

APPENDIX E: DATA MANAGEMENT PLAN

Sabine River Authority of Texas

Information Resources Department
Information Management Plan
FY 2002 -- 2003

PREFACE.....55

I. PLANNING.....56

- A. Data Inventory56
 - 1. Databases.....56
 - 2. Internet Data Resources59
 - 3. Geographic Information System (GIS) layers62
- B. Personnel66
 - 1. Data Resources Manager66
 - 2. Website Administrator67
 - 3. Data Analyst.....67
 - 4. GIS Analyst.....68
 - 5. Information Resources Support Specialist68
- C. Proposed Personnel Needs.....68
- D. Data Needs69

II. INFORMATION MANAGEMENT SYSTEM DESIGN.....69

- A. Computer Hardware and Software Requirements69
- B. Database Table Documentation.....75
- C. Timeline.....76
- D. Budget76

III. INFORMATION MANAGEMENT PLAN IMPLEMENTATION.....76

- A. Functions for acquiring data sets76
- B. QA/QC on acquired data sets77
- C. Report generating capabilities77
- D. Overview of services currently available on the World Wide Web site.....77
- E. Services planned for the World Wide Web site84

IV. INFORMATION RESOURCES ADMINISTRATION84

- A. Security.....84
- B. Software Management85
- C. Quality Assurance/Control85
- D. Programming Processes86
- E. Genealogy86
- F. Documentation.....86
- G. Archives/Backup86
- H. Disaster Recovery.....87
- I. Migration/Transfer/Conversion87
- J. Client Service87

Appendix 189

Appendix 290

Appendix 391

Appendix 492

Appendix 595

PREFACE

The Sabine River Authority Information Resources Department (IRD) Information Management Plan (SRAIMP) is designed to document the Information Resources Information Management System as it exists at Sabine River Authority and to target areas for growth and future development.

The SRAIMP is the reference for Information Resources information management at SRA and is thus a living document. The document and related appendices are be posted on SRA's World Wide Web (WWW) site, <http://www.sra.dst.tx.us> and revised as the Information Management System evolves.

SABINE RIVER AUTHORITY OF TEXAS INFORMATION MANAGEMENT PLAN

I. PLANNING

A. Data Inventory

1. Databases

SRA's FY2002-2003 Texas Clean Rivers Program (TCRP) work plan committed SRA to maintain a watershed inventory of a specific list of data sets. Each listed data set and the status of each is listed below:

- *Surface water quality monitoring data from all available sources*
Surface water quality monitoring data, located at \\SRA_IR1\VOL3\SRWMP\SWI\WQ\SWQM.MDB, contains the tables tblDscrptns and tblData. The structure of these two tables is based on the original tables SWEVNT05 and SWRSLT05 and GSEVNT05 and GSRSLT05 received from the Texas Natural Resources Conservation Commission (TNRCC) in 1993.

The database is updated monthly with new water quality data from SRA. Monthly updates may also include data from Sabine Basin entities covered under the SRA's Quality Assurance Project Plan (QAPP). Periodic updates are received from TNRCC and these are also appended to the database. Update files from TNRCC are archived in \\SRA_IR1\VOL3\SRWMP\IMPORT\SwqmUpdtes\<DatedDirectory>. SWQM.MDB currently contains data from as early as 1968 and is current with the last completed data set from the Environmental Services Division (ESD) Laboratory.

Reference tables associated with SWQM.MDB are found in \\SRA_IR1\VOL3\SRWMP\SWI\COMMON\LIBRARY.MDB. These reference tables include tblStns, LatLong, SW_PARM, and SW_ParmCodesAsText. Tables tblStns and LatLong are based on the table SWQM_STN originally received from TNRCC and SW_PARM and SW_ParmCodesAsText are based on the original TNRCC table SW_PARAM.

Tables LatLong and tblStns are updated when new station identification nuMBers are requested and received from TNRCC. SW_PARM and SW_ParmCodesAsText are updated when a new SW_PARAM table is requested and received from TNRCC.

Data Dictionaries for tblDscrptns, tblData, tblStns, LatLong, SW_PARM, and SW_ParmCodesAsText are found in Appendix 1.

Water quality data from SWQM.MDB can be accessed through SRA's World Wide Web site at <http://www.sra.dst.tx.us/data/wq/swqm/> and is updated at the same frequency as the source data.

- *Special investigations of spills, fish kills, complaints*

A Special Investigations database for the storage and access of information about complaints, spills, fish kills, biological or botanical surveys has been designed and population of the database continues. The master database series is located at:

\\SRA_IR1\VOL2\SHARED\DATA\SUBWSINV\SPECINV.

The database series consists of:

SIENTRY.MDB: data entry, report generation, data export; access to data via linked tables

SILIB.MDB: reference tables

SIEXPORT.MDB: tables designed to export data

SPECINV.MDB: the data

Data dictionaries of the tables in the Special Investigation databases are found in Appendix 1.

- *Events affecting water quality*

The Events database stores information about events that have occurred in the Basin that could have impacted the water quality. Such events include drought, flood, spills, issuance of new permits, construction of new waste water treatment plants, and changes in legislation to name a few. Population of the database has been limited due to shortage of personnel and time to do so.

\\SRA_IR1\VOL2\SHARED\DATA\SUBWSINV\EVENTS\EVENTS.MDB.

Data Dictionaries of the tables in the Events database are found in Appendix 1.

- *Landuse information*

Landuse information exists as GIS coverages instead of in database format. These coverages, obtained from the USGS and the EPA, are maintained on Internet GIS Clearinghouse (<http://www.sra.dst.tx.us/data/gis/>).

- *Municipal and industrial permits*

The municipal and industrial permits database, located at

\\SRA_IR1\VOL2\SHARED\DATA\TNRCC\DISCHARGE\DISCHARGE.MDB, contains tables WWFacNoDups, WWEXTNoDups, WWPARAM_NoDups, WFlocNoDups, WxLoc, WWPteeNoDups, and WxRqtCorrNoDups. The structure of these tables is based on the original Paradox tables received from TNRCC in 1993 with only minor changes and corrections.

Periodic updates to this database are requested from TNRCC and these are appended to the database. Update files from TNRCC are archived in

\\SRA_IR1\VOL2\SHARED\DATA\TNRCC\DISCHARGE\DischrgUpdtes\<DatedDirectory>. Data Dictionaries for the tables in DISCHARGE.MDB are found in Appendix 1.

Certain information from DISCHARGE.MDB can be accessed through SRA's World Wide Web site GIS Clearinghouse and is updated at the same frequency as the source data.

DISCHARGE.MDB is a difficult database to use because of the large number of tables, the large number of records in each table, the database is not normalized, many of the fields appear to be filled with erroneous data, and some of the data imported from TNRCC and the documentation originally provided by TNRCC is out of date.

- *Stormwater permits*
A stormwater permits database is not currently available, but a request for available Basin 5 data from TNRCC is planned.
- *Population data by basin, subwatershed, and municipality*
SRA originally received 1990 Census data for Basin counties in printed report format. In addition, SRA also has obtained 1999 population estimates. Population data is found at \\SRA_IR1\VOL2\SHARED\DATA\TNRCC\POPLTN\POPDMS.MDB.

Information regarding data dictionaries for the tables in this database are found in Appendix 1.

- *Petroleum storage tanks that could lead to pollution*
SRA received petroleum storage tank data from TNRCC in February of 1996 in the form of Paradox tables named P_AST, P_FAC, P_OWNER, and P_UST. These tables were imported from the Paradox tables into Access tables of the same name. The table structures of all were conformed to the structure outlined in PST_STRU.TXT, a text database dictionary. The database dictionary did not mention, however which field(s) were the keys so none were assigned

The database is found at \\SRA_IR1\VOL2\SHARED\DATA\TNRCC\PST\PSTDATA.MDB with the data dictionaries for these tables included in Appendix 1.

