution of large scale numerical ocean models across different locations on the network
L2-CU-RQ-173	The CI shall support workflows for automated numerical model execution, including just-in-time input data preparation, model computation, output post-processing, and publication of results
L2-CU-RQ-174	The CI shall enable the one-time and recurring execution of numerical models on any net-worked computational resource with quality-of-service guarantees based on contracts and policy.
L2-CU-RQ-175	The CI shall provide interfaces to compose workflows
L2-CU-RQ-176	The CI shall provide services to execute workflows on computational resources with varying characteristics
L2-CU-RQ-177	The CI shall provide services to chain a plurality of workflows
L2-CU-RQ-178	The CI shall provide services to monitor and control instantiated processes
L2-CU-RQ-179	The CI shall provide actors with estimated performance/turnaround for instantiated proc-esses
L2-CU-RQ-180	*The CI shall provide event-triggered workflow execution services*
L2-CU-RQ-181	The CI shall provide real-time access to high performance computation resources
L2-CU-RQ-182	The CI shall provide process support for the planning and operation of observational pro-grams
L2-CU-RQ-183	The CI shall provide process support for the coordination of instrument recovery, mainte-nance and replacement
L2-CU-RQ-184	The CI shall support, automate and combine workflows of shipboard observers
	4.7 Sensors and Instrument Interfaces
L2-CU-RQ-186	*The CI shall provide a real-time communication interface for remote resources*
L2-CU-RQ-187	*The CI shall support robust instrument development, operation and maintenance processes*
L2-CU-RQ-188	The CI shall support discovery of the characteristics of sensors deployed on an instrument platform
L2-CU-RQ-189	The CI shall support adaptive observation resource control
L2-CU-RQ-190	The CI time standard shall be NIST traceable
L2-CU-RQ-191	The CI shall provide a synoptic time service with an accuracy of 1 microsecond to all re-sources connected to the OOI observatories
L2-CU-RQ-192	The CI shall serve synoptic time throughout the observatory using Network Time Protocol
L2-CU-RQ-193	The CI shall provide services to correct remote clocks to a synoptic standard
L2-CU-RQ-194	The CI shall provide services to synchronize remote clocks relative to each other with an accuracy of 1 microsecond
L2-CU-RQ-195	Upon receipt, the CI shall synoptically timestamp message headers with an accuracy of 1 millisecond
L2-CU-RQ-196	The CI shall provide robust instrument access protocols
L2-CU-RQ-197	The CI shall provide direct bidirectional communications to resources that preserves their

ID	Requirement / Category Heading
	native functionality
L2-CU-RQ-198	The CI shall provide remote desktop access to resources that preserves their native functionality
L2-CU-RQ-199	The CI shall automatically close down inactive direct access sessions
L2-CU-RQ-200	The CI shall provide interactive web-based configuration of instrument platforms, instruments and sensors
L2-CU-RQ-201	The CI shall provide capabilities and interfaces for monitoring of resource-specific operational and environmental parameters
L2-CU-RQ-202	The CI shall provide services for positioning of mobile assets with a precision commensurate with the location technology
L2-CU-RQ-203	The CI shall support automated docking of mobile resources, including power management and high speed data down and up load
L2-CU-RQ-204	The CI shall be capable of triggering instrument measurements
	4.8 Mission Planning and Control
L2-CU-RQ-206	The CI shall support swarm-based deployment patterns for mobile instruments
L2-CU-RQ-207	The CI shall provide a repository for instrument behaviors
L2-CU-RQ-208	The CI shall provide a repository for observation plans
L2-CU-RQ-209	The CI shall provide shore-side and on-vehicle control capabilities for autonomous observational resources
L2-CU-RQ-210	The CI shall support observational resource control at different user-selected levels
L2-CU-RQ-211	The CI shall integrate environment and vehicle behavior models for event detection, coordinated control and adaptive sampling
L2-CU-RQ-212	The CI shall provide capabilities and interfaces for planning longitudinal observations
L2-CU-RQ-213	The CI shall provide capabilities and interfaces for planning objective-driven observations
L2-CU-RQ-214	The CI shall provide capabilities and interfaces for ad-hoc interactive and automated modification of ongoing observations
L2-CU-RQ-215	The CI shall provide capabilities and interfaces for simulating and verifying observation plans
L2-CU-RQ-216	The CI shall provide resource provisioning calculations from observation plans
L2-CU-RQ-217	The CI shall support observation planning and scheduling decisions based on the opportunity cost of observations and resource provisioning
L2-CU-RQ-218	The CI shall provide graphical user interfaces for planning observations and missions with spatial and temporal visualization of observation parameters
L2-CU-RQ-219	The CI shall provide spatial visualization of observation data overlaid with observation plans
L2-CU-RQ-220	The CI shall support tasking, deployment, mission control and retrieval of mobile and fixed instruments
L2-CU-RQ-221	The CI shall provide capabilities and interfaces for the simulation of observational infrastructure
	4.9 Application Integration and External Interfaces
L2-CU-RQ-223	The CI shall provide documented resource-data connectors for all services
L2-CU-RQ-224	Conditional on OOI policy, the CI shall not impose specific processes, tools and formats on resource providers for the operation and control of their OOI-connected resources
L2-CU-RQ-225	The CI shall interface with external resource monitoring, operation and control systems
L2-CU-RQ-226	*The CI shall provide a Web 2.0 environment*
L2-CU-RQ-227	The CI shall support interfacing with web service-accessible resources
L2-CU-RQ-228	The CI shall interface to live video feeds during instrument operation and maintenance
L2-CU-RQ-229	The CI shall provide interface support for Java-based tools and scripting languages

