COL/SE/02

A Powerful New Step for NOAA's Integrated Ocean Observing System

FY 2005 Proposal to the NOAA HPCC Program

September 13, 2004

| Title Page | Proposed Project | Budget Page |

Principal Investigator: **James C. Hendee, Ph.D.**

Line Organization: OAR/AOML
Routing Code: Ocean Chemistry Division
Address:

Atlantic Oceanographic and Meteorological Laboratory
National Oceanic and Atmospheric Administration
4301 Rickenbacker Causeway
Miami, FL 33149-1026

Phone: (305) 361-4396
Fax: (305) 361-4392
E-mail Address: jim.hendee@noaa.gov

Other Investigator 1	Other Investigator 2	Other Investigator 3
John.Proni@noaa.gov	Kevin.Wong@noaa.gov	Mark.Eakin@noaa.gov

Proposal Theme: **Technologies for Collaboration, Visualization or Analysis**

Funding Summary:

James C. Hendee, Ph.D.	Cathy B. Steward	Judith Gray
Research Oceanographer	Administrative Officer	Acting Director, AOML
NOAA/OAR/AOML/OCD	OAR/AOML	OAR/AOML

A Powerful New Step for NOAA's Integrated Ocean Observing System

Proposal for FY 2005 HPCC Funding

Prepared by: James C. Hendee, Ph.D.

Executive Summary: NOAA has committed to integrating ocean data from a variety of sources under a project called the Integrated Ocean Observing System. NOAA's Coral Reef Conservation Program has also committed to integrating coral data from a variety of sources for the benefit of coral reef researchers and Marine Park Area (MPA) managers. Herewith is a plan for achieving ocean data integration without reinventing the wheel, and using the world's most powerful data integration and inferencing tool, something NOAA has heretofore done without.

Problem Statement: In his *NOAA 2007 Annual Guidance Memorandum (*September 03, 2004), Vice Admiral Conrad Lautenbacher said, "The Integrated Ocean Observing System (IOOS) must be developed as a major component of the US contribution to GEOSS [Global Environmental Observing System of Systems]. Our contribution to IOOS should focus on connecting and strengthening existing observing systems and developing new techniques and measurements to improve spatial, temporal, and spectral density." Similarly, the National Office for Integrated and Sustained Ocean Observations (Oceans.US) has made at least three recommendations that call upon NOAA to play a major role in integrating ocean data:

Recommendation 26–9. Congress should fund the Integrated Ocean Observing System (IOOS) as a line item in the National Oceanic and Atmospheric Administration (NOAA) budget, to be spent subject to National Ocean Council direction and approval. IOOS funds should be appropriated without fiscal year limitation. NOAA should develop a streamlined process for distributing IOOS funds to other federal and nonfederal partners.

Recommendation 15-4. The National Oceanic and Atmospheric Administration, U.S. Geological Survey, and U.S. Environmental Protection Agency, working with other appropriate entities, should ensure that water quality monitoring data are translated into timely and useful information products that are easily accessible to the public and linked to output from the Integrated Ocean Observing System.

Recommendation 21–4. The U.S. Coral Reef Task Force should identify critical research and data needs related to coral reef ecosystems. These needs should guide agency research funding and be incorporated into the design and implementation of the Integrated Ocean Observing System.

NOAA's Coral Reef Conservation Program (CRCP) is in the process of establishing a Coral Reef Ecosystem Integrated Observing System (CREIOS). As part of CREIOS, the CRCP is building an array of in situ oceanographic monitoring stations to observe ocean conditions and processes influencing reef health. At present, the Atlantic Oceanographic and Meteorological Laboratory (AOML) and Coral Reef Ecosystem Division (CRED) are the principal programs building this in situ observing network. This network, which uses specialized software developed by AOML, is called the Coral Reef Early Warning System (CREWS) Network, and is unique among *in situ* monitoring networks in that it uses a suite of expert systems (an artificial intelligence tool) to monitor the quality of data in near real-time, and to advise of conditions in the marine environment that are conducive to coral bleaching, as well as any other marine behavioral event that is dependent upon the physical environment, and which can be measured

by an unattended sensor. Alerts of drifting or malfunctioning instruments are sent to a field crew for the proper attention; conditions conducive to coral bleaching are sent to researchers and Marine Park Area (MPA) or Sanctuary managers (for instance, Billy Causey of the Florida Keys National Marine Sanctuary). The CREWS Network consists of seven monitoring stations in the Florida Keys, five stations of the Australian Institute of Marine Science's Weather Station Network on the Great Barrier Reef, 27 buoys of the NMFS Coral Reef Ecosystem Division network in the Pacific, and two (of 18 planned) AOML custom monitoring stations in the Atlantic.