Updates to these tables are to be requested from TNRCC.

- *Solid waste information is available for household solid waste and industrial solids waste management facilities.*
Solid waste management facility data tables are found in \\SRA_IR1\VOL2\SHARED\DATA\TNRCC\SLDWST\SldWst.MDB. The tables in this database include information for counties outside the Sabine Basin and many fields have no data. Due to these problems, this data set has limited usefulness.

Data dictionaries for the tables in SldWst.MDB are found in Appendix 1.

Updates to these data sets along with the accompanying data dictionaries are to be requested from TNRCC.

- *Superfund*
SRA does not currently maintain a database of Superfund sites in the Sabine Basin but rather uses EPA's and TNRCC's databases which are accessible via Superfund site at <http://www.epa.gov/superfund/> or via the Surf Your Watershed site at <http://www.epa.gov/surf/> as well as the Texas Superfund Program site at <http://www.tnrcc.state.tx.us/permitting/remed/superfund/>.

- *Water rights permits*
A data table containing information about water rights permitted diversions in the Sabine Basin is found in \\SRA_IR1\VOL2\SHARED\DATA\TNRCC\WTRRIGHTS\WTRRIGHTS.MDB. The database contains the table ZDRVNPTS which is a derivative from the original data table received from TNRCC. A data dictionary for table ZDRVNPTS is found in Appendix 1.

An update to this data set along with accompanying data dictionary is to be requested from TNRCC.

- *Wells that could lead to pollution*
Although this data set is available from the Texas Railroad Commission (RRC) for a fee, SRA has attempted to obtain it through TNRCC at no charge under the data sharing provisions of Senate Bill 818, but has not been successful.
- *On-site disposal facilities*
SRA does not currently have a database of on-site disposal facilities. Under Task 3 of the 1998 TCRP workplan, SRA will begin acquiring this data for selected sites in Subwatershed 1.02 (Cow Bayou, Segment 0511), Subwatershed 1.03 (Adam's Bayou, Segment 0508), and Subwatershed 7.07 (Cowleech Fork, Segment 0507)

Data planned for incorporation into the Sabine River Data Clearinghouse include:

- *Receiving Water Assessments (RWA)*
This data is currently available and those parameters with storet codes are currently incorporated into the SWQM.MDB database. Presentation of the remainder of the data and interpretation of the results is planned for addition to the World Wide Web site.
- *Rapid Biological Assessment (RBA)*
This data is currently available in paper format. The development of an appropriate database structure to facilitate presentation of the data on the World Wide Web site is planned.
- *Ambient Aquatic Toxicity (AT)*
This data is currently available and those parameters with storet codes are currently incorporated into the SWQM.MDB database. Presentation of the remainder of the data and interpretation of the results is planned for addition to the World Wide Web site.
- *Water quality data from cities and industries in the Sabine Basin*
SRA has received data from the City of Kilgore and Texas Eastman. SRA currently receives data from the City of Longview. The data received from these entities is included in SWQM.MDB. Data from Texas Utilities and Dallas Public Utilities is available on paper. Electronic copies of this data as well as data from other entities in the Basin is to be requested and incorporated into SWQM.MDB as appropriate and used for development of the events affecting water quality database.
- *GPS verification of existing locational information*
Beginning with on-site disposal sites (septic systems), SRA will begin population of data tables of GPS locations. Other site verifications will include permitted wastewater outfalls and solid waste facilities. Existing data for these sites must be provided by TNRCC before GPS verification can begin.

2. Internet Data Resources

The explosive growth of the World Wide Web in the past several years has allowed SRA to utilize the Internet to access, share, use, and provide data from other entities without having to maintain copies of the data in SRA's data repository. The many links to water quality and related data on SRA's Web site (<http://www.sra.dst.tx.us/>) include the following:

<i>Site</i>	<i>Address</i>	<i>Description</i>
EPA Index of Watershed Indicators	http://www.epa.gov/iwi/	The Index of Watershed Indicators (the IWI or Index) is a compilation of information on the "health" of aquatic resources in the United States.
EPA Surf Your Watershed	http://www.epa.gov/surf2/	A service to help locate, use, and share environmental information
Federal Geographic Data Commission Clearinghouse	http://www.fgdc.gov/clearinghouse/clearinghouse.html	The Clearinghouse is a distributed discovery mechanism for digital geospatial data. Using the data elements defined in the Content Standards for Digital Geospatial Metadata, governmental, non-profit, and commercial participants worldwide can make their collections of spatial information searchable and accessible on the Internet using free reference implementation software developed by the FGDC. SRA is a node in the Clearinghouse.
National Response Team	http://www.nrt.org/	The National Response Team and the Regional Response Teams are the federal component of the National Response System (NRS). The NRS is the federal government's mechanism for emergency response to discharges of oil and releases of chemicals.
Texas Natural Resource Conservation Commission	http://www.tnrcc.state.tx.us/	The Texas Natural Resource Conservation Commission strives to protect our state's precious human and natural resources, consistent with sustainable economic development.
Texas Natural Resources Information System	http://www.tnr.is.state.tx.us/	The Texas Resources Information System (TNRIS) is the state's clearinghouse and referral center for natural resources data. Its primary purpose is to make data available to data users quickly and reliably.
The Hydrologist's Home Page	http://www.thehydrogeologist.com/i	This page is a collection of hundreds of links to hydrogeological organizations, software and data repositories, publications, and other resources of potential use to hydrogeologists.
U.S. Geological Survey Near Real-Time River and Reservoir Data for the Sabine River Basin	http://www.sra.dst.tx.us/basin/lake_and_river_conditions.asp?page=real-time	Stream flow conditions, daily-mean stream flow, reservoir conditions, water quality conditions, and daily stream loads

		for 25 locations throughout the Sabine River Basin updated as frequently as every 15 minutes.
<i>Site</i>	<i>Address</i>	<i>Description</i>
Weather	http://www.sra.dst.tx.us/basin/weather.asp	Near real-time weather resources including current conditions, forecasts, and images

Recent advances in Internet technology allow for moving beyond just links to Internet-based distributed database applications. A prototype has been developed for a Distributed Water Quality Database Query. The purpose of the prototype is to serve as a starting point for development of a *Common Geospatial Gateway* to water quality data. The prototype demonstrates applying Active Server Pages (ASP) and eXtensible Markup Language (XML) to querying a distributed water quality database and generating a standardized representation of the data that can be manipulated in a client-side Internet application.

The prototype queries the Texas Natural Resource Conservation Commission (TNRCC), U.S. Geological Survey (USGS)1 and Sabine River Authority of Texas (SRA) water quality data databases over the Internet and creates an XML file that combines the data from each server into the common format shown below:

```

...<?xml version="1.0" ?>
-<WQDATA>
  -<EVENT>
    <TAG>J003131</TAG>
    <DATE>4/27/83</DATE>
    <TIME>9:30:00 AM</TIME>
    <SOURCE Source1="SR" Source2="-0" Program=""/>
    <STATION TNRCCID="10468" USGSID="08019500"/>
    <DEPTH>0.3</DEPTH>
  -<SAMPLES>
    -<SAMPLE>
      <PARAMETER ParameterCode="00010" ParameterDesc="TEMPERATURE, WATER (DEGREES
CENTIGRADE)"/>
      <GTLT />
      <VALUE>18</VALUE>
    </SAMPLE>
    -<SAMPLE>
      <PARAMETER ParameterCode="00020" ParameterDesc="TEMPERATURE, AIR (DEGREES
CENTIGRADE)"/>
      <GTLT />
      <VALUE>20.5</VALUE>
    </SAMPLE>
  -<EVENT>
</WQDATA>

```

Though the prototype simply prints the XML data out in a Hypertext Markup Language (HTML) table, the data can easily be attached to a Java or ActiveX object for more interesting client-side applications such as sorting by different fields, graphing and summaries.

