Gulf of Mexico Regional Workshop
on an Integrated Data System for Oceanography

31 October - 2 November 2000
Stennis, MS

Meeting Report

Introductions and Background

The workshop was called to order at 0910 on 31 October 2000, by W. Nowlin. Provisional agendas and participant lists were distributed. The agenda is given in Appendix 1. Each attendee gave a self introduction. Attendees and affiliations are given in Appendix 2.

Nowlin gave an overview of the Virtual Ocean Data System Hub project funded by the National Oceanographic Partnership Program (NOPP). This presentation, given in Appendix 3, included:
(1) The overall goal of the project and specific research objectives identified in the NOPP Broad Area Announcement to which the project was a response;
(2) The organizational structure of the project;
(3) The phases of the program;
(4) Issues to be considered at regional workshops;
(5) The objective, approach, timing, and participants for the national workshop; and
(6) An outline of the implementation expected to emerge and guide the out year proposal.

The Distributed Oceanographic Data System

 D. Holloway (University of Rhode Island) gave a presentation describing the Distributed Oceanographic Data System (DODS), including real time demonstrations of DODS capabilities using the internet. This presentation is summarized in Appendix 4. Included were discussions of
>   What is DODS?
>   Interoperability and metadata
>   What is required to serve data via DODS-enabled servers?
>   What is required to locate and access data?
>   DODS and web browsers
>   DODS-aware applications (MATLAB, Excel, IDL,..)
>   What types of data can/cannot be served?
>   Future of DODS applications/servers

This session continued through the morning of the first day with many questions and comments.

Several general points of agreement emerged from this session. DODS is principally a bottom up data system allowing the user easy data access from within his application provided the data are on a DODS server and the user has knowledge of the URL for the data. DODS does not provide a data catalog; that will be a necessary part of the regional and national data systems. Not all users have need for DODS. As examples, modelers do not normally work in a DODS-enabled application, and many analysts prefer to examine the data and metadata carefully prior to its use in an application. However, DODS can be a very useful element of the required regional and national data models. DODS service of data is to be encouraged, but should not be mandatory.

Another key point of agreement is the necessity for data originators to freely share data if they are to be part of an observing system, regional, national or international. The tendency of oceanographers to treat data as proprietary should cease.

A third point on which there is general agreement is the need for adequate metadata to accompany the data sets. It is particularly important to have adequate use metadata, and to maintain quality flags as are available. This point was revisited later in the workshop.

Presentations Regarding Regional Data and Products

Representatives of each participating regional institution were offered time to make a presentation regarding data holdings, ongoing and planned observational programs, and models. It was suggested that each presenter include indications of data sets which they would be willing to share, whether they could be served via a DODS server, and data sets they would like to receive. Presentations continued during the afternoon of day one and the morning of the second day. Considerable discussion ensued as part of the presentations, which are summarized as Appendices 5-19. Presentations were made by:

John Blaha (NAVOCEANO) speaking on the Northern Gulf Littoral Initiative

John Lever (NAVOCEANO) speaking on the willingness of NAVOCEANO to provide 24 hr by 7 d server capability for U.S. ocean data

J. J. O'Brien (FSU, COAPS) speaking on SEVEER, SeaWinds from QuickScat, and the NCOM under development for the Gulf of Mexico

Nan Walker (LSU Coastal Studies Institute) speaking on real time and archived data from the Earth Scan Laboratory and CSI, BAYWATCH, and WAVCIS

Frank Muller-Karger and Doug Myhre (USF) speaking on fields from SeaWifs and AVHRR satellite data now being served by USF

George Ioup (U. New Orleans) speaking on data sets from Lake Ponchartrain, Barataria Basin, and other littoral waters that could be made available

Jim Corbin (MSU Engineering Research Center) speaking on the Distributed Marine Environmental Forecasting System

Worth Nowlin (TAMU) speaking on data holdings and ongoing projects of Texas A&M University