Although CREWS is the only network configured with a suite of expert systems for screening data, it is like most, if not all, other *in situ* observing systems in that it is not integrated with the other systems, and as a result, Marine Park Area (MPA) decision makers and researchers can not get the benefit of "the big picture." This is a problem facing the implementation of a truly integrated observing system: collating the huge amounts of data in a timely and efficient manner. Another problem is what to *do* with all that data once it is received. Historically, the data is handed over to myriad researchers to do their analyses on it, and they, in turn, make their recommendations to the policy makers and MPA managers. This of course takes a lot of time and, because of the time involved, critical, time-sensitive decisions may not be made in time (e.g., as in the case of a harmful algal bloom). Finally, there is the problem of time-consuming custom programming to be attended to each time a new station is added to the Network.

The Fit with the NOAA HPCC Program Objectives: The present proposal provides an extremely powerful, state-of-the-art, major off-the-shelf software solution for sharing and integrating data of many types across all available platforms connected by the Internet, for the purpose of providing data in a variety of formats, for providing split-second decisions as encoded by knowledge engineers in cooperation with problem domain specialists, and provides customizable Web interfaces to the data from remote locations. The proposed project will be of use across all Line Organizations in NOAA engaged in coral research, but the process will provide a path to follow for any data integration or real-time decision making problems within NOAA. The proposed solution also provides a tool for disaster planning, coordination of mitigation strategies, scheduling of response and recovery to disasters, and can provide intelligent agents at an unlimited number of locations to help provide redundancy and ease and speed of response.

Proposed Solution: There exists in other U.S. agencies a powerful expert system (and other artificial intelligence tools) integration and inferencing platform that is used for missions so critical they actually require immediate life and death decisions. This integrating and inferencing platform has the capability of integrating data from essentially all commercial databases, so long as the databases are connected by the Internet, or physically housed within the server, and joins them all together where they can be acted upon by decision making (i.e., inferencing) software in split-second time (i.e., hundreths of a second). This platform architecture, evolving since 1986, is called G2 (by Gensym, Inc.) and is currently being used by NASA, CIA, FBI, NSA, NRO, DOD, FAA, USAF, USN, DOE, the Army's Knowledge Engineering Group, the Joint Chiefs Decisions Support Group, the Defense Information Systems Agency, the Joint Intelligence Center, Boeing, the European Space Agency, Inmarsat, Intellsat, Lockheed, Satcomm, Iridium, and other groups and agencies. But not by NOAA.

Because of our proven expertise in *knowledge engineering* (i.e., the building of expert systems) in the successful building of the unique and powerful CREWS inferencing system, we at AOML are poised better than anybody in NOAA to take advantage of using this much more powerful software application development architecture in the development of a first phase of a truly *integrated ocean observing system* for the purposes of utilizing the multifarious sources of data in understanding the dynamics in the lives of coral reef organisms.

- The architecture can integrate data, images and documents from all *in situ,* biological and satellite stations, as well as from historical data (e.g., paleoecological data).
- Integration can occur through storage of data on various server applications (e.g., Oracle, Sybase, SQLServer, GemStone, dBase III+, or other media, as long as they are attached to the Internet or the Web.
- Data can be served on the Web via custom applications developed at AOML, or, (more importantly) can be developed and manipulated remotely (e.g., from Boulder, the Florida Keys National Marine Sanctuary, Silver Spring, etc.) to provide for *custom* views. (This capability has been called the Holy Grail of IOOS by one colleague who sits on the Oceans.US committee.)
- Predictions, reports, alerts, etc. can be made using neural network and/or expert system inferencing (using data and/or images) within this architecture, in real-time, or not, as necessary.
- The knowledge base upon which real-time alerts and recommendations are made, will be developed through an iterative process with all the major experts in coral reef ecosystem research.