1 The USGS data query is turned off due to a discrepancy in how USGS data is returned from the USGS server (not all data values have dates). A minor re-write of the code should allow for this discrepancy.

A recently proposed XML Query Language (XML-QL, <http://www.w3.org/TR/NOTE-xml-ql/>) promises to facilitate development of this project by providing a standard query language (SQL) for XML. XML-QL can extract data from XML documents and construct new XML data.

The next step of this project is to organize a work group from the Technology Subcommittee of the Data Collection Networks Focus Group to work towards accomplishing the goal of having a state-wide XML-based Common Geospatial Gateway application on-line for demonstration at a future *Texas Water Monitoring Congress*.

3. Geographic Information System (GIS) layers

A large number of GIS layers have been incorporated into the Sabine River Authority Data Clearinghouse or are planned to be incorporated. The Sabine Basin Geographic Information System (GIS) Clearinghouse can be accessed on the World Wide Web at <http://www.sra.dst.tx.us/data/gis/>. Links to GIS layers on other Web servers is also included at this site. The data sets currently incorporated are listed below:

- Aquifers, Major
 - Carrizo
 - Gulf Coast
 - Trinity
- Aquifers, Minor
 - Queen City
 - Sparta
- Basemaps
 - Basin Counties (line and polygon)
 - Reach X Hydrology (line; X = 1 to 7)
 - River Map (line)
 - Subwatersheds (polygon)
 - Subwatersheds (line)
 - Texas Counties and Louisiana Parishes (line and polygon)
- Boundaries
 - Texas Counties and Louisiana Parishes
 - Basin boundary (line and polygon)
 - Reaches (line and polygon; 1 - 7 with subwatersheds)
- Cities
- Contours Sabine Basin TX portion TNRIS
- Councils of Government (COGs)
- Certificates of Convenience and Necessity Areas (CCN) from TNRCC (polygon)
 - ❖
- Environmental Monitoring Sites in the Sabine Basin
 - Fiscal Year 2000 [point]
 - Fiscal Year 2001 [point]
 - Fiscal Year 2002 [point]
 - Sampling Sites all Historical to current [point]
- Environmental Protection Agency (EPA) Landuse 1:250,000 Scale, 1994
 - Alexandria
 - Beaumont
 - Dallas
 - Lake Charles
 - Port Arthur
 - Sherman
 - Shreveport
 - Texarkana
 - Tyler
- Federal Lands
- Federal Emergency Management Assistance

- Calcasieu Parish LA [polygon]
- Cameron Parish LA [polygon]
- Collin County [c48085; polygon]
- Jasper County [c48241; polygon]
- Natchitoches Parish LA [polygon]
- Orange County [c48361; polygon]
- San Augustine County [c48405; polygon]
- Shelby County [c48419; polygon]
- Flood Plain data below Toledo Bend to Orange County line from 1993 Sabine River Flood Study prepared by Brown and Root
 - 100 year flood plain (polygon)
 - Flood way fringe (polygon)
 - 10 year flood plain (line, incomplete)
 - 50 year flood plain (line incomplete)
 - 500 year flood plain (line incomplete)
- Hydrologic Unit Codes (USGS)
- Hydrology
 - Tiger [line; total hydrology for the Sabine Basin 21 counties in Texas & 8 parishes in LA]
 - Texas Department of Transportation (TxDOT)
 - ❖ Collin County
 - ❖ Franklin County
 - ❖ Gregg County
 - ❖ Harrison County
 - ❖ Hopkins County
 - ❖ Hunt County
 - ❖ Jasper County
 - ❖ Kaufman County
 - ❖ Newton County
 - ❖ Orange County
 - ❖ Panola County
 - ❖ Rains County
 - ❖ Rockwall County
 - ❖ Rusk County
 - ❖ Sabine County
 - ❖ Shelby County
 - ❖ Smith County
 - ❖ San Augustine County
 - ❖ Upshur County
 - ❖ Van Zandt County
 - ❖ Wood County
- Imagery
 - Orange County (USGS Monochrome 1:24,000)
 - LA DOQQ for Sabine Basin in LA at these become available (<http://atlas.lsu.edu>)
 - Land Sat image of LA, Louisiana GIS CD
 - Sure Maps 24k USGS quads mosaiced of the Sabine Basin
 - USGS 24k quad sheets scanned with collars
 - Texas Orthoimagery Program resampled data:
 - ❖ 2.5 meter, 3.75 X 3.75 minutes 8-bit color
 - ❖ 10 meter 7.5 X 7.5 minutes 24-bit color
 - ❖ 30 meter 7.5 X 7.5 minute 24-bit color
- Landfills [point] Texas Natural Resources Conservation Commission (TNRCC) 1996
- Landuse 1:250,000 scale USGS/Census
- Minerals
 - Coal [point and polygon], Comprehensive Sabine Watershed Management Plan (CSWMP)

- Gas field [polygon] CSWMP
- Minerals [point and polygon] CSWMP
- Oil Reservoirs [polygon]
- Oil Wells TX partial CSWMP, LA Complete Louisiana GIS CD
- Plays CSWMP
- Priority Sites CSWMP
- Natural Regions
 - Major Natural Regions [polygon]
 - Mitigation Banks [polygon] CSWMP
 - Minor Natural Regions [polygon]
 - Federal Lands [polygon]
 - Refuges in LA [polygon], Louisiana GIS CD
 - Wildlife Management Areas for LA Louisiana GIS CD
 - Bottomland Hardwoods CSWMP
- Outfalls
 - Permitted Hazardous Waste TNRCC
 - Sabine Basin Industrial and Domestic Permitted Outfalls SRA
- Political State / US
 - House
 - Senate
- Population Data
 - 1990 Census Regions and Blocks for Counties and Parishes in the Sabine Basin
- Reservoirs
 - Reservoirs Texas Water Development Board 1997
 - LA Reservoirs Louisiana GIS CD
- Roads
 - Tiger (all Basin combined)
 - TxDot (each county individually)
- Sabine Basin Outline [polygon]
- Senate Bill 1 Regional Water Planning Areas [polygon]
 - RWPA C
 - RWPA D
 - RWPA I
- Schools District Boundaries TNRIS
- Soils Data TX and LA
- SRA Parks
- Texas Air Monitoring Sites

- Texas Department of Transportation
 - Collin County
 - Franklin County
 - Gregg County
 - Harrison County
 - Hopkins County
 - Newton County
 - Orange County
 - Panola County
 - Rains County
 - Rockwall County
 - Rusk County
- Topographic Data Sets for Texas By River Basin (U.S.G.S.) 1:250,000 scale, 1997
 - Sabine River Basin Elevation
 - Sabine River Basin Flow Accumulation
 - Sabine River Basin Flow Direction
 - Sabine River Basin Hill Shade
 - Source data sets for Sabine and other Texas Basins on CD-ROM server at <http://www.sra.dst.tx.us/cd-server/CDVOL000/>
- Toxic Release Inventory EPA 1990
- United States Geological Survey (USGS) Quads Map Index [polygons; points]
 - Sabine Basin [7.5 minute]
 - Orange County [7.5 minute]
 - USGS [1 to 250K]
 - USGS quads [points; polygons]
- Vegetation [polygons; Sabine Basin]
- Water Supply
 - Sabine Basin public water supply [points] (TNRCC) 1997
 - Concerns from 1996 Data Analysis
 - Possible Concerns from 1996 Data Analysis
 - Not Concerns or Possible Concerns from 1996 Data Analysis

Federal Government Data Clearinghouse (FGDC) compliant metadata for each of the above GIS data sets can be found at <http://www.sra.dst.tx.us/data/gis/>. See also Appendix 2.

GIS data sets that Sabine River Authority (SRA) plans to acquire and incorporate into the Sabine River Data Clearinghouse include:

- Oil and Gas Well (Texas Railroad Commission [RRC])
- DRG maps of Texas and Louisiana portions of Sabine Basin
- 2000 Census block and Region Data.

