DEVELOPMENT OF THE COOK INLET INFORMATION MANAGEMENT/MONITORING SYSTEM (CIIMMS)

Kelly Zeiner[1], Russell Kunibe[2], Carol Fries[1], Greg Kellogg[3], Leslie Patrick[4], Jeff Hock[2], Lowell Suring[5], Dann Sklarew[6], Cary Gaunt[6], Michael Fene[6] and William Samuels[6]

[1] AK Dept. of Natural Resources
550 W. 7th, #706, Anchorage, AK 99501
[2] AK Dept. of Environmental Conservation
410 Willoughby Ave., Juneau, AK 99801
[3] US Environmental Protection Agency
[4] US Geological Survey
[5] US Forest Service
[6] Science Applications International Corporation

ABSTRACT

The Cook Inlet Information Management/Monitoring System (CIIMMS) provides an interactive website that links to a geographically distributed system of information providers. Through the CIIMMS website, users are able to identify and access (e.g., download and print) information ranging from primary data (geospatial and tabular) to reports, project descriptions, and other documents across a variety of themes, such as habitat, land use, resource management, pollution, and water-quality information. CIIMMS will also provide (on-line) tools to make it easy to contribute information to the CIIMMS network. In the long term, CIIMMS will use integrated information resources and tools to create a virtual community center for Cook Inlet learning, resource management, and related activities.

CIIMMS not only comprises the hardware, software, and information contact components of the Cook Inlet information management system; it also establishes a framework for managing information resources more efficiently. Two important components of CIIMMS development are (1) establish an advisory group made up of data providers and users to oversee implementation and (2) publish guidelines on how to implement various aspects of the framework (e.g., compile metadata). By providing the tools and the framework that enable a unified approach to information management throughout the watershed, CIIMMS will help agencies and organizations use existing resources more effectively.

INTRODUCTION

The CIIMMS is a three-year project, funded by the *Exxon Valdez* Oil Spill Trustee Council, which will deliver a website with tools for accessing, contributing and sharing of information for the Cook Inlet watershed. This site and its associated graphical user interface will support monitoring, management and restoration of resources and services injured by the *Exxon Valdez* Oil Spill, as well as provide access to data sets and tools valuable to addressing the ecological health of a watershed. This system is unique in that it will provide access to a broad spectrum of timely data and visualization tools via the Internet. CIIMMS will allow regulators, resource managers, planners, and others to approach decision-making from a watershed perspective.

The vision for CIIMMS is to develop a comprehensive information system based on the latest Internet technologies that will enable a wide range of users to contribute, identify, share, and access valuable information about the Cook Inlet watershed and Cook Inlet related activities. This vision was established after extensive background investigations and culminated in a 2-day user needs workshop held in Anchorage, Alaska, on January 26-27, 1999. Two reports (SAIC 1999a and 1999b) describe the results of the background investigations and workshop outcomes.