We have garnered support for this effort from our coral colleagues at AOML (Ocean Acoustics Group—Dr. John Proni), in NOS (Coral Biogeography Group—Mark Monaco), NMFS (Coral Reef Ecosystem Division—Rusty Brainard and Kevin Wong), and NESDIS (Palaeoclimatology Group—Mark Eakin): in all instances they will support our efforts at providing data, and in select cases (i.e., co-PIs), will participate in the training and development of the Web-based interface. We propose to develop this integrative capacity as part of NOAA's new Coral Reef Ecosystem Integrated Observing System (CREIOS), then provide the expertise to others in NOAA for the benefit of their individual missions (e.g., storm and hurricane forecasting, the PORTS program, etc.). The procedures developed can be used for any other projects seeking to integrate ocean data and develop models for the purpose of decision support in real-time.

Implementation Plan and Activities

1) The G2 software will be purchased via the GSA contract, then will be installed on AOML's 8-node Linux Beowulf Cluster which currently conducts data screening and inferencing tasks for the CREWS Network.

2) Representatives from OAR/AOML, NMFS/CRED and NESDIS/ORA (the "G2 Group") will undergo two week's training from Gensym, Inc. at AOML. The Gensym representative will also be available during the installation of the software on the server.

3) The G2 Group will develop a common interface and Web presentation of benefit to the CREIOS group, and other coral research and MPA users.

4) A Telewindows tool (provided by Gensym) will be installed at each LO's site so that remote modifications of Web interfaces can be implemented to suit each user's needs.
5) All CREWS Network data, NESDIS satellite data and NMFS buoy and fisheries data will be integrated and made available to the G2 inferencing system and stored on the AOML Beowulf server, utilizing specific database bridges (e.g., an Oracle Bridge).
6) Interviews with domain problem experts will be conducted to properly encode the G2 inferencing system, e.g., for coral bleaching alerts, specific weather patterns, fish spawning predictions, etc.
7) Instances will be developed within G2's powerful object-oriented (OO) architecture to provide for OO Class-level monitoring stations, with the ability to modify for site-specific needs at the monitoring sites (OO instances).
8) The system will be tested, with feedback sought on corrections and enhancements from coral researchers and MPA managers. Figure 1, developed by Gensym, Inc., shows a possible Web interface at the end of the project.

Analysis:

Rationale. The present approach saves time and money to reach the IOOS and CREIOS goals by not reinventing the wheel, and taking an off-the-shelf proven capability for integrating all major types of data, and even better, actually doing something with the data in real-time (i.e., alerts to environmental conditions conducive to specific biological and other events).

Comparison Alternatives. CREWS was constructed to review data and predict coral bleaching in near real-time—and it is the only expert system that reviews marine ecosystem data in near real-time (hence there is no alternative)—but it has been known all along that the expert system shell utilized (CLIPS by NASA) would be labor intensive for programming if the Network ever grew (as it now has). G2's OO environment provides for quick development of instances of stations, as well as powerful data integration tools.

Benefits of this Solution. The proposed solution integrates data across multiple Line Organizations and carries with it a purpose beyond just slapping all the data together in one spot: it conducts real-time inferencing on the data as it is received, and sends out alerts to specified conditions of use to coral researchers and MPA users. The CREWS program ($500K/year) is leveraging significant personnel and hardware resources toward this goal.

Performance Measures:

Milestones

Month 02 – Software purchased and installed at AOML
Month 04 – Meeting convened with LO representatives.
Month 05 – Prototype Web interface developed and early skills developed.
Month 09 – Data integration into central spot at AOML concluded.
Month 09 – Interviews with problem domain specialists and knowledge encoding
Month 10 – Testing and refinement of the system.

Deliverables

- A pool of integrated data of various sources, available via a common interface.
- A Web-based CREIOS/IOOS Decision Support interface
- Configurable Web pages from remote sites to the central data collection pool.
- Automated alerts of coral bleaching and other marine environmental events.

