

***APPENDIX I:
H-GAC DATA MANAGEMENT PLAN***

Data Management Plan

August 2011



HOUSTON-GALVESTON AREA COUNCIL
Community & Environmental Planning Department

*Prepared in cooperation with the
Texas Commission on Environmental Quality
under the authorization of the Texas Clean Rivers Act*

Table of Contents

Section	Page
Introduction	3
1 System Resources	3
1A Hardware	3
1B Software	3
1C Personnel	3
1D Data	4
1E Budget	4
1F System Schematic	4
2 Data Maintenance, Manipulation and Use	4
2A Quality Assurance/Quality Control	4
2A1 Data Limitations	4
2A2 Data Entry Protocol	4
2A2a Data Input	4
2A2b Data Dictionary and Metadata	4
2A2c Data Conversion	5
2A2d Coordinate Systems	5
2A3 Data Validation	5
2A3a Data Quality Control	5
2A3b Equipment Quality Control	5
2B Genealogy	6
2C Migration/Transfer	6
2D Data Security	6
2D1 Archives/Backup	6
2D2 Disaster Recovery	6
3 Client Services	6
3A Programming	6
3B Training	6
3C Data Access	6
3D Documentation	6
 Appendices	
Appendix A1 Data Source Information Sheet	7
Appendix A2 Data Log Sheet	8
Appendix B1 Hardware	9
Appendix B2 Software	9
Appendix C1 Data List	10
Appendix C2 Data Dictionary	11

Introduction

The Data Management Plan (the Plan) outlines the standard policies and procedures for data management within the Community and Environmental Planning (C&E) Department. The Plan covers the management of both tabular (non-geographic) and spatial (geographic) datasets. Its primary purpose is to ensure the efficient access and maintenance of these datasets within the Department's Geographic Information Systems (GIS) and Data Clearinghouse environments.

GIS technology provides a systematic means to capture, manipulate, analyze, store and display spatially referenced data. GIS supports a wide variety of applications ranging from site assessments, environmental planning, urban planning, and spatial analysis to support organizational strategies. In general, GIS supports the overall departmental goals of guiding regional planning, enhancing the quality of the region's natural environment and public education through outreach programs.

The H-GAC C&E department maintains both an internal spatial data warehouse (SDE) and an external site for downloading publicly available data (Data Clearinghouse). The SDE serves as the primary repository for data, metadata and other information relevant to the activities and goals of the C&E department. Datasets determined to be viable for public use are exported to the Data Clearinghouse website, thereby allowing the general public widespread access to this information via the internet. Members of the public may view and/or download any of the datasets that are posted to the Data Clearinghouse without the limitations imposed by hardware, software packages and organization barriers. In some instances these datasets are used in web-based mapping applications that are also accessible via the clearinghouse.

The Plan is considered a dynamic working document which responds to changing technology, funding, staffing, and project requirements. Consequently, the Plan is reviewed on an annual basis and amended as necessary.

1. System Resources

Following is an explanation of the System Resources required to support data management efforts.

1A. Hardware

The configuration of the hardware used by staff that perform GIS and data Management work is a "distributed network." This network consists of several PC's which are connected to central file servers. The department also uses a central web mapping server for online mapping applications.

The hardware includes three Windows Servers, one for Data Clearinghouse efforts, one for long-term Data Storage for in-house access and the third one for GIS web applications. A complete listing of departmental hardware is found in Appendix B1.

1B. Software

The software products currently used to accomplish the department's data management objectives are listed in Appendix B2.

1C. Personnel

The Data Management staff will be responsible for the maintenance and development of both the SDE database and Data Clearinghouse. These data management responsibilities cover a wide range from original data creation, acquisition and integration, data archiving and distribution. Additional responsibilities include enhancing the geographic extent, feature attributes, and metadata of the datasets and managing data distribution via options including the Internet, File Transfer Protocols (FTP), and CD-ROM technology.

The program is supported by three full-time and two hourly staff members, which support the program as well as other C&E programs. These staff members are part of the Socio-Economic Modeling program within C&E and provide data development, analytical, and web-based applications development support for program initiatives.

H-GAC's Data Services Department plays an indirect role in the implementation and maintenance of The Plan. The Data Services Department is responsible for managing the underlying hardware and network upon which C&E stores GIS data and implements GIS-based applications.

