CPRA DATA MANAGEMENT PLAN

January 31, 2013

THE WATER INSTITUTE
OF THE GULF

Table of Contents

List of Acronyms iii
Figures vi
Tables vi
Acknowledgements 1
1. Overall Data Management Plan Objectives 2
1.1 Plan Components 2
1.2 Standards 2
1.3 Policy 3
1.4 Resources 3
1.5 Data 3
1.6 User Interfaces 3
1.7 Security 4
1.8 Architecture 4
1.9 Implementation Activities 4
2. Standards 6
2.1 Standards Group 6
2.2 Enforcing the Standards through Policy 8
3. Policies 10
3.1 Policy Benefits 10
3.2 Major Components 10
3.3 Capturing Data in Accordance to the Plan 12
4. Resources 14
4.1 Personnel 14
4.2 Roles 14
4.3 Systems Capability and Recommendations 16
5. Existing or Anticipated Data 17
5.1 Data Streams 17
5.2 Data Organization 23
6. User Interface 29
6.1 Wiki or Portal Features – Usability 29
6.2 Portal or Wiki Features - Extensibility 30
6.3 Extensible Architecture 30
6.4 Success Metrics 31
6.5 Feedback or Alert Mechanism 31
6.6 Dashboards 31
6.7 Citation Tool 35
6.8 Project Reports / Summaries 37
6.9 Query Functionality 37
7. Security 39
7.1 Assignment of Authority 39
7.2 Portal Security - Best Practices to Secure Portals 40
8. System Architecture 41

8.1 Open vs. Proprietary Software: Considerations 41
8.2 GIS Specific Software Considerations 42
8.3 Hosting Environments: Considerations 46
8.4 Web Services 48
8.5 Selecting Storage Media for Long-Term Preservation 51
9. Implementation Activities 53
9.1 Short-Term Goals 53
9.2 Long-Term Goals 55
APPENDIX A Preliminary Inventory of Data Management Systems in Coastal Louisiana 57
PROGRAMS 57
ORGANIZATIONS 64
APPENDIX B Guide to the 2012 Coastal Master Plan Modeling Data 68

List of Acronyms

API	Application Programming Interface
ASCII	American Standard Code for Information and Interchange
BICM	Barrier Island Comprehensive Monitoring
BLOB	Binary Large Object
CERP	Comprehensive Everglades Recovery Plan
CPRA	Coastal Protection and Restoration Authority
CRMS	Coast-wide Reference Monitoring System
CSDGM	Content Standard for Digital Geospatial Metadata
CSV	Comma-Separated Values
DHS	U.S. Department of Homeland Security
DMP	Data Management Plan
DOD	Department of Defense
DOI	Digital Object Identifiers
DQO	Data Quality Objectives
ENC	Electronic Navigational Charts
EPA	U.S. Environmental Protection Agency
ESIP	Embedded Computer Resources Support Improvement Program
ESRI	Environmental Systems Research Institute
ESRL	Earth System Research Laboratory (NOAA)
FEMA	Federal Emergency Management Agency
FGDC	Federal Geographic Data Committee
FTP	File Transfer Protocol
GB	Gigabyte
GCMD	Global Change Master Directory (NASA)
GIS	Geographic Information System
GOHSEP	Governor's Office of Homeland Security and Emergency Preparedness
GUI	Graphical User Interface
HTTP	Hypertext Transfer Protocol
HAZMAT	Hazardous Material
IHO	International Hydrographic Organization
I/O	Input/output
ISO	International Organization for Standardization
IS&T	Information Services and Technology
KML	Keyhole Markup Language
KPI	Key Performance Indicator

