SIOExplorer: Digital Library Project

Stephen P. Miller, Geological Data Center, SIO

John Helly, San Diego Supercomputer Center

Anthony Koppers, IGPP, SIO

Peter Brueggeman, Dir., SIO Library

University of California, San Diego
9500 Gilman Dr.
La Jolla, California 92093-0220 USA

Abstract-The newly-launched SIO Ocean Exploration Center is a joint effort between oceanographers, librarians and computer scientists to integrate access to data, imagery and publications, using modern database technologies and metadata standards.

I. INTRODUCTION

We are faced with a number of real challenges, and of course opportunities, with the flood of data coming from today's oceanographic sensors, and the backlog of valuable information housed in local archives of varying condition. Modern search engine tools, metadata standards and storage technologies can be harnessed to solve some of the problems, but careful collaboration will be required to insure that the results will have an impact on a wider set of users, especially for educational purposes. To tackle this problem we have brought together a working group from three organizations: the Scripps Institution of Oceanography (SIO), the UCSD Library which includes the SIO Library, and the San Diego Supercomputer Center (SDSC).

The group has chosen to focus its efforts on **"bridging the gap between libraries and data archives."** Libraries make published scientific information and historical archives available to a wide range of users, including both scientists and the general public. However, the raw material on which scientific literature is based is normally unavailable in traditional libraries. Those data are found instead in scientific data archives, which are generally quite separate from libraries and focus on meeting the research interests of a narrow group of expert users. Data archives tend to be difficult to search, poorly documented, lacking in long-term sustainability, and frequently without an intuitive interface for the non-expert user.

We are taking advantage of recent developments in Information Management technology, such as the portable **Storage Resource Broker** software system of the SDSC, to pose long-term, scalable, reliable and cost-effective solutions to these problems. Currently, our primary digital data holdings are 440GB in size, growing at a rate of about 100 GB per year. We will develop modular approaches that will allow domain-specific metadata to be embedded into a broader metadata catalog. Reusable general tools will perform geospatial and expert-level searching, calling upon modularized domain-specific tools to deal with the special complexities of multibeam seafloor mapping, for example.

Our central theme is an Ocean Exploration Center. We will display a global map of the thousands of ship tracks of oceanographic cruises in our holdings. Clicking on particular ship tracks or outlining areas of the ocean map will lead the visitor to scanned images of ship logs, descriptions of the early voyages, or continuously updated maps of the oceans, data on climate, oceanography or marine geology and geophysics (Fig 1.)

Figure 1. Roger Revelle, Gulf of California Expedition, 1939, on the new *R/V E.W. Scripps*.

We are developing interactive procedures to extract information from our global and local databases and create custom maps on demand, not just the display of pre-existing images. Our goal is to stimulate inquiry-driven discovery, which is important for the researcher as well as the student. We will also make a special effort to capture key historic documents associated with Scripps' voyages of discovery and place them online, along with the cruise data. The SIO Archives within the SIO Library record the human side of oceanography, with a collection of photographs and documents going back to 1904.

II. DIGITAL LIBRARY OVERVIEW

The 1998 NSF Workshop on the SMETE Library (for Science, Mathematics, Engineering and Technology Education) articulated a vision of "a new type of library that would provide a comprehensive collection of digital resources and services that are available for undergraduate education." A primary goal of the National SMETE Digital Library (NSDL) is to provide integrated and effective access to a wide range of materials, including but going well beyond the types of materials found in a conventional library. The NSDL is designed to include not only the publications found in a traditional academic library collection but also the tools and resources that provide for interactive learning and data sets for analyses.

Currently these materials are dispersed. Research publications are found in libraries, along with reference materials, maps, manuscripts and photographs, all of which are organized and preserved for generations. Visitors to libraries normally see only the end results of scientific methods. Visitors to historical and scientific archives, on the other hand, have access to the data on which discoveries are based. These collections are a rich source of materials that can communicate the excitement of scientific discovery. However, archives are generally designed only for highly trained users. While they support research and instruction for a specialized user group, they are often difficult to search, lacking in long-term sustainability, and frequently without an intuitive interface for the non-expert user. The result is that the typical user community is small and restricted to research specialists.