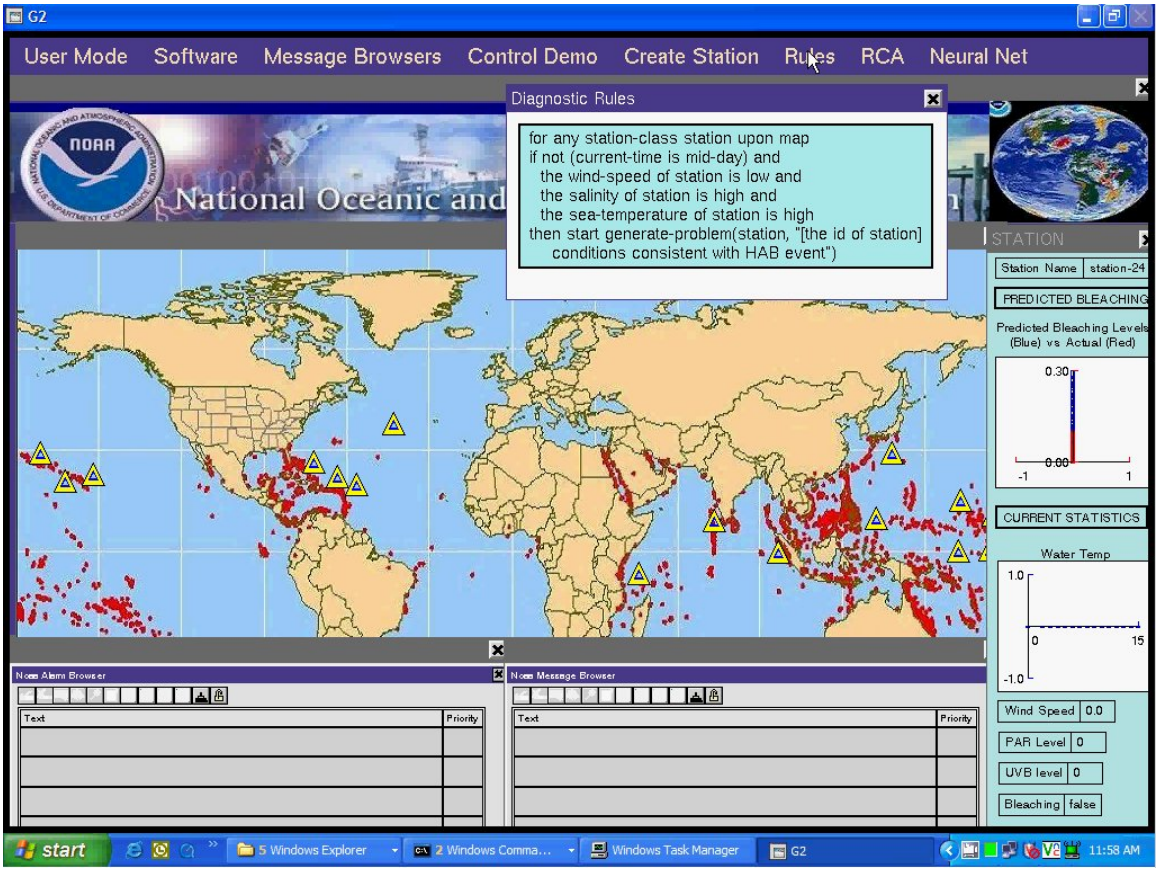

Figure 1. Prototype G2/CREWS interface, developed by Gensym, Inc., in cooperation with the Principal Investigator. The interface is designed to be Web-based and shows all major coral areas (red dots), CREWS Network monitoring stations (yellow triangles), two configurable Alert Browsers (for drifting sensors or other data inconsistencies, and coral bleaching alerts), an example (hypothetical) heuristic/production rule for Harmful Algal Blooms), and two graphing widgets for automated display (when desired) of parameters (e.g., sea temperature) vs. time.

Selected References

Hendee, J.C., and R. Berkelmans. Expert system generated coral bleaching alerts for Myrmidon and Agincourt reefs, Great Barrier Reef, Australia. Proceedings, 9th International Coral Reef Symposium, Bali, Indonesia, October 23-27, 2000. Indonesian Institute of Sciences, 1099-1104 2003

Hendee, J.C., G. Liu, A. Strong, J. Sapper, D. Sasko, and C. Dahgren. Near real-time validation of satellite sea surface temperature products at Rainbow Gardens Reef, Lee Stocking Island, Bahamas. Proceedings, Seventh International Conference on Remote Sensing for Marine and Coastal Environments, Miami, FL, May 20-22, 2002. Veridian Systems Division, CD-ROM, 9 pp. 2002

Hendee, J.C., E. Mueller, C. Humphrey, and T. Moore. A data-driven expert system for producing coral bleaching alerts at Sombrero Reef in the Florida Keys. Bulletin of Marine Science, 69(2):673-684 2001

Hendee, J.C. An environmental information synthesizer for expert systems: A framework for use in near real-time detection of harmful algal blooms. Proceedings, 17th International Conference of The Coastal Society: Coasts at the Millennium, Portland, OR, July 9-12, 2000. The Coastal Society, 233-241 2000