B. Personnel

1. Data Resources Manager

The Data Resources Manager is responsible for ensuring that the Data Management and Watershed Inventories (Task 4) Objectives of the current active TCRP work plan are achieved. These objectives include:

- ensuring that Basin water quality data are collected, managed in an efficient manner
- ensuring that such data is transferred to TNRCC as per the deliverable schedule in the current active work plan
- ensuring the accessibility of all data collected through the CRP to the public via the internet as well as providing hard copy and/or digital formats of data sets upon request
- ensuring that data is collected for a database containing information on sites that have the potential to adversely affect water quality.
- ensuring that location coordinates for all sampling sites are collected and stored in an accurate and accessible database table
- ensuring the maintenance of a watershed inventory consisting of data sets of factors which can influence water quality
- ensuring that data submitted to TNRCC for inclusion in the SWQM database has been verified by a quality assurance procedure as outlined in the Information Management Plan (SRAIMP).

The Data Resources Manager also:

- is responsible ensuring the data system is documented and that the SRAIMP is updated accordingly
- is responsible for management of SRA's local area network (LAN) and wide area network (WAN) computer systems (except for the Accounting System subnetwork).
- is responsible for ensuring and maintaining data security and integrity
- oversees development of new database applications
- oversees revisions of existing database applications
- is responsible for ensuring that applicable data requests by the public and other entities are fulfilled
- is responsible for specification and purchase of new information resource-related equipment
- is responsible for overseeing the maintenance and repair of computer resources

2. Website Administrator

- The Website Administrator is responsible for ensuring the Web Site Requirements listed on pages 6.6 and 6.7 of the Program Guidance FY2002-2003 are met.

In fulfillment of and in addition to the above, the Website Administrator also:

- is responsible for creating and maintaining documentation of the World Wide Web site
- is responsible for creating and maintaining Web pages for SRA's World Wide Web site
- is responsible for development of projects related to the World Wide Web site
- coordinates with the Data Resources Manager and GIS Analyst in posting data sets and GIS layers or coverages on the World Wide Web site
- responsible for security of Web Server System
- responsible for pursuing and implementing compatible links with the Web sites of other entities for the purpose of sharing data across the Internet

3. Data Analyst

The Data Analyst is primarily responsible for performing whatever duties are required to meet the objectives of Task 5 of the Program Guidance, Data Analysis and Reporting. These duties include:

- performing data analysis functions according to the methodology prepared by the CRP Data Analysis Workgroup
- performing the Integrated Trend Analysis using SRA's water quality database, Special Investigation information, and other data sets that could impact water quality
- using GIS and Subwatershed Inventory data sets to produce comprehensive subwatershed maps to aid in the determination of sources of impairments

Other duties of the Data Analyst include

- participation in the preparation of the annual Basin Highlights Report in the areas related to the data analysis as stated in Task 2
- participation in the preparation of the Basin Summary Report in areas related to data analysis and comprehensive map production as stated in Task 2
- assisting the QA Officer with QA/QC review of data sets and other data management duties as delegated by the Data Resources Manager

4. GIS Analyst

The GIS Analyst responsible is for fulfilling the objectives of Task 4 of the Program Guidance, Data Management and Watershed Inventories, which include:

- developing spatial and descriptive data sets related to basin, reach and subwatershed area spatial hierarchies into full GIS coverages
- coordinating with the Data Resources Manager and the Website Administrator in providing these data sets to interested parties through the World Wide Web
- procuring GIS data sets necessary for CRP applications by:
 - networking GIS data and technology through individual contacts and local and regional GIS consortiums having common data procurement needs
 - hosting and participating in local and regional GIS users group meetings where specific GIS data types are discussed and exchanged
 - pursuing cost sharing arrangements in the procurement of large and complex GIS data sets
- participating in specific subwatershed intensive studies where numerous GIS data layers will be applied to water quality specific relationships

The GIS Analyst is also responsible for:

- system administration and software maintenance of the GIS workstation
- data conversion tasks requiring the functionality of the workstation and/or ArcInfo software
- digitization functions

Responsibilities of the personnel listed above are contingent on revisions to the TRCP work plan and SRA's Data Management System.

5. Information Resources Support Specialist

The Information Resources Support Specialist performs computer and software support duties.

The duties of the Information Resources Support Specialist include:

- Troubleshoot and maintain computer systems hardware
- Coordinate with consultants on specialized repairs and maintenance
- Assist with maintenance and development of network operating systems, software, and other major software packages in use at the SRA
- Identify and correct hardware and software issues
- Assist in the development, implementation, training, and distribution of database applications throughout the Sabine Basin
- Provide software training for SRA staff
- Identify various group/users needs and locate existing training opportunities or develop in house programs to meet those needs
- Organize workshops as needed
- Any other duties required by the Data Resources Manager

C. Proposed Personnel Needs

1. Student interns on a temporary or seasonal basis to:
 - Enter data or other information into databases
 - To perform research required to populate databases
 - To reorganize storage of information in physical or electronic file systems
 - To convert hard copy information into an electronic format
 - To assist the IRSS in systems maintenance or special projects

D. Data Needs

Data needs in addition to the planned data acquisitions listed above include:

- updates to and subsequent regular updates to the data sets received from TNRCC as listed in Section I. A., Data Inventory, above. Expansion and Growth.
- water quality monitoring data from industry and municipalities in the Basin in digital format
- Texas Parks and Wildlife Department fisheries and aquatic life data
- improved landuse data (SRA does not currently have the resources to develop this data)
- oil well data from the RRC
- Digital Ortho Quarter Quad (DOQQ) coverages for entire Basin
- U.S. Census Block and Regional data for 2000 <http://factfinder.census.gov/servlet/BasicFactsServlet>
- National Response Center oil and chemical spill database (data available on the Internet but on-line search not suitable for SRA's needs, e.g. Events Affecting Water Quality Database and Subwatershed Inventory; data downloadable in relational format, just a matter of importing it into a database)

II. INFORMATION MANAGEMENT SYSTEM DESIGN

A. Computer Hardware and Software Requirements

The Sabine River Authority computer system consists of a Wide Area Network (WAN) combining the Local Area Network (LAN) at the Authority General Office (AGO) and the ESD LAN via a 56 Kbps leased line. The AGO LAN consists of a two Novell NetWare 5.1 Servers and a Novell NetWare 4.2 Server providing storage, database access, as well as file and print services to a 10MBps shared Ethernet network. The AGO LAN also includes a Unix Server hosting the Accounting Department software and data files as well as Unix print services. The ESD LAN consists of a NetWare 5.1 server and a NetWare 4.2 server providing storage, database access, and file and print services to its 100 MBps switched Ethernet network. The file servers, workstation nodes, and printers communicate with each other and the rest of the network via hubs and various routers. The hub also allows access to the Internet through SRA's connection with Texas Higher Education Network (THEnet, <http://www.the.net/>). The Gulf Coast Division Office has a 10MBps Windows 9x peer-to-peer network with dial-up access to the Internet via AGO. Networks are planned at the other SRA Divisions. Each currently has dial-up access to the Internet via local Internet Service Providers. Details of hardware and software composing the ESD subnet, the Information Resources Department (IRD) subnet, and the Web site are listed in the table below. Hardware and software used within the Accounting subnet and by other SRA Divisions are not currently included.