Figure 1. CIIMMS conceptual design structure.
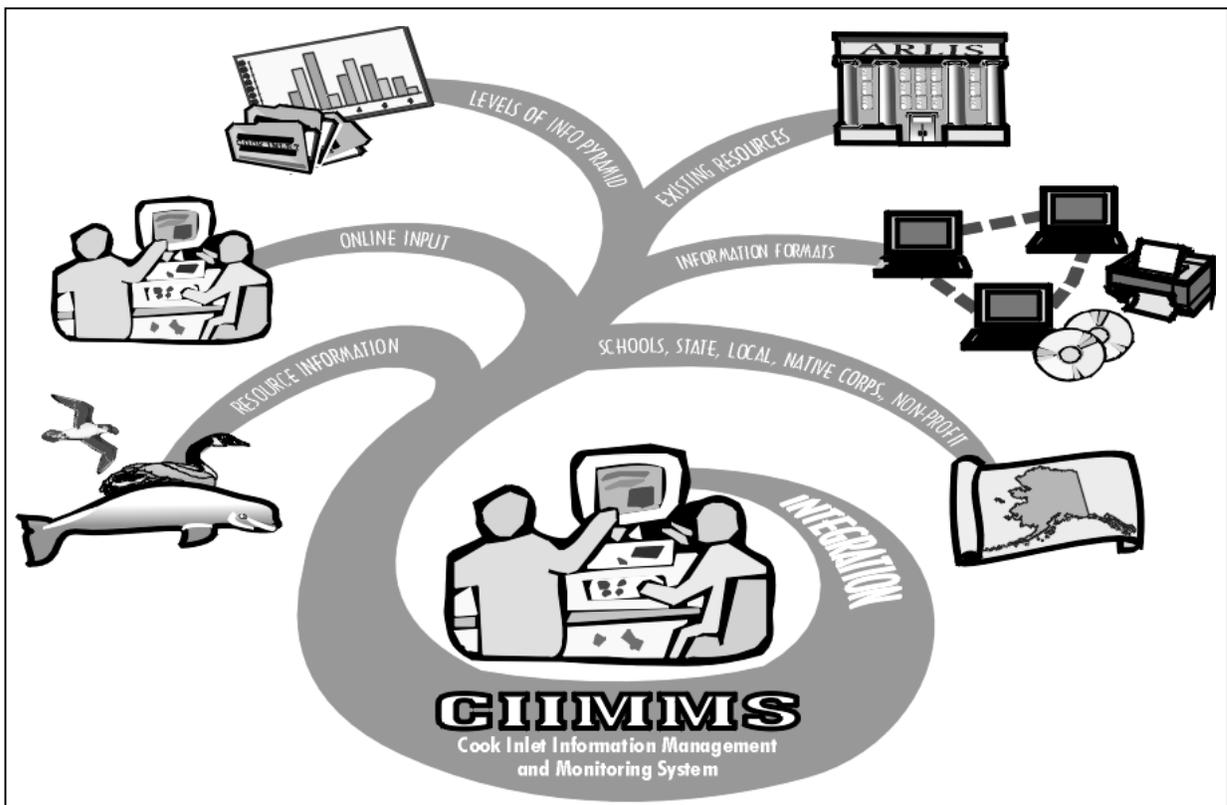
BACKGROUND

The Cook Inlet watershed is a large and complex ecosystem containing a diverse and abundant biota that is subject to intense physical forces as well as to increasing human influences. Each year, industry, government, the scientific community, and citizen watchdog groups generate and use large quantities of information about this area and its resources. Typically this information is used to focus on a single resource, issue, or problem, and data management techniques are used that are specific to that need. These actions may lack a coordinated, comprehensive ecosystem approach. Ecosystem or watershed level management requires integration and access to data from many different projects and initiatives.

Watershed management has a scope that requires evaluation of a much broader spectrum of factors within a defined geographic area. In most large, intensively used and managed watersheds, such as Cook Inlet, some stakeholders collect and analyze samples and generate data, while others rely on data to monitor resources, conduct research, or make management and policy decisions. All gain through sharing resource information.

Information sharing has become an expectation rather than an anomaly. Attitudes, capabilities, and expectations regarding information are evolving. Even so, it is currently difficult to develop a comprehensive picture of existing data resources when working to understand the health and needs of a watershed. Past practices have often ended in data and results being pigeonholed and lost because of various situational restrictions such as funding, perceived propriety, or technological capabilities at the time of the investigation. It is time to use existing data more efficiently, to make funding more effective, and to stem the loss of historic knowledge. The CIIMMS project aims to do just that: provide a large and diverse user base with access to the necessary data and tools to provide a more unified approach towards resource management within the Cook Inlet Basin.

IMPLEMENTATION METHODOLOGY

In October 1998, the Alaska Department of Natural Resources (ADNR) and the Alaska Department of Environmental Conservation (ADEC) launched the CIIMMS project. A multi-agency project team was formed and presently includes members from U.S. Environmental Protection Agency, the U.S. Forest Service, the U.S. Geologic Survey, and the ARLIS Library (Alaska Resources Library and Information Services). Science Applications International Corporation (SAIC) provides technical guidance and support to the CIIMMS team.

CIIMMS' implementation is occurring in three phases:

Phase I: The Pilot Phase, generating a functional prototype (completed December 1999);
Phase II: Initial Production Phase (January 2000 - April 2001) ; and
Phase III: Final Implementation Phase (Years 3 through 5).

Phase I: The CIIMMS Prototype

CIIMMS' Phase I encompassed the following activities:

- Assessment of user needs through surveys, focus groups, and a 100-person stakeholder workshop, which also evaluated current information management approaches within the Cook Inlet community.