Fortunately, the advent of digital library technology allows us to unite these disparate entities into a single resource that can be accessed by a wide audience. A digital library now can go far beyond the traditional library to provide direct, immediate access to scientific literature and to a phenomenal range of historical and data archives, creating an organized body of knowledge accessible to all learners. Using digital libraries to bridge the gap between published literature, historical archives and scientific archives is of interest to a broad community. Academic research libraries are redefining their role in an age of digital libraries. At the same time, managers of scientific data archives are constantly finding ways to organize and maintain an ever-increasing amount of data. The merging of library collections with historical and data archives offers the promise of expanding opportunities for learning, and at the same time insures the stability of the historical scientific record.

The Report of the President's Panel for Ocean Exploration [1] proposes a new era of ocean exploration, recalling the earlier expeditions of discovery, and makes specific recommendations on the required database infrastructure:

"An essential element of this program is data management and dissemination. The database must incorporate a variety of data types ... and use widely accepted standards ... Flexibility, however, is the key to integrating a broad range of data ... The Program should incorporate older, historic data sets relevant to exploration into the new database. ... so that discoveries can have the maximum impact in the research, commercial, regulatory and educational realms."

It is our intent to provide a general model for the integration of such materials in other disciplines, such as medicine, chemistry, and engineering. The ability to link scientific publications with historical archives and data archives will be useful for researchers, educators, students, policy makers, regulatory agencies and the public. The key steps to our solution involve:

- **Collaboration** between scientists, librarians, and archivists in the creation of a digital library.
- **Specification of metadata** to link scientific data, scientific publications and historical archival materials through a relational database.

  A modular approach will allow domain-specific metadata to be embedded into a broader metadata catalog that can retrieve digital materials from disparate sources. We will define a metadata interchange to act as the link between the libraries and the archives and to facilitate the sharing of data across a future federation of libraries and archives.
- Development of a **search and delivery interface** that will present materials from data archives and libraries in an integrated format that promotes discovery and learning.

  We will develop interactive procedures to extract information from our global and local databases and create custom displays on demand, not just display pre-existing static images. Reusable general tools will perform geospatial and expert-level searching, calling upon modularized domain-specific tools to deal with the special complexities of multibeam seafloor mapping, for example. The results will

address total "end-to-end" data management and long-term stewardship requirements. This development process will clearly require step-wise phases of concept formulation, prototype design and evaluation.

In the text to follow we present our strategy for "Bridging the Gap between Libraries and Data Archives." We begin by describing a proposed user-interface to our digital library – the Ocean Exploration Center. A brief overview of the underlying system architecture is provided. The digital library will integrate a number of diverse collections, from the SIO Library, the its historical SIO Archives, the GDC shipboard data archives, and collections residing in an on-line database, EarthRef.org, whose development has been spearheaded by researchers at SIO. We describe these collections and how they will contribute to our digital library. Construction of the digital library also requires that we develop the tools to access, catalog and manage our collections.

III. ACCESS AND DISCOVERY OPERATIONS

The **Ocean Exploration Center** will support inquiry into the history of ocean exploration, as well as today's state-of-the-art databases. As a centerpiece, we will display a global map of all available SIO and other expeditions that cover the oceans. The current tracks of the SIO vessels will be shown, updated every few hours. Selecting particular ship tracks, or outlining areas of the map, will lead the visitor to descriptions of voyages, literature, maps of the depth of the ocean, age of the oceanic crust, sediment thickness, and other oceanographic data. The OEC will allow users to access content which has been classified in one of probably four expert levels, such as an entry level (K-6), student level (grade 6-12), college, and research level. In such an expert-level access approach, users will be able to choose an entry level and advance to state-of-the art science, including original scientific papers or data. This advance is not limited by the age or the grade of a particular student, but rather by their curiosity and the excitement of discovery.