LAN	Local Area Network
LASARD	Louisiana Sand Resource Database
LDEQ	Louisiana Department of Environmental Quality
LDHH	Louisiana Department of Health and Hospitals
LDWF	Louisiana Department of Wildlife and Fisheries
MRHDMS	Mississippi River Hydro-dynamic and Delta Management Study
MTBF	Mean Time Before Failure
NAP	North American Profile
NASA	National Aeronautics and Space Administration
NCDC	National Climatic Data Center (NOAA)
NESDIS	National Environmental Satellite Data and Information Service (NOAA)
NetCDF	Network Common Data Form
NTFS	New Technology File System
NWIS	National Water Information System (USGS)
OGC	Open Geospatial Consortium
OMB	Office of Management and Budget
OS	Operating System
OSS	Open-source Software
PH	Hydrogen Ion Activity (in relation to water quality)
PI	Performance Indicator
Portal	A Web Site that Brings Information or Links from Diverse Sources and Presents in a Uniform User Interface
PSD	Physical Science Division (NOAA)
QA/QC	Quality Assurance / Quality Control
RDBMS	Relational Database Management System
RDF	Resource Description Framework
SDI	Spatial Data Infrastructure
SDSFIE	Spatial Data Standards for Facilities, Infrastructure, and Environment
SI	Service Interfaces
SLA	Service Level Agreements
SOAP	Simple Object Access Protocol
SONRIS	Strategic Online Natural Resources Information System
SQL	Structured Query Language
SSO	Single Sign-On
TB	Terabyte
TSM	Tivoli Storage Manager
UCAR	University Corporation for Atmospheric Research
UCP	UCAR Community Programs
UDDI	Universal Description, Discovery, and Integration

UI	User Interface
UIC	UI Components
UIP	UI Process Components
UNF	Unique Numeric Fingerprint
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
USACE	U.S. Army Corps of Engineers
USGS	U.S. Geological Society
Users	Contributors, Consumers, and Managers of a System
VDS	Virtual Dedicated Server
VPS	Virtual Private Server
W3C	World Wide Web Consortium
WCS	Web Coverage Service
WFS	Web Feature Service
WDC	World Data Center for Meteorology (NOAA)
wiki	Website which allows users to add, modify, or delete content
WSDL	Web Services Description Language
WMS	Web Map Service
XML	Extensible Markup Language

Figures

Figure 1. Example diagram of client-to-portlet information exchange. 30
Figure 2. Example operational dashboard. 33
Figure 3. Example analytical dashboard. 33
Figure 4. Example analytical dashboard. 34
Figure 5. Example of a digital object identifier (DOI). 36
Figure 6. Example architecture of a simple website that uses the built-in functionalities and data provided by the Maps APIs. 50

Tables

Table 1. Data management workgroup membership, roles and responsibilities. 1
Table 2. Data streams typically incorporated during emergency response efforts. 20
Table 3. Scorecard approach for evaluating currently available media against selection criteria. 52

Acknowledgements

This effort was funded by the Coastal Protection and Restoration Authority (CPRA) Cooperative Agreement No. 2503-12-58, Task Order #2 (Data Management). The Water Institute of the Gulf would like to acknowledge and thank the participants on the Data Management Workgroup for their efforts, as well as the efforts of the Weston team for their work on drafting this document. Data Management Workgroup Membership, Roles and Responsibilities are listed in Table 1. It is envisioned that the plan described herein will serve as the starting point for the implementation of a data management system for CPRA. This document was designed to capture new and innovative ideas that can be applied to the development of a broad and highly functional data management system for coastal Louisiana.

Table 1. Data management workgroup membership, roles and responsibilities.

Membership	Role	Responsibility
The Water Institute		
Lead: Ehab Meselhe Support: Denise Reed and Alaina Owens	Principle Investigator	• Direct and coordinate the development of the Data Management Plan. • Provide advice and guidance. • Facilitate communication. • Ensure task progression and completion.
Coastal Protection and Restoration Authority		
Lead: Ed Haywood Support: Christopher Robertson	Project Manager	• Provide local data management expertise and guidance. • Provide oversight and review. • Approve the final plan.
Weston Solutions		
Lead: David Bordelon Support: Bradley Morgan, Eric Bay, Kenneth Westerman, and Curtis Vaughn	Primary Sub-contractor	• Facilitate the workshop. • Draft, revise, and finalize the Data Management Plan.
Deltares		
Lead: Edwin Welles Support: Fedor Baart	Advisor	• Provide national and international perspective. • Advise on Dutch and other coastal data management efforts. • Review and comment on the plan.
USGS		
Lead: Craig Conzelmann Support: Christina Hunnicutt	Advisor	• Provide a federal perspective. • Advise on standards and integration of model-related data. • Review and comment on the plan.
Shaw		
Lead: Melany Larenas Support: Garvin Pittman	Advisor	• Advise on data collection and storage standards, including LASARD and CRMS. • Review and comment on the plan.