Norman Guinasso (TAMU, but representing the TX General Land Office) speaking on the Texas Automated Buoy Project

Ruben Solis (Texas Water Development Board) speaking on hydrological and environmental data being monitored in Texas and on models of Texas estuaries run by his agency

Patrick Michaud (TAMU Corpus Christi) speaking on the TCOON and on present and planned CODAR observations along the coastal Gulf

Mark Luther (USF) speaking on the Coastal Ocean Monitoring and Prediction System on the Florida shelf, including the Tampa PORTS, and developments of the USF Center for Ocean Technology Development

Robert Molinari (NOAA/AOML) speaking on archived ocean station, ship-of-opportunity, and drifter data that could be made available, the real time and delay mode data from the NOAA GOOS center, and the Florida Bay project

Tony Amos (Univ. of Texas Marine Science Center) speaking on the many long-term, multidisciplinary data sets available from the Institute and of his real time data and tidal predictions

Susan Starke (NOAA Coastal Data Development Center) speaking on plans of the center to facilitate access to "coastal" data, from the EEZ to 300 km inland. She gave a brief overview of Data Exchange Interfaces, a middleware under development for Navy and intended for use by the Center.

Richard Campanella (Tulane Univ.) speaking on the Long-term Ecosystem Assessment Group

Table 1 indicates the various data sets offered by the presenters for sharing. The descriptions are necessarily brief (perhaps schematic in some cases). These data sets will be more fully described on a web site to be maintained by TAMU for reference of all interested users of regional data sets from the Gulf of Mexico. Tentative commitments were made toward the end of the workshop to which of these data sets would be served initially, by what time and whether it seemed likely that they would be offered via a DODS server. That information also is indicated in Table 1 and will be available on the TAMU web site.

General Issues for DODS Regional Workshops

The workshop was asked to consider five general issues. That was done.

1. Is the DODS data model adequate for data sets to be served in the region? What additions are required?
2. What are important interface issues for regional users? From data discovery to use?

Key to any successful data and product exchange system is the willingness of data originators to openly share data in a timely manner. Participants in the distributed data system for the Gulf of Mexico region must be willing to do so.

It was agreed that DODS is useful as one basis for data sharing in a distributed mode. Of course, it is very useful for users wishing to operate within DODS-enabled application software. However, it is not adequate, or intended, for all regional needs:
• Data location is essential (catalogs are needed);
• Many clients use applications not supported by DODS;
• Large archives now serving data need not necessarily change to DODS mode;
• Distribution of data and products in real time to users (especially public) may be best done via other methods.

Encouragement of new data servers to use the DODS mode is very desirable and will enhance data utilization by many users. It was agreed that Gulf of Mexico regional data providers/users will serve data via DODS servers as feasible. For some applications this will not be the most expeditious or logical method. Open exchange of data and products via the internet is the desired outcome.

Preparation and maintenance of a catalog of information now on DODS servers, as well as on other servers of ocean data, in the region is one essential initial step. (See issue number 5 below.)

A concern expressed by workshop participants was the need for a continuing archive for data and product sets considered to be of long-term value. Many such sets are compiled by individual researchers, and on their retirement or death may no longer be accessible. NAVOCEANO is providing a server for non-classified ocean data. It will be available for national as well regional use. In cases where data sets are used frequently, and especially if they are large, they can reside at NAVOCEANO.

The need to preserve real-time data streams as time series was stressed. Real-time data streams should be further quality controlled in delayed mode and aggregated to produce as complete series of quality data as possible. This is already the situation for many data sets (e.g., drifter, ship-of-opportunity, and Argo data).

DODS offers password protection. However, it was agreed that data confidentiality will be up to the data provider who serves the data.