The focus will be to provide a comprehensive view of the oceans, intelligible to a non-specialist, with direct access to state-of-the-art databases and selected websites. The OEC will serve as an almanac and atlas of the oceans and the ocean floor. For this, we will use the concept of an *information access console* (Fig. 2). This console will be arranged into different sections, including navigation and

geology/geophysics, and ultimately more sections on other aspects of oceanography, text, photographs, and other information.
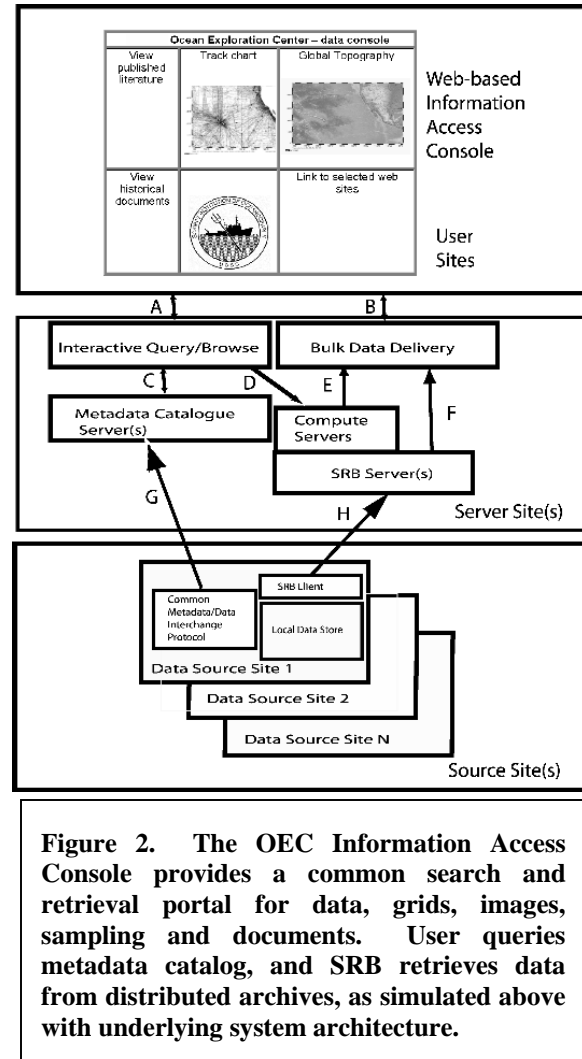
Figure 2. The OEC Information Access Console provides a common search and retrieval portal for data, grids, images, sampling and documents. User queries metadata catalog, and SRB retrieves data from distributed archives, as simulated above with underlying system architecture.

We will make sure that different types of data are more easily accessed, and viewed simultaneously next to each other. The geological/geophysical section of the console will contain maps of the ocean floor, the age of the ocean crust, the sediment thickness, locations of earthquakes, volcanoes, tectonic plate boundaries and other information. There will be chemical or physical data of geological samples taken at particular locations.

IV. TECHNICAL APPROACH: INFORMATION ARCHITECTURE

The flow of information among users, metadata catalogs, servers and data source sites will be managed by the flexible client/server software of the SDSC Storage Resource Broker (SRB), (Fig. 2). The

relationship with user communities is defined by a **conceptual model**, which will be the basis for the user-interface design for access and discovery. Conceptually, metadata, data and derived products will be organized into classes that are oriented to the skill level of the selectable user class.

The **logical model** is the basis for the design of the metadata catalogue and will determine the type of queries that can be applied to the digital library, as discussed more fully in section VII. A research expert working with raw data will need access to a different family of queries, compared to a K-6 student exploring a thematic map of a region. Building on our experience with other projects at the SDSC, we will use **arbitrary digital objects** (ADOs) to organize data into distinctly identifiable objects. As data and relationships evolve, it has proven to be very useful to maintain the traceability of the data to its source, to accommodate the versioning of ADOs, and their association with prior versions. Our multi-level ADO naming convention begins with an original persistent name, and adds other names as needed.

The **physical model** defines collection management. It governs access methods and the relationship of the distributed data sources to the data repository at SDSC as well as how the resources held at SDSC will be provided to other digital library services such as DLESE. Physically, data will be stored in files in computer systems and metadata will be contained within some type of searchable catalogue system typified by a database management system but not limited to it. Splitting the data and metadata systems into separate systems enables **flexibility** in media migration, arbitrary and appropriate selection of client and server platforms, reduces single-point failures and encourages proliferation of participating metadata and data server sites. Media migration is a problem all computer systems face and it is particularly acute for server sites with long-term storage requirements.

