SWOT Results – Software and Data Systems Group --

Process

SWOT analysis is an initial step in the process of developing a strategic plan. The analysis is divided into an internal analysis (strengths and weaknesses), and an external analysis, (opportunities and threats).

This approach to strategic planning is based on the notion that it is somewhat easier for a strategic pan to address the internal issues than external issues. Internal issues are, in principle, more under management control. A basic management strategy is to build on strengths while mitigating weaknesses. The external issues, however, require strategic principles that can address the changing condition that are not under MBARI's control. In summary a SWOT analysis, provides starting points for the plan itself.

Drawing from the group of staff members assigned to the SDS planning process by Jim, we formed into four subgroups. Generally each person was on three of the four subgroups, so there was ample interaction between teams.

The groups each created a comprehensive list of items to consider then synthesized this into a few categories, followed by a verbal explanation of the important categories. You will note some inconsistencies in what follows since, in effect, a glass that is half empty, is also half full.

The subgroups:

Strengths – John Graybeal, Mike Godin, Dan Davis
Weaknesses – Dan Davis, Duane Edgington, Kevin Gomes
Opportunities - David Caress, Dan Davis, John Graybeal
Threats - Kevin Gomes, Mike Godin, Duane Edgington.

- Strengths –

People With Experience

MBARI has been working on software and data management problems in the oceanographic setting for a long time. The organization has built many working (and a few non-working) solutions for both observing systems and domain-specification applications, and has learned from this experience.

MBARI and its SDS team also have experience leading large external collaborations with technology/engineering components, AOSN and MMI being two recent examples with heavy SDS involvement. Our demonstrated accomplishments give us the experience and respect to lead other initiatives in the future.

Diverse Data Sets

MBARI has one of the most comprehensive and diverse data sets for any region in the world (and access to even more data in collaborating institutions). Our existing data sets include both unique data types, and most data types (via MOOS and MARS) that will be generated by large observing systems. We can and should be an exemplar of data integration for ocean observing systems, just using our own data.

Environment—People Interactions
The SDS can build teams internally with all kinds of experts (operators, developers, technicians, scientists, and outreach specialists) needed for creating and operating the systems, and using and evaluating the data produced by the systems. In data management this advantage is critical for producing working, near-optimal comprehensive solutions that in fact solve most data management problems, for MBARI and the community.

Environment—Policies/Practices
MBARI has an institutional commitment to disseminate data which can be used as the basis of a comprehensive data management approach. In addition to ready access to and publication of core stream data, MBARI can integrate even more regional data sets by embracing the enlightened self-interest of the science teams, who will certainly appreciate the value of their data being widely available in an integrated and professional way.

Environment—Education/Outreach
MBARI can use its in-house and closely related education and outreach facilities to build academic, commercial, regulatory and personal collaborations, and to enhance our visibility and reputation. We must develop a focused plan, in collaboration with those programs, to maximize the benefit of our education and outreach resources.

Environment—Stable Resources
MBARI's long-term and stable funding enables approaches and attitudes not available to other institutions. We can uniquely apply that funding by hiring highly qualified engineers, supporting relatively high risk or delayed payoff projects, and leading community-scale projects with relatively little direct benefit (monetary or staff size) to MBARI. By being open to community-based approaches (not limited by a not-invented-here attitude), SDS will extend MBARI's reputation and collaborative opportunities even further.

External Connections
The SDS team is well coupled to the major existing observing efforts (MARS, NEPTUNE, ORION, IOOS, MOOS) and several campaigns such as AOSN. We have established ties to the Monterey Bay oceanographic community, the wider oceanographic community, and the international metadata community, as well as to non-oceanographic data management and software communities like Silicon Valley, NASA, university research labs, and the astronomy community. We can and should leverage these connections into collaborations, proposals, and important data integration efforts.

Very Cost Effective
In distinction to developments of new platforms or new instruments, which often require additional costs to the organization once transferred to operational status, investment in software and data systems can lead to decreased organizational costs through increased efficiency and more product (science) from the same observations (data) at small incremental costs.

- Weaknesses –

Limited resources
There is not enough staff to be connected to both commercial software and academic research communities, to have a critical mass of capabilities in many areas, and to develop any single large system while also supporting a set of broad and diverse internal needs. Also there are areas, such as modeling, artificial intelligence, database and data management technology, that are not well represented on our staff. We have a low quality connection to the Internet.