3. What types of semantic metadata will be required? (Focus of search and use metadata.) What standards will be used?

The workshop considered what metadata of this type will be required. It was agreed that at a minimum the following information must be included:
Where do the data reside?
4-D location information,
Definitions of parameters,
Units (including standards where appropriate),
Accuracy and precision (as available), and
Flags from originators and secondary reviewers.

Very desirable additional information that should be included includes:
Methods,
Platform,
Instrument type, model, band,
Calibration data, and
Reference to algorithms used.

Other useful metadata that is recommended includes:
Supporting descriptions,
Data originator and contact information,
Source of funding for data collection, and
Reports in which data are included.

In discussing desirable formats for metadata, consideration was given to the desirability of having some degree of uniformity and the fact that data served by federal servers must use a standard format for metadata. It was agreed to adopt for the region FGDC. The amount of metadata *required* by FGDC is actually less than agreed necessary by the workshop. Moreover, there appear to be software packages to assist the data originator in preparing metadata in FGDC format.

4. What data sets will be served initially as part of this effort in order to seed the system? What assistance is needed?

The representatives present were queried as to what data sets they would agree to serve initially, by when would the initial data sets be served, whether they would enable a DODS server, and whether they would require assistance to do so.

Shown in Table 1 are the institutes holding Gulf of Mexico data and product sets they are willing to share. As indicated some already are served via web sites. Those marked with the priority "initial" will be served first, by the dates shown and by DODS server if so indicated.

This represents a significant commitment to the establishment of a Gulf of Mexico regional data system, the first step toward a model-based Gulf observing system.

5. Is a regional node needed for coordination, including data location, reference for user support, etc.?

A number of activities were recognized as necessary first steps in the implementation of a Gulf of Mexico regional ocean data system as part of a national system.

- COAPS at FSU agreed to set up and maintain a list serve to enable ease of communication among regional participants.

- NAVOCEANO will procure and maintain a server for distribution of non-classified ocean data, model output, and products via the internet.

- TAMU will set up and maintain a web site indicating data, model output and products that participants are prepared to serve. It will include links to web sites where data are now being served. It will include a catalog of regional data on DODS-enabled servers.

- The formation of several ad hoc groups was agreed to.

   1. A working group on bathymetry and coastlines with the objective of specifying the best currently available data sets for bathymetry and coastline locations for the Gulf of Mexico. Membership will include representatives from COAPS, TAMU (Wm. Bryant's group), NGDC, and NAVOCEANO. (Bryant to chair?)

   2. A working group to consider and recommend the best approach to serving all real time data now being collected for the Gulf by non-federal institutions. This method should allow a user access to the intersection of the data sets. If possible it should be accessible through a DODS server. Membership will include representatives of TCOON, TABS, COMPS, WAVCIS, Florida Bay, NAVOCEANO, BAYWATCH, and URI (Dan Holloway). (Michaud to chair?)

   3. An ad hoc group to consider research needed to evaluate regional satellite and model-derived data fields produced by more than one organization, with a view to identifying characteristics and recommending specific fields for use. Such fields include SST, color, wind, and sea surface height. Membership was not determined, but O'Brien might be willing to chair.

Other General Considerations

It was hoped that industry participants could be entrained into the system. As an example of the potential information that might be shared, Nowlin reviewed data sets and model output held by the oil and gas producers joint industry projects EJIP and CASE. This includes current meter time series, CODAR data, synoptic oceanographic survey data, rather fine resolution circulation model output, and atmospheric storm data and analyses.

It was agreed that Nowlin would send to the workshop participants a copy of the project budget as submitted to NOPP. Those budgets for years after the initial planning year were purposely vague. The reason for distribution is to allow regional participants a say in determining priorities for items to be included in out-year budget requests.

To increase basic data coverage in the Gulf, it was considered desirable that provisions be made for increasing ship-of-opportunity data from both commercial and research vessels regularly traversing this area. It was suggested that such vessels might be equipped with XBT launchers, ADCPs, and Improved Meteorological (IMET) packages.