ID	Requirement / Category Heading
L2-CU-RQ-230	The CI shall provide standalone installations that may have no or intermittent connection to the OOI network
	4.10 Presentation and User Interfaces
L2-CU-RQ-232	The CI shall provide annotation, commenting, ranking and rating services for CI-managed resources
L2-CU-RQ-233	*The CI shall provide user and group workspace capabilities*
L2-CU-RQ-234	The CI shall provide capabilities to personalize user and group workspaces
L2-CU-RQ-235	*The CI shall provide social networking capabilities*
L2-CU-RQ-236	*The CI shall provide an intuitive interface to access the functionality of all CI services and resources*
L2-CU-RQ-237	*The CI shall present the full CI functionality at a single access point with a single dashboard*
L2-CU-RQ-238	The CI shall provide services to make OOI-standard metadata human readable
L2-CU-RQ-239	The CI shall provide a resource monitoring and control interface
L2-CU-RQ-240	The CI shall provide an adaptive, simple-to-use interface for data access
L2-CU-RQ-241	The CI shall provide transparent access to heterogeneous, large-scale computational re-sources
L2-CU-RQ-242	The CI shall provide transparent access to heterogeneous, large-scale storage resources
L2-CU-RQ-243	*The CI shall provide a single user interface that supports observatory operators, science and engineering users, the education community and the general public*
L2-CU-RQ-244	*The CI shall provide dialog box interaction for operations requiring the input of more than two parameters*
L2-CU-RQ-245	*The CI shall provide input screens that include tabs for any process that requires users to input more than five parameters*
L2-CU-RQ-246	*The CI shall provide a common font set for all screens*
L2-CU-RQ-247	*The CI shall employ a common look and feel based on a standard screen design*
L2-CU-RQ-248	*The CI shall employ a standard set of colors for use in all user interface presentation screens*
L2-CU-RQ-249	*The CI shall employ a standard workflow for all user interface screens*
L2-CU-RQ-250	*The CI shall employ a common navigation scheme that is consistent from application to application*
L2-CU-RQ-251	The CI shall provide visualization and metadata browsing of the processing pipeline
L2-CU-RQ-252	The CI shall provide checklists for standard instrument operations
L2-CU-RQ-253	The CI shall provide capabilities and interfaces to capture structured input, feedback and results from analysis processes on data
	4.11 Security, Safety and Privacy Properties
L2-CU-RQ-255	The CI shall authenticate and authorize all resources connected to an OOI observatory
L2-CU-RQ-256	*The CI shall authenticate all observatory actors*
L2-CU-RQ-257	*The CI shall provide different levels of access to actors with different levels of authorization*
L2-CU-RQ-258	The CI shall enforce user privacy policies
L2-CU-RQ-259	The CI shall be capable of auditing all services and resources under CI governance
L2-CU-RQ-260	The CI shall trace resource utilization to the initiating actor
L2-CU-RQ-261	The CI shall support different levels of access for resources and their metadata
L2-CU-RQ-262	*The CI shall protect physical resources from damage and misuse by enforcing resource use policies*
L2-CU-RQ-263	The CI shall provide interfaces to define security and policy for information managers at participating institutions