- Development of the "Cook Inlet Information Management Requirements and Recommendations" post-workshop report (available online at http://www.dec.state.ak.us/ciimms/project/about.htm). More than 100 Cook Inlet stakeholders contributed directly to this document's evaluation of CIIMMS priority information, functions, system design, and user interface.

- Design and implementation of the CIIMMS prototype (available online at http://www.dec.state.ak.us/ciimms).

The Phase I deliverable was the CIIMMS prototype website. The site acts as a clearinghouse node for searching and reviewing Cook Inlet related metadata (i.e., an electronic card catalog that characterizes specific Cook Inlet information resources). Metadata records characterize data sets and project descriptions, and a web crawler provides a fast index to Cook Inlet related web sites. CIIMMS provides a common interface for searching through bibliographic metadata, geospatial metadata, and keyword-based metadata on the CIIMMS server and other Internet-linked information servers. CIIMMS-accessible servers include Alaska's geospatial clearinghouses; state and local libraries; local, state and federal agencies' websites; non-governmental websites, such as those of local newspapers, research centers, and citizens' groups.

Wherever possible, metadata records contain hyperlinks to the information resources that they characterize. Such resources will include spatial and tabular data, maps, project descriptions, technical reports, organizational points of contact, and other useful documents. Specific information include ecosystem assessment data, land ownership maps, government reports on contaminated sites, or descriptions of salmon restoration projects. Once accessed, the information can be viewed, downloaded, and printed by the user.

Evaluation criteria were developed for the CIIMMS prototype, and these criteria were used to build evaluation forms, which were handed out (and filled out by evaluators) at all the CIIMMS outreach and demonstration workshops. To allow evaluators to provide feedback while accessing the site, some simple, straight-forward entry forms were provided on-line, to collect more spontaneous feedback. General evaluation criteria included ease of use, response time, usefulness of information, visual appearance, and other appropriate criteria.
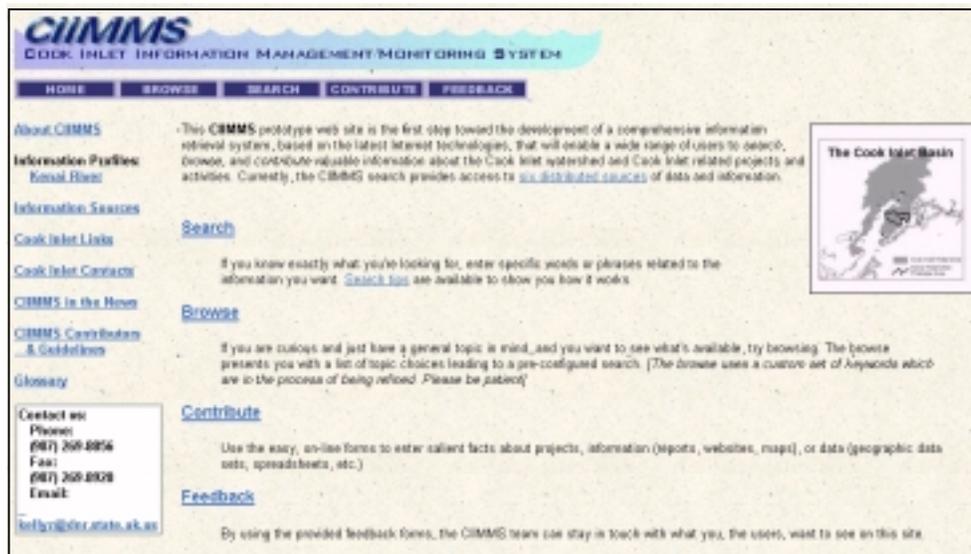
This approach provided for rapid collection of reviewer feedback during the Pilot Phase. User feedback and reviewer comments were used to build the specifications for the Initial Production Phase.

The prototype area of interest, CIIMMS home page and functionality delivered by the CIIMMS Pilot Phase is shown in Figures 2, 3 and 4 respectively.

