

Chapter 13

Communicating Results: Data and Information Management

Priority Problem

One of the limitations of estuary monitoring systems across the country, including Galveston Bay, is that results from different monitoring programs are not easy to compile for ecosystem analyses. Agencies maintain different data bases and report formats, acquisition of data can be time-consuming, and no centralized data management system is currently available to report on overall trends. To alleviate these problems, a Data and Information Management System (DIMS) for Galveston Bay has been developed as an integral part of the Regional Monitoring Program.

DIMS Objectives

The Galveston Bay DIMS must operate on several levels. At one level, the program must be concerned with the management of a system which will accommodate the data to be generated throughout the Galveston Bay monitoring effort. On a higher level the program must work to facilitate exchange of a wide variety of data types between state, local governments and organizations, and federal agencies. Critical among these data types will be the development of retrieval and storage systems which will allow the exchange of geospatial information. To address these needs the Galveston Bay DIMS has been structured to perform the following functions:

- Ensure the long-term integrity, storage, and accessibility of data collected by Galveston Bay's Regional Monitoring Program,
- Ensure data quality,
- Improve the access to information at various decision-making levels,
- Facilitate the integration and analysis of existing physical, chemical, and biological data to generate information useful to resource managers,
- Support statistical, graphical, spatial analysis and mapping of monitoring data, (e.g. power analyses, computer-compatible geographic information system format) ,
- Facilitate access to large amounts of monitoring data from disparate data bases located throughout the state,
- Allow multiple levels of user access to raw monitoring data, data quality information, summary statistics, and maps
- Integrate geographic information system (GIS) functions with appropriate databases.

Summary of Agency Data Management Systems

There are 19 programs presently conducting monitoring in Galveston Bay. In most cases the data are stored 1) on in-house computers under a variety of formats, or 2) on paper. Although most data are made available to the public, access is often difficult. There is no central data storage system that would allow easier access for the public or the agencies presently concerned with monitoring Galveston Bay. Some duplication of effort is noted, particularly for point source monitoring. Most efforts are directed at fulfilling specific agency mandates and have not been geared to ecosystem scale assessments. Ward and Armstrong (1992) cite numerous challenges encountered in compiling 26 data sets for such an analysis. McFarlane, (1991a and 1991b) also documents monitoring deficiencies encountered in obtaining and compiling historical data sets.

To address some of these problems during the development of *The Plan*, several GBNEP projects were conducted to compile data sets from diverse sources and to allow easy exchange of existing bay information. For example, the Galveston Bay Information Center was developed to serve as a clearinghouse for all types of literature about the bay. A number of these same data sets were also distributed to the Texas Natural Resource Information Service (TNRIS).

State-wide Data Integration and Exchange Efforts

There are several existing database management systems (DBMS) and Geographic Information Systems (GIS) running on various platforms at the local, state, and federal agency level. The diversity of existing and planned DBMS applications, GIS applications, and hardware platforms at the agency level reveals a determined use of best available technology. However, these conditions have made it difficult for agencies to access, query, transmit, and analyze resource data in an efficient and timely manner. Getting different DBMS on different hardware platforms to communicate is technically challenging. Currently, no statewide computer network system exists to quickly and easily share data among local, state and federal resource agencies.

In 1989 the Texas Legislature enacted legislation which requires that state agencies share information and information resources. In the same year, the Department of Information Resources (DIR) was established to provide the leadership role in this area. The Texas Geographic Information Systems Planning Council was formed to coordinate an interagency effort to improve and expand the development of geographic information systems and to make recommendations to DIR concerning GIS policies to achieve this goal. Members include representatives from over 20 state and local agencies (Table 13-1). A number of committees and sub-committees have been formed to deal with issues such as: development or acquisition of geospatial data; data standards, including output format standards and spatial information standards; improving network data accessibility among member agencies; development of global positioning systems; remote sensing and TIGER updates. For example the Standards Committee has recently proposed standards and Guidelines for Geographic Information Systems in the State of Texas (TGISPC, 1992), which specifies standards related to :