V. THE COLLECTION

A. Historical Content from SIO Archives
The SIO Archives record the human endeavor in oceanography, with a collection of documents and photographs of work at sea going back to 1904. By following the threads of careers, project documents and publications, we will show how expeditions have been a fundamentally collaborative, interdisciplinary, and evolving process. These historical archives reveal a wide range of information that can make scientific data archives "come alive" for scientists

and non-scientists alike. When these materials are combined in a database that also allows users to view seafloor topography and associated data underlying ship tracks, educators will have a resource that will easily communicate the excitement of ocean science to a broad audience. A hint of the content of the full SIO Archives (Fig. 3), housed in the SIO Library, may be seen at http://scilib.ucsd.edu/sio/archives/. The SIO Archives is a participant in the California Digital Library's Online Archive of California, and has contributed numerous encoded finding aids to that resource.
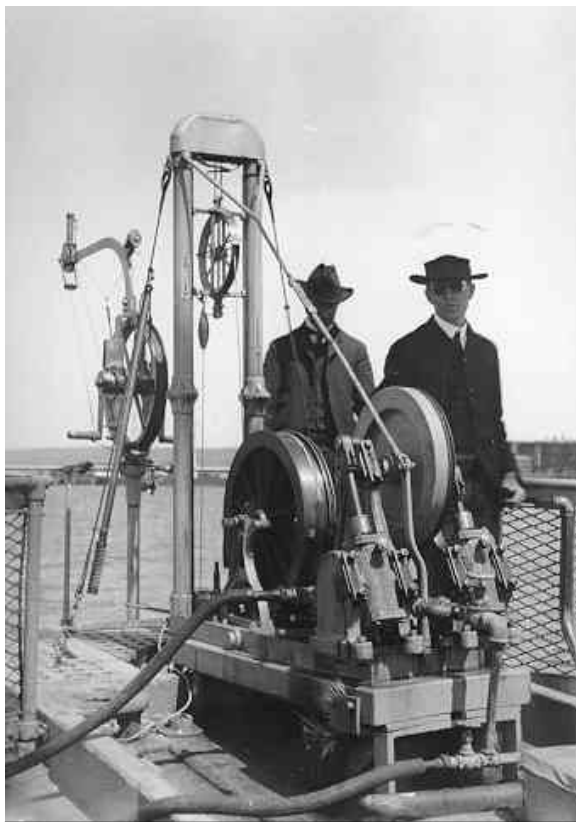
Figure 3. Sigsbee Sounding Machine, R/V Albatross, March, 1904.

The Geological Data Center (GDC) Archives date back to the 1950's, and data gathering activities at sea have certainly evolved. As data continue to age, and scientists retire or pass away, researchers and students scientists will need to recover data and companion information, including processing steps and techniques. We are fortunate in having an intact series of standard GDC Expedition Reports for each cruise leg, going back about 30 years, which include the participant list, a scientific overview, track charts, and plots of underway data profiles. Some of the older reports will need to be scanned, but this series

of documents will become a very practical starting place to follow the threads of GDC data, SIO Archive documents, and the literature, linked by our database through time and space.

The Gulf of California 1939 Expedition was the first comprehensive hydrographic survey of the Gulf, using the newly acquired ship, *E.W. Scripps* (Fig.1). "Historically Gulf of California 1939 was a forerunner of the Revelle Era of Pacific Exploration" (Walter Munk, pers. comm., 2001). The MidPac Expedition of 1950 was the first great SIO postwar oceanographic expedition, a 25,000-mile voyage through the central Pacific to the Marshall Islands. In the words of Jerry Winterer (pers. comm., 2001),

"Without question, I regard the MidPac Expedition to be the most important, in the sense that its findings completely changed our world view. The dredge haul from guyot summits contained shallow-water reef fossils of Early Cretaceous age, thus killing the prevailing theory of the permanence of continents and ocean basins, and establishing the youthfulness of the oceans. Youth demands mobility, and thus the door to plate tectonics was opened. A revolution, like that of Copernicus."

The Capricorn Expedition of 1952-53 covered wide reaches of the South Pacific and included a number of cultural contacts between scientists and the local island communities. We are particularly intrigued with the members of the scientific party, all of whom went on to become stars in the field of oceanography over the next half century. The Pascua Expedition to Easter Island in 1983 with the new SeaBeam swath mapping sonar system rode down the axis of the East Pacific Rise for thousands of kilometers. When the axis mysteriously died out at 9° N, the plan was altered to search for it, thereby discovering the first "Overlapping Spreading Center", which led to a host of other expeditions and publications in marine geophysics, geology, geochemistry, and biology.