Lack of consensus and focus on what is important

There is a lack of MBARI consensus and long term commitment on what data problems we should focus on. There is a need to focus on problems with a high science and community impact. A demand on our group to focus on internal data management problems and science whose focus may not be representative of important community wide problems, as well as a lack of visibility of our work to the outside community may have the consequence that our solutions to problems do not have community wide relevance. A failure to have consensus on this results in an inconsistent understanding of the problem and a lack of focus.

Internal users also may not recognize the importance of the visibility of our data and data management solutions to the outside community.

Inconsistent engineering process

Historically we have not executed crisply. Across the department and within the group *we apply* processes and discipline and teaming inconsistently which does not lead to a timely execution of projects not *even* when we have decided what to do. We need consistency across the core engineering/organization on processes and systems engineering to enable effective execution of projects.

We also need to carefully consider COTS and open source solutions and components. Inadequate and objective systems engineering focus, given the large number of problems and opportunities, leads to a further diffusion of effort and fragmented results which become a further obstacle to productivity. A failure to adequately address the quality control of data, inadequate metadata and interface standards, etc., and access to our data externally, is not only a symptom of limited resources. It is compounded by an inadequate engineering process.

-Opportunities –

Leadership

MBARI is well positioned to provide leadership to the oceanographic community in Data Systems technologies.

The oceanographic community is increasingly focused on observing systems such as those championed by the MBARI MOOS project. The NSF and other funding agencies recognize the importance of data systems problems to the community. MBARI operates (or will operate) and derives data from most of the classes of observing platforms used within oceanography (e.g. ROVs, AUVs, moorings, cabled observatories), and so efforts to work with MBARI's large, diverse datasets are directly relevant to the problems faced by the oceanographic community as a whole.

The ocean technology community benefits from and welcomes leadership on truly collaborative projects. Many technologists and developers at commercial and academic institutions are eager to partner on real, interesting problems. Through community collaborations, MBARI is capable of serving as an integration house for both point (specific) technologies and broad initiatives (e.g., observing systems, system of systems architectures, standards and protocols). MBARI is already involved with most or the entire major ocean observing initiatives.

Large external developments exist in industry (XML, grid computing, Cyberinfrastructure, semantic frameworks) that can be exploited in the oceanographic setting. Several MBARI engineers have connections to people and activities in Silicon Valley.

Observatory Data Logging and Archiving

Solutions for the logging, real-time distribution, archiving, and dissemination of data from ocean observing systems, including long-term observatories, are an essential component of platform and observatory development efforts. MBARI has invested considerable effort into handling the data from

cabled observatories, moorings, and other platforms, and is thus in a good position to complete and disseminate solutions to the community.

Integrating Modeling with Observations

Many oceanographic topics involve complex phenomena and systems that are being investigated using large and diverse datasets, and modeling is an essential component of the science. However, no satisfactory system exists for connecting real-time data to models, or vice-versa.

Integrating Diverse Data Types into Useful Data Products

MBARI produces and archives a diverse group of large datasets. In many cases, integrated views or visualizations derived from disparate data enable new insights and new science. In order to take full advantage of our data, MBARI needs to develop interfaces to the data that integrate all types (e.g. a common portal for all datasets using consistent interfaces for each) and provide access to useful data products.

Data Mining and Query Algorithms

MBARI and the rest of the oceanographic community hold vast and diverse stores of data but use crude methods to find and extract information relevant to any particular topic. MBARI has an infrastructure (in SSDS) that can be used as a basis for developing more advanced, useful, and easy to use systems. MBARI can also take advantage of the considerable progress in this topic that has been made external to oceanography.

Integrating Data Archives from Multiple Institutions

MBARI has considerable data complementary to other institutions' holdings and of interest to the oceanographic community. However, MBARI, like most other institutions, operates its data archives with little linkage to external archives. The increasing Internet bandwidth available now makes it possible to integrate data archives from multiple institutions so that data products derive from all of the sources without a need for duplicating data between the archives. Developing these sorts systems will require standard system interface protocols and application of semantic computational tools now starting to mature (e.g. Marine Metadata Interoperability effort). Some of the needed infrastructure software has been developed externally, but only limited useful implementations exist. MBARI also has some reference software (VARS, SSDS, possibly SamplesDB) that can take advantage of existing or proposed interface protocols.