1. Overall Data Management Plan Objectives

The State of Louisiana Coastal Protection and Restoration Authority (CPRA) creates, consumes, and disseminates large amounts of electronic data during its day-to-day operations. As such there is a need for a Data Management Plan which provides guidance regarding a data management system and the data contained within. This document communicates the resource requirements, policy needs, data handling procedures, and storage structures necessary to develop and maintain a data management strategy that facilitates access to data associated with the state of Louisiana's coastal protection and restoration efforts.

Through the combined efforts of the scientific community and data managers, a well-designed, forward-looking data management system will increase efficiency, facilitate awareness and communication of data resources, and create a lasting secure repository for data. The standards and policies developed by data stakeholders will enable a complete catalogue of existing data streams to be made available in a scalable and adaptable fashion. This Data Management Plan has been created to provide the consumers, contributors, and managers (the users) with guidance to maintain the organization and accessibility of information associated with coastal protection and restoration. The primary goal of the Data Management Plan is to expand the use of existing data by creating a system that the end users value. Therefore, the system's success will be measured by the user group's utilization of the available data and the scientific community contributing to the system. A dedicated team of personnel and access to the resources necessary to develop and maintain the system will be required to achieve these objectives.

1.1 Plan Components

The Data Management Plan is designed to provide the user with a set of guidelines that will facilitate the establishment of a data management system and the procedures and policies necessary to support and maintain the system. The success of a data management system is dependent upon each of the primary components complimenting one another to form a cohesive system that provides secure archival of information, user access, and reliable data. To establish such a system, the stakeholders and the implementation team must work together in alignment with the overarching goal of effectively managing and providing access to the data.

The Data Management Plan will focus on two areas of concern, data integrity and data access. Data integrity will be managed through a set of policies, standards, and resources (personnel, hardware, software), to ensure quality and organizational structure. Access to the data will be achieved through architecture, user interfaces, and security policies. Sections 2 through 8 will provide guidance to the system administrators and users to design and maintain these components in a manner that will maintain the integrity and access, as well as the flexibility to manage ever increasing and complex data in the future. Information below is provided to the reader as an overview of the components.

1.2 Standards

Data standards will promote efficient sharing of information among the targeted user groups through the establishment of well-defined and formatted data elements. Standards will be detailed and enforced through policy. Common principles and standards must be applied uniformly and as part of a

coordinated effort between all users and administrators to maintain the integrity of the data. A standard can be defined as something agreed upon as a model, example, or rule for the measure of quantity, weight, extent, value, or quality. Data standards are documented agreements on representations, formats, and definitions of common data. Details regarding data standards are provided in Section 2 of this plan.

1.3 Policy

Policies will focus on establishing the principal requirements for building and maintaining the data management system. Policies are the requirements necessary to maintain a secure, organized, and high quality system. The requirements detailed within the Policy section will require means of enforcement such as automation of backups and entry filters to reject non-compliant data. The section will also provide an estimated level of effort, title and position descriptions, and the requirements necessary to conduct reviews and confirm quality control systems are effectively addressing the review processes such as formatting and submission standards. Details regarding data policies are provided in Section 3 of this plan.

1.4 Resources

The operation and maintenance of the data management system will rely on system hardware and software, a dedicated group of personnel, as well as adequate financial resources. Once the extent and general architecture of the system is known, the hardware requirements for the hosting system (cloud server, routers, local area network [LAN], etc.) will be determined and submitted to the CPRA management for consideration. Personnel associated with the data management system will include one or more Administrator(s), Quality Control Analyst(s), Programmer(s), and position(s) associated with user support. A workgroup established to act as a gatekeeper for data streams is also recommended to expedite inclusion of new sources and to develop methods to facilitate seamless integration of data. Details regarding necessary resources are provided in Section 4.

1.5 Data

The primary contributor to the data management system will be the CPRA, with additional data streams provided by various federal and state agencies, academia, and private contractors. Within these groups, existing programs such as the Coast-wide Reference Monitoring System (CRMS), Louisiana Sand Resource Database (LASARD), Strategic Online Natural Resources Information System (SONRIS), and data associated with the CPRA Master Plans will be the building blocks of the system. Integrating these programs into a single clearinghouse will provide users with the ability to access and utilize data from across the state and country, eliminate duplication of efforts, and derive the most value from the resources expended to develop the data. Section 6 provides current data streams and an overview of the scope and purpose.