B. Publications from the SIO Library
In order for our collection to be a source of high quality information suitable for scientific teaching and research, published journal articles must be included. Bibliographies of relevant materials are essential, but use of the literature will be far greater if the actual text is available online. We will locate published materials related to cruises on our web site, seek publisher permission, and then create metadata and images of the content so that site visitors have access to them.

C Multibeam and other mapping-related content from the GDC
The foundation of our unique collection will be the multibeam bathymetry data collected and archived in the SIO Geological Data Center (Fig 4). In 1981 the new SeaBeam swath mapping sonar on Scripps' R/V Thomas Washington marked the beginning of a new era in US academic exploration of the seafloor. The new Simrad EM120 on the R/V Roger Revelle routinely maps a swath 25 km wide in deep water. Furthermore, the new systems return high resolution backscatter images of the seabed, which allows scientists to detect faults and map contrasts in bottom type, such as the extent of lava or debris flows, manganese deposits, and hydrocarbon seeps. SIO currently operates two deep water systems that will generate approximately 100 GB per year of new data, to be added to the existing archive holdings of more than 400 GB. In addition to multibeam data, we will access the GDC underway holdings of 646 cruise legs since 1952, as well as nearly 2400 additional cruise legs of data that have been assembled as part of the Global Topography model [2; http://topex.ucsd.edu/marine_topo/mar_topo.html].

D Sampling and Modeling from EarthRef.org
Additional supporting materials will come from an expansion of the collections now residing at EarthRef.org. The core of this website is an online relational database that provides available geological and geochemical information on the world's major chemical reservoirs, e.g. core, mantle, crust. Within the Earth science community this website has become well utilized, with an increasing number of scientists downloading data from it (~5000 hits per year), for both research and teaching. Data files may be archived as strictly peer-reviewed contents, or as an open exchange of useful resources shared between scientists, teachers, and students. The digital archives are designed to be a medium for the exchange and access to a rich and diverse range of geo-information. It has a working data upload capability allowing registered users to add content.

The types of data archived in EarthRef.org span a wide range from individual data points (e.g. chemical analyses, age determinations) to any other arbitrary digital objects. For Earth scientists and teachers we developed the EarthRef Digital Archive (ERDA) that includes any "digital" contribution ranging from data tables to diagrams to reports to geological maps to videos. In the same category we have developed a **Seamount Catalog** containing bathymetry maps, dredge locations and data on seamount morphology (http://EarthRef.org/databases.htm).

VI. ACCESS AND DISCOVERY METHODS:
BUILDING ON CURRENT EXPERIENCE

These existing EarthRef.org online databases are designed in Oracle and they are maintained by Anthony Koppers (SIO) at the SDSC in consultation with John Helly (SDSC) with a flexible architecture that allows for the archiving, searching and downloading of any type of Earth science digital data [3]. We are linking these data sources directly with on-line modeling tools that allow for a variety of calculations, such as the TnT2000 global geochemistry toolbox [4,5; http://EarthRef.org/tools/tnt.htm]. The existing EarthRef.org Seamount Catalog system, created by Anthony Koppers, provides an excellent model for developing our more general bathymetric interface. The content creation for the Seamount Catalog begins with multibeam swath mapping sonar data, using the tools of the MB-System [6,7].

To generate complete map coverage, since multibeam surveys often have gaps, the approach carefully embeds the detailed multibeam data, at typically 100m lateral resolution, into the surrounding Global Topography data, which has 2-minute latitude, longitude resolution. The user can view or download map images, search metadata content, and download gridded bathymetry or the original multibeam files. We will expand the current database content and the metadata catalog to include other seagoing-related data archived in the GDC and the SIO Archives, as well as the publications in the Library, linked through metadata such as latitude, longitude, date, and keywords, and managed with the SRB.