Developing Standard Interfaces for Interoperability

MBARI has invested considerable effort into developing standard interfaces to allow interoperability of instruments and data systems (ISI and SIAM and SSDS). Interoperability is necessary in order to facilitate multi-institution use of high cost platforms and observatories.

Developing Quality Control

Effective quality control processes are necessary for handling vast data streams deriving from long-term observatories. MBARI needs to develop capabilities for integrated, automate real-time quality control checks on data streams. Tools for manual quality control are also required, including improved interfaces for identifying and flagging errors for future users.

Understanding Optimum Sampling Scales in Time and Space

MBARI holds datasets from many platforms that sample processes over a variety of time and space scales. These datasets are appropriate for developing optimal strategies for using our limited platform resources.

Automated Data Interpretation

Advances in Artificial Intelligence technology external to oceanography provide an opportunity for MBARI to advance automated data interpretation techniques. Use of vision technology for automated interpretation of video data could enhance use of the data and provide for automated control of vehicles

(ROV and AUV). There is also interest in real-time analysis of observatory data relative to historical archives and data from other platforms. The community's long term goals for observatories include oceanographic event detection and response applications. Finally, there is a need for adaptive quality control software.

Virtual Observatories ("Point-and-click Oceanography")

Echoing the astronomical community's efforts towards Virtual Observatories ("Point and Click Astronomy"), advances in Cyberinfrastructure along with Ocean Observatories enables an oceanographic data avalanche. Breakthroughs in cabled and moored observatory platforms, sensors, and computer technology allow ocean scientists to produce terabytes of data and images. These datasets will cover the ocean from different perspectives – overhead satellites, surface observations, coastal arrays, and deep sea installations, all arriving to shore in near real time. In a few years it will be easier to "dial-up" a part of the ocean than wait many months to access a UNOLS ship. "Point-and-click oceanography" could greatly enhance the scientific output from the expensive platforms and instruments by bringing more PI eyes onto the same data sets. With the advent of inexpensive storage technologies and the availability of high-speed networks, the concept of multi-terabyte on-line databases interoperating seamlessly is no longer outlandish. More and more catalogs will be interlinked, query engines will become more and more sophisticated, and the research results from on-line data will be just as rich as that from "real" ocean observations. These technological developments will fundamentally change the way oceanography is done. These changes will have dramatic effects on the sociology of oceanography itself. *(Shamelessly adapted from the NSF National Virtual Observatory project).*

- Threats –

Competition and barriers to collaboration

Some oceanographic research institutions and corporations may see MBARI as a competitor, which limits opportunities for both collaboration and external adoption of MBARI technologies. Other research organizations may be more successful than MBARI at capturing resources and mindshare, due to funding sources' preferences for commercial, academic, or other particular research groups. The wide array of ocean data management efforts, including several high-profile efforts (ROADNET, IOOS, and Neptune Canada) may lead to a few de-facto standard approaches, making MBARI's approaches irrelevant. In addition to research organizations, if we do not stay connected to, and abreast of, commercial solutions, they may render MBARI's solutions irrelevant as well.

Disconnection from evolution of community/industry standards and technology

Developments that are external to MBARI and the SDS group in the areas of standards and technology could render internally developed products to be inferior, not acceptable, or not applicable for the oceanographic community. This is also true for the case when MBARI adopts external standards and technology, which then evolve in a direction that does not benefit MBARI or becomes proprietary.

Current status of our reputation in the community

Our current reputation and prominence in the oceanographic community could be insufficient to overcome biases (both intentional and unintentional) that would preclude us, and our developments, from participating in collaborations and implementations in the oceanographic community.

Both the instability of, and our inability to leverage, external resources

The changing landscape of agencies and funding structures could make it difficult to get the resources needed to complete projects. Also, if we depend on external collaborations to complete projects, problems with those collaborations would jeopardize the success of the projects.

Risks in sharing MBARI's data and technology

Upon opening up our data and technology to the external community, there are risks that both can be used without attributing intellectual rights to MBARI, or in some cases, even damaging MBARI's reputation and prominence in the oceanographic community.