TABLE 13-1. TEXAS GIS PLANNING COUNCIL MEMBERSHIP

Advisory Commission on State Emergency Communications (State 911)
Comptroller of Public Accounts
Department of Information Resources
General Services Commission
Legislative Council of Texas
Lieutenant Governor's Office
Office of Attorney General
Office of Court Administration
Office of the Secretary of State
Public Utility Commission of Texas
Railroad Commission of Texas
Secretary of State
Texas Department of Commerce
Texas Department of Criminal Justice
Texas Department of Human Services
Texas Department Of Transportation
Texas Education Agency
Texas General Land Office
Texas Historical Commission
Texas Health and Human Services Commission
Texas Natural Resource Conservation Commission
Texas Natural Resources information Service
Texas Parks And Wildlife Department
Texas Rehabilitation Commission
Texas Water Development Board

State Universities

The Bureau of Economic Geology, the University of Texas at Austin, Texas Agriculture and Mining University (TAMU)

Ex-officio memberships— Regional and Private Sectors

Texas Mapping Advisory Committee
Texas Association of Regional Councils
The Texas Association of Appraisal Districts

Ex-officio memberships— Federal Sector

US Geological Survey- National Mapping Division

- Cartographic standards,
- Data dictionary,
- Data interchange, and
- Data layer classification.

The GIS Policy Council will continue to work to provide the leadership at the statewide level to assist in the development of GIS technology and data networks for the cost-effective development of geospatial data applications. Currently, statewide efforts are in early states of planning, with no computer network system existing to quickly and easily share data among resource agencies and

organizations. Because there is no existing system which would accommodate GBNEP's data information needs, GBNEP has planned a regional DIM system that conforms to existing and planned local, state-wide and agency data information management plans wherever possible.

Current Activities in Data Networking

Several pilot and developmental programs are currently underway in which the Galveston Bay DIM system could participate in the design and implementation of a state-wide data integration and sharing system.

One such pilot program currently under development is the Wetland Resource Database which has been developed as a joint project of the Texas GLO and the Texas Natural Resources Information Service (TNRIS). This project funded by a grant from the USEPA-Region 6 Wetlands Program is a distributed data management model which uses InterNet connections indexed on Mosaic software. This system provides real-time connections to client state and federal agency databases. This one year program is currently completing a peer review process and has a prototype server running on a limited basis. Clients currently participating in this pilot program are GLO, TPWD, TNRCC and DIR on the state level and the USEPA Region 6 office. The NMFS and USFWS are limited users of the system.

If funding is approved for the second phase of this project, efforts will be directed at a more formal implementation of the network: including formalizing current connections; formalizing TNRIS as the hub of the network; and continuing to work on resolving identified problems. Included in the new workplan will be plans to broaden the scope of involvement in the project. Specifically targeted groups include the Galveston Bay and Corpus Christi Bay National Estuary Programs and coastal universities and libraries.

Another effort currently underway is the Gulf of Mexico Information Sharing Network for Ecological Protection. Through this program the GBNEP will be provided computer hardware and software support to link the Gulf NEP's through the InterNet. This will be a network designed to address sharing of information pertaining to ecological protection which will allow Gulf of Mexico NEPs and the Gulf of Mexico Program to share information with citizens, universities and other governmental agencies.

Design of Galveston Bay DIMS

Overall Systems Design

To meet the demands of identified uses, a distributed data management model will be used to develop the DIM system. Such a system will allow integration of Galveston Bay monitoring data and will allow access to available information from various other data systems. As Figure 13-1 illustrates, a distributed data management model assumes that data are maintained in several remote databases, which are linked through a network. Each participating agency maintains regional