We will leverage this approach and apply it to several other important global databases from SIO, thereby extending the benefits of this development. We will explore databases on the continental crust and the age of the ocean floor (http://omphacite.es.su.oz.au/StaffProfiles/Dietmar/Agegrid/agegrid.html). Our IGPP/SIO colleagues Gabi Laske and Guy Masters can provide global sediment thickness (http://mahi.ucsd.edu/Gabi/sediment.html), and multilayer crustal thickness models of velocity and density (http://mahi.ucsd.edu/Gabi/crust.html). We will also access global measured heat flow data, and the present day Earth's magnetic field. A great deal of effort has been made by the research community to compile these global data sets and models. Each model has its own resolution and format. Our contribution will be to provide a common information access console to select geographical areas from these databases, and then display and compare the results at a common scale.
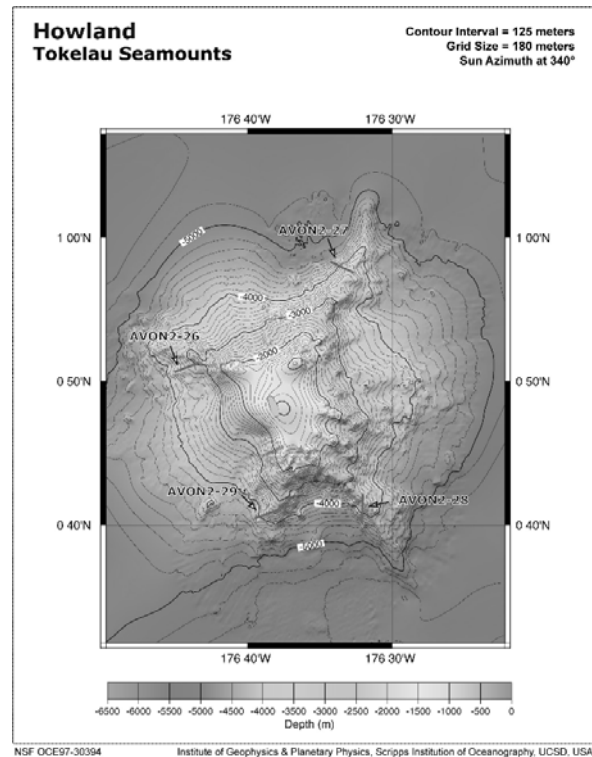
Figure 4. Example of multibeam bathymetry (seafloor depth) map with dredge locations, as created by created fromEarthRef.org database.

VII. CATALOGING THE COLLECTION:
METADATA

At the SDSC there is considerable experience in developing systems that create and maintain metadata [8,9,10]. Metadata are distinct from the data holdings themselves and will be designed to be **interoperable across digital library catalogues** by building the necessary import/export tools for converting metadata into a flat, ASCII format suitable for parsing by import tools. We are proposing an architectural approach that splits the digital library function from the storage of data, thus enabling the replication of metadata sites for ease-of-access and enhanced reliability, maintainability and availability on an arbitrary number of hosts. This architecture is scalable in both the metadata and data domains and permits the establishment of metadata catalogues without regard to how or where data and associated computing is located.

The quest for metadata standards and tools is a topic of current activity in a diverse set of communities, and collaboration will be the key to interoperability. While we will follow the XML methods currently enjoying widespread pursuit, we will also develop a **metadata interchange format** following the approach developed for the CEED (Caveat Emptor Ecological Data) Repository [9,10; http://ceed.sdsc.edu]. This is essentially a flat ASCII format in which each record in the file is defined by three fields, 1) a human interpretable, meaningful keyword, 2) a value for the field corresponding to the keyword, and 3) a label to be used in applications displaying the metadata. This flexibility will ensure the preservation of the invaluable metadata content in a highly portable, self-describing format, and will enable easy ingestion by other metadata systems whether XML-based or not. Readers are invited to test drive the CEED site and examine its functionality, which includes geospatial search, journal archive access and the uploading of contributed data.

Fortunately, MB-System provides us a good start on an automatic method for creating and updating metadata content for our multibeam holdings. The command mbinfo returns an ASCII file with about 40 useful attributes, such as the latitude, longitude, depth and time bounds of a data file, as well as quality attributes, such as the number of good or bad soundings. We will create streamlined procedures to harvest this information from the data and insert it into our general metadata catalog, which will be built to interface with library and DLESE metadata search standards.