Figure 2. Map of the Cook Inlet Basin and CIIMMS Prototype study area (Kenai River Watershed).
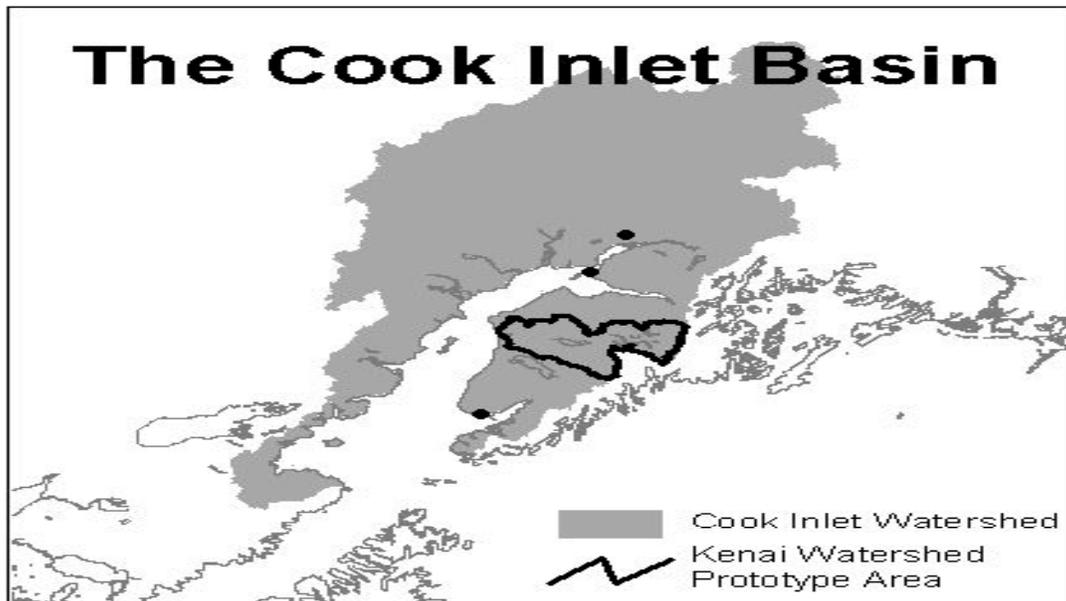
Figure 3. CIIMMS prototype website homepage.

Figure 4. Design Summary for CIIMMS Prototype

Geographic Scope	Kenai River Watershed
Information Types	Anadromous streams/species, wetlands, ground water well logs, EVOS data, etc.
Information Sources	ADEC, ADF&G, ADNR, Kenai Peninsula Borough Planning Department, USGS, and others
Functions	Search , browse, on-line metadata entry capability
Information Features	
-- Identifying	• Categorical indexes for Cook Inlet information inventory • Keyword and advanced metadata searching • Restoration project activities
-- Accessing	• Ability to view, download, and print static maps and web documents • Metadata records linked to actual data and summary information • Hotlist of all related offsite links
-- Contributing	• Form for suggesting information and links to add to clearinghouse • Metadata entry tool for data/projects (short form) to populate clearinghouse
Evaluation Tools	• Feedback forms online

Phase II: Initial Production Phase

The pilot phase has been completed and an operational prototype has been developed. The initial production phase is underway. This phase will include expansion of access to knowledge, information, and data, including *traditional ecological knowledge*, building on the datasets available via the prototype. Information accessible via CIIMMS includes on-going projects taking place within the basin. The search capabilities will be refined to include a map-based and advanced, fielded search tools, and refinement of the interface will build upon the browse and metadata entry tools delivered in Phase I. Coordination and outreach continue to be a vital part of CIIMMS, and the team cooperates with the Alaska Geospatial Data Committee, and many other large state and local data management initiatives.

Data Resource Expansion

There is no shortage of data for the Cook Inlet community, but as mentioned earlier, much of it is difficult to access because users don't know it exists. Data collected as part of the *Exxon Valdez* Oil Spill restoration will be tracked down and made accessible wherever possible. The Alaska Department of Fish & Game, in cooperation with the CIIMMS project, will convert biological information from the Alaska Habitat Management Guides to a format that can be used with desktop GIS applications. The geographical focus will be Cook Inlet, and the data will be provided on-line via the ADF&G web server. Metadata for each theme will be created, and will reside at the most appropriate clearinghouse, if not on the ADF&G server.

The CIIMMS team, during the Initial Production Phase, will add several new data stores to the list of Z39.50 servers, including the Anchorage Municipal Libraries (AML), and the Capital City Library in Juneau, Alaska. AML includes the ARLIS Library, whose catalog includes bibliographies from most of Alaska's natural resource agencies and organizations, as well as all the reports produced as a result of the *Exxon Valdez* Oil Spill. The Capital City Library catalog includes all the State of Alaska agencies' holdings.