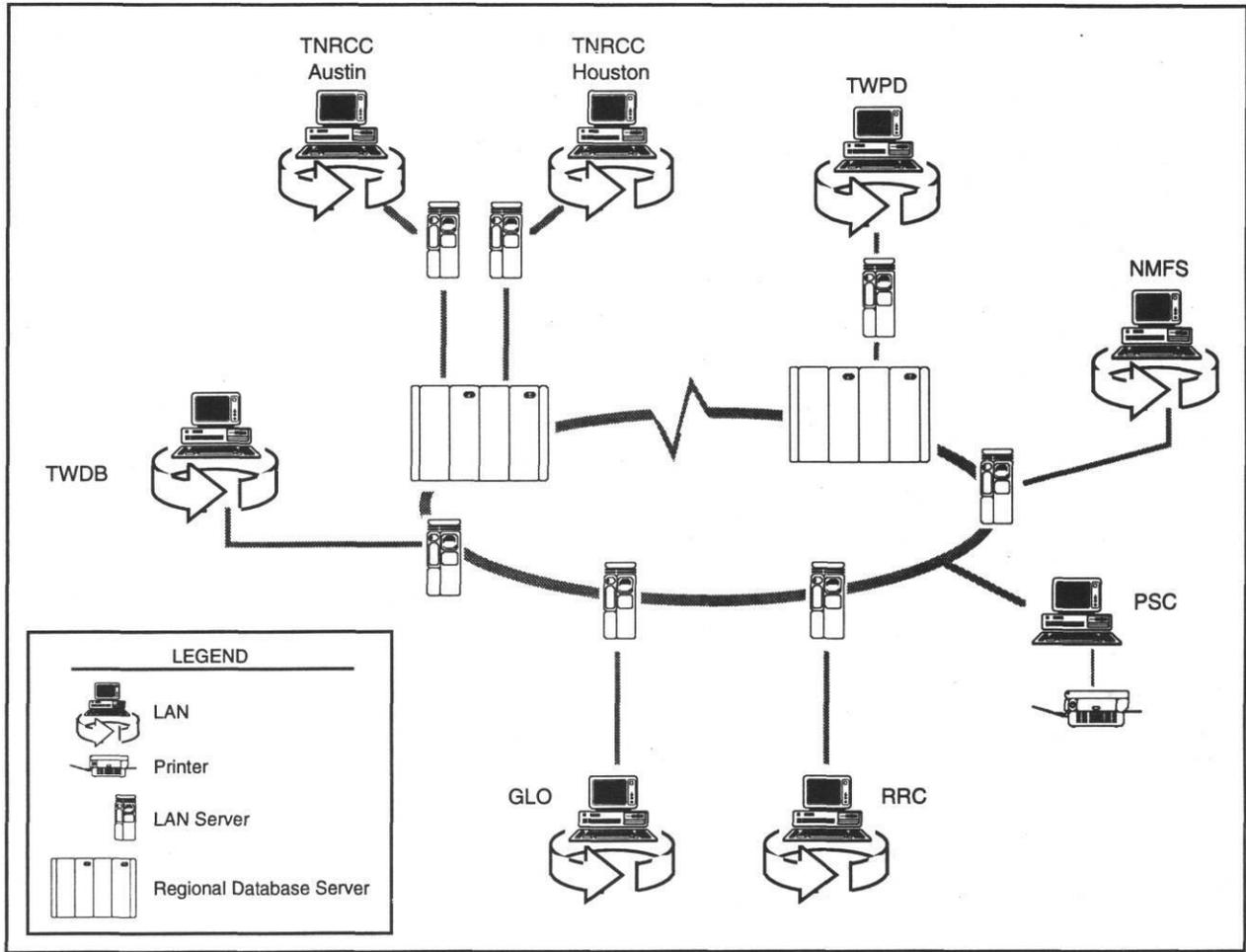


Figure 13-1. Distributed Data Management Model. (For illustrative purposes only, not intended to show all participating organizations.)

monitoring data that it has collected on its own database server. All participating agency database servers are linked over a WAN. A centralized data dictionary lists the location and contents of distributed databases.

Advantages and disadvantages of distributed systems are listed in Table 13-2. Standardization of data and information management protocols and communication among database managers are crucial to this strategy's success since distributed databases tend to diverge—in structure and function—in response to diverse needs of the primary agency.

Local Network Design

The local element of the Galveston Bay DIM system will be implemented in phases. In the first phase, Galveston Bay's DIMS will emphasize adoption of standardized data file structures and implementing data storage and retrieval from a centralized system. Centralizing data in one database server will simplify the tasks of storing, maintaining, locating, querying, and retrieving regional monitoring data. Under

TABLE 13-2. ADVANTAGES AND LIMITATIONS OF DISTRIBUTED DATA MANAGEMENT MODELS

Advantages

- Data are maintained by those who are primarily responsible for their collection and use
- Data are “local” to those who use the data most often— on average, 90% of the time data are used by local users; 10% of the time data are used by other users
- Design facilitates timely data checks by persons who are most familiar with the data and associated common data errors
- Design allows local autonomy and facilitates rapid system evolution in response to user needs