1.6 User Interfaces

The data management system will be comprised of a portal or wiki which will serve as the entry point for the user. Through this platform, the user will have access to programs, data streams, dashboards, and documents associated with Louisiana's coastal restoration and protection programs. The design

and functionality will be determined by the CPRA and stakeholder group to maximize access in an intuitive manner. Dashboards will be incorporated to provide users with summaries of projects and programs, and can be tailored to the primary user groups. User interfaces are described in more detail in Section 6.

1.7 Security

System security will include both the methodology to maintain data in long-term storage and the access rights for individuals and groups. The stakeholder group and contributing agencies will need to develop a matrix of expected user groups and data streams and the access rights granted to each group. Restricted data will be protected using system architecture and passwords. Details regarding security are provided in Section 7.

1.8 Architecture

The system architecture will serve as the means to unify the various components into a functional system. The architecture will consider the hosting method, software, web services, format requirements, and archiving. The architecture will need to consider the exponential nature of data growth and complexities. Available options regarding the system architecture are found in Section 8.

1.9 Implementation Activities

Section 9 provides a general list of both short- and long-term activities that the stakeholder group identified as the steps necessary to approach the development of the data management system.

Short-Term Goals

To expedite the benefits of a comprehensive coastal protection and restoration data management system, short-term goals have been established to create the structure and measure the effectiveness of the system. The objective is to achieve these goals within 6 months to 2 years; however, as priorities change and technology advances, the plan will be updated as needed. The initial short-term goals were provided by CPRA to meet their immediate needs (details in Section 9):

- **Establish standards and policies** (details in Sections 2 and 3) to which data will conform. These policies and standards will be made available to the users in a Standards Library. Create a Standards Group, selected by the CPRA, to maintain and continually develop these standards.

- **Create a complete catalog of existing and anticipated data streams** (details in Section 6).

- **Identify a Pilot location (Basin) with a reasonable historical dataset** to use as a testing vehicle for the Data Management System being developed herein. Specifically, the following tasks will be applied to this pilot Basin:

    a. **Develop system architecture (details in Section 8) to include interfaces such as a portal and dashboards** designed to efficiently guide users to their interests. Dashboards are expected to include items such as executive level progress reports, project specific data and summaries, status of programs, and scientific research initiatives. The CPRA has

identified the three dashboards that should be developed in the short term: Executive Summary; Operational; and Analytical.

b. **Establish training workshops for internal and external users for the Pilot Basin**. This would provide direct and early feedback on the Data Management System developed herein. This pilot study would provide valuable information on the frequency at which trainings should occur and identify the training staff. It would also facilitate developing guidelines and tutorials for each of the data streams.

- **Identify personnel and equipment resources required to develop and maintain the data management system** (details in Section 4).

- **Utilize the 2012 Coastal Master Plan modeling inputs, outputs**, etc. to develop an initial component of the data management system. Developers will coordinate their efforts with the modeling data contributors and users to create a secure repository for future use and accessibility. The system will catalog and make such model outputs available for re-consumption. Such an effort will establish and provide standards for future Master Plans and other modeling efforts.

- **Develop a method to measure system usage and success** to demonstrate the value of the efforts put forth in the development of the data management system and continued improvement of the users' experience.

Long-Term Goals

Long-term goals are targeted for completion within two to five years, and will continue to be modified as needed (see Section 9 for more details). Long-term goals include:

- **Extend the system established for the Pilot Basin to other basins across the coast of Louisiana**. Eventually, the entire Louisiana Coast should be incorporated into the Data Management System.

- **Implement a method to capture, catalog, and make available final outputs from the 2017 Coastal Master Plan** data within the data management system.

- **Incorporate fragmented or disparate databases and their data into the data management system**.

- **Identify legacy data and develop the tools to capture the information** in the formats specified in the Data Management Plan.

- **Develop standards for the incorporation of new data** previously not addressed.

- **Develop automated tools for communicating data updates** to registered users.

- **Develop and implement a plan for the operation and maintenance of the data management system**.

- **Develop personnel and equipment resource projections to accommodate the anticipated growth of data** and information relevant to the Coastal Protection and Restoration Authority.

2. Standards

Data standards are a set of guidelines developed to ensure consistent collection and exchange of information. Standards are definitions designed to seamlessly incorporate data from different user groups and are flexible enough to cover a wide range of data types and methods of collection. Standards describe in detail, the Data Quality Objectives (DQOs), and the format in which data will be delivered to the group.