Table 1. Computer Workstation Hardware/Software Configuration				
Configuration	Current		Planned	
Type	Hardware/Software	Date	Hardware/Software	Year
PC Workstation Hardware	Pentium, Celeron, Pentium II, or Pentium III processors, 64+ MB RAM; 1.44 MB floppy and CD-Rom drive; EIDE, ATAPI, or SCSI hard drives; 10/100Mbps Ethernet cards; SVGA Monitors; mouse; UPS; some systems also have removable storage (Jaz, Zip), PD-CD, or tape capability.	Lab subnet: 22 IRD subnet: 11	Best available Intel based processor; 256+ MB RAM; 1.44 MB floppy; CD-Rom; 10+ GB drive; ZIP or Tape drive; 10/100 baseT Ethernet card; SVGA monitor (18+ " VIS); UPS. Modems where applicable.	Target: Replace Pentium workstations.
Portable PC Hardware	Pentium II 233 MHz with 128 MB RAM; 1.44 MB floppy, 4GB HD; CD-Rom; 56Kbps modem; 10/100 NIC.	One system currently available.	Best available Intel based processor; 128+ MB RAM; 1.44 MB floppy; CD-Rom; 6+ GB HD; 10/100 Mbps Ethernet NIC; 56 Kbps modem	Add at least one portable PC in FY2002
PC Workstation Software	MS Windows NT 4.0, 2000, or Windows 9x; MS Office Pro and Std 2000; other software is department or role specific.	Current	New PC's purchased with WinNT, 2000 OS plus other applicable software	FY 2002

Table 2. Intel GIS Workstation Hardware/Software Configuration				
Configuration	Current		Anticipated	
Type	Hardware/Software	Date	Hardware/Software	Year
Intel based GIS Workstation Hardware	SGI Dual Pentium II 500; 512 MB RAM; 14 GB EIDE internal disk; 21" color monitor w/ XMB video card; 3.5" 1.44MB PC-compatible floppy drive; 12/24 GB Travan tape drive; UPS and surge protector; Hewlett Packard 1050 Large Format color plotter;	One system installed	No anticipated upgrades to this system planned for FY2002	
GIS Workstation Software	WinNT 4.0 w/ SP6; ArcINFO 8.x; GRID x.x; ArcVIEW 3.2; ArcPress x.x;	Current	Upgrades to ArcINFO and ArcVIEW as available	

Table 3. IRD Server Hardware/Software Configuration

Configuration	Current		Anticipated	
Type	Hardware/Software	Date	Hardware/Software	Year
SRA_IR1 Server Hardware	Dell PowerEdge 4400 Dual Intel Pentium III 600 MHz Xeon; 1 GB RAM; 1.44 MB floppy and CD-Rom drive; RAID 5 with 9 GB 10,000 RPM SCSI HD; 12/24 GB DDS3 DAT tape drive; two 10/100 Ethernet NIC; SVGA Monitor; Smart-UPS 1400	Current	No upgrade or replacement planned	
SRA_IR1 Software	Novell NetWare 5.1; Backup-Exec 8.5 Multi- server edition; Oracle 8i	Current	Upgrades as available via Novell MLA with State of Texas	
SRA_IR2 Server Hardware	Pentium 90 MHz Intel based PC; 64 MB RAM; 512 Kb Cache; 1.44 MB floppy and CD- Rom drive; Dual 9 GB SCSI HD; 4/8 GB DAT tape drive; 10Mbps Ethernet card; SVGA Monitor; Smart-UPS	Current	No upgrade or replacement planned.	
SRA_IR2 Software	Novell NetWare 4.2; ArcServe 6.1 SP2	Current	No upgrade planned	

Table 5. ESD File Server Hardware/Software Configuration				
Configuration	Current		Anticipated	
Type	Hardware/Software	Date	Hardware/Software	Year
SRA_ESD1 Server Hardware	Pentium III 933 MHz Intel based PC; 512 MB RAM, 1.44 floppy and CD-Rom drive; 4 X 9GB HD in Raid 5 array; tape backup drive; dual 10/100 base T Ethernet cards; Smart-UPS	Current	No upgrades planned	
SRA_ESD1 Software	Netware 5.1; Backup-Exec 8.5; Sample Master LIMS; print services	Current	Upgrade to Netware 6.X upon release	
SRA_ESD2 Server Hardware	Pentium II 300 MHz Intel based PC; 256 MB RAM; 512 Kb Cache; 1.44 MB floppy and CD-Rom drive; Duplexed 9 GB SCSI HD; 4/8 GB DAT tape drive; 10/100baseT Ethernet card; SVGA Monitor; Smart-UPS	Current	No upgrades planned	FY 2001
SRA_ESD2 Software	Novell Netware 4.2; hosts backup of Sample Master LIMS database	Current	No upgrades planned	

Table 6. World Wide Web Servers Hardware/Software Configuration				
Configuration	Current		Anticipated	
Type	Hardware/Software	Date	Hardware/Software	Year
PowerApp Web 120 Primary Web Server Hardware	Dual PIII 1 GHz/133 bus; 1Gb RAM; 256K Cache; 3 X 36 Gb Raid 5; Cd-Rom; dual NICs;	In progress	Add USB tape drive and USB external modem	FY2002
PowerApp Web 120 Primary Web Server Software	Windows 2000 Server; IIS 6; iCal	In progress	No upgrades planned	
WEB4100 (Secondary) Web Server Hardware	Pentium 200 MHz Dual Intel based PC; 512 MB RAM; 512 Kb Cache; 1.44 MB floppy and CD-Rom drive; 9 GB Mirrored SCSI HD; 4/8 GB DAT tape drive; Ethernet NIC; SVGA Monitor; Smart-UPS	Current	No upgrades planned	
Web4100 Web Server Software	MS Windows NT 4.0 Server;; MS Access 2000; MS Front Page 2000; ; IIS 5.0; ArcIMS	Current	No upgrades planned	
Web CD-Rom Server Hardware (vendor software)	Pentium 100 MHz Intel based PC; 32 MB RAM; 512 Kb Cache; 1.44 MB floppy; 2 GB HD; 7 8X SCSI CD-ROM readers on 5 CD changer with expansion unit of 7 8X readers on 5 CD changers (70 CD's total capacity), expandable.	Current	Additional RAM; larger HD; possible motherboard and processor upgrade; Win2K compatible software	FY2002

Table 7. World Wide Web DNS/Email Server Hardware/Software Configuration				
Configuration	Current		Anticipated	
Type	Hardware/Software	Date	Hardware/Software	Year
Dell Poweredge 1400 SC Email Servers Hardware	PIII 866MHz/133; 256K Cache; 256 Mb RAM; 2 X 9 Gb HD in Raid 1; DDS4 tape	In progress	No upgrades planned	
Dell Poweredge 1400 SC Email Servers Hardware	Redhat Linux 7.0; sendmail 8.11.x	In progress	Upgrades to sendmail as required; Web-based email	FY2002
PII 333 DNS Server Hardware	PII 333; 128 Mb RAM; 4Gb SCSI HD; zip; cd- rom	In progress	No upgrades planned	FY2002
PII 333 DNS Server Software	Redgat Linux 7.0; BIND 9.x	In progress	Upgrades to BIND as required	FY2002
SonicWALL XPRS2 Firewall Appliance Hardware/Software	133MHz Toshiba CPU; 8 Mb RAM; 3 100 BaseT ports	Current	Upgrades to firmware as available	FY2002
Sun Netra DNS/Mail Server Hardware	Sun Netra Sparc 6; 85 MHz RISC processor; 32 MB RAM; 1 GB HD; SVGA monitor; 3 button mouse; 2 ea 10baseT Ethernet cards; UPS	Current	Replace with Linux based systems	In progress
Netra Server Software	Sun Firewall-1 Light; Sun Internet Server	Current	Replacing with Linux based systems.	In progress

Other hardware includes hubs, switches, routers, transceivers, and CSU/DSU's required for the LAN, WAN, and Internet site.

Other planned upgrades include:

- FY2002 – FY2003 Upgrade software and RAM on Cisco routers; replace hubs with switches; upgrade AGO network to Fast Ethernet; continue replacement of Sun Netra DNS/Email services.
- FY2002 – FY2003 Replace 56 kb leased line WAN link with T1, or Frame Relay with committed information delivery rate (CIR) at >= 384 kb speed, or other cost effective broadband solution.