Limitations

- Greater potential for multiple versions of the data leading to loss of data integrity and inconsistent analyses
 - Extensive transformations of data formats are usually required before data may be analyzed — NEPs have found that up to 40% of the data analysis budget is spent transforming data
 - To date, environmental data have not been readily accessible, which has led to long delays in data analyses and reporting
 - Costs may be higher due to maintenance of multiple systems and staff, and re-standardization of diverging distributed systems
-

this DIMS the centralized data base will be housed at the Houston-Galveston Area Council (H-GAC). H-GAC has been named as the regional provider in the state data information system and is responsible for conducting of the Texas Clean Rivers Program (TCRP) in the Galveston Bay area. Through the TCRP, H-GAC is responsible for completing a comprehensive assessment of water quality in the basins surrounding Galveston Bay. One component of the TCRP is to act as a central clearing house for water quality information. In this role, H-GAC will serve as custodian and repository for all Galveston Bay regional monitoring data.

Later phases will include modifying the system to include linking all local participating agency database servers through the distributed type data management system currently being developed by the DIR. In this system all participating agency servers are linked through the InterNet. The H-GAC server will serve as the link from the local network to the statewide WAN. In this system, as in the previously described local network, each participating agency maintains monitoring data that it has collected on its own database server. A central index or data dictionary at H-GAC will list the location of custodial databases of interagency interest. The inquiring agency's server will locate and retrieve requested data from

the appropriate data base server. Examples of data that may be accessed through this system are core base maps developed by other state/federal agencies, TPWD wetland classification maps, state-wide digital orthophoto quarter-quads, regional land-use maps from the GLO or other agencies, the TNRCC's state Surface Water Quality Assessment Data base, and other environmental data.

Network Architecture

Network media will be required to connect database servers regardless of the data management model selected. Media that link remote agency databases range from telephone lines to satellite networks. Establishing a network requires more than a cable linking two or more servers— choosing the appropriate network architecture also involves considering factors such as distance, amount of data transferred, transmission speed, and cost (Figure 13-2).

The network architecture of the Galveston Bay DIMS is also planned to be implemented in a phased approach. The initial phase will be a direct link between the H-GAC and the Program Office. The link will utilize integrated Services Digital Network (ISDN) technology. The ISDN program digitizes the telephone system and eliminates analog voice lines. This process divides the available bandwidth into three data channels: two move data at 64 Kbps and one moves data at 16 Kbps. A pilot project to evaluate such a connection is currently underway. This pilot involves establishing direct links between the H-GAC, Texas Department of Transportation- District 12 office, and the Metropolitan Transit Authority (Metro). This pilot will serve as a model for the H-GAC - Galveston Bay Program Office link.

Such a direct link would be critical in the Program's development of GIS capabilities. GIS files are usually large and require extended transmission periods to transfer the data and would require large amounts of storage at the program for these coverages if operated on a file transfer basis. Development of GIS capabilities is seen as an important tool, but personnel and budgetary requirements to develop and maintain comprehensive GIS analytical capabilities would be prohibitive at this time. Development of a direct link to H-GAC will enable the Program Office to access the broad range of GIS capabilities and extensive GIS coverages already in place at H-GAC on a "real-time" basis. H-GAC runs Arc/Info[®] GIS software on UNIX workstations. To take advantage of the H-GAC GIS system the DIMS recommends Arc/Info[®] ARCVIEW 2 software electronically linked to the main server at H-GAC as the level of entry into GIS. ARCVIEW 2 will allow the integration and manipulation of intermediate and final GIS products needed for this program without the prohibitive costs of operating and maintaining a comprehensive GIS system.

The next phase of the DIMS network architecture to be developed involves providing access to other monitoring agencies in the Galveston Bay area to the Galveston Bay DIM system. This access will be provided initially through dial-up connections using slip technology to increase data transmission rates. If remote data access activity increases dramatically, dedicated or direct lines will be established to link data base servers. This link will use the InterNet to network the servers and their associated data bases.

Another consideration, for this program, is the availability of external sources of information and the network that will best provide that access. Two of the existing network systems previously discussed, the Wetland Resource Database and the Gulf of Mexico Information Sharing Network, utilize the InterNet for data transmission and retrieval. Connectivity to the InterNet and hardware/software to operate on the InterNet will be sponsored by the Gulf of Mexico program. Having InterNet Connectivity will also provide access to the Wetland Resource Database.