The value derived by following a well-defined set of data management standards will benefit the contributors, end users, and data managers. DQOs and pre-established formats provide the contributor with the requirements necessary to achieve the program objectives, reducing the level of effort needed to develop these standards on their own or to correct non-conforming data at a later date. When the contributors follow the established standards, the data managers will be more efficient with the data intake and categorization. Data will be well defined and organized to facilitate the use of established querying tools developed by the data managers. The end user gains value through efficient searches, known quality control standards, and trust in the system due to the front end requirements sifting out data which does not meet the standards.

Applying such data standards across information stored in the data management system will allow disparate data sets to become comparable and integrated. As data becomes normalized, the user's ability to manipulate datasets for comparison, validate similar studies, conduct quality control analysis, and realize additional value from models developed for the datasets, will increase substantially. Standardized data will also increase the data's use by the scientific community outside of those involved in the state's coastal restoration and protection programs. This increased access to the information will allow visibility of the data by these experts and, potentially, new funding, ideas, or processes that could benefit the state.

The Spatial Data Standards for Facilities, Infrastructure and Environment (SDSFIE) http://www.sdsfie.org/ is another data model standard that has been published. It is the Department of Defense Standard but is open to all users.

2.1 Standards Group

The first step toward establishing the standards will be to convene a Standards Group that will have the authority to make decisions regarding the standards and policies to be utilized for data to be incorporated in the data management system. The Standards Group will create a Standards Library to document the system's requirements and DQO's. This group will be responsible for reviewing existing standards currently employed by the State, and either incorporating or modifying these requirements for data. The group will also review standards utilized by other agencies, international agencies, and academia to mine existing endeavors to establish data management standards. When previously established standards do not apply, the Standards Group will develop data specific requirements and publish them to the library. The group may convene a panel of project scientists and engineers with similar study outputs to discuss the most comprehensive, efficient manner of standards to utilize.

The Standards Group will develop type-specific nomenclature for all data submitted to the agency. This nomenclature will be made publicly available through the Standards Library and should be included in

any agreement in which data is to be collected. CPRA policy will require adherence to the naming convention developed by the Standards Group for all delivered data.
Examples of established standards:

- The Federal Geographic Data Committee (FGDC) is an interagency committee that promotes the coordinated development, use, sharing, and dissemination of geospatial data on a national basis. This nationwide data publishing effort is known as the National Spatial Data Infrastructure (NSDI). The FGDC develops geospatial data standards for implementing the NSDI and in consultation and cooperation with state, local, and tribal governments, the private sector and academic community, and to the extent feasible, the international community. The FGDC develops geospatial data standards only when no equivalent voluntary consensus standards exist in accordance with Office of Management and Business (OMB) Circular A-119.[1]

- The largest and most followed industry standard for Open GIS is created by the Open Geo-spatial Consortium (OGC) http://www.opengeospatial.org/ogc. This consortium has developed and maintains standards that are followed by many agencies and organizations. The OGC includes international members, as well as many United States businesses, universities, and government agencies (Lockheed Martin, University of Texas, Department of Defense, etc.) They offer a robust offering of standards that can be implemented as part of an overall data model including standards that cover most of the data streams identified in this plan.

- An International Organization for Standardization (ISO) standard is developed by a panel of experts, within a technical committee. Once the need for a standard has been established, these experts meet to discuss and negotiate a draft standard. As soon as a draft has been developed it is shared with ISO members who are asked to comment and vote on it. If a consensus is reached the draft becomes an ISO standard, if not it goes back to the technical committee for further edits.[2]

- The Spatial Data Standards for Facilities, Infrastructure, and Environment (SDSFIE) is a community standard, registered in the U.S. Department of Defense (DoD) Information Technology (IT) Standards Registry and recognized as the enterprise standard for geospatial data across the entire DoD business mission area. It is being managed by the Defense Installation Spatial Data Infrastructure (DISDI) Group. Collectively, the DISDIG is responsible for the development and maintenance of SDSFIE with the focus on defining the data architecture necessary for Installation Geospatial Information & Services (IGI&S) capabilities and requirements which support DoD installations, environment, and civil works missions.[3]

- NOAA data that is available for download is provided in appropriate industry standard formats. The following is a list of common standards organizations that adopt, develop, and endorse standards related to geospatial data access, data publication, and data formats. Those organizations are:

