Self-Guided Tour of the SCOOP Infrastructure

This document takes you on a click-by-click tour of different web sites that make up the public face of the SURA Coastal Ocean Observing and Prediction (SCOOP) program, an initiative of the SURA Coastal Research Committee. The major subsections can be visited out of sequence.

Introduction

We organize this document by the three priorities areas in the SCOOP program:

1. Community standards,
2. Ocean Observations Interoperability Demo,
3. Coastal Prediction Distributed Facility Prototype.

Originally we called these (1) Data Standards, (2) Data Grid and (3) Model Grid. We abandoned the term Grid in deference to SURAgrid, which provides increasingly strong infrastructure support for SCOOP. Now, we refer to the SCOOP program as prototyping a distributed facility for coastal research and applications based on a "Service-Oriented Architecture" (SOA). The concept supports a bold and new "Cyberinfrastructure Vision for the 21st Century Discovery" (NSF Cyberinfrastructure Council, March 2007), and advances the only sensible architecture for an Integrated Ocean Observing System (IOOS) that supports national need. Thus, the SCOOP program demonstrates how SURA members can leverage SURA's strength in IT to assure that that the research community plays strong and substantive roles in these visions for the future.

1. Community Standards

The SCOOP program was an early adopter and advocate of standards from the Open Geospatial Consortium (OGC, www.opengeospatial.org). The OGC is a non-profit, international industry consortium of 339 companies, government agencies and universities participating in a consensus process to develop publicly available web-service interface specifications. With early support from SEACOOS and GoMOOS, SURA helped create www.OpenIOOS.org, which continues to serve as a proof-of-concept portal to OGC-compliant data integration and interoperability demos. Community involvement continues to grow, especially with SURA's role as co-PI on an NSF-grant for the marine metadata interoperability project (MMI, www.marinemetadata.org).

2. Interoperability Demo

SURA (an OGC member) is leading the OGC-sanctioned Ocean Science Interoperability Experiment (OCEANS IE; http://www.opengeospatial.org/projects/initiatives/oceansie). Co-initiators include Texas A&M, NCAR/Unidata, MBARI and GoMOOS. The Ocean Science IE is the OGC-member component of the open-access OOSTethys activity (www.oostethys.org). OOSTethys combines two formerly distinct interoperability demos: OpenIOOS and MMITethys (hence the name OOSTethys). OOSTethys is SURA's responsibility on the NSF-funded MMI grant. SURA also supports OOSTethys with substantial ONR and NOAA funds. OpenIOOS.org is an aggregator/portal for OOSTethys. International contributors are growing in number, and include private-sector and government partners. All software & cookbooks are open source.

3. Distributed Facility Prototype for Coastal Prediction

The SCOOP SOA for coastal prediction is described in a publication to appear this summer in the MTS Journal special issue on inundation. This is by far the most sophisticated and technically complex initiative of the SCOOP program, and provides a proof of concept for the notion of a Distributed Coastal Laboratory being advanced by SURA's Coastal Committee.

Table of Contents

Start: SCOOP Program Web Site (http://scoop.sura.org)

NOTE: For April 17th, please use the development version, http://scoopdev.sura.org, of the program web site, http://scoop.sura.org. We have yet to verify and migrate all the new capabilities to our "production" site: http://scoop.sura.org.

This site links to many of the key elements of the tour:

1. For an overview of all the different oceanographic and atmospheric models used within SCOOP, click the "**Models**" menu item in the top navigation bar. This page lists the wind, wave, and surge models that are used within SCOOP for environmental prediction.
2. To get more detailed information on any of the models, click the "**Usage**" button to the left of the model name.
   a. From the Usage page, to view the details on a particular collection of actual model run results that are available within the SCOOP Archives, simply **click on any of the collection names**.
3. The OpenIOOS Interoperability Experiment provides visualization of many of the SCOOP data products. This application is accessible at the "**OpenIOOS**" menu item on the top navigation bar (see the full tour of OpenIOOS below). OpenIOOS.org also serves as a portal to OOSTethys.org.
4. The SCOOP Data System is comprised of a centralized data catalog, supporting several distributed archives hosting SCOOP data products. Data discovery, search and ordering is available at the "**Data**" menu item on the top navigation bar (see the full tour of the SCOOP Data System under "SCOOP Catalog" below).
5. SCOOP partners are investigating and applying the use of Grid computing technologies for

enhanced computational capabilities. Some of our participants are international leaders in the field. For information about some of the SCOOP-specific activities, click on the "**Grid Computing**" menu item on the top navigation bar. (Note: Text on this page is slightly out of date.)