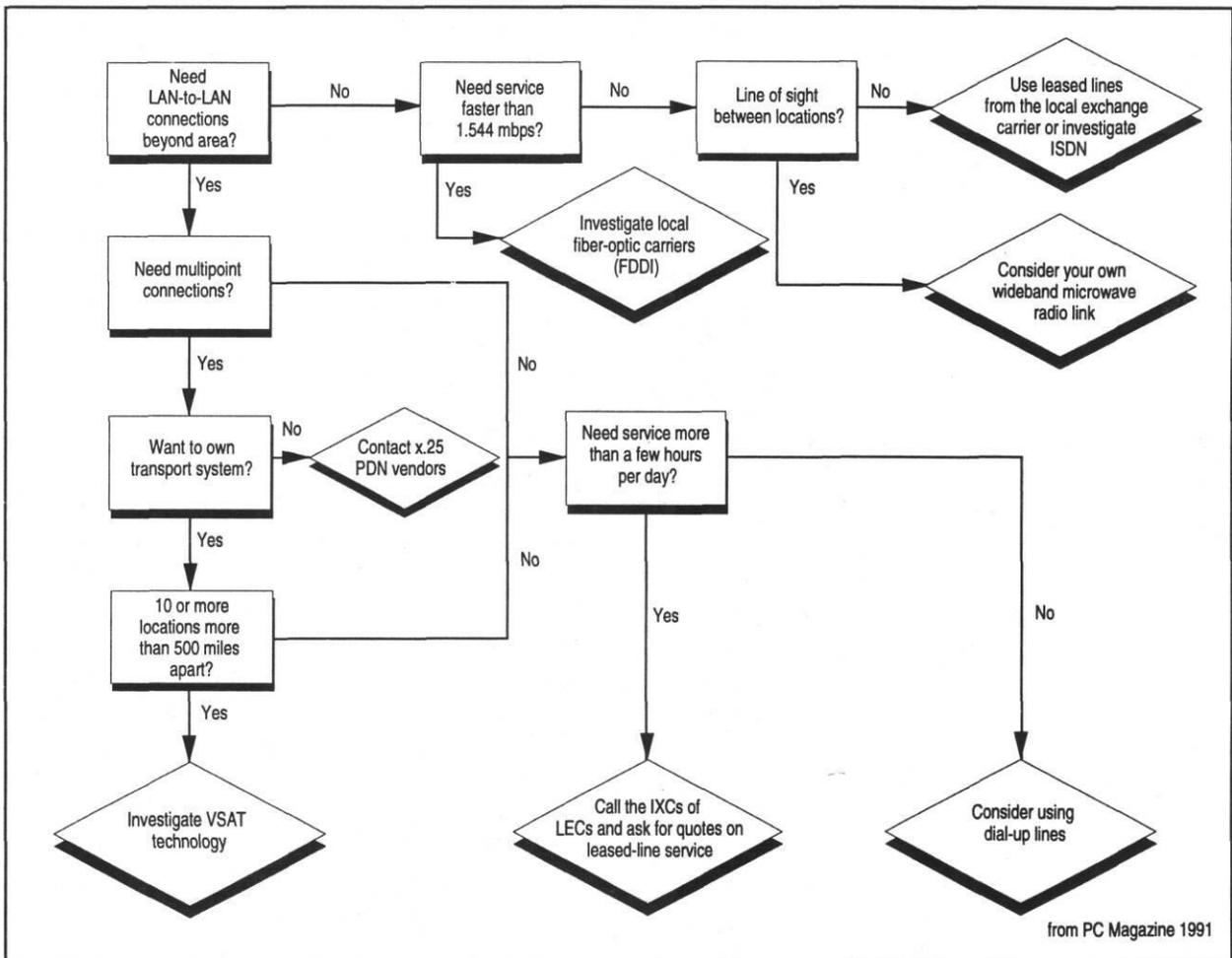


Figure 13-2. Decision chart for selection of network media.

