Lowell Bahner

Chesapeake Information Management System and the World Wide Web

The Chesapeake Bay Program is making information about the Bay and its tributary rivers readily accessible and useful using the Internet and the World Wide Web (http//www.chesapeakebay.net). Citizens, schools, local governments, federal agencies, and non-governmental organizations are all large user groups interested in current information about the health of the Bay and its tributaries. Since the Bay Program encompasses work conducted by many agencies and organizations over a six-state area, data tend to be located all across the Bay basin. The Chesapeake Information Management System (CIMS) is an initiative to organize this large amount of information, using distributed Internet technology.

CIMS: Re-engineering Data Management for Chesapeake Bay

The growth of the information age has pushed the Bay Program's data center to re-examine how it does its business. In June 1996, the Program's Information Management Subcommittee conducted a workshop to gather and discuss user information needs. Approximately 100 people, representing most Chesapeake Bay Program agencies, participated to make their needs and expectations known.

The major findings of the workshop were:

- Many users or potential users find it difficult or impossible to get the data they need, and what are available are often outdated, too broad geographically, or insufficiently documented, increasing the amount of time needed to perform even routine procedures;

- Environmental Indicators, an important priority for the Chesapeake Bay Program, are difficult and inefficient to produce with the current system;

- To undertake complex analyses, different types of data need to be integrated;

- There are important changing directions in Chesapeake Bay Program investigations that will require new types of data and new ways of managing them; and

- There is difficulty summarizing and analyzing data spatially (e.g., loadings related to sources in a specific geographic location).

The major recommendations of the workshop were:

- Promote a Chesapeake Executive Council directive on data management;
- Develop standards for data providers, including a requirement that all Bay Program data be accompanied by standardized metadata;
- Make as much Bay Program information available over the Internet as is feasible; and
- Develop the CIMS, making it appropriate for the needs of a diverse user group, so that it is easy to use, allows users to query both spatial and textual information, and is partly decentralized.

The immediate response to the workshop's findings and recommendations was to improve access to information held by the Bay Program. For years, raw data had been managed, but with access to the Internet, documents, press releases, fact sheets, graphs, charts, environmental indicators, and data interpretations could all be made available for public access at a fairly low cost. Each Bay Program subcommittee was given the responsibility of publishing and maintaining its information using agreed-upon formats. A management team was created to oversee Web site development and maintenance. Soon partner agencies were creating their own Web sites. This ad hoc approach gained much popularity as a means of making information available to the public, students, scientists, and government agencies. However, the rapid creation of many different Web sites in the Chesapeake region, while a major improvement in obtaining information instantly, would also be a detriment when an integrated assessment was required by partner organizations.
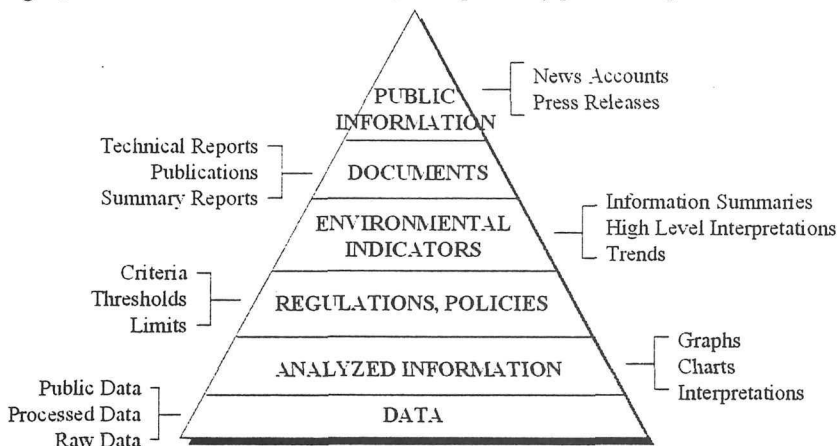

Figure 17. CIMS information pyramid

CIMS Implementation

In October 1996, the Chesapeake Executive Council (comprising the governors of Pennsylvania, Maryland, and Virginia; the mayor of the District of Columbia; the chairperson of the Chesapeake Bay Commission; and the administrator of the U.S. Environmental Protection Agency, representing all federal agencies) signed the "Strategy for Increasing Basin-wide Public Access to Chesapeake Bay Information" (http://www.chesapeakebay.net/bay-program/pubs/96cims.htm). This document directs the Bay Program to maintain a coordinated data management system, which not only provides timely information on the progress of the restoration program, but promotes greater understanding among citizens about the Bay, its problems, and the policies and programs designed to help it—thereby fostering individual responsibility and stewardship of the Bay's resources. The strategy directs the implementation of the basin-wide CIMS as a coordinated, user-friendly, Internet-based system, designed as a distributed network among participating organizations throughout the watershed and nationwide. Participating agencies must ensure that data and information products that are created or procured through grants or contracts can be loaded easily and maintained on the network.