It was generally agreed desirable to include representatives of Cuban and Mexican institutions in this Gulf of Mexico regional data center.

Table 1. Institutes holding Gulf of Mexico data sets they are willing to share. Some are now available via web site as indicated. Those marked as initial priority will be served first, by dates shown and by DODS server if so indicated.

Institution	Data or Program	Priority	Date of service	DODS
NOAA/AOML	• Florida Bay project • Data from NOAA GOOS Center, available now by web site • Archived drifter and station data, available on request	initial	03/01	yes willing depends on $ "
FSU	• SeaWinds from QuickScat: operational and research products available via web site; 6-hr winds at 0.5 degree resolution from July 1999 • NCOM for northern Gulf	initial	02/01	yes
LSU	• WAVCIS • Earth Scan Lab archived and real time data: AVHRR images for LATEX period • CSI archived data • BAYWATCH program • LATEX B data	initial	03/01 ? ? ?	perhaps yes yes yes yes
MSU	• Offer computational power • DODS server • Investigate link to MEL for service	initial		yes
NDBC	• Buoy data, available now via web site	initial		yes, if support available

NAVO	• New server for U.S. data sets
	• Selected archived data sets:
	  Gridded global DBDBV bathymetry
	  Gridded global T-S data base
	• Selected operational products:
	  MODIS fields for Gulf
	  COAMP for Amer. Med.
	  Global SST
	• Selected 3-dimensional POM output
	• Selected data from Northern Gulf of Mexico Littoral Initiative
		initial
		initial
		
		initial
		initial
		initial
		initial
		initial
			procure
			03/01
			
			03/01
			03/01
			?
			03/01
			03/01
				yes
				yes
				
				yes
				yes
				?
				yes
				?

TAMU	• LATEX A data • Historical daily river discharge from major U.S. rivers • Analyzed wind fields for LATEX period • LATEX C data • NEGOM data • Historical archives of MBT, XBT, ocean station, drifter, moored current meter, and ADCP data • MAMES I & III; CHEMO I & II, available now on request via FTP site	initial initial second second second	03/01 03/01	yes yes

	• NOAA Status and Trends, available now on request via FTP site • EPA EMAP data, available now via web site • National Estuarine Programs data from Galveston and Corpus Christi Bays			
TAMU Corpus Christi	• TCOON data and products, available now via web site • CODAR observations	initial	02-03/01	perhaps yes
Tulane				
Texas General Land Office	• Texas Automated Buoy System data, real-time available now via TAMU web site	initial	03/01	yes
Texas Water Development Board	• TX estuarine hydro. surveys, 1987-97, available now via web site • Sonde estuarine water quality data, available now via web site • TX coastal hydrology records, available now via web site • Model output for TX estuaries	initial initial	early 01 early 01	yes yes
University of Colorado	• Daily fields of sea surface height anomaly • Model average sea surface height for GoM	initial initial		yes yes
University of New Orleans	• Lake Pontchartrain project • Barataria Basin project • Coastal Data, Northern GoM ☺ Acoustics Data, Northern GoM	initial initial available initial	? 5/01 8/01	perhaps perhaps perhaps perhaps
University of South Florida	• Tampa PORTS data, available in real time now via web site • COMPS: archived data • ECOHAB • Archived satellite data (AVHRR, Sea WIFS): Reduced resolution AVHRR since 1993 • Current satellite data streams (Sea WIFS, CZS, AVHRR, MODIS)	initial initial	03/01 03/01	yes yes password for SeaWIFS
USM	• TBD			
University of Texas Marine Science Institute	• Archived ocean station data • Long records of tidal height, surface T & S • Tidal predictions, available via web site • Records of effects of storms on sea level, available via web site • Marine mammal, turtle, and bird stranding data	initial	03/01 03/01	yes yes no