DIMS Systems Administration

Galveston Bay's DIMS Steering Committee will be responsible for overseeing the implementation of the DIMS, including; approving all DIMS system modifications, securing future funding sources, and making data management recommendations to H-GAC and the Galveston Bay Program office. The H-GAC will be responsible for

administration of the Regional Monitoring Data system. These responsibilities will include:

- Providing DBMS technical support to agency database managers and system users
- Designing and implementing user interfaces and/or other system applications
- Securing sources of financial support for the system

Additional specific responsibilities will include supervision of system quality assurance and implementation of system upgrades. Future responsibilities will include; implementing and maintaining the WAN and implementing and managing Public Service Centers (PSCs). It is recommended that the H-GAC become a member of the Texas GIS Planning Council, GIS Standards Committee and GIS Managers Committee.

Database Server Managers

In the future phases of DIMS implementation, as other agency servers are linked to the system, agencies will have certain responsibilities. Agencies with primary responsibility for housing regional monitoring data are responsible for updating and adding new data sets to their database server system. Database server managers are responsible for conducting standard data QA/QC checks established as part of the DIM strategy. They are also responsible for ensuring that any upgrades of their DBMS does not disrupt transparent querying and access to regional monitoring data stored on their database server. Furthermore, database server managers are responsible for correcting and updating data sets as specified by the submitter.

It is highly recommended that the H-GAC and all database server managers consistently meet to review system maintenance activities. The systems administrator and servers managers will produce an annual report describing:

- Present status of the system
- Problems encountered and how they were resolved
- Next year's proposed goals and how they will be achieved
- Estimated maintenance and enhancement costs

Data Types

The DIMS system will support the following data types:

- Discrete and continuous numeric monitoring data
- Nonparametric monitoring data (e.g., presence/absence data)
- Text or memo formats
- Maps and charts, i.e., geographically referenced data.

Sufficient information (i.e., metadata) must be associated with the monitoring data to ensure that secondary users can correctly use and interpret the data. These metadata include:

- Quality assurance/Quality control data (e.g., blanks, spike recoveries),
- Measurement units (e.g., mg/kg, ug/l)— not ppm or ppb which are ambiguous,
- Detection levels for chemical data (e.g. minimum analytical levels), and
- Data qualifiers such as “non-detected” and “not analyzed”.

The use of QA abstracts will be implemented which will be directly linked to the monitoring data. Each data QA abstract will summarize information that secondary users need to know when deciding the value of a particular data set, such as contact person, date of survey, list of stations, sampling methods, analytical methods, summaries of QA/QC data, and a brief description of important or anomalous conditions pertinent to the collection of the data.

Core Base Maps— Galveston Bay’s DIMS system will store core base maps in a central location in accordance with the state GIS Planning Council’s recommendation. Participating agencies may request copies of specific core base maps and have them mailed on magnetic disks or optical CDs if electronic transmission is not possible. Core GIS data sets will be kept and maintained on H-GAC servers.

Standard File Structures and Formats

Currently, participating agencies store the same type of data in dissimilar file structures making it difficult to transparently query and retrieve data. The Galveston Bay DIMS will overcome this difficulty through the use of standard file structures and a standard database access interface.

A standard file structure will be established for each data type (e.g., water quality data, population abundance data, toxicity testing data). The Galveston Bay monitoring Work Group working together with all participating agencies will develop or adopt a standard file structure for each data type. Examples of information to be developed are:

- Numbers and names of data fields
- Appropriate field formats (e.g., numerical, alphanumeric),
- Key fields that link relational databases
- Data codes

Database Queries/Transfer

Data Queries— Currently, different agencies use different Data Base Management Systems (DBMS) running on different platforms making it difficult to seamlessly query the database. Initially, the centralized data base will utilize a standard documented database access vehicle such as Standard Query Language (SQL) for data queries. In later phases, the distributed databases will all use the same access vehicle.

Data Transfer— The system administrator will work with the data managers at each participating agency to adopt the standard DBMS or develop translation programs which will 1) translate from agency file structures to Galveston Bay’s

standard file structure for data storage and 2) provide translation from Galveston Bay's standard structure to agency file structures for retrieval (Figure 13-3). This will allow agencies to manipulate, analyze and display all data residing in the central database initially and distributed databases ultimately using familiar software applications available through their agency.

Data Transfer Formats— The database must have the capability to download data easily to other data analysis and presentation programs. The Texas GIS Standards Committee recently recommended that Spatial Data Transfer Standard (SDTS) format be used to facilitate the exchange of GIS information (TGISPC, 1992). SDTS format will be adopted in accordance with this committee's recommendation. Other GIS standards developed by the GIS Standards Committee will be incorporated into Galveston Bay's DIM strategy as appropriate.