B. Database Table Documentation

Each permanent database table in the Sabine River Authority Data Repository for Watershed Inventory must have an accompanying database table documentation table stored in \\SRA_IR1\VOL3\SRWMP\SWI\COMMON\DATDICT.MDB. Within each individual database file, the required data dictionaries are available through a table link to DATDICT.MDB. Each record of the data dictionary contains the following information:

Field Name	Each field of the database table is listed
Data Type	The data type of each field
Field Size	The size of the field
Is Key?	Is the field a key field?
Contents	What type of data should the field contain
Required?	Is an entry required in the field
Comments	Other explanatory information about the field

Changes to a database table must be accompanied by subsequent changes to the database table documentation table dictionary. See Appendix 1.

GIS data as presented on the World Wide Web site will be accompanied by a metadata file, which meets the standards of the Federal Government Data Clearinghouse (FGDC). See Appendix 2.

C. Timeline

The Sabine River Authority Information Resources Department Data Management System (SRADMS) is in place and further development will be needs-driven either by requirements of the Sabine River Authority, the Texas Clean Rivers Program, or new developments in technology. Timelines specific to implementation of specific items are included within the section wherein that item is described as applicable.

D. Budget

The Information Resources Department at SRA is budgeted through the Environmental Services Division and the Water Resources Branch.

III. INFORMATION MANAGEMENT PLAN IMPLEMENTATION

A. Functions for acquiring data sets

1. Routine SRA water quality monitoring data sets are acquired and integrated according to procedures in “*SRWMP Water Quality Data From Collection to Data Repository*” [See Appendix 3].
2. Reference data sets from TNRCC are usually received after making a formal request via email or formal letter. The updated data sets either replace or are appended to the existing data sets as per instructions received from the TCRP Data Manager. Data dictionaries for each data set will also be requested. Data paths for updates of specific data sets will be documented as each data set is updated.
3. GIS data sets are updated as new information becomes available with the previous data set being replaced with the new. Metadata must accompany any GIS data set update.
4. New data sets, both database and GIS, are acquired on a “needs-driven” basis following the creation of a new SRA project or the assignment of a TCRP task. Data that can be obtained free of charge is preferred over data that must be purchased.

B. QA/QC on acquired data sets

1. QA/QC on routine SRA water quality monitoring data sets is outlined in “*QA/QC Procedures for SRWMP Data*” [See Appendix 3]. All incoming water quality data must go through these procedures before being posted on the Web site, uploaded to TNRCC for inclusion in SWQM, or appended to the SRA Water Quality Information Clearinghouse.
2. Reference data sets from TNRCC are checked for structural integrity before being appended to the existing reference tables. Data dictionaries will be requested with all future data updates to facilitate this task.
3. GIS data sets are reviewed for the existence and accuracy of metadata.
4. QA/QC for new data sets is dependent on the data type and the source. Data from entities in the Sabine Basin must be covered by an approved QAPP before it can be included in the data clearinghouse. Other data sets are reviewed on a case-by-case basis depending on the source and the intended end use of the data. QA/QC considerations for all data include QA/QC practices of the provider of the data set and the existence of a data dictionary or metadata.

C. Report generating capabilities

1. The World Wide Web site:
 - Adobe Acrobat PDF file (free download at <http://www.adobe.com/products/acrobat/readstep2.html>)
 - Printable Web page
 - ASCII text
 - Maps
2. The SRA LAN
 - Printed text (within reason based on the size of the file)
 - Microsoft Word document via diskette, ftp, or email
 - ASCII text via diskette, ftp, or email
 - Maps (large format capable)
 - Microsoft Access database report or Excel spreadsheet report via diskette, ftp, or email

In general, report generation is governed by the data set to be reported, the software to be used to generate the report, and the skill of the software user.

D. Overview of services currently available on the World Wide Web site

Services available on SRA’s WWW site are outlined in the **Table of Contents** at http://www.sra.dst.tx.us/about_this_site/navigation/toc.asp.

These include:

- [About SRA](#)
- [Authority General Office](#)
- [Board of Directors](#)
- [Contact Us](#)
- [Office Locations](#)

[Staff e-mail Addresses](#)

[Employment Opportunities](#)

[Environmental Services Division](#)

[Gulf Coast Division](#)

[History of the Sabine River Authority](#)

[Iron Bridge Division](#)

[Lake Fork Division](#)

[Public Calendar](#)

[Toledo Bend Division](#)

[Yahoo!Groups](#)

[SRATXINFO](#)

[About this Site](#)

[Site Navigation](#)

[Index](#)

[Search this Site](#)

[Table of Contents](#)

[SRA Privacy Policy](#)

[Data](#)

[GIS Clearinghouse](#)

[Texas Strategic Mapping Program](#)

[Water Quality Information Clearinghouse](#)

[Surface Water Quality Monitoring](#)

[Newsfeeds](#)

[Moreover Environmental News Feed](#)

[Moreover Utilities News Feed](#)

[Press Release Sources](#)

[Site News](#)

[SRA Press Releases](#)

[Projects](#)

[Iron Bridge Project](#)

[Lake Fork Project](#)

[Toledo Bend Project](#)

[Sabine River Basin](#)

[Community Profiles](#)

[Beckville](#)

[Big Sandy](#)

[Bridge City](#)

[Canton](#)

[Carthage](#)

[Center](#)

[East Tawakoni](#)

[Edgewood](#)

[Emory](#)

[Gladewater](#)

[Greenville](#)

[Hallsville](#)

[Hawkins](#)

[Hemphill](#)

[Jasper](#)

[Joaquin](#)

[Kilgore](#)

[Kirbyville](#)

[Lindale](#)

[Marshall](#)

[Mineola](#)

[Newton](#)

[Orange](#)

[Overton](#)

[Point](#)

[Quinlan](#)

[Quitman](#)

[San Augustine](#)

[Sulphur Springs](#)

[Tatum](#)

[Tenaha](#)

[Timpson](#)

[Tyler](#)

[West Tawakoni](#)

[White Oak](#)

[Wills Point](#)

[Lake and River Conditions](#)

[Legislative and Government Information](#)

[Links](#)

[Cities](#)

[Councils of Government](#)

[Professional Groups](#)

[East Texas GIS and GPS Users Group](#)

[Orange County, Texas, GIS Consortium](#)

[Sabine Basin Economic Development](#)

[Maps](#)

[Recreation Guide to the Sabine River Basin](#)

[Regional Planning](#)

[Sabine River Basin Overview](#)

[Water Conservation](#)

[Weather](#)

[Sabine River Watershed Management Program](#)

[Comprehensive Sabine Watershed Management Plan](#)

[Project Team](#)

[Scope of Work](#)

[Tasks](#)

[Glossary](#)

[Mr. Water Wizard](#)

[Special Reports](#)

[Summary Report of MtBE Contamination in Lake Tawakoni from Gasoline Pipeline Rupture](#)
[Texas Clean Rivers Program](#)

[Overview](#)

[State of the Basin](#)

[Basin Highlights](#)

[Monthly Water Quality Monitoring Reports](#)

[Quality Assurance Project Plan](#)

[Reach Summaries](#)

[Sabine Basin Action Summary](#)

[Sabine Basin Currents](#)

[Special Studies on Priority Watersheds](#)

[Summary Reports](#)

[TCRP Meetings](#)

[Sabine Basin 2001 Coordinated Monitoring Meetings](#)

[Sabine Basin Steering Committee Meetings](#)

[Texas Clean Rivers Program Calendar](#)

[Texas Clean Rivers Program Partners](#)

[Watershed Management in Texas: The Clean Rivers Program](#)
[Trans-Texas Water Program](#)
[Southeast Study Area Preliminary Findings](#)

[Services](#)

[Economic Development](#)

[Environmental](#)

[Biomonitoring Laboratory](#)

[Field Services](#)

[Water Quality Laboratory](#)

[Information Resources](#)

[Permits](#)

•

E. Services planned for the World Wide Web site

Services planned for the World Wide Web site include:

- Receiving Water Assessment data and information
- Aquatic Toxicity data and information
- Rapid Biological Assessment data and information
- Expansion of interactive mapping to include Digital Ortho Quarter Quad layers for the entire Sabine Basin
- Trend analysis graphs
- Current data for each parameter compared to historical mean/min/max information
- Comparison of current subwatershed water quality inventory status as compared to historical standards
- Definitions of water quality parameters in layman's terms

IV. INFORMATION RESOURCES ADMINISTRATION

A. Security

1. *Virus protection:* Norton Antivirus Corporate Edition 7.5 is installed on ESD, IRD, AGO, GCD, LFD, and IBD computers. Installation on TBD computers is planned. Updated virus signatures are retrieved from the Internet by each LAN's file server. PCs on a LAN download the updates as available with each login. PCs not on a network receive updates directly from the Internet.