Appendix 1: Agenda

Gulf of Mexico Regional Workshop on an Integrated Data System for Oceanography
31 October - 2 November 2000
Naval Oceanographic Office
Stennis, Mississippi

31 October 2000
Opening and welcoming remarks

Objectives of Gulf Component of an Integrated Ocean Data System
- Enhance data set discovery, sharing, and access through the use of web-based DODS servers and DODS-enabled applications
- Increase the number of Gulf of Mexico data sets available in a readily usable format
    - Historical
    - Real time
- Identify suite of data types of special interest to Gulf workers
- Foster collaborations on Gulf of Mexico region specific issues

DODS as a capability
- Briefing and demonstration of DODS server capabilities by University of Rhode Island personnel
    - What is DODS?
    - Interoperability and metadata
    - What is required to serve data?
    - What is required to locate and access data?
    - DODS and web browsers
    - DODS-aware applications (MATLAB, Excel, IDL,..)
    - What types of data can/cannot be served?
    - Future of DODS applications/servers

Gulf Regional Activities—Ongoing and Planned Activities
Presentations by participants representing institutions.
- Naval Oceanographic Office
    - Northern Gulf Littoral Initiative
    - Web-based server initiative
- Florida State University
    - Special Quikscatt products for the Gulf
    - High-resolution Gulf model using NCOM
- Louisiana State University
    - Earth Scan Laboratory remote sensing capabilities
    - WAVCIS- Wave-current surge information system
    - BAYWATCH: The Vermilion-Cote Blanche Bay Physical Measurements Program
- Mississippi State University
    - Distributed Marine-Environment Forecast System
- Texas A&M University
    - MMS-sponsored activities and data archive
    - Modeling activities
- Texas General Land Office
    - Texas Automated Buoy System
- Texas Water Development Board
    - Bays and Estuaries Program - data collection and dissemination
    - Oil spill and other modeling activities
- University of South Florida
    - Remote sensing capabilities

Tampa PORTS
Coastal Ocean Measurement and Prediction System
- AOML
- University of South Florida
- University of Alabama
- Eddy Joint Industry Project
- University of Texas, Marine Science Institute
- TAMU Corpus Christi
- Tulane
- University of Southern Mississippi
- University of New Orleans
- Minerals Management Service
- National Coastal Data Development Center

1 November 2000
Gulf Regional Activities (Continued)

Regional Partnerships
- Motivation for becoming a partner
- NOPP-DODS support for regional partners
- Summary of datasets to be shared (which are of interest; which do you have to share?)
- General versus region-specific datasets
  Selection of data types of special regional interest (perhaps, fresh water distributions or measures of hypoxia) and what metadata are needed
- Interests in cooperative studies

DODS Questions
Having considered data sets that are available and might be served, this is an opportunity to return to specifics of DODS. Issues such as the following may need further consideration:
Central/distributed servers (resources)
Data quality control
Metadata
Recognition for data contributions
Security
Sub sampling and bandwidth for large datasets
Catalog and data discovery
Continued data availability

2 November 2000
Decisions regarding future activities
- Assign DODS server installations
  Some support is available via the NOPP project.
- Assign data sets to be served as practical demonstration of regional interest and participation
- Regional organization needed for future development of an integrated regional-national international data system for oceanography
- Needs for regional activities (e.g., web-based data information center giving pointers to data sets, or assistance with server installations)
- Select representatives to national workshop