Currently, there are no standard data transfer formats for parametric data. The GIS Standards Committee has recommended that flatfile, non-compressed ASCII format data interchange be employed for the transfer of parametric data.

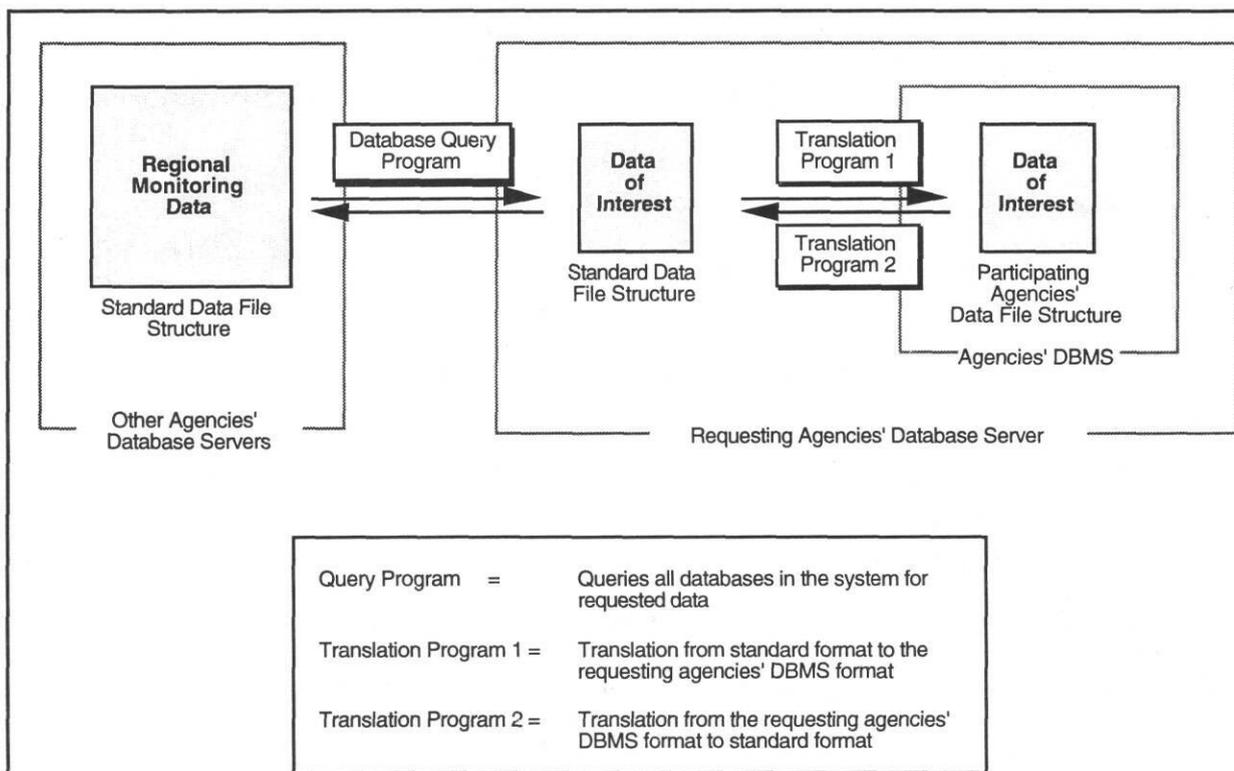


Figure 13-3. Three types of data processing programs.

Data Quality Assurance / Quality Control (QA/QC)

Data Accessibility— Maintaining the integrity of data stored in the system is critical to ensuring user confidence in the system. Data stored in the DIM system can be queried, read, copied, and downloaded to workstations for local manipulation and analyses. Initially this will be available only through requests to H-GAC or the Program Office. As direct access to the system is expanded users of the system will have read only access. Additions and updates to the data will only be made through standard quality control protocols established at the outset of the system's implementation.

Data Submission— Standardized procedures for checking submitted data will include: computerized code and range checking, technical data review, and preparation of a data QA abstract. The data abstract is for describing sampling and analytical methods, QA/QC information, and any other pertinent metadata information needed to assess data quality. Data submitters are responsible for compiling regional monitoring data, conducting data QA/QC checks, and submitting both their data and data QA abstract to the H-GAC for entry into the Galveston Bay database.