2. *LAN security:*

The local area network is protected from intrusion via the Internet by a SonicWALL XPRS2 firewall appliance. The Checkpoint Firewall-1 Light firewall installed on the Sun Netra server continues to protect the Netra on weekends and after normal business hours. Only HTTP is allowed inside the firewall from the Internet.

Dial-up access via modem is username and password protected and allows only Internet access by default. Access to servers or routers is governed by internal LAN username and password. Dial-up access to individual PC's using PC-Anywhere is governed by user level access rights, username and password, and PC-Anywhere level encryption.

Each LAN user is granted a password, which must be changed every 90 days. It is against IRD policy for a user to use another individual's username and password to access the network. Users are advised to keep their passwords private and to log off the network when they are not present at their workstation. Logins to the NDS8 (eDirectory) is governed by Secure Sockets Layer at the NetWare client.

LAN users are allowed access only to those resources to which they are granted rights via assigned NetWare rights stored in the NDS tree.

The IRD administrative workstation is protected by Windows NT 4.0 security and is kept locked when not occupied by the Network Administrator. The console of the SRA_IR1 and SRA_IR2 servers are kept locked and password protected, as is console access to the Web servers, the DNS/Email/Firewall server, the SonicWALL firewall, and the CD-Rom server.

Administrative passwords are severely restricted. Each department container of the NetWare NDS tree has a container administrator with administrative rights to that container only. The NDS root administrator password is limited to key personnel, as are the root password of the Sun Netra, the administrator password of the Web Server, the admin password of the firewall, and the enable passwords for the Cisco routers.

3. *Software security:*

The Internet and list services are monitored by IRD personnel for information about security bugs in Web browser and email software. Security patches are applied ASAP after notification.

4. *Email security:*

Users are advised that email is not a secure means of communication. Users who wish to keep their email secure from internal viewing are advised to password protect their screen savers, to lock NT workstations, or to use a password protected Outlook Express identity.

5. *Planned improvements for security system:*

- Implementing Novell Zenworks to lock down workstations to approved software only.
- Replacing Windows9x workstations with WindowsNT or 2000 workstations
- Replacing dial-up access with secure VPN

B. Software Management

Legal usage of software within the *authoritative domain* of IRD is governed by centralized storage of software, installation of software by IRD personnel, and periodic software audits using the Business Software Alliance's Soft Scan. The licensed software for each workstation and server is stored in boxes labeled for that workstation and a software-licensing database is in development. Most software is purchased through the Texas Department of Information Resources (DIR, <http://www.dir.state.tx.us/>) at discounts off retail. DIR assigns control numbers to each Purchase Order purchased and sends out periodic statements of licensing. These statements as well as other licensing documentation is kept on file.

Novell NetWare software is obtained from DIR under the Master License Agreement (MLA) Novell has with the State of Texas. SRA's Novell MLA documentation is stored in SRA's fireproof vault.

Plans to implement software that will "lock down" workstations to licensed software only are underway. The purpose of the "lock down" is to prevent users from installing software downloaded off the Internet or installing personal software on Authority computers.

Installation of software not legally licensed for use on Authority systems is prohibited on IRD-managed systems.

C. Quality Assurance/Control

"Cradle to Grave" data quality assurance is primarily governed by procedures stated in the following documents:

1. ***"Quality Assurance Project Plan for the Sabine River Authority for Environmental Monitoring and Measurement Activities Relating to the Water Quality Monitoring Program"***. The latest approved revision is available on the SRA Web site at http://www.sra.dst.tx.us/srwmp/terp/state_of_the_basin/qapp/
2. ***"Sabine River Authority Quality Assurance Plan"***. The latest approved revision is available by contacting SRA's Quality Assurance Officer.
3. ***"SRWMP Water Quality Data From Collection To Data Repository"*** [Appendix 3]
4. ***"SRWMP Water Quality Data QA/QC Procedures"*** [Appendix 3]

The SRA field biologists are responsible for maintaining quality assurance procedures in the field with the Laboratory Supervisor and Quality Assurance Officer overseeing quality assurance of laboratory analyses and data entry. The Data Resources Manager is responsible for verifying data outliers and overseeing corrections to data in the data clearinghouse. The Website Administrator is responsible for correcting mistakes on the Web site with the GIS Analyst and Data Analyst being responsible for correction of GIS data layers under the direction of the Data Resources Manager. The Data Resources Manager is responsible for data systems at all levels.

D. Programming Processes

In-house programming largely consists of Visual Basic and HTML programming specific to development of Web site applications. Web programming is being expanded using XML (eXtensible Markup Language), Active Server Pages, and VB Script. The Website Administrator is the primary programmer of the Web site and is responsible for testing, debugging, and documentation of all aspects of Web development programming.

Application and database programming with C++, Java, and Visual Basic for Applications will increase in the coming year.

E. Genealogy

Currently, the genealogy of each data set is documented in bound books labeled "Data Management Log Book." More standardized procedures for documenting genealogy are under development.

F. Documentation

The primary means of data set documentation are written data paths, database table documentation tables, and metadata. See Appendices 1, 2, and 3.

G. Archives/Backup

Backup

The SRA_ESD1, SRA_ESD2, SRA_IR1, and SRA_IR2 servers are backed up using 4mm DAT tape using either Backup-Exec 8.5 or ArcServe 6.1 for Netware. The following tape rotation system is used:

Monday Full backup on Monday tape
Tuesday Full backup on Tuesday tape
Wednesday Full backup on Wednesday tape
Thursday Full backup on Thursday tape
Friday 1 (1st Friday of month) Full backup on Friday 1 tape
Friday 2 (2nd Friday of month) Full backup on Friday 2 tape
Friday 3 (3rd Friday of month) Full backup on Friday 3 tape
Friday 4 (4th Friday of month) Full backup on Friday 4 tape
<MonthYYYY> Last Friday of month Full backup on <MonthYY> tape (month tapes are kept for six months)

The SRA_IR1 and SRA_IR2 tapes are stored in SRA's fireproof vault and the most recent tape is carried "off site" overnight or until a more recent tape is available. The SRA_ESD1 and SRA_ESD2 tapes are stored in the Quality Assurance Officer's office with the most recent tape being carried "off-site" overnight or until a more recent tape is available.

The Web Server is backed up using 4mm DAT tapes with a similar tape rotation system using Microsoft NT Backup 4.0. The tapes are stored in SRA's fireproof vault. Archived data is written to CD-R disks.

GIS data is backed up weekly or as needed on 20 GB Travan tape and stored in the fireproof vault. Permanent data is written to CD.

A backup/restore test is performed with each Backup-Exec job on the SRA_IR1 server and periodically on other systems.

The General Manager's Secretary uploads critical PC files daily to the SRA_IR1 file server that are backed up with the server's backup.