The Bay Program's Information Management Subcommittee has developed policies and guidelines for directing CIMS-related activities ("Chesapeake Bay Program Guidance for Data Management," http://www.chesapeakebay.net/cims-pol4.pdf). This document provides guidance for all participating organizations on locational accuracy, map coordinates, metadata, station names, common data dictionaries, database design, calendar dates, method codes, and data reporting and deliverables. These policies set the baseline rules so that a coordinated, yet distributed, information system can be created and maintained.

The success of a distributed information system depends on each partner organization doing its share. Memoranda of understanding are being signed by partner organizations to state what that organization's role will be in helping to implement CIMS. The process of getting these memoranda signed serves the purpose of focusing the upper management of federal, state, and local governmental agencies, universities, and non-governmental organizations on their responsibility properly manage information, publish it in a timely manner, and give the public access to it. Historically, the Bay Program's data center was one of the largest repositories; however, there was a large cost to maintaining a central repository because data in computers at the center would soon get "out-of-sync" with those held at state agencies due

to edits and updates not being made to all copies. These problems were very expensive to fix, and fixes were soon out-of-date anyway. By changing the process so that the generating organizations become the owners and keepers of the data, quality will increase while the overall cost decreases. Significant savings are already being reported due to publication of Bay Program data and information products over the Internet.

Timely access to information was another issue that is solved by having the data generator publish over the Internet. The generator, being the owner of the information, is more likely to improve the quality of the information through pride-of-ownership, and the information can be made available over the Internet in a timely manner since there are fewer layers of handling. This is important when quick management decisions need to be made in response to an environmental hazard.

Making a distributed information system work relies on the partners following the same road map. The importance of the policies and guidelines comes into play—i.e., the rubber meets the road—when data from one location need to be merged with those from another. Since this is an everyday occurrence for analyzing the interactions of air, land, and water quality with living resources within the region, it is crucial that all partners use the same language—that is, the data dictionaries for all the various databases must be consistent. The Internet has provided the means for sharing information, but the largest challenge is for each organization to publish its information in the same language so others can access and use it.
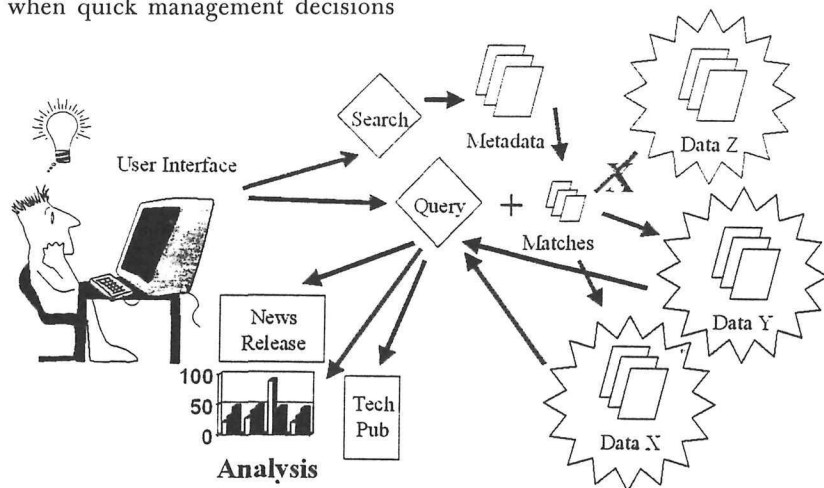

Figure 18. CIMS at the desktop

The vision for how all of this will work is becoming clear: a user can select specific information of interest through a "smart" user interface. The user interface engine will take this search request and send it to the metadata search engine, which acts like a library card catalog. The metadata search engine will search the metadata holdings (the extended documentation for every document or data set) and identify where (i.e., on whose computer) the actual information is housed. The query portion of the search request is then forwarded to those remote computers that are identified as having the requested information. If they do, the information is sent to the requesting computer. Once all of the information is received, it is packaged into a user-friendly format and delivered to the user's computer. This "CIMS at the desktop" goal is to provide data, graphs, charts, maps, and documents to the user in a manner that is most useful, much like the daily newspaper or the six o'clock news, but with the extended capability of "drilling down" into more detailed information when required.

Lowell Bahner, Chesapeake Bay Program Office, National Oceanic and Atmosphere Administration, 410 Severn Avenue, Suite 109, Annapolis, Maryland 21403

If you would like more information on the Chesapeake Bay Program on-line, visit its Web site at http://www.chesapeakebay.net. You will find detailed information about the program's partners, activities, publications, maps, and monitoring data.