Open Discussion and Wrap up

Adjourn Meeting

Appendix 2: Meeting Attendees and Affiliations

Tony Amos   University of Texas Marine Science Institute
Landry Bernard   NAVOCEANO
John Blaha   NAVOCEANO
Jim Bonner   TAMU-CC/TEES
Jim Braud   NAVOCEANO
Richard Campanella   Tulane/Xavier Center for Bioenvironmental Research
Jim Corbin   MSU ERC/CCS
Steve Foster   MSU ERC/IDSL
Jim Fritz   TPMC
Mike Garcia   SAIC/NDBC
Norman Guinasso   GERG/Texas A&M University
Martha Head   NAVOCEANO
Dan Holloway   University of Rhode Island
Matthew Howard   Texas A&M University
Stephan Howden   University of Southern Mississippi
George Ioup   University of New Orleans, Stennis
Peter Lessing   NDBC
John Lever   NAVOCEANO
Alexis Lugo-Fernandez   Minerals Management Service
Mark Luther   University of South Florida
Melanie Magee   Gulf of Mexico Program
Robert "Buzz" Martin   Texas General Land Office
Eugene Meier   Gulf of Mexico Program
Patrick Michaud   TAMU-CC/CBI
Bob Molinari   AOML/NOAA
Steven Morey   COAPS/Florida State University
Frank Muller-Karger   University of South Florida
Doug Myhre   University of South Florida
Worth Nowlin   Texas A & M University - NAVOCEANO
Jim O'Brien   COAPS/Florida State University
George Rey   LEAG/CBR
Reyna Sabina   AOML/NOAA
Mitch Shank   NAVOCEANO
Ruben Solis   Texas Water Development Board
Susan Starke   NCDDC/NOAA
Vembu Subramanian   University of South Florida
Molly Sullivan   Tulane University
Jack Tamul   NAVOCEANO
William Teague   NRL
Nan Walker   Louisiana State University
Patti Walker   DATASTAR/NDBC

Appendix 3: Introduction to NOPP-sponsored National Data Hub project by Worth Nowlin

Project Goal and Objectives

Goal

Plan and implement a "Virtual Ocean Data Hub (VODHub) system that will provide users with the ability to "easily access certain data types in specified locations/times, regardless of data source, and without special efforts or insights on the part of the user about the data source(s)".

Specific research objectives identified in the BAA are:

1. Identify individuals/organizations that will take the lead to foster development of community-based conventions for specific data types.
2. Partner with local institutions (public and private) to improve access to coastal and regional data via community-based conventions.
3. Enhance connections to existing national and international archives of ocean data as well as the program offices of major ocean programs (e.g., WOCE, JGOFS) via developed community based conventions.
4. Partner with international groups to foster a world-wide "ocean data dictionary" to further commonality of access for all sources of ocean data.
5. Work with national and international user standards groups (e.g., International Hydrographic Organization and their Electronic Navigation Charts) to foster access to ocean data via a growing number of user interfaces.

Partnership Structure
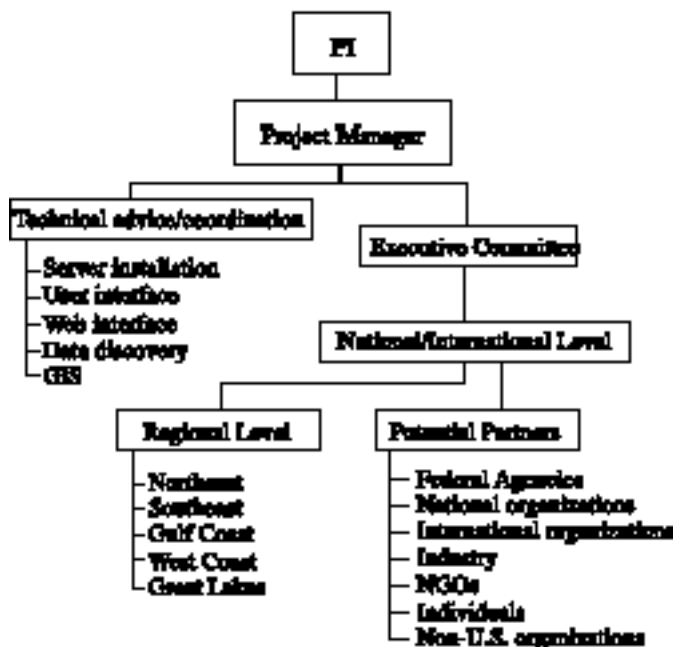
Phases

Year 1: The Planning Phase

* Regional Workshops (5)