1D. Data

Department staff members will be consulted annually to determine priority needs for data management. Based on this consultation, specific data sets will be acquired or further developed for the various program areas represented in the department. The current list of department-specific data sets is shown in Appendix C1.

1E. Budget

Budgetary requirements to sustain data management efforts will be reviewed annually.

1F. System Schematic

The C&E Socio-Economic Modeling Program maintains a schematics representation of the data and applications architecture. This includes a list of datasets stored in the SDE and the Data Clearinghouse as well as access privileges to those datasets.

H-GAC's Data Services Department maintains a schematic representation of the agency-wide network. This includes network layout with basic hardware and software configurations for every personal computer in the agency.

2. Data Maintenance, Manipulation, and Use

2A. Quality Assurance/Quality Control(QA/QC)

QA/QC is designed to standardize screening, documentation, entry, output, analysis, correction, and updating of data in the system. QA/QC will document those responsible for data and system maintenance.

2A1. Data Limitations

Prior to the integration of data within the SDE and/or Data Clearinghouse, a review of the data set will be completed to determine predefined data limitations such as missing values, different sampling frequencies, multiple measurements, analytical uncertainty, censored or unavailable data, and duplicated data with existing data sets. After review of the data set, a report will be generated which records any errors detected and any corrections that may be necessary.

2A2. Data Entry Protocol

Specification of appropriate protocols for data entry, including standardization of data input and conversion.

2A2a. Data Input

Standard conventions for data input will be determined on a per project and or individual data set basis. To ensure Year 2000 Compliance, all data sets with date/time fields will include a four-digit year (YYYY). Either of the following formats will be used: International Standard Date notation where the date field is represented as MM/DD/YYYY (Month/Day/Year), or an ordinal format where the date field is represented as YYYYDDD.

2A2b. Data Dictionary and Metadata

Dictionary of data definitions and descriptions of all data sets, and attribute items.

A Data Dictionary and Metadata resource will be established to provide any data user with a full description of the data sets within the SDE/Data Clearinghouse. These resources provide detailed information such as data completeness, currency, intended use, coverage precision and projection system, annotation, and item types with definitions and related codes.

2A2c. Data Conversion

A standards method for converting data to agency formats.

Data to be imported into the GIS/Data Clearinghouse from hard copy, digital or by manual data entry, will follow a uniform conversion protocol to comply with the structure of current data sets. The type of data being converted will determine the protocol.

2A2d. Coordinate Systems

The Texas Stateplane Coordinate System, North American Datum 1983 (NAD83) will be the standard for geographic data at H-GAC. This coordinate system is based on the Cartesian coordinate system, or rectangular coordinates. When receiving geographic data from other sources the data will be transformed into the Stateplane Coordinate System to ensure compatibility with current data sets.

When publishing mapping services for use in web-based GIS mapping applications, the Web Mercator Auxilliary Sphere projection is used for all Data Frame projections. However, the underlying GIS data within these mapping services still use the Texas Stateplane Coordinate System, North American Datum 1983 (NAD83) projection.

2A3. Data Validation

2A3a. Data Quality Control

When data are received from any source, documentation will be created to include the source name, date received, format of data and a brief description of the contents. Data will be loaded onto the system from the media received and a review of the data will be made along with any corrections being made to the source documentation. An analysis will be made in order to determine the means of data entry into the system whether it is only a stand-alone database, a number of linked tables, or a geographic database. The data will be converted to the appropriate format for integration with the current system whether it is a conversion into MS Access, Excel, or ESRI ArcGIS. The data will be visually examined to determine its validity and accuracy. If the data is invalid it will be corrected (if possible) otherwise the data will be incorporated into the SDE/Data Clearinghouse and used in conjunction with existing data. A QA/QC report of all procedures and a detailed description of how the data was incorporated into the current system (from the date received to the date of integration) will be generated.

2A3b. Equipment Quality Control

One Windows 2000 Server

This server houses a portion of the Department's data clearing house and serves as an ftp server.

Two HP1055CM Plotter

The Data Services Departments of H-GAC maintains an HP1055CM plotter which is available for use by C&E Department staff. Cleaning and maintenance are completed on an as needed basis.

HP2500CM and LaserJet 4M Printers

The C&E Department maintains both the HP2500CM and LaserJet 4M Printers. Cleaning and maintenance are completed on an as needed basis.