ID	Requirement / Category Heading
L2-CU-RQ-264	The CI shall support the diversion, filtering and sequestering of raw data streams at the acquisition point
	4.12 Quality Properties
L2-CU-RQ-266	The CI infrastructure shall deliver messages with reliability that is comparable to that of the Internet
L2-CU-RQ-267	The CI shall provide robust, reliable remotely deployed components
L2-CU-RQ-268	The CI shall provide services with reliability and accuracy that is comparable to those of distributed Internet applications
	4.13 Education and Outreach
L2-CU-RQ-270	The CI shall provide numerical ocean models with a limited set of modifiable parameters for educational purposes
L2-CU-RQ-271	The CI access point shall provide educators with instructions about data usage
L2-CU-RQ-272	The CI access point shall provide the educator with a list of projects and their attributes
L2-CU-RQ-273	The CI access point shall provide the educator with a means for social networking.
L2-CU-RQ-274	The CI shall provide a discoverable repository for educator-provided tools
L2-CU-RQ-275	The CI shall provide versioning and citation for educator assets
	4.14 Documentation
L2-CU-RQ-277	The CI IO shall make all source code for the OOI Cyberinfrastructure implementation and drivers publicly available, subject to applicable licenses
L2-CU-RQ-278	The CI IO shall document all external interfaces
L2-CU-RQ-279	The CI IO shall document all device drivers
L2-CU-RQ-280	*The CI shall provide discoverable web-based documentation for all services*
L2-CU-RQ-281	*The CI shall utilize a naming scheme that is compliant with OOI naming conventions*
L2-CU-RQ-282	**4.15 Development Process**
L2-CU-RQ-283	The CI IO shall seek to influence the direction of CI standards to effectively meet the needs of OOI users
L2-CU-RQ-284	The CI shall utilize open standards and open source software to the maximum possible extent
L2-CU-RQ-285	The CI IO shall accommodate local innovation that can be scaled to the community level
L2-CU-RQ-286	The CI IO shall support the verification of hardware and software components that will be deployed on OOI infrastructure
L2-CU-RQ-287	The CI shall support modular components
L2-CU-RQ-288	The CI implementation shall be platform-independent
L2-CU-RQ-289	CI service interfaces and capabilities shall maintain backward compatibility as the services evolve
L2-CU-RQ-290	The CI architecture shall be scalable to accommodate an increasing range of actors, resources, and services
L2-CU-RQ-291	The CI shall be extensible to allow the addition of new resources, services and applications to the OOI infrastructure
	4.15.1 Other
L2-CU-RQ-293	The CI shall provide process support for "dry" observational infrastructure development, verification and simulation
L2-CU-RQ-294	*The CI IO shall provide technically-qualified user care support and assistance through a human actor*
L2-CU-RQ-295	*The CI shall provide capabilities to maintain contact between users and user care*
L2-CU-RQ-296	*The CI shall provide capabilities to initiate and track trouble tickets*

ID	Requirement / Category Heading
L2-CU-RQ-297	*The CI shall provide tools for observatory operators to communicate with users*

6 Workshop Conclusions

6.1 Feedback from the Participants

The following list contains feedback statements from the workshop participants that were provided during and at the end of the workshop in the feedback sessions. The statements are listed anonymously and in no given order. Statements might be redundant, overlapping and contradictory due to the fact that they originate from different individuals.

- It was a good workshop
- The CI could be the face of the integrated observatory. This actual decision about the extent of this statement has not been made. It is necessary within OOI to come to an agreement of the extent of the CI
- What is the decision procedure of what capabilities will be in which release of the CI
- The level of effort for carrying out the workshops was commendable. It goes beyond what other projects have done regarding requirements
- A lot of the concepts captured on the whiteboard were lessons learned. It will be interesting to see how the current day issues apply to the situation in 5 years.
- Appreciated the questionnaire. There is a much valuable operational material in there to be mined. Will be interesting to see how this will be put into results.
- A different way to get further information is to go into details of the questionnaire and ask the contributors of the respective topics
- The workshop was one of the best opportunities so far to talk about what the marine observatories are doing
- Enjoyed the process of workshop facilitation and learning about the different observatories
- It was very important to learn about other groups views and discussions
- The participation from the Marine IOs was very valuable
- Really educational about role of MARS with CI. Provides many opportunities to work with CI
- Interface agreements have a firm deliverable date June 30. Hope that a good amount of material will get into the interface agreements.
- From an operator's point of view it was surprising that there was not so much discussion of how to interface with other observatories.