The SCOOP program employs a service-oriented architecture in which interoperating system components are loosely coupled through standardized interfaces, and interconnected over the Internet. The concept is increasingly being implemented in other scientific disciplines (Ian Foster, *Science*, 308: 814-817.) Information on public components of the SCOOP architecture are available for others to utilize and investigate by selecting the "**Documents**" menu item on the top navigation bar and then clicking on "**Documentation for SCOOP web service interfaces**" at the bottom of the page (or go directly to http://scoopdev.sura.org/services.html) (See Bogden et al. 2007; MTS Journal, in press.)

1. Community Standards (OpenIOOS.org & OOSTethys.org)

The OGC Ocean Science Interoperability Experiment (OCEANS IE)

The OGC Ocean Science IE participants must be OGC members. The OCEANS IE is described here: http://www.opengeospatial.org/projects/initiatives/oceansie. A relatively large group is involved in the companion, informal, open-access OOSTethys IE (www.oostethys.org), which includes important contributions from participants in the MMI (notably, Luis Bermudez [MMI/MBARI]; Gerry Creager [Texas A&M University]; Helen Conover [University of Alabama at Huntsville]). Both IEs have had as their primary foci datasets and descriptions typical of ocean sensors that produce time series of point/station observations.

The goals of both initiatives can be characterized using the three basic "sub-systems of the Integrated Ocean Observing System (IOOS) as described in Ocean.US planning documents (www.ocean.us), namely, (1) data acquisition, (2) data management and communications (DMAC) and (3) modeling and analysis. The conceptual architecture of the OCEANS IE (and OOSTethys) is attempting to blur the distinction between the data acquisition and DMAC subsystems. In an abstract sense, sensors and numerical models are both data sources: we are already considering these techniques to describe numerical models and their output in addition to the more conventional *in situ* sensors.

The OpenIOOS Demo End-User Portal

OpenIOOS.org demonstrates how standards enable innovation. The web site "serves" OGC-compliant integrated data products. But the real novelty is the integration capacity behind the scenes, which leverages the web site's role as a "client" of OGC-compliant data services. In this role, OpenIOOS.org implements a service-oriented architecture along with the data providers, all of whom use a wide variety of open source and proprietary software implementations of the OGC specifications. OGC-compliant data providers actually do most of the work. OpenIOOS.org demonstrates the power of standards-based integration when a community of providers complies with industry standards for open-access web services. Those data-service providers deserve most of the credit for OpenIOOS.org.

To get there, click on the "**OpenIOOS**" menu item in the top navigation bar on the SCOOP web site (http://scoop.sura.org) or visit http://www.openioos.org/ directly.

The OpenIOOS site is the public portal demonstrating many of the information products being generated by the SCOOP infrastructure. This tour will walk through a sampling of the available products.

1. **Portal Interface –** The OpenIOOS portal shows data for three different operational modes: 24/7, Event Driven and Retrospective:
    a. **24/7 mode** is the normal day-to-day ocean observations and forecast models that run within the SCOOP infrastructure which include water level and wave forecasts for the east coast of North America.
    b. **Event-Driven mode** begins when the National Hurricane Center issues a Forecast Advisory indicating that there is a severe Atlantic storm that may threaten North America.
    c. **Retrospective mode** is used to visualize and download past storm data. The sections below will describe the different products available on OpenIOOS for the 24/7 and Event Driven modes.
2. **24/7 Operational mode -** During normal day-to-day operations, the primary products available on OpenIOOS include salinity and sea surface temperature observations (which will be discussed below in the OOSTethys.org section of the tour), water level observations and forecasts, and wave observations and forecasts.
    a. Click on the "**Water-level Models**" link in the left-hand navigation. This will display an animation showing the ADCIRC water-level model 6-hour hindcast and 30-hour forecast. You can click on any of the NOAA tide stations on the map and bring up a graph comparing model forecasts to actual observed data.
    b. Click on the "**Wave Models**" link in the left-hand navigation. This will display an animation showing the WaveWatch3 wave model 6-hour hindcast and 30 hour forecast. You can click on any of the buoy locations on the map and bring up a graph comparing model forecasts to actual observed data.
3. **Event-Driven (hurricane) mode -** http://www.openioos.org/hurricane/
    a. When the National Hurricane Center issues a Forecast Advisory, the OpenIOOS home page automatically changes to show this page, highlighting whatever the current storm in the Atlantic basin is. Since there is no active storm at this time, it is showing the first storm from last hurricane season, Alberto.
    b. This display defaults to showing the **water-level forecasts** generated during Alberto. During this animation you can see the storm track being refined as time goes by and the storm progresses along its track. In addition, you can click on any of the NOAA tide stations (colored squares) on the map to see a graph comparing forecasted water level with observed water level during the storm.
    c. You can also click on "**Wave Models**" under the "Alberto" subheading in the left-hand navigation to see a similar animated display for the wave forecasts done during Alberto. In addition, you can click on any of the buoy locations (red triangles) on the map to see a graph comparing forecasted waves with observed waves during the storm.
4. **Retrospective mode –** As the products available in Retrospective mode are very similar to those in Event Driven mode, we will not cover Retrospective mode. Retrospective mode is initiated with a request for reprocessing of data. Insertion of a dataset into the data transport system initiates a process essentially identical to the Event-Driven mode.
5. **Dynamic Mapping –** From the "Wave Model" page on (http://www.openioos.org/hurricane/), click on the link in the upper right labeled "**Dynamic Map**". This page displays a GIS map generated with an OGC-compliant Web Mapping Service (WMS) from the SCOOP Archive. Click on the **"i"** icon on the upper-left of the