Global Positioning System (GPS) Units

The C&E Department possesses two GPS units.

HP Scanjet 7400c

The CEP Department owns one network-accessible HP scanner.

Brother Intellifax 4750e

The C&E Department owns one fax machine

2B: Genealogy

Upon receipt of data from outside sources, all data will be screened for integrity and completeness. After the preliminary evaluation of the data, a log of the data source, type and completeness is created and maintained with the associated data. A description of the data and the responsible personnel are documented.

2C. Migration/Transfer

A copy of every C&E generated GIS dataset will be housed in the C&E SDE which C&E GIS staff manage the contents and structure of datasets. The underlying hardware and network connections for the SDE are maintained by the Data Services Department. Datasets that are of public interest will be placed in the Clearinghouse for public access. Transfer from the SDE to Data Clearinghouse will occur on an as needed basis following department QA/QC measures.

2D. Data Security

GIS and tabular data will be secure through directory permissions. H-GAC will employ Firewall or Proxy Server Technology to filter and severely restrict access to internal networks and database systems. Virus protection will be implemented to ensure system and data integrity.

2D1. Archives/Backup

H-GAC's Data Services Department will backup and archive C&E data at regular intervals.. A backup will be performed daily and the tapes will be maintained for 8 weeks before they will be recycled. Every six month, a complete system backup will be performed and the tapes will be archived and kept for five years off-site for security.

2D2. Disaster Recovery

H-GAC's Data Services Department will be responsible for Disaster Recovery.

3. Client Services

3A. Programming

Programming services will be provided on an as needed and resource available basis. All programming efforts will follow a standard procedure from needs assessment, program planning, development and testing, to refinement and documentation. The principal programming languages to be used in task automation and project customization will depend on the nature of the need and the current state of the technology. At this time, all web-based GIS applications are developed using the ESRI ArcGIS Server platform, and user interface components to that platform are developed using the Adobe Flex API.

3B. Training

Training for all users of the system is a critical part of *The Plan*. C&E staff directly responsible for data management will attend conferences, seminars, and software/hardware training courses as needed. H-GAC users of the system will be trained and/or receive technical support by the data management staff.

3C. Data Access

Data placed on the Data Clearinghouse will be available to those with Internet browsing and/or FTP capability. Data requests from staff from other agencies and the general public will be evaluated on an individual basis. When the data requests are received, a preliminary evaluation of the deliverable will be determined and a timeline and cost if applicable will be provided to the requesting agency or individual.

3D. Documentation

Documentation related to data management efforts such as system evolution, structure, and procedures for use will be compiled and made available for the end user. Documentation will be made available online and in hard copy format.

Appendix A1

SAMPLE DATA SOURCE INFORMATION SHEET

Data Title:

Source Agency:

Contact:

Title:

Address

Phone:

Data Description:

Data source:

Date created:

Accuracy:

Media:

Data items:

Description of data:

Format (specify what software)

Map:

Tabular:

Image:

Text:

Retrieval Procedure:

Command(s):

Appendix A2

Data Log Sheet

Date received: _____

Report Prepared by: _____

Source Name and Phone: _____

Format: _____

Media: _____

Check the following steps to determine the validity of the data:

1. What is the extent of the geographic area? _____

2. Structure (Circle One) *Vector* *Raster*

3. Scale? _____

4. Projection and Datum? _____

1. Do any of the key fields have missing values? If so which parameters have missing values? Yes ___ No ___

2. Any known duplicate records? Yes ___ No ___

Appendix B1

HARDWARE

This is a listing of a Departmental Windows-based Server Hardware.

1. NTCEIS01 - Windows 2008 GIS Web Application Server

Model: HP Proliant BL460c G6 Blade
CPU: Quad-Core Intel Xeon X5560 (2.80 GHz, 8M Cache)
Memory: 8GB
Internet Address: 204.65.99.189
Serial #: USE936RV4S
Hard Drive: 300GB
OS: Windows 2008
Purchased: January 2010.

2. NTIS04 – Windows SQL Server

Model: HP Proliant DL 380 G3
CPU: Single Intel Xeon 2800
Internet Address: 204.65.99.240
Memory: 1GB
Serial #: D313LDN1L122
Hard Drive: C = 16 GB, D=66 GB
OS: Windows 2000 SP 4
Purchased: April 2003.