6.2 Next Steps and Action Items

Next steps include:

- Consolidate requirements from all user requirements workshops into a consistent list of CI user requirements.
- Prioritize and rank all user requirements, leading to a selection of baseline requirements for the construction of the OOI, to be reviewed during FDR.
- The user community will be asked to validate the requirements as well as the prioritization and selection during various phases before and after FDR.
- The requirements validation and community involvement process will continue past FDR throughout the entire OOI design and construction program

6.3 Conclusions from the Organizers

This workshop the fifth OOI CyberInfrastructure Requirements Workshop, hosted by Scripps Institution of Oceanography, UCSD, was instrumental in providing use case scenarios for operations and management of observatory systems. Listening to scientists and engineers who are on the ships, on the pier or in the field with a network of sensors was crucial in gaining understand and insight into the problems facing these people in their work. If the CI is going to be a transformative system, it must start wit the users and this workshop did just that. Several sessions were devoted to listening to stories told by the scientists and engineers doing the work of collecting data on our environment, both in the sea and on dry land.

This workshop was very successful in advancing the CI requirements definition and validation efforts, for refining and complementing the CI architecture and design, and in further fostering the mutual understanding of prospective CI user communities and the CI design team. Direct outcomes include jointly developed domain analyses and several extensive current day and future use scenarios. Each will contribute to complementing and refining the OOI requirements and design efforts in preparation of the upcoming Final Design Review in November and the time afterwards.

All presentation materials can be found on the workshop website [CI-IOM-WEB]. The CI workshop overview page [CI-WS-WEB] provides a more general context for all the CI requirements and design workshops to be scheduled and completed before FDR, with detailed background and accompanying material.

Appendices

A Workshop Participant Questionnaire

The CI ADT identified several relevant categories for the CI science user requirements; for each of the categories, a number of questions were identified, which when answered could lead to new and refined CI science user requirements. All questions in the respective categories together with an introduction and context setting were compiled as a slide set presentation. The workshop participants received a significantly shortened version of this questionnaire prior to the workshop.

Intent of this template

- We are collecting information about the users for the OOI CI. In order to get good information, we are sending this slide set template to all participants in the works. This template will be used:
  - For presentations during the workshop
  - To capture relevant information in a structured way
- Goals of this exercise are
  - To capture as many CI-relevant details as possible before the workshop
  - To capture structured, relevant information for use during and after the workshop
  - To enable quick information access for domain modeling during the workshop
  - To provide you some ideas about the expected outcome and materials covered during the workshop from the perspective of the CI design team
- We ask you to please fill it out to the degree possible/applicable. Please try to provide answers to as many (relevant) questions as you can
- You can use this template as you like. You can modify it, take only parts of it, add your own slides, copy/paste from it, or use it to structure your own text/spreadsheed/slideset documents …

General Goals for the Requirements Analysis

- Describe your current situation. We want to know how you work, what do you do? Please describe your work:
  - Definition of basic terms: instrument, platform, data, etc.
  - Tools, technologies, processes, data used and/or available
  - Organizational details (e.g. responsibilities, roles in team, workflows, policies)
  - What works well?
  - What are your biggest challenges?
- Determine short-term improvements
  - What would make every-day observatory management tasks easier and more effective?
  - What problems are causing delays or other issues and need to be addressed ASAP?
- Identify CI transformative vision and requirements
  - Assuming there is a transformative community CI in place, what are the expectations for an “ideal CI”?
  - Capabilities, interfaces, necessary guarantees, resources provided, etc.
- Scope
  - As relevant to the OOI CyberInfrastructure
  - From the viewpoint of your community

Current Situation and Expected Changes

- What capabilities and properties do you require from a cyberinfrastructure that supports your current or anticipated ocean observatory? Please rank.
- What capabilities and properties do you expect from a next generation cyberinfrastructure in the oceanographic domain that would benefit you and the community in the next decade? Please rank.
- What works particularly well in your domain? Exemplary standards, tools, platforms, portals, technologies, etc?
- Please list the biggest impediments that currently exist for your work and/or the community. Please rank and explain.

Ocean Observatory Management

- What tools do you currently use and/or are you developing to manage the resources in your ocean observatory? Please explain some important specifics of these and any related tools.
- Please describe a typical every-day scenario operating your ocean observatory. Example pictures, configurations, documentation etc. are always helpful. Please attach, if available.
- What would make your ability to manage your ocean observatory more effective?