    - Open Geospatial Consortium
    - Federal Geographic Data Committee
    - International Organization for Standardization[4]

[1] http://www.fgdc.gov/standards/
[2] http://www.iso.org/iso/home/standards_development.html
[3] https://www.sdsfieonline.org/default.aspx
[4] http://www.cio.noaa.gov

- Network Common Data Form (NetCDF) is a data model / file format for array-oriented scientific data, a freely distributed collection of access libraries implementing support for that data model, and a machine-independent format. Together, the interfaces, libraries, and format support the creation, access, and sharing of scientific data.[5]

- CPRA "A Contractor's Guide to Minimum Standards for Contractors Performing GPS Surveys and Determining GPS Derived Orthometric Heights within the Louisiana Coastal Zone."[6]

- CPRA "A Standard Operating Procedures Manual for the Coast-Wide Reference Monitoring System-Wetlands: Methods for Site Establishment, Data Collection, and Quality Assurance/Quality Control."[7]

- CPRA "Report on Methodology to Catalogue Geoscientific Data for the Louisiana Sand Resources Database (LASARD)." LASARD Data Specification, a guide is presented that describes individual data formats for various survey data, which will be included in the data management system.[8]

2.2 Enforcing the Standards through Policy

Along with the authority to develop the Standards Library, the Standards Group must also have the authority to enforce data compliance to the standard. Whether through a simple filtering tool utilized to reject nonconforming data or through QA/QC reviews, information not following the established standards will not be allowed to be stored and made public. The data contributors will be given access to tools at the Standards Group's disposal which would facilitate updating the data to the systems requirements. Similar issues of nonconformance are expected to be encountered when examining legacy data. The Standards Group will be tasked with making a cost/benefit determination regarding the level of effort needed to update legacy data to current data requirements. Further discussion of legacy data issues are presented in Section 5.

One approach by which the standards and policies adopted by the CPRA can be enforced is to specify the requirements in contracts with data collection and reporting in the scope of work. The clear, contractual specification should leave no room for varying interpretation and include consequences for nonconformance. This will ensure that all future updates, enhancements, or alterations of the data management system will consistently follow the overall policies.

Validators

All data sets (and changes to existing data sets), whether submitted manually or through an automated system, should be subjected to a validation process before being published to the repository. Data validation should start with the development of rules or requirements. Rules can be collected through a requirements capture exercise conducted by the Standards Group. This exercise is designed to categorize each type of data in the system and to summarize the required properties of that data category. For instance, a common data category is position data. By establishing the required properties of position data, the validator can verify the data has latitude, longitude, and elevation

[5] http://www.unidata.ucar.edu/software/netcdf/docs/standards.html
[6] http://sonris-www.dnr.state.la.us/dnrservices/redirectUrl.jsp?dID=4075091
[7] http://sonris-www.dnr.state.la.us/dnrservices/redirectUrl.jsp?dID=4175395
[8] http://sonris-www.dnr.state.la.us/dnrservices/redirectUrl.jsp?dID=3904015

values, as well as a coordinate system reference. This type of process can be conducted for all data categories to determine their validation requirements.

Validation processes may include such methods as:

- **Allowed Character Checks** verify that only expected characters are present in a field.
- **Cardinality Checks** verify that each record has a valid number of related records.
- **Consistency Checks** do comparisons that verify consistency between fields. For instance, if the data record Title field contains the data "Ft.," then the check will fail if the Distance field does not contain the data "Feet."
- Other simple validation methods are: Data Type Checks, Limit Checks, Presence Checks, Range Checks, Table Lookup (Valid Value List) Checks, and Uniqueness Checks.

Post-Validation Actions

When the data set is submitted for validation, the system must take action when the validation fails. Depending on the type and severity of the failure, the required action falls into two general categories: Enforcement Actions and Advisory Actions.

Enforcement action typically denies accepting or updating the data set and requires fixes to the data set before it can be accepted into the system. Another form of enforcement action involves automatically changing the data and saving a conformant version instead of the original version. This is most suitable for cosmetic changes, but is not appropriate in situations where the enforcement leads to loss of information, such as truncating strings if the length of the string is longer than the data structures allow.

Advisory actions typically allow the data to be accepted into the system unchanged, but send a message to the submitter or source actor indicating the validation issues encountered. This is most suitable for non-interactive systems, for systems where the change is not technically relevant, for cleansing existing data, and for verification steps of an entry process.