map, then click on any of the buoy locations. This will bring up the observation from that buoy at that point in time in the Alberto storm. If you click the **"graph"** icon on the upper-left of the map, then click one of the buoy locations, you'll see the model data to observation graph. You can also click on any of the "**twisting hurricane icons**" along the storm track to bring up the details of the National Hurricane Center's forecast advisory. Lastly, if you wish to shift the map forward or backward in time along the storm track, you can click the **"clock"** icon on the upper-left of the map, you can then click anywhere along the storm track and the model will shift forward or backward to the time that it was when the storm was located at the spot you clicked on. You can also add additional layers to the map by clicking on the checkboxes on the right then clicking "**Redraw**". One interesting layer to add is the "**NASA (GSFC): Aqua MODIS Daily Composite**" satellite image, which will show you a satellite image of what the storm looked like at that point in time.

2. Interoperability Demo

The OOSTethys Project - http://www.oostethys.org/

OOSTethys is a grassroots community initiative led by SURA partners (notably, Texas A&M, UAH, VIMS and GoMOOS) and other participants from across the globe. Many of these individuals come to OOSTethys from the MMI community initiative, www.marinemetadata.org. OOSTethys is serving as the interoperability demo for both SURA and MMI. SURA – a co-PI on the NSF-funded MMI initiative – has primary responsibility for the MMI interoperability demo.

The OOSTethys activity is an open-source, software development and SOA implementation effort that includes partners who have aligned to make some concrete choices that will advance the "system of systems" concept for ocean sciences. OOSTethys partners are applying interface specifications from the Open Geospatial Consortium (OGC), and standards from ISO, OASIS and the W3C consortium. OOSTethys is currently focusing on: (1) OGC Sensor Web Enablement (SWE) initiative, Web Feature Services (WFS) and related OGC specifications, (2) semantic mediation services as defined by the Semantic Web W3C group, and (3) OASIS catalogs.

1. The right side of the home page (http://www.oostethys.org) lists the participating **SURA partners** and other contributors. These individuals have weekly conference calls and an active listserv. The group is open to new participants, and we encourage participation, but there tends to be a self-selection process for determining leadership. Anyone interested in contributing to leadership of the initiative should get engaged with this group. This kind of organization is typical of any open-source, software development activity. For OOSTethys, this dedicated involvement includes several SURA partners, Luis Bermudez from MMI and the Monterey Bay Aquarium Research Institute (MBARI), and a growing number of regional data providers in the list.
2. There are two smaller maps that appear side-by-side. Each of these maps is a link that takes you over to the OpenIOOS site. This is because the "SOS Salinity" map on OpenIOOS is the testbed for the use of the OGC standards adopted by OOSTethys. There are two different maps available. The **left-hand map** brings up a map based on **Google Maps** technology. The **right-hand map** displays a map based on the **Open Source MapServer** system. Either map then allows you to click on any of the observing platform locations on the map and get current observation data. This data is retrieved from distributed data providers using the Sensor Observation Service (SOS) protocol from the OGC.
3. Back on the home page of OOSTethys, if you click on the link below the two maps that says "**How are we building these maps? Click for details**" it will take you to a detailed description of exactly what is taking place behind the scenes to retrieve the real-time

observation data using the SOS protocols.