Management Tasking

- How do you plan ocean observatory resource management?
- Do you modify observatory resources while they are operating? Please describe how this is done.
- How do you assess observatory operation? Please describe how metrics are collected and assessed.
- How do you store/visualize the results from observational programs?
- How does virtualization and simulation help you in managing observatories and their resources?

Operation and Maintenance

- What are the standards for operations and maintenance you need to get clean data? How do operation and maintenance requirements affect the design of your ocean observatory?
- What importance does operation and maintenance have in your overall work?
- How do you manage changes to instruments and observatory resources such as power and bandwidth?

Security, Privacy, Policy

- How do policy and security concerns affect the operation of your observatory? Which technologies, mechanisms, procedures and restrictions do you have to employ to guarantee secure operations?
- Which domains of authority, and security roles and responsibilities exist in your observatory?
- Please explain the relevant security and policy guarantees that you and/or your organization require. This includes authentication mechanisms, authorization (access control) and resource access policy strategies, privacy needs, intellectual property issues, etc.

Interfaces

- What application interfaces, user interfaces and visualization support do you envision and/or require of an effective and easy to use community cyber-infrastructure?

- Which processes and procedures facilitate marine operator/user interactions in your case? Which of these have proven particularly important/effective?
- What instrument interfaces (both sensor and actuator) do you envision and/or require of an effective and easy to use community cyber-infrastructure?

Education and Outreach

- How do education and outreach concerns affect your ocean observatory and the presentation of the results?
- How do you make observation program results available for education and outreach purposes?
- What would make these tasks easier?

Comments, Expectations, Suggestions

- What do you expect from the upcoming OOI CI requirements workshop?
- What have we missed? What didn't we ask you about that we should have?

Additional reading materials, References

- Are there any similar projects/communities that you like and/or that are technology-wise exemplary?
- Are there standards, other national or international efforts that the OOI design team should consider/evaluate?
- Anything you think is relevant that you want to add to this questionnaire?
- Further reading materials
- References

B List of Previous User Requirements

The following table provides the list of CI science user requirements as of May 2008, resulting from the first two requirements workshop. For detailed explanations with each requirement, please refer to [CI-RWS2].

Cat.	Req-ID	Requirement
Resource Management		
	RWS2-R1	The CI shall notify registered users and applications when new resources are added to the system.
	RWS1-R3	The CI shall be extensible to allow the addition of new resources and applications to the OOI infrastructure.
	RWS1-R9	The CI shall provide a catalog listing all resources under CI governance.
	RWS1-R9A	The CI shall enable users to discover observatory resources together with their meta-data based on resource characteristics and user-defined search criteria.
	RWS1-R11	The CI shall catalog physical samples in the CI resource catalog.
	RWS1-R12	The CI shall support cross-referencing from CI governed resources to external re-source catalogs and metadata.
	RWS1-R16	The CI shall bind metadata to all resources under CI governance throughout the re-source life cycle.
	RWS1-R18	The CI shall provide standard OOI metadata descriptions that include, but are not limited to, a complete description of resource behavior, content, syntax, semantics, provenance, quality, context and lineage.
	RWS1-R19	The CI shall allow the discovery of all information resources that are based on a given original information resource.
	RWS1-R20	The CI shall provide information resource subscribers automatic and manual fallback options with similar characteristics in case the original resource becomes unavailable.
	RWS1-R26	The CI shall provide notification of resource state change to all resource subscribers.
	RWS1-R33	The CI shall collect and provide resource access statistics.
Data Management		
	RWS1-R21	The CI shall be capable of archiving all data and data products associated with an OOI observatory or other CI-governed information resource.
	RWS1-R22	The CI shall support the publication, distribution and archiving of different versions of the same data product.
	RWS1-R23	The CI shall ensure the integrity and completeness of all data products throughout the OOI life cycle.
	RWS1-R24	The CI shall ensure that all archived data products can be restored in their complete and most recent state.
	RWS1-R30	The CI shall publish new data products resulting from processing of existing data products.
	RWS1-R31	The CI shall enable users and applications to subscribe to information resources in the form of data streams.
	RWS1-R47	The CI shall provide a topic-based (publish-subscribe) data distribution infrastructure that supports real-time and near real-time delivery, guaranteed delivery, buffering and data streaming subject to resource availability.
Science Data Management		
	RWS2-R2	The CI shall interface with, ingest, and distribute data from external data sources, databases, and data distribution networks of related scientific domains.
	RWS2-R3	The CI shall provide interactive and automated data quality control (QC) tools.