Data and metadata will be supplied by participating Data Source Sites, initially the GDC, the SIO Library, and IGPP. These sites will publish their data, on their own schedules, to a Storage Resource Broker (SRB) server site (see section VIII). The first of these will be at SDSC and eventually at federated sites that will act as regional centers focusing on data within their region or as mirror sites established to facilitate reliable and speedy access by the user community. The interposing of SRB server sites between Data Source Sites and users will relieve the research laboratories of the burden of direct support of a user community beyond their resources while permitting those that care to participate at that level of commitment to do so by installing a server at their location. It enables the leveraging of the NSF investment in Supercomputer centers as well as providing an open architecture that encourages active participation by both the observational research and the information technology communities.

Metadata technology will be the key to bridging the gap between the library and our data archives. We will embed the domain-specific metadata catalog into a more general catalog, conforming to the digital library standards of DLESE and other organizations. Access tools will allow a combination of keyword, geospatial, temporal, and expert-level searching of the catalog and retrieval of data, imagery or text.

The metadata for library and historical archive materials will include latitude and longitude, as well as place names, so that geospatial searches will discover both library and data. Some entries may be defined by a single point, and others will require a polygon. We will be guided as well by the developing metadata standards of the NSDL, particularly those of DLESE and the Core Integration Systems projects, as well as by established library metadata standards such as AACR2 and Dublin Core and developing metadata standards in the Earth sciences community such as the ESML (Earth System Markup Language).

Specifically, we will coordinate our ongoing metadata design efforts with DLESE, which has adopted a strategy to apply "required" metadata to all resources in their searchable collection (http://www.dlese.org/Metadata). These include approximately ten information fields that enable searching/browsing functions in the discovery system. We anticipate also adopting such core metadata. DLESE is actively working to develop additional extensions to their metadata collection including a) geospatial footprints b) elements of the Earth system (processes and concepts rather than disciplinary keywords), and c) mapping onto the National Science Education Standards for K-12 education. We will both contribute to and use these new metadata as our own project proceeds. A necessary part of our project will be the development of additional metadata extensions to cover the special fields of expertise encompassed by our project. Feedback of these metadata into DLESE will be mutually beneficial.

VIII. COLLECTION MANAGEMENT: THE STORAGE RESOURCE BROKER

To manage our collections, we will call upon the Storage Resource Broker (SRB), a product of the San Diego Supercomputer Center. The SRB is public domain software, can be loaded on most Unix workstations, browsed on PCs, and is currently being used at approximately 150 sites. Our proposed digital library information architecture is shown

above, with major interfaces labeled alphabetically and shown by arrows (Fig. 2).

Just as materials are shelved in a traditional library by a small army of students wheeling **book carts**, the SRB will take care of loading digital content into our collection. It will take the place of the **shuttle van**, hauling content back and forth to the off-site storage facility, as it transfers files from robotic tape storage to disk farms. There are approximately 60 TB available on the SDSC High Performance Storage System (HPSS) tape system, and 20 TB at IGPP's Digital Library. The **resource sharing** function will be performed at Internet speed, as the SRB works with distributed archives. At the moment, the SRB is used extensively to manage large-scale archives at a number of Universities and National Laboratories. The SRB was designed to work with peta-byte (1000 tera-byte) scale databases, with hundreds of millions of files, across such distributed archives. A distributed archive need not be in a Supercomputer Center. It could easily be in a researcher's office, just so long as the local workstation is running the SRB server software. Significantly, this is an example of a **federated system** that can place the data archive at the point in the organization where the expertise about the data resides. Thus, the SRB supports both centralized and distributed facilities. The **metadata** for our collection will be managed by the SRB's modern MCAT metadata catalog. The SRB can maintain the currency of our **digital collection** as it examines various versions of documents across distributed archives. There is even a synchronize command to automatically update multiple collections, and make them all consistent. Since it has been designed for distributed archives, it inherently provides excellent **backup** capabilities, as well as mirroring, allowing convenient monitoring of the time history of all the content. The SRB can manage **collection access** with an extended set of file permission attributes. Beyond the usual access control, it will allow a data owner, to authorize a key for another individual user to access selected content for a specific amount of time. The SRB can also deal with **intellectual property issues**. It can be made aware of the publisher licensing agreements of the client's institution to facilitate full text viewing of published articles, as appropriate. Further information may be found at www.npaci.edu/online/v5.4/srb118.html and www.npaci.edu/DICE/SRB/.