* Synthesis of regional workshop results

* National Workshop

* Prepare Final Recommendations

Years 2 & 3: Implementation

Regional Workshops

Asked to consider the following issues:

1. Is DODS data model adequate for datasets to be served in region? What additions are required?

2. What are important interface issues for regional users? From data discovery to use.

3. What types of semantic metadata will be required? What standards will be used? Focus on search and use metadata.

4. What datasets will be served initially as part of this effort? Seed the system? What assistance is needed?

5. Is a regional node needed for coordination? Data location, reference for user support, etc.

National Workshop

Background
Synthesize results of regional workshops. Study by Executive Committee.

Objective
Develop the implementation plan for years 2 —>.

Approach
Address same issues as at regional workshops, with synthesis as background.

Timing
Nine months after project began

Participants
Project management
Regional coordinators
Representative from federal agencies with holdings, e.g., NAVO, NODC
Representative from national organizations of marine institutions, e.g., NAML, Sea Grant Program
Representative from international organizations, e.g., Australian BMRC
Representative from industry with significant data holdings
Experts with data systems/networks

Representative from other NOPP projects developing model/assimilation capabilities

Review
    Workshop recommendations to be reviewed by community

Implementation Plan

Three principal areas will be:

Population

Population refers to addition of data to system. Focus on in situ data sets — particularly regional.

System Core

Areas of expansion considered necessary are:
- Development of GIS clients
- Work with virtual data sets
- Web interfaces
- Improve Data Access Protocol

System Maintenance

Including system documentation and user support

Appendix 4: Dan Holloway – University of Rhode Island

(http://po.gso.uri.edu/~dan/dods-regional-workshops/dods-regional-workshops.html)

The URL listed above is a link to the presentation given by the DODS group at the Gulf of Mexico NOPP regional workshop.

Following is a summary of a number of important points discussed during the presentation.

Introduction to DODS

- DODS is an open source software project designed to allow users to easily access and move data over the network.

- The DODS project has two underlying principles it adheres to:

- Anyone willing to share data should be able to do so via DODS; .e.g., the scientist, a state agency, a private company, or a federal data center.

- Users should be able to use the application package with which they are most familiar to examine or analyze the data of interest.

- How these principles are incorporated into the design are as follows:

- Data providers must not be required to store their data in any special format.

- The data system must require a minimum set of metadata from the data provider.

- The DODS core software must be easy to interface to existing applications so that scientists can use the packages they are most familiar with.

- The data system must provide all the metadata that is required to effectively use the data in the client application.

The requirement to provide the minimal set of metadata required to use the data directly competes against the requirement to not burden the data provider with meeting specific metadata requirements.

The DODS project has taken a bottom-up approach toward solving the distributed data access problem. Rather than focus on the directory level, or data location aspect of the distributed data access problem, DODS has focused on the data-level interoperability. As part of that effort, the DODS project has delimited the metadata requirements for distributed data access into 'use' and 'search' metadata. To achieve data-level interoperability a definition of 'use' metadata was formulated by segregating the data-level metadata requirements into syntactic and semantic 'use' metadata. The DODS core software provides a strict syntactic 'use' metadata representation in the data transmission component of the software. This is required in order for software components on the client-side to be able to decipher the encoded binary data stream. The DODS core software supports but does not require additional semantic 'use' metadata from the data provider. However, as the level of metadata increases, both 'use' metadata and 'search' metadata, the degree of interoperability that can be attained with remote datasets increases. It is important to note that all of the more advanced client applications using DODS to access remote data have specific semantic 'use' metadata requirements, but none of these go so far as to require FGDC CSDGM metadata fields.

During the presentation two DODS client applications were demonstrated.