Computerized code and data range checks will be performed on the data prior to its entry into the central database. Any errors or discrepancies will be resolved before data is loaded onto the system. Users will not be permitted to make *ad hoc* modifications to data stored in the system. Additions and updates to the data will only be made through standard quality control protocols established at the outset of the system's implementation.

Communicating Monitoring Results

The system will support the following user groups:

- Galveston Bay Council,
- Technical and scientific staff of participating agencies,
- Technical and scientific staff of non-participating agencies,
- Private industry,
- Public interest groups,
- Schools, and the
- General public

Technical staff of non-participating agencies, private industry, public interest groups, schools, and the public will have access to raw monitoring data. In addition there will exist the ability to retrieve selected summary statistics and display these data on core base maps. Requests for data will be handled through H-GAC or the Program Office. In the future the Program Office will support evaluation programs for development of direct link public service centers (PCSs) at strategic locations. These will be established to support direct access to the system for non-participating agencies, private industry, public interest groups, schools, and the general public. Selected bay-wide summary statistics will be available to the public

for downloading at the service centers. Requests for raw regional monitoring data will continue to be made through the H-GAC or the Galveston Bay Program Office.

Information from The Galveston Bay Regional Monitoring Program will be available in two formats: technical reports for the scientific community and non-technical briefs for the lay public. Programs will be written to automatically conduct the appropriate data queries, data retrievals, data analyses, and data presentations (e.g., graphs, maps). Ad hoc analyses may be used, as needed, to supplement these fundamental data analyses. A set of most requested bay information may be published in an annual report as well as made available on-line in the system.

The primary purpose of the DIMS is to provide data which can assist in establishing the link between management goals and objectives and environmental results. The Galveston Bay Program Office will be responsible for evaluating and analyzing the results from the monitoring program as they relate to Plan goals and objectives. Evaluation of monitoring program results will provide information for feedback to the program on two levels. Such evaluations of the data will establish whether the monitoring program is providing the expected information for assessing plan actions and if changes to the program are needed to obtain the necessary information. Secondly, evaluations will determine if data supports a conclusion as to whether resource management goals are being met.

This information will be disseminated to provide information to resource managers, scientific and technical sources, and the general public. Publications such as the Galveston Bay *Bay Line*, a quarterly newsletter published by the Galveston Bay Program, will inform the public on Bay issues and generate public interest and support for program initiatives. This information will also be made available to the scientific community in technical publications and through papers and poster presentations at scientific and technical meetings.

To supplement publications and to allow additional data presentation the Galveston Bay Program will continue to host the biennial Galveston Bay Symposium. The goals of this symposium are to: identify Bay projects being conducted by institutions other than the Program; to promote peer interactions among scientists involved in this research; to improve our understanding of estuarine problems; and to encourage project coordination in an ecosystem context. (GBNEP, 1993)

Bay Barometer

A potential tool for communicating the status of the health of the ecosystem is the development of a Galveston Bay barometer similar in approach to the one used in the Chesapeake Bay Program. For example, such a barometer could include several easily measured components of the Galveston Bay system. This information would be published as a regular feature in the Galveston Bay *Bay Line*. The Bay Barometer concept will be developed by the Galveston Bay Program Office.

Sources of Financial Support

Several candidate sources of financial support for the implementation and maintenance of the DIMS system have been identified, including

- Corporate sponsors,
- Other state-wide or gulf-wide data sharing projects, and
- Line-item support from the State.

Private sector corporations and software/hardware vendors will be sought out to sponsor, in part, the cost of hardware and software needed to implement the DIMS system. Candidate corporate and vendor sponsors will be aggressively pursued by the chair of the Steering Committee, H-GAC, and the Galveston Bay Program Office.

Furthermore, the implementation of Galveston Bay's DIMS system may be partially supported through other state-wide or gulf-wide data management projects including the Coastal GIS Initiative, the Natural Resources Inventory, the Texas Clean Rivers Program, and the Gulf of Mexico Program. The DIMS Steering Committee, H-GAC, and the Galveston Bay Program Office will pursue all state-wide and gulf coast financial sources. They will aggressively approach the legislature to seek line-item status for Galveston Bay's DIMS system.