To illustrate the need for improvement over current practices with local data archives, we shall consider a typical multibeam data access scenario. A student or researcher picks a geographical area of interest for a study and runs local programs to produce a list of all the available multibeam and underway data that cross through it. The resulting file list needs to be compared to the list of cruises on "proprietary hold." For areas a few degrees square, the search may turn up hundreds to thousands of multibeam data files from various vessels and cruises, which may be found anywhere among the million or so files and 400 GB of the SIO archives. Although about 100 GB of commonly-accessed data are held online, literally days may be spent on site finding archived Exabyte, DAT or CD media in file cabinets and loading the appropriate files onto temporary disk storage.

Using the SRB, a student could open a browser as a client on their own PC or workstation from anywhere across the web and search the SRB's central metadata catalog for this project, discovering the files on distributed SRB servers for the area of interest. They could also discover all the associated information for that area in the same search, such as existing data grids and images, published reports, and archival photographs. Alternatively, they could restrict the search for a particular vessel, span of dates, or cruise name. With modifications to our search engine and metadata content, users could also search for a named area, such as the "Juan de Fuca Plate." Upon discovery, images could be viewed immediately and existing grids or other objects downloaded directly. If they have the SRB running on their own computer, they could simply copy files or whole collections from the SDSC SRB archive to their own. They could request that a custom container be made from the thousands of files, and appropriate HPSS or IGPP robotic tape system would scan various volumes into temporary storage on a disk farm, and send an e-mail to the user that their data collection is ready to be picked up by FTP.

REFERENCES

[1] McNutt, M., V. Alexander, J. Ausubel, R. D. Ballard, T. Chance, P. Douglas, S. Earle, J. Estes, D. J. Fornari, A. L. Gordon, F. Grassle, S. Hendrickson, P. Keener-Chavis, L. Mayer, A.E. Maxwell, W. J. Merrel, J. Morrison, J. Orcutt, E. Pikitch, S. Pomponi, U. Sexton, J. Stein, G. Boehlert, J.

Cleveland, T. Curtin, R. Embley, E. Lindstrom, M. Purdy, M. Reeve, W. Schwab, M. Sissenwine, and R. Spinrad, The Report of the President's Panel for Ocean Exploration, Discovering Earth's Final Frontier: A U. S. Strategy for Ocean Exploration, http://oceanpanel.nos.noaa.gov/panelreport/ocean_panel_report.html.

[2] Smith, W. H. F., and D. T. Sandwell, Global sea floor topography from satellite altimetry and ship depth soundings, *Science*, 277, 1956-1961, 1997.

[3] Koppers, A.A.P, H. Staudigel, and J. Helly, EarthRef.org Database Development, *EOS* 81, F1285, 2000a.

[4] Koppers, A.A.P., H. Staudigel, and J. Phipps Morgan. TnT2000: A Toolbox for Modeling the 1,176, *EOS* 80, 1, 176, 1999.

[5] Koppers, A.A.P., H. Staudigel, and J. Phipps Morgan. Contrasting Mantle Convection Models by modeling their Geochemical Evolution with the Terra Nova Toolbox (TnT2000). Goldschmidt Conference, Oxford, UK, *Mineralogical Magazine*, 2000b.

[6] Caress, D. W., and D. N. Chayes, New software for processing sidescan data from sidescan-capable multibeam sonars, *Proceedings of the IEEE Oceans 95 Conference*, 997-1000, 1995.

[7] Caress, D. W., and D. N. Chayes, Improved processing of Hydrosweep DS multibeam data on the R/V Maurice Ewing, *Mar. Geophys. Res.*, 18, 631-650, 1996.

[8] Michener, W. K., J. W. Brunt, J. J. Helly, T. B. Kirchner and S. G. Stafford, *Nongeospatial Metadata for the Ecological Sciences*, Ecological Applications, 7, pp. 330-242, 1997.

[9] Helly, J., T. T. Elvins, D. Sutton and D. Martinez, *A Method for Interoperable Digital Libraries and Data Repositories*, Future Generation Computer Systems, Elsevier, 16 (1999), pp. 21-28, 1999.

[10] Helly, J., T. T. Elvins, D. Sutton, D. Martinez, S. Miller, S. Pickett and A. M. Ellison, *Controlled Publication of Digital Scientific Data*, CACM (accepted October 3. 2000) (2000).