In addition to Z39.50 server access, CIIMMS would like to be able to access the EPA Envirofacts data warehouse, as well as several of the Alaska Department of Environmental Conservation databases (solid waste sites, water quality, and contaminated sites, just to name a few). The data and information mentioned here is only a small representation of the data CIIMMS plans to make available for the Cook Inlet.

Search and Interface Enhancements

The search capabilities developed in the prototype will be expanded to include a map-based search tool for finding different types of information for one geographic area. The map-based search will use latitude/longitude coordinates and geographic keywords to locate information for a certain "area of interest." Refinement of the interface will include the addition of an originator browse tree, a "checkbox method" for defining canned searches, and other easier, more intuitive methods of searching (by watersheds, or borough boundaries, etc.).

CIIMMS will provide two ways in which to assist information providers and data managers in creating and maintaining their own metadata. One, via easy-to-use web-accessible metadata entry tools; and two, by developing and publishing data provider/contributor guidelines, in cooperation with the AGDC. Several versions of metadata entry forms will be available to cover the broad range of data provider capabilities. A special tool, formulated with a series of specific questions, will be developed for capturing and maintaining traditional ecological knowledge. The on-line tools support development of both spatial and non-spatial metadata housed either on the CIIMMS server or on contributors' websites.

Coordination and the CIIMMS Advisory Group

Achieving the CIIMMS vision will require the cooperation of many entities. The project team will provide coordination, outreach, and user training to CIIMMS contributors in order to help build links necessary for success. The CIIMMS Advisory Group, made up of a broad group of data providers and stakeholders, will ensure that there is a bridge between technical, management, and end-user concerns. This group will help ensure that the system will remain usable in the future.

The CIIMMS project is represented on the Alaska Clearinghouse Working Group, a subcommittee formed under the auspices of the Alaska Geospatial Data Committee (AGDC). The AGDC is responsible for overseeing and coordinating geospatial data activities in the Alaskan community, and for carrying out the broad objectives of the Federal Geographic Data Committee (FGDC). CIIMMS supports the expansion of FGDC-compliant metadata clearinghouses, and is providing more robust tools for accessing and contributing to Alaska's clearinghouse nodes. Currently, there are two FGDC clearinghouse nodes in Alaska: the Alaska Geospatial Data Clearinghouse (AGDC), and the Alaska State Geospatial Data Clearinghouse (ASGDC). The Municipality of Anchorage Geospatial Data Clearinghouse node is currently under development.

Phase III: Final Implementation Phase

CIIMMS will mature into a geographically distributed decision-support system, complete with new tools for data visualization and interactive mapping, data filtering to minimize download times, and integration of online-generated graphs and maps into electronic documents. In addition, CIIMMS will provide opportunities for users to collaborate in creating new documents, to coordinate restoration projects and publicize activities, and to contribute and access traditional ecological knowledge about Cook Inlet.

In the long term, CIIMMS will use integrated information resources and tools to create a virtual community center for Cook Inlet learning, resource management, and related activities. Throughout its development, CIIMMS aims to remain user-driven. CIIMMS will continue to enable users to submit evaluations and recommendations regarding products and services, to ensure that the tools and resources continually meet the evolving needs of the Cook Inlet community.

SYSTEM DESIGN AND SPECIFICATIONS: FROM PILOT TO PRODUCTION

System Functionality

Search Tools

CIIMMS will provide users three mechanisms to actively search for information resources. These include a basic keyword search, an advanced search for terms in specific fields, and a map-based search for information within specific geographic boundaries.

The results of two-term Boolean search currently found on the prototype interface are shown in figure 5.   In order for the CIIMMS search to locate data and information, metadata must exist for that resource. Metadata in this context can be any reference made to the data or information that is accessible via CIIMMS. The reference to the information could "live" in any of the following locations (see figure 6):

♦   CIIMMS Local Database
♦   CIIMMS Project Database
♦   Web pages
♦   ARLIS or any other bibliographic catalog which is accessible via Z39.50 protocol
♦   Alaska Geospatial Data Clearinghouse
♦   Alaska State Geospatial Data Clearinghouse
♦   Municipality of Anchorage Geospatial Data Clearinghouse

Browse Tools

An alternative to the CIIMMS search will be the CIIMMS browse tree. A browse tree is a set of hierarchically organized lists of terms that provide a semantic guide for finding information resources according to theme or subject, geographical location, or information provider or originator. CIIMMS will have two browse trees, one organized by subject, and one by originator. The browse will present the user with a list of selections. Based on what is selected the user will be presented with another list.   The lists will be collapsible and eventually will lead to a category that returns a list of results. As the data accessed by CIIMMS is expanded, the number of browse levels will be expanded in order to keep the number of results returned to the user manageable.