Cat.	Req-ID	Requirement
	RWS2-R4	The CI shall provide standard and user-defined methods to assess the quality of data.
	RWS2-R5	The CI shall facilitate the moderation and auditing of published data.
	RWS2-R6	The CI shall act as a broker for CI-managed data products.
	RWS2-R7	The CI shall provide access to CI-manage data products in standard formats and subsets.
	RWS2-R8	The CI shall act as a broker between information and processing resources.
	RWS2-R9	The CI shall make unprocessed raw sensor data available on request.
	RWS2-R10	The CI shall track data provenance and correspondence.
	RWS2-R11	The CI shall credit data publishers when data products are accessed.
	RWS2-R12	The CI shall create and distribute related data products from a given source data product that have different characteristics, such as resolution, level of detail, real-time, form and quality.
	RWS2-R13	The CI shall flag data stream state change.
	RWS2-R14	The CI shall support the provision of complete metadata by users.
	RWS1-R4	The CI shall support a standard set of data exchange formats.
	RWS1-R4a	The CI shall translate between the standard data exchange formats without loss of information.
	RWS1-R5	The CI shall allow the addition of user-defined data exchange formats and translators.
Research and Analysis		
	RWS2-R15	The CI shall provide capabilities and user/application interfaces for researching scientific materials and OOI-governed resources across disciplines.
	RWS2-R16	The CI shall suggest suitable data products, data transformations, observation resources, analysis tools, visualization tools and other OOI resources based on user-specified research questions in domain language.
	RWS2-R17	The CI shall support interactive and iterative analysis and visualization through infrastructure, tools and user interfaces.
	RWS2-R18	The CI shall provide tools, user interfaces and visualization for the analysis, combination and comparison of disparate, heterogeneous data sets..
	RWS1-R25	The CI shall provide a standard, extensible set of data product processing elements that provide data assimilation, alignment, consolidation, aggregation, transformation, filtering and quality control tasks.
Ocean Modeling		
	RWS2-R19	The CI shall enable the efficient configuration, execution, debugging and tuning of numerical ocean models.
	RWS2-R20	The CI shall support the interaction of model developers and non-expert model users.
	RWS2-R21	The CI shall provide facilities to develop and tune numerical models and their parameters.
	RWS2-R22	The CI shall provide a virtual model environment and simulator to determine optimal model inputs, parameterizations and outcome qualities.
	RWS2-R23	The CI shall enable the sharing of ocean modeling, data assimilation and visualization components, including the extension of models with new model components.
	RWS2-R24	The CI shall provide a repository and sharing capabilities for numerical model algorithms, model configurations, data processing tools and documentation.
	RWS2-R25	The CI shall archive numerical models under configuration control.
	RWS2-R26	The CI shall recompute model data products using archived models and workflows.
	RWS2-R27	The CI shall enable the modification of archived numerical models and workflows.

Cat.	Req-ID	Requirement
	RWS2-R28	The CI shall provide an environment for the development of community numerical models under community process support.
	RWS2-R29	The CI shall provide a non-restricted environment for the development of independent numerical models.
	RWS2-R30	The CI shall support nesting of ocean models at different geographical scales.
	RWS2-R31	The CI shall provide a framework for the adaptation of model resolution to the available resources.
	RWS2-R32	The CI shall support model ensemble definition, execution and analysis.
	RWS2-R33	The CI shall publish both elements and aggregated ensemble data products.
	RWS2-R34	The CI shall support flexible high performance model execution.
Visualization		
	RWS2-R35	The CI shall provide a uniform and consistent for numerical model output visualization and analysis in 2D, 3D and 4D.
	RWS2-R36	The CI shall provide interactive visualization of the 3D and 4D ocean.
	RWS2-R37	The CI shall support the integration of external visualization and analysis tools.
Computation and Process Execution		
	RWS2-R38	The CI shall support the execution of large scale numerical ocean models across different locations on the network.
	RWS2-R39	The CI shall support workflows for automated numerical model execution, including just-in-time input data preparation, model computation, output post-processing, and publication of results.
	RWS2-R40	The CI shall enable the one-time and recurring execution of numerical models on any networked computational resource with quality-of-service guarantees based on contracts and policy.
	RWS1-R27	The CI shall provide uniform and easy-to-use interfaces to computational resources with varying characteristics to define executable processes.
Sensors and Instrument Interfaces		
	RWS2-R41	The CI shall provide flexible and reliable access to remote resources.
	RWS2-R42	The CI shall provide real-time monitoring of remote sensors.
	RWS2-R43	The CI shall provide continuous collection of scientific data during extreme weather events.
	RWS2-R44	The CI shall provide discovery for the number and characteristics of sensors deployed on an instrument platform.
	RWS2-R45	The CI shall support adaptive observation.
Mission Planning and Control		
	RWS2-R46	The CI shall provide capabilities and user/application interfaces for mission planning and control.
Application Integration and External Interfaces		
	RWS1-R1	The CI shall provision an integrated network comprised of distributed resources, applications and users.
	RWS1-R2	The CI shall enable non-persistent connection of resources, users and applications.
	RWS1-R6	The CI shall provide application program interfaces (APIs) to all CI services.
	RWS1-R7	The CI shall provide a synoptic time service with an accuracy of TBD to all resources connected to the OOI observatories.
Presentation and User Interfaces		
	RWS2-R47	The CI shall provide "one stop shopping" interfaces that provide and collocate relevant information regarding scientific research using OOI resources.