4. The "**Downloads**" link in the left-hand navigation takes you to a page where you can download the various cookbooks that have been created as part of the OOSTethys project. Each cookbook is designed to allow a data provider to begin supplying its observation data to OOSTethys using SOS with a minimum of effort. Most of the data providers currently involved with OOSTethys used one of these cookbooks as a starting place.

OGC Ocean Science IE

The OOSTethys effort has been expanded into an officially sanctioned "**Interoperability Experiment**" (IE) within the OGC. The IE is just getting underway and you can read about it on the OGC web site at http://www.opengeospatial.org/projects/initiatives/oceansie .

3. Prototype Distributed Facility for Ocean Prediction

NOTE FOR APRIL 17 DEMO: Although the primary SCOOP web site is http://scoop.sura.org, all links below reference the "development" version of this site, http://scoopdev.sura.org. As new functions are evaluated and vetted, they will be migrated to the "production" web site.

The SCOOP Catalog

A user-friendly, human interface to the SCOOP Catalog, http://scoopdev.sura.org/Catalog/, can be found by clicking the "**Data**" link in the top navigation bar on the SCOOP program web site (http://scoopdev.sura.org). This is a general-purpose interface for metadata-based discovery of all the information in the system, including archived model output, ocean observations, and web-service interfaces.

1. The SCOOP Catalog is the repository of information (i.e., metadata) about all of the environmental data that has been generated by the SCOOP program and is stored in our distributed Archives. The Catalog interface supports searching by Collection Data keywords and Storms. From the Catalog page, you can generate a listing of all the storms with data in the SCOOP system by clicking on the "**Go!**" button within the blue box entitled "**QUICK SEARCH BY STORM**". Alternatively you can enter a storm name and/or year to narrow the listing of storms that results.
    a. From the **STORM SEARCH RESULTS** page you can get four types of information as shown in the key at the top of the page. By clicking on the four icons to the left of any storm name, you can get:
        i. A list of all data collections in the SCOOP Archives for the time period of this storm
        ii. An interactive tool that shows the storm track and allows you to draw a bounding box around any portion of the storm track then list data collections in the SCOOP Archives for the selected storm track time period.
        iii. An animation of the storm track showing the storm category over time with associated wind and pressure.
        iv. A Google Earth data file (KML file) that displays all the storm tracks for the year of the storm in Google Earth and shows an animation of the storm progression with wind and pressure data **(NOTE: This requires that you have the free Google Earth installed on your computer – download it from http://earth.google.com/download-earth.html).**
    b. Back on the Catalog home page, by selecting the "**Advanced Search**" button under "QUICK SEARCH BY STORM" you can specify more storm details to refine the search, such as date, wind speed, pressure, etc.

2. To search the data collections available in the SCOOP Archives based on keywords associated with the data, go back to the Catalog home page http://scoopdev.sura.org/Catalog/, and specify a keyword search context in the blue box entitled "QUICK SEARCH BY COLLECTION". For this example, click on the "Go!" button within the "**QUICK SEARCH BY COLLECTION**" to display all the collections.
    c. Scroll down through the list of data collections and check "NHC Forecast Advisory"
    d. Scroll to the bottom of the page and click "Submit" – this will take you to a page where you can view the date ranges that this data is available at the individual archives.
        v. From the resulting "**SCOOP DATA AVAILABILITY**" page there are two ways to actually download the data files from the SCOOP Archives. First, click the blue arrow icon to the right of the word "LSU" – this will bring up a listing of all the files in this collection available at our LSU Archive. You can click on any individual file to download and view it.
        vi. The second way to download data files is by ordering them. This allows you to order large groups of files packaged as a single compressed archive file instead of having to download them individually. To see how this works, go back to the "**SCOOP DATA AVAILABILITY"** page, and check off only the box next to the date range shown for the Texas A&M Archive, shortening the date range if desired. Then click the "**Order**" button. This takes you to a page where you will enter your name and email address. Once the order is completed by the Archive, you will receive an email telling you where to download the order.

Infrastructure Portal – Login: "guest" with password "scoop1"

All links in the remainder of this section require that you first log into the SCOOP Infrastructure Portal, which is the primary interface to event-driven, disributed computing:

http://portal.scoop.sura.org/gridsphere/gridsphere.