Once the user has reached the bottom level of the browse hierarchy, the user will be given the option to turn on additional filters for data source (type), and date. The number of results will be displayed in a table next to a description of the data source. When the user selects a data source a brief listing of the results will be displayed. The user will then be able to select a single result for more information or if the result is a web resource, the link will be directly to the described web page.

Figure 5. Example results from CIIMMS search.

Figure 6. CIIMMS metadata sources.

CIIMMS Local Database

This database is internal to CIIMMS, and right now, contains the Data Inventory that the CIIMMS Project Team produced as part of the User Analysis Phase of CIIMMS. It only houses references to data not found at other clearinghouses (!). CIIMMS metadata structure is flexible enough to include both FGDC-compliant and non-compliant records.

If you'd like to enter information about data that you would like to make 'discoverable' via CIIMMS, you too can contribute to CIIMMS using an on-line data entry form.

CIIMMS Project Database

The CIIMMS team has generated this unique database to store information on resource management projects across the Cook Inlet basin. Projects taking place in the Kenai River watershed of Cook Inlet are already being included, with those from other locations to follow soon.

To enter information about a particular project within the Cook Inlet basin, please feel free to contribute to CIIMMS via the on-line data entry form.

WorldCat (Bibliographic Resources)

The WorldCat™ database, a.k.a., the OCLC Online Union Catalog, contains more than 39 million records describing items owned by libraries around the world -- including ARLIS, the Alaska Resources Library and Information Services. Every record in WorldCat™ describes a holding at one or more of these libraries. WorldCat™ will be replaced by ARLIS as soon as the ARLIS server is up and running (mid-November, 1999)!

WorldCat™ helps users identify library materials (as opposed to metadata, spatial data, or web resources). For each material, a WorldCat record can provide information on its author(s), title, publication, series information, subject headings, notes, call numbers, languages, geographic location, or other elements selected by the CIIMMS Project Team.

ASGDC (Alaska State Geospatial Data Clearinghouse)

AGDC (Alaska Geospatial Data Clearinghouse)

The ASGDC and the AGDC are part of a growing clearinghouse infrastructure that give Internet users access to search against (FGDC-compliant) geographic metadata databases nationwide. The AGDC uses the same clearinghouse search software as CIIMMS, but targets a more technical audience, and those familiar with the analysis and manipulation of geographically referenced datasets. The CIIMMS project includes these two clearinghouse nodes in its search in order to deliver a 'comprehensive' accounting of all the data and information available on a given topic. As of October 1999, the CIIMMS search engine queries only the theme keywords section of the metadata records.

CIIMMS Web Harvest

The CIIMMS web harvest is a database of web resources, collected weekly by CIIMMS database administrator, using Blue Angel Harvester software. The harvest focuses on collecting information from specific sites and pages, using customized software, so that the user gets more relevant results, or hits, from a search. The Harvester is currently being tested with a Cook Inlet-specific set of web addresses (URLs) to search.

Contribute Tools

The Contribute functionality is the key to the long-term sustainability of CIIMMS because it will provide the Cook Inlet stakeholders with the necessary tools for keeping CIIMMS up-to-date. "Contribute" will provide the following tools to users:

a) step-by-step instructions, in the form of a tutorial, to assist users in making their data/information accessible via CIIMMS, for first-time contributors;
b) CIIMMS Contributor Guidelines, an on-line publication of technical specifications for advanced contributors
c) direct links to all on-line metadata entry forms and upload tools, for repeat customers.

System Architecture

In order to accomplish the system functionality prioritized at the CIIMMS user needs workshop, the system architecture was designed and implemented as follows:

- A client-server architecture was implemented, based on an Internet browser (i.e., a thin client)

- The data and information is served in a hybrid environment — both centralized and distributed. For example, during the pilot phase, a centralized server was provided for metadata, summarized data, and some primary data, with links established to distributed servers for sources of the majority of primary data.