Cat.	Req-ID	Requirement
	RWS2-R48	The CI shall provide annotation, commenting, ranking and rating services for resources.
	RWS2-R49	The CI shall provide project and user workspace capabilities and user interfaces.
	RWS2-R50	The CI shall provide long-term and ad hoc social networking and collaboration capabilities.
	RWS1-R34	The CI shall provide homogeneous, intuitive, easy-to-use web-based interfaces to all CI services and resources.
	RWS1-R35	The CI shall provide the capability to make OOI-standard metadata human readable.
	RWS1-R38	The CI shall provide extensible configurable visualization capabilities for selected types of data streams.
	RWS1-R49	The CI shall provide real-time analysis and visualization for data resources.
Security, Safety and Privacy Properties		
	RWS2-R51	The CI shall provide interfaces to define security and policy for information managers at participating institutions.
	RWS2-R52	The CI shall provide secure operations.
	RWS2-R53	The CI shall only permit authenticated and authorized users to access OOI resources.
	RWS1-R43	The CI shall provide mechanisms to enforce user privacy policies.
	RWS1-R44	The CI shall enable any authenticated party to share their resources.
	RWS1-R44A	The CI shall grant or restrict resource access subject to use policy.
Quality Properties		
	RWS1-R46	The CI infrastructure shall provide services and deliver messages with reliability and accuracy that is comparable to that of distributed Internet applications.
Education and Outreach		
	RWS2-R54	The CI shall facilitate the creation of publicly available idealized numerical ocean models with a limited choice of modifiable parameters for educational purposes.
Documentation		
	RWS1-R41	The CI IO shall make all source code for the OOI CyberInfrastructure implementation and drivers publicly available, subject to applicable licenses.
	RWS1-R42	The CI shall provide documentation for all components of the CI, including all application program interfaces (APIs) to CI services.
	RWS1-R39	The CI IO shall provide all documentation in web-based formats.
Development Process		
	RWS2-R55	The CI IO shall circulate CI requirements and designs within and outside the OOI community so that comparable infrastructures can adopt them.
	RWS1-R8	The CI shall utilize open standards and open source software to the maximum possible extent.
	RWS1-R40	The CI IO shall provide a process for submitting and incorporating user-suggested changes to the CI.
	RWS1-R48	The CI shall provide for the flexible and transparent extension of CI services and interfaces to incorporate user-provided processes, user and application interfaces, applications and resources.

C Workshop Agenda

Day 1, May 28, 2008 (Wednesday)

Time	Presenter(s)	Topics
08:30 AM	Frank Vernon	Welcome & Introductions
08:40 AM	Frank Vernon	Purpose and intent of this meeting, outcome expectations; OOI CI user requirements elicitation process
09:10 AM	Eric Terrill	Background Presentation: SCCOOS
09:40 AM	Mark Moline	Background Presentation: Glider network management
10:25 AM	Paul Macoun	Background Presentation: VENUS observatory management
10:55 AM	Craig Dawe	Background Presentation: MBARI Operations
11:25 AM	Jennifer Eakins	Background Presentation: ANF, EarthScope
01:00 AM	CI ADT	Present day observatory Management use cases scenarios. "A day in the life of an observatory operator". Capture two example scenarios trying to cover the main focus topics of the workshop such as domains of authority, security concerns, roles and responsibilities, processes, interfaces and interactions.
03:15 AM	CI ADT	Detailed analysis of one example scenario in each of the two groups.
04:45 AM	CI ADT	Presentation of breakout session results; feedback,