Data Conventions

The methods and principles that dictate the ways in which data can be stored, retrieved, and manipulated follow what are called "conventions." As technology advances, the conventions evolve over time and new species of conventions come to life. Data conventions that have not yet become standard for a large number of users will be something that this data management system must be flexible enough to allow for in the future. Generally speaking, when these new conventions are adopted and data sets adhering to them are allowed to be uploaded into the data management repository, the conventions should be documented according to the same stated policies for documentation as the organization's Standards. The International Standard ISO 11179[9] defines how to develop and implement naming and metadata conventions.

[9] http://metadata-standards.org/11179/

3. Policies

Policies will provide the procedures and identify the staff resources necessary to execute the Data Management Plan. Setting policies ensures a structured and uniform approach to managing data, and clearly assign responsibilities to personnel. An established guide for creating, storing, and disseminating data will provide those involved with the data management system the ability to plan the data collection components of contracts, projects, and studies in advance and to create efficiencies in the data intake and distribution. The policies should be dynamic in adapting to various data streams, while providing definitive guidance regarding the quality, format procedures, schedules, etc.

3.1 Policy Benefits

Systems backed by established policies provide a much higher return in the quality of data presented as well as a fundamental approach to designing solutions for the issues at hand. The policies should be flexible enough to work through any issues that may arise such as setting a procedure for a data quality objective that may be unattainable in real world collection. For example, setting a procedure of collecting minute-by-minute readings may greatly increase the size of the data. However, setting a 15-minute average can still fulfill the reporting requirement without putting an unneeded burden on the database.

As in the case of standards, a means to enforce policy is necessary to maintain the quality and accessibility of the data. The enforcement of policy needs to be specific and maintained throughout the life of the system. The system administrator will utilize constraining filters and periodic reviews of all data to ensure the information is being captured in accordance with the Data Management Plan. Any discrepancies can then be rectified early in the process so as not to further delay submittal of the data.

3.2 Major Components

Data Ownership

Depending on the contractual language, all data submitted to the CPRA through contractual obligations becomes the property of CPRA. If a shared ownership agreement is reached, the details in storage and security of the data will be evaluated and formulated prior to the signing of the agreement.

Enterprise data is owned by the CPRA along with supporting metadata. The data will be modeled, named, and defined consistently according to the standards. It will be stored close to the source as possible. Any post processing that is performed will be captured in the metadata, and the raw data will be maintained alongside any processed data.

Shared data is partially owned by the CPRA as well as any other party involved. Shared data ownership agreements will clearly define the collection methods, processing, storage as well as the agreed upon level of dissemination.

Quality Assurance/Quality Control

A QA/QC policy will accompany every contract that will produce data deliverables. The policy will state the goals for the data (formats, QC requirements, delivery method, etc.) as well as a process for addressing data determined to be noncompliant.

Security

A security policy will be setup in regards to the display of data on any portal or dashboard as well as the dissemination of the data to any third party or to the public in general. Differing levels of security will be provided to various groups wishing to incorporate the data for their specific use.

Accessibility

An accessibility policy will be setup that dictates accessibility to the system. The State of Louisiana already has published web access standards and mandates that adhere to the Electronic and Information Technology Accessibility Standards (Section 508) which is the federal accessibility standard designed to ensure access to all users.[10] The accessibility policy for the proposed system should follow all state and federally mandated guidelines.

Metadata Requirements

A metadata requirements policy should be set up that will dictate that metadata is a requirement for all data deliverables. Metadata is defined as information about data providing context and detail as to how the data was collected. The metadata should include collection methodology, when it was collected, under what meteorological conditions, and any other information that could possibly skew the data. Citation information should also be included. The use of Digital Object Identifiers (DOI) can enhance the use of citations and provide further metadata information. Requiring metadata in a standard format can greatly enhance the usability of data associated with the system because it provides context to why the data was collected and any assumptions used when collecting the data. This will in turn increase the confidence with which the data can be applied to a research problem.

Collaboration Agreements

Collaboration agreements between the CPRA and other parties will provide documentation of the communication and understanding of the system requirements (standards and policies). Agreements will be pursued to ensure complete data sets are made available to the CPRA and to provide formal documentation of the transfer. The methodology of transfer, security measures, purpose, and documentation of transfer will be agreed upon by all parties prior to the transfer of data. Agreement forms will