Users are advised not to store critical data on a local HD without a regular backup procedure in place.

Archive

Archive data is stored in dated directories on the SRA_IR1, SRA_IR2, SRA_ESD1, and SRA_ESD2 servers and backed up as per the schedule above. Archive GIS data sets are written to CD-R.

H. Disaster Recovery

Disaster recovery procedures consist of reinstalling the operating system and software from the original software media and restoring data from tape. The original software for all systems is stored in the fireproof vault at SRA's Authority General Office Building. Disaster recovery procedures were tested in-house during the month of August of 1998 when a lightning strike damaged the SRA_ESD server hard drive and again in September of 1998 when the HD of the SRA_DB server (predecessor to SRA_IR1) was replaced. Disaster recovery methods were successful.

Planned improvements to disaster recovery procedures include the purchase and installation of Hewlett-Packard DAT tape drives with the One Button Disaster Recovery feature and Veritas Intelligent Disaster Recovery module of Backup-Exec. Other improvements to disaster recovery procedures include writing restore images of each PC to bootable CD.

I. Migration/Transfer/Conversion

Standard procedures for migration, transfer and conversion of data are found in data paths and metadata related to the specific data type or data set. See Appendix 3 for links to data paths.

J. Client Service

Public access of data is available through the SRA World Wide Web site in a variety of formats, the specific format being dependent on the type of data:

- Adobe Acrobat PDF
- Delimited and fixed format ASCII text
- Printable Web pages
- MS Excel spreadsheets
- ArcView .SHP files
- GIS .EEO files
- Data download routines incorporated into the Web site

Data not available on the Web can requested on a case-by-case basis from the Data Resources Manager. Data will be made available to the public in a format SRA can provide without resorting to custom programming or purchase of new software.

Appendices

Appendix 1	Sabine River Data Clearinghouse Database Table Documentation
Appendix 2	Sabine River GIS Clearinghouse Metadata
Appendix 3	Data Path Documents
Appendix 4	Detail of Data Transfer to TNRCC under Texas Clean River Program
Appendix 5	Normalization

Appendix 1

Sabine River Data Clearinghouse Data Dictionaries

Database table documentation in the Sabine River Data Clearinghouse is available in an Access 2000 database at \\sra_ir1\vol3\srwmp\swi\common\DatDict.mdb.

Appendix 2

Sabine River Data GIS Clearinghouse Metadata

Links to metadata files for GIS coverages in the Sabine River GIS Clearinghouse are found at <http://www.sra.dst.tx.us/data/gis/>.

Appendix 3

Sabine River Data Clearinghouse Data Paths

The data path documents are available upon request by contacting Mary S. Vann at mvann@sra.dst.tx.us.

Appendix 4

Detail of Data Transfer to TNRCC under Texas Clean River Program

Data updates will be transferred to TNRCC under the Texas Clean Rivers Program (CRP) at the frequency specified in the current CRP work plan. The preferred method of transfer will be to post the files on SRA's ftp site (<ftp.sra.dst.tx.us>) and notify the Sabine Basin CRP Project Manager that the files are available for download. The data files will be provided in ASCII text with each field delimited by the pipe character (|) in the field format detailed below. Note that the alphanumeric (A) field sizes indicate the maximum number of characters allowed and do not indicate nor imply a fixed size field in the delimited ASCII text file.

Events file:

Tag_id A7 This field is the key between the event and results tables and is 7 characters long. The first character(s) is the prefix code for the submitting agency.

Station A9 This is an obsolete field and should be left blank.

Stationid A5 This is a unique id that identifies each sampling station. This number is generated by the TNRCC.

Enddate A10 The date the sample was collected in the form of MM/DD/YYYY

Endtime A5 The time the sample was collected in military format HH:MM)

Enddepth A6 This is the depth in meters at which the sample was collected.

Startdate A10 This field is only required for composite samples and is the beginning date in the form of MM/DD/YYYY

Starttime A5 This field is only required for composite samples and is the beginning time (in military format) at which the sample was collected (HH:MM)

Startdepth A6 This field is only required for composite samples and is the depth nearest surface (in meters) at which the sample was collected.

Category A1 This field is only required for composite samples and should correspond to the following codes:

- T is for time composites
- S is for space composites (i.e.depth)
- B is for both space and time composites
- F is for flow weighted composites

Calculation A1 This field is no longer used and should be left blank

Type A2 This field is only required for composite samples and should correspond to the following codes:

- CN for continuous
- ## where ## is the number of grab in the composite
- GB where the number of grabs is unknown

Comment A135 This is a text field where record of any observational data is included with the sample

Source1 A2 The TNRCC assigned code for the submitting agency.

Source2 A2 The TNRCC assigned code for the entity performing the actual sample collection.

Program A2 A field that further identifies the sample. This field may be used to tie targeted monitoring to specific permits.

Results File:

Tag_id A7 This field is the key between the event and results tables and is 7 characters long. The first character(s) is the prefix code for the submitting agency.

Enddate A10 The date the sample was collected in the form of MM/DD/YYYY

Storecode A5 This is a five digit code which identifies the substance or measurement.

Gtlt A1 If the value is above the detection limit then this field should contain an >.

Value A8 This is the test result and should be reported in units according to the storet description.

Associated files included with data upload will include:

1. Error log of data points which fell outside the minimum or maximum limits of the SW_PARM table or outside the limits of SRA's historical data set. The error log will indicate the verification status of outlier data points.
2. Readme file listing all files included in the upload and their contents.

Appendix 5

Normalization

This Appendix provides a brief discussion of the normalization component of relational database design. It is technical in nature and consists mainly of definitions of terms, concepts, and rules. It is taken mostly from WERTZ92 which provides numerous examples not repeated here.² Normalization is best explained in terms of a series of rules that address a particular type of problem and the rule's resulting *normal form*.

First some definitions of general relational database concepts:

Definition 1. A *relationship* is an association or connection between two different things.

Definition 2. A *relation* is a particular collection of data organized according to certain rules, represented in a relational database by a table.

Definition 3. An *attribute* is a data element represented in a relational database by a field.

Definition 4. The *primary key* for a relation is the attribute or combination of attributes used to differentiate one record from another within a table.

Definition 5. A *functional dependency* is a relationship between two attributes such that an attribute B is functionally dependent on an attribute A if, and only if, there must be a single B value associated with a given A value.

²Future versions of the Data Management Plan may include a more exhaustive discussion of normalization.

Definition 6. In a *multivalued fact* a single A value determines a specific set of associated B values.

Definition 7. A *multivalued dependency* exists when a relation contains more than one independent multivalued fact (or both a functional dependency and one or more independent multivalued facts).

Now the series of rules which when applied successively to a database yield a normalized database. The rule for first normal form requires relations defined within the system (tables) to conform to the definition of a proper relation given above.

Definition 8. The *Rule for the First Normal Form* is: eliminate repeating groups.

The rule for the second normal form eliminates problems associated with adding records (insertion anomaly), changing records (update anomaly) and removing records (deletion anomaly) when attributes are only partially dependent upon the primary key.

Definition 9 The *Rule for the Second Normal Form* is: all non-primary key attributes of a relation must be functionally dependent on the entire primary key.

The rule for the third normal form eliminates insertion, update, and deletion anomalies when attributes are functionally dependent on non-key attributes.

Definition 10. The *Rule for the Third Normal Form* is: each non-primary key attribute of a relation must be functionally dependent on the key and nothing else.

The rule for the fourth normal form further eliminates anomalies by eliminating multivalued dependencies.

Definition 11. The *Rule for the Fourth Normal Form* is: eliminate multivalued dependencies.

The fifth normal form (and beyond) is currently of academic interest only does not merit discussion here.

As stated in the body of the Information Management Plan, applying the normalization procedure described above yields a logical, robust, database design that eliminates redundant data and is not susceptible to insertion, update, and deletion anomalies.