Appendix B2

SOFTWARE

Word and Data Processing

Microsoft Office Pro (2007) - *Word, Excel, Access, Powerpoint, publisher, Infopath and Outlook.*

Graphics and Desktop Publishing

Macromedia Fireworks 4
Adobe Illustrator (v8.01) – *Graphics.*
Adobe Photoshop (v5.0) – *Graphics.*
Corel Draw (v7.0) - *Graphics.*
Quark Express (v5.0) - *Desktop Publishing.*
Paintshop Pro (v. 4.12)

World Wide Web Browsing and Development Software

Internet Explorer (v7) – *Primary Development Tool.*

Programming

Visual Basic (v6.0) – *Web Mapping Development Tool.*
MS Active Server Pages (v2.0) – *Web Database Development Tool.*
Adobe Flex Builder (v4.0) – *Web-based GIS application development tool.*

Geographic Information System

ESRI ArcGIS desktop (v10) – *Computer mapping and database manipulation capable of using ArcView, ArcInfo, and ArcEditor licenses as needed.*

ESRI ArcGIS Server (v10) – *Internet Mapping Application Server.*

Data Management

Access (2007, 2010) - *Relational Database.*

SQL Server(2000) - *Relational Database.*

Operating System

Windows XP - *PC working environment/Operating System.*

Windows 7 - *PC working environment/Operating System.*

Windows 2003 & 2008 - *Server Operating Systems.*

Appendix C1 Data List

Tabular Data Sets

Ambient Surface Water Quality Monitoring
Wastewater Self-reporting Data
Parcel-Based Land Use, Attributes, and Valuation (9 counties)
Census Data

Base Data Layers

Incorporated Cities & Census Designated Places
U.S. Census Blocks, Block Groups, Tracts, Urbanized Areas, and MSA
Election Precincts
City Ordinance, Zoning, Comprehensive Plans
County Boundary
Major Roads and Highways
Local Streets & Roadways (StarMap product)
Current Land Use (9 counties)
Forecasted Parcel-Based Land Use (7 counties)
Parks and Natural Areas
Eco-Logical Features
USFWS Wetlands
Farmland
Parcel Boundaries (9 counties)
TIRZs
Zip Codes
School Districts
Rivers (Surface Hydrography, Linear)
Lakes (Surface Hydrography, Polygon)

Recycling and Solid Waste Management Data Layers

Solid Waste Management Facilities
Closed Landfill Inventory
Recycle Center Locations

Water Data Layers

Watershed and Basin Boundaries
CRP Monitoring Station Locations
FEMA Designated Floodplains
USGS Flow Gages
OSSF Permit Locations
Wastewater Treatment Plant Outfall Sites
Wastewater Treatment Plant Service Area Boundaries
Aquifer Recharge Zones
Bio Monitoring Sites
USGS Steam Gauges
Sea Level Rise Model (1 to 35 Feet)
MS4 Permit Areas

Raster Datasets

Landcover 2008
Landcover 2002
NLCD Land Cover 2001, 2006

Appendix C2

Data Dictionary

<p>Data Dictionary Houston-Galveston Area Council Community and Environmental Planning Department</p>
--

General Information		
Thematic Layer Name		
Feature Class		
Topology		
Table Name		
Data Source		
Report Prepared by		
Phone	Fax	E-Mail

Attribute Table				
Variable	Begin Column	Item Name	Alternate Name	Item Definition

Data History
Source Agency
Originating Date
Originating Scale

Status Information
Percentage Complete
Planned Completion Date
Geographic Extent
Planned Enhancements
Known problems or limitations

Maintenance Information

Maintaining Office/Division/Section
Contact Name
Contact Telephone Number
Type of updates performed
Frequency of Updates

Data Format Information	
Data Format	
Software/Version	
Number of features/records	
Total File Size	

Projection	
Geographic Projection:	
Spheroid:	
Zone:	
Datum:	
Units:	
Fips Zone:	
Quadrant:	
X Shift:	
Y Shift:	
1st Standard Parallel:	
2nd Standard Parallel:	
Central Meridian:	
Lat. of Projection Origin:	
False Easting:	
False Northing:	

Additional Documentation	
Quality Assurance Quality Control	
Attribute Reports Available	
Additional Documentation Available	