The CIIMMS Project Team evaluated several options in developing the CIIMMS system architecture. These included the Chesapeake Bay system design (specifically, its metadata entry tool), the Federal Geographic Data Committee (FGDC) clearinghouse design, and several commercial-off-the-shelf (COTS) products (Compusult Metamanager, Blue Angel Technology). The Team decided to base the architecture on a COTS solution. A COTS solution has the advantage of providing the team with access to software upgrades, as technology changes, and access to a vendor-supplied "Help Desk" for aiding the CIIMMS administrator with system maintenance. A COTS solution would also 'free' the Project Team from being dependent on a technical contractor once development was complete.

Based on its evaluation, the Team selected the COTS software solution, MetaStar Enterprise, developed by Blue Angel Technologies for the CIIMMS Prototype system architecture. MetaStar Enterprise integrates database, search engine, and web technologies in a single solution that provides for the search and discovery of metadata and information via the Internet. Enterprise is a bundle of the MetaStar Data Entry, Repository, Gateway, and Server components. The architecture also includes the Metaharvester, which is a software robot that gathers selected information from designated websites. This tool gathers and parses XML and HTML, extracting

designated elements such as HTML tags and META tags (e.g., title, body, meta, etc.). The robot will allow CIIMMS to include web-page information in its metadata database and search engine.

During the Initial Production Phase, a map server component will be added to provide the map-based metadata search function, mentioned previously. This tool will display local GIS layers, selected boundary files (HUC, Quad, Borough) and a digital gazetteer to generate place name keywords and bounding rectangle coordinates. These criteria will be passed to the Blue Angel Technology MetaStar Enterprise to facilitate the search and discovery of metadata.

Figure 7 shows a high-level schematic of the prototype system architecture. The system is structured into three tiers, as follows:

- Tier I: End user machine with an Internet browser, such as Netscape or Internet Explorer

- Tier II: Web Server, which houses the MetaStar Enterprise software for discovering and accessing metadata and information from all levels of the CIIMMS Information Pyramid, and the Map server, which provides the map-based metadata search function. This function displays local GIS layers and a digital gazetteer. The map presentation and GUI allow for bounding coordinates, place names and subject keywords to be passed to the MetaStar Enterprise software.

- Tier III: SQL Server, which houses the CIIMMS metadata, projects, and contacts database.

Identifying and accessing metadata and information will be accomplished by using the Z39.50 standard client-server protocol for information retrieval and for focused web crawling of non-Z39.50 sites that contain web pages and hypertext relevant to Cook Inlet. A map-based query tool will be provided to facilitate the search and discovery of metadata. The system architecture also will provide for online linkages to the data and information through data downloads (via ftp), document viewing via portable document formats (PDF) files, and image viewing through standard formats, such as graphics interchange format (GIF) files.

The CIIMMS system utilizes a hybrid centralized/distributed design, meaning that some CIIMMS metadata and information resources will be housed on the (centralized) CIIMMS server, while other resources will be housed on servers located elsewhere (e.g., at different agencies) throughout the watershed. Metadata collected using the CIIMMS on line contribution tools will be centralized, on the CIIMMS server. However, data housed on CIIMMS in the early stages will eventually be migrated toward a more distributed design. CIIMMS aims to help agencies and organizations become more successful at providing access to their own data and information using emerging web technologies, by providing web-published guidelines and technical support. This type of guidance and support will enable migration toward a more distributive system. To test the

distributive concept, a few distributed sites are currently accessible via CIIMMS. These pioneer sites will provide lessons learned and expertise to organizations interested in providing data and information access via CIIMMS.