Please use login: "guest" with a password of "scoop1"

This is the portal that allows SCOOP scientists to monitor activities and workflow within the SCOOP infrastructure. It includes "portlets" (represented by tabs within the portal) that allow you to monitor SCOOP computing resources, model runs ("jobs") and data flowing through the system, and allows you visualize model output.

Monitoring Portlets

Once logged into http://portal.scoop.sura.org/gridsphere/gridsphere:

1. Click on the tab called "**Resource Monitoring**", then click on the sub-menu link called "SCOOP Resources" to see a map showing all of the SURAgrid computing resources along with their current status.
2. Click on the tab called "**Job Monitoring**" to see what models are currently running and have run earlier today. The columns show the model name, what wind forcing is being used, the forecast cycle, the grid being used for the model, what time the model run started, how many CPUs are in use for this run, what computing resource the model is running on, and what the current status of the model run job is. The Katrina demonstration runs a 10 track ensemble of ADCIRC and WaveWatch3 jobs forced by a set of synthetic winds called Analytic Winds.

3. Click on the tab called "**Data Flow Monitoring**". This portlet shows the data flows within the SCOOP system for any day and cycle time. The Katrina demo generates a artificially short forecast from the 27th of August, 2005 during hurricane Katrina, so to view the data for the Katrina demo, enter "**20050827T00**" into the "**Forecast Cycle**" box and click "**GO**". The 2nd row in the display show the wind forecasts being generated at University of Florida (UFL), being sent to the University of North Carolina (UNC) to be used as the forcing for the ADCIRC water level model, being sent to Louisiana State University (LSU) to as the forcing for the WaveWatch3 (WW3) model, and being sent to the SCOOP Archives at LSU and Texas A&M University (TAM). The 3rd row shows UNC sending completed ADCIRC model data to the two archives and the University of Alabama, Huntsville (UAH) data translator which extracts geographic point forecast data for pre-established locations for use on OpenIOOS.org. The 4th row shows LSU sending the completed WaveWatch3 model data to the two archives and to the UAH data translator for point forecast extractions for OpenIOOS.

Advanced User Interface (AUI)

The AUI provides a high functionality, OGC-compliant interface and visualization tool for scientists using the infrastructure, and for system administrators who want to preview archive contents. To get there from within the portal http://portal.scoop.sura.org/gridsphere/gridsphere,

1. Click on the tab called "**Model Visualization**" to view the **SCOOP Advanced User Interface (AUI)** for visualizing model ensemble output (note – if the portlet display appears tall and narrow, click the "Maximize" box two icons to the right of the "Model Visualization" label).
   e. The AUI is a visualization tool that displays animations of the SCOOP forecast ensemble running against hurricane Katrina data. It includes visualizations of the ADCIRC and CH3D water level models, as well as the WaveWatch3 wave model.
   f. Click on the "**ADCIRC Water Level**" in the gray box on the left labeled "Select a Quick View". This will bring up a display of the ADCIRC ensemble 96 hour forecast starting at midnight UTC on 8/27/2005. On the righthand side, there is a "**Play**" button below the word "Animate" Clicking on this "Play" button will start the animation.
   g. By clicking one of the "**Wind Forcing**" links on the left, then pressing the "**Play**" button again, you can see the forecast for each individual track in the ensemble.
   h. If you scroll down the page, on the lower left you'll find additional analysis tool links. These tools allow you to view:
      vii. **Maximum of Ensemble (MOE) –** This shows the maximum water level projected across all of the ensemble storm tracks for each point in time in the forecast. After clicking this link, you can click the "**Play**" button to restart the animation showing the MOE.
      viii. **Maximum of Forecast (MOF) –** This shows the maximum water level projected for a single track across the entire forecast period. After clicking this link, click a track listed under "**Wind Forcing**" to see the MOF for that track.
      ix. **Maximum of Maximum (MOM) –** This shows the maximum water level for all ensemble storm tracks across the entire forecast period.
   i. You can click on additional models from the "Select a Quick View" list at the top left to repeat the above steps for the WaveWatch3 wave model and the CH3D water-level model. Note that the Maximums are not currently available for the CH3D model.

2. Click on the "**Visualization Output**" tab to view three single-frame visualizations of the short forecast wind, wave, and water level model runs just completed during the current demo run.