 - NOAA's Live Access Server

The Live Access Server (LAS) is a web-server application that provides an interface to data stored at NOAA's Pacific Marine Environmental Lab in Seattle, and a number of DODS served datasets located at NOAA's Climate Diagnostic Center in Boulder, and the International Research Institute at Columbia's Lamont-Doherty Earth Observatory.

This web-server uses a DODS-enabled client application, Ferret, to retrieve data from these remote sites and process it into one of the available representations defined in the interface. Ferret is a notable example of a scientific application whose functionality has been extended by enabling it to access remote datasets via DODS.

The strength of this interface is that it provides a standardized interface to a relatively large number of gridded datasets. Additionally, it can be easily customized for use at other sites.

The weakness of this approach is that it predefines the range of analysis possible on the remote datasets. The interface does permit the data to be easily downloaded, but the data must then be ingested into the scientist's application for further study.

 - DODS Matlab GUI

The Matlab GUI was developed as a testbed application to better understand the problems associated with building applications based on distributed data access. The primary result of this effort has been to better understand the semantic 'use' metadata requirements for data-level interoperability. The Matlab GUI has limited, though strict semantic 'use' metadata requirements.

The Matlab GUI provides an interface to two sets of oceanographic data, global datasets providing Sea Surface Temperature, Winds, etc., and local in-situ datasets for the Gulf of Maine GLOBEC project. These two sets present different problems to building applications using DODS.

The strength of this interface is that it provides direct access to the data in the scientist's analysis application. Once the GUI has selected and retrieved the remote data, the data is now located in the scientist's workspace for further analysis, or saving to local files, no additional steps are required.

The weakness of this interface is that it can require a high-level of effort to add new datasets to the interface. However, this constraint has led to better understanding of the various aspects of 'use' level metadata requirements.

DODS current development activities.

1. The DODS core software has been ported to Java.
2. A JDBC server has been built and is currently in beta test at OSU and MBARI.
3. A GrADS server has been built using the Java port by the COLA group. This server can provide access to GRIB data via DODS. Additionally, this server can be used to process data remotely, returning the results of those operations as DODS datasets.
4. There is a native Windows port of the DODS core, as well as a Cygwin port of the software for PCs.
5. There is a WMT-DODS gateway prototype in development at NASA's Goddard Space Flight Center.
6. The DODS Matlab GUI is being rewritten to facilitate adding new and n-dimensional datasets, and to allow the user to easily customize the interface for their use.

DODS future development activities.

- The project is actively pursuing migrating the core's data access protocol(DAP), to use XML as an encoding scheme.
- DODS is working with NGDC and ESRI to make DODS data accessible to GIS clients, and GIS data accessible to DODS clients.
- The project is investigating designs to support tertiary metadata servers, as well as client-side metadata support, to permit interoperability for datasets with limited associated metadata.
- A web-crawler will be designed and implemented as part of the NOPP project to assist data location.

Comments on DODS

The DODS project has taken a bottom-up approach in proposing solutions to the distributed data access problem. At it's current stage of development the project is focused on data-level interoperability in that problem domain. Goals of the project have been to permit scientists to easily serve, and access data over the internet. To accomplish this the software must not require any reformatting of data by providers, and permit direct access to data within the scientist's existing analysis environment. To reach these goals the project has selected several client applications to demonstrate the concept, Matlab, idl, and netCDF. The project also identified a number of commonly used data formats, as well as packages which support access to PI help datasets. Arguably, pre-built servers for every data format used in the science community are not currently provided by the DODS project, but as has been demonstrated by various groups writing DODS servers can be a realistic alternative.

At it's current stage of development DODS does provide direct access to data within a number of scientific analysis environments. Unfortunately, given the lack of standardized interfaces for searching for data within the community at large, or within specific data repositories, formulating DODS URLs to access datasets can be difficult. One goal of an integrated data system should be to work toward standardized mechanisms for accessing distributed data, which includes both ease of location and use of those data.