Day 2, May 29, 2008 (Thursday)

Time	Presenter(s)	Topics
08:30 AM	Mike Kelly	Regional Observatory Infrastructure
09:00 AM	Bob Collier	Coastal and Global Observatory Infrastructure
09:30 AM	Matthew Arrott	Proposed CI infrastructure for the OOI
10:15 AM	CI ADT	Future OOI observatory management user case scenario. Observatory management processes involving the coastal, regional and global components and their interfaces to the CyberInfrastructure, forming an Integrated Observatory.
01:00 PM	CI ADT	Future OOI observatory management user case scenario Data integration and instrument access involving the coastal, regional and global components and their interfaces to the CyberInfrastructure, forming an Integrated Observatory
02:45 PM	CI ADT	Domain modeling and prioritization session
04:45 PM	Frank Vernon	Wrap up and feedback
05:00 PM		Workshop adjourns

D List of Participants

Name	Organization	Project Role
Matthew Arrott	UCSD/Calit2	CI Project Manager
Jim Christman	Ocean Leadership	
Bob Collier	Oregon State University	
Craig Dawe	MBARI	
Jennifer Eakins	UCSD/SIO	
Duane Edgington	MBARI	
Emilia Farcas	UCSD/Calit2	CI System Modeler
Mike Harrington	UW APL	
Mike Kelly	UW APL	
Rosie Lunde	Ocean Leadership	
Paul Macoun	University of Victoria	
Steve Meacham	NSF	
Michael Meisinger	UCSD/Calit2	CI Requirements Analyst
Matt Moldovan	UCSD/SIO	
Mark Moline	CalPoly, San Luis Obispo	
Don Peters	WHOI	
Lloyd Regier	UCSD/SIO	
Elizabeth Rosenzweig	Bubble Mountain Consulting	CI Information Architect, Consultant
Alex Talalayevsky	Ocean Leadership	
Eric Terrill	UCSD/SIO	
Frank Vernon	UCSD/SIO	CI Deputy Director

E Abbreviations

Abbreviation	Meaning
CI	OOI CyberInfrastructure
CI ADT	OOI CyberInfrastructure Architecture and Design Team
CI IO	OOI CyberInfrastructure Implementing Organization
IOOS	Integrated Ocean Observing System
NetCDF	Network Common Data Form
OOI	Ocean Observatories Initiative
PDR	Preliminary Design Review
SCCOOS	Southern California Coastal Ocean Observing System

F References

Reference	Citation
[CI-CARCH]	CI conceptual architecture and initial requirements, available at http://www.orionprogram.org/organization/committees/ciarch
[CI-DPG-WEB]	OOI CI Requirements Elicitation Workshop, Data Product Generation. Website available at: http://www.ooici.ucsd.edu/spaces/display/WS/RWS-DPG
[CI-IOM-WEB]	OOI CI Requirements Elicitation Workshop, Integrated Observatory Management. Website available at: http://www.ooici.ucsd.edu/spaces/display/WS/RWS-IOM
[CI-OOP-WEB]	OOI CI Requirements Elicitation Workshop, Ocean Observing Programs. Website available at: http://www.ooici.ucsd.edu/spaces/display/WS/RWS-OOP
[CI-PAD]	OOI CI Architecture Document, PDR Final version, 16-Nov-2007
[CI-RWS1]	OOI CI First Science User Requirements Elicitation Workshop Report, OOI CI, Final version 1.0, 08-Nov-2007, available at: http://www.ooici.ucsd.edu/spaces/download/attachments/10453181/OOI-CI-ReqWS1-Report-FINAL.pdf?version=1
[CI-RWS2]	OOI CI Second Science User Requirements Elicitation Workshop Report, OOI CI, Final version 1.0, 09-May-2008, available at: http://www.ooici.ucsd.edu/spaces/download/attachments/10453181/OOI-CI-ReqWS2-Report.pdf?version=2
[CI-WEBSITE]	OOI CI Website, available at http://www.ooici.uscd.edu
[NORIA]	Network for Ocean Research, Interaction and Application (NORIA) Proposal, 22-Dec-2006
[OOI-CU-REQ]	OOI Cyber User Requirements, exported from OOI DOORS requirements database. Version of 7/31/08. Available at: http://www.ooici.ucsd.edu/spaces/display/WS
[SCIPROSP]	OOI Science Prospectus, Dec 2007, available at: http://www.oceanleadership.org/files/Science_Prospectus_2007-10-10_lowres_0.pdf