Figure 7. CIIMMS high-level system architecture.
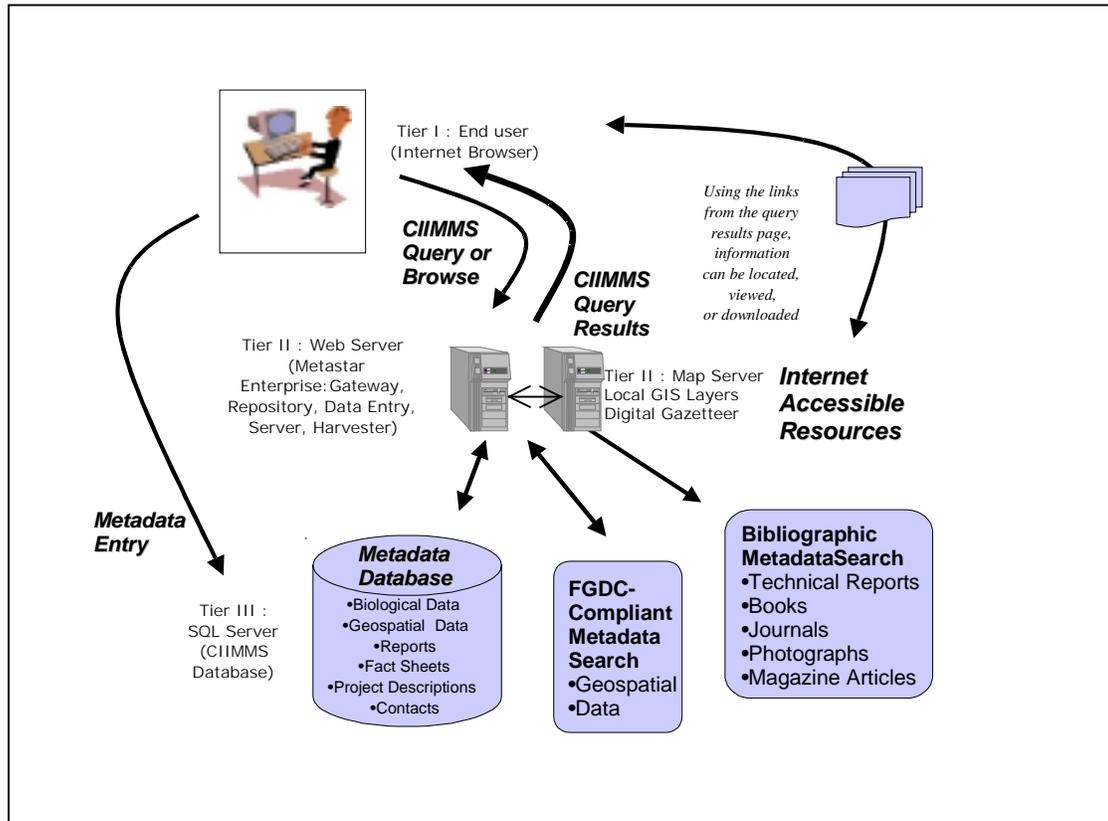
LESSIONS LEARNED

The process of accessing information and building a distributed network of data and information providers, via the web, is iterative. The speed with which technology is advancing, and the building of interest from multiple sectors, necessitates the practice of adaptive management and sliding targets. As the CIIMMS project has evolved, the Team has narrowed the scope to address specific needs of Cook Inlet users, related to data and information access and discovery, as opposed to the actual integration of data, which relies heavily on data standards.

Technical roadblocks were small compared to cooperation and coordination issues. Data dissemination by agencies on a broad scale is happening more slowly than was perceived at the beginning of the project, and reluctance to share certain types of data is still widespread among data providers.   Although participating Cook Inlet stakeholders agree that an information system that reaches out and accommodates the needs of the wide range of users throughout a watershed is a good idea, there currently is no long-term

commitment to maintain CIIMMS, outside of the State agencies who initiated the project. There is also no charter or MOU in existence for the CIIMMS project, although the Project Team hopes to initiate those actions in the coming year.

ACKNOWLEDGEMENTS

The Project Team is responsible for ensuring that the implementation of CIIMMS fulfills the CIIMMS vision. This includes the following activities:

- Implementing all phases of CIIMMS, starting with the pilot phase
- Soliciting input on user needs and considering user feedback

REFERENCES

SAIC (Science Applications International Corporation). 1998. "SAIC Metadata Entry Tool Survey Results." Submitted to EPA's Chesapeake Bay Program. March, 1998.

SAIC. 1999a. "Cook Inlet Information Management Requirements and Options—Draft Workshop Report." Submitted to Alaska Department of Natural Resources and Alaska Department of Environmental Conservation. January, 1999.

SAIC. 1999b. "Cook Inlet Information Management Requirements and Recommendations: Results from the User Needs Workshop." Submitted to Alaska Department of Natural Resources and Alaska Department of Environmental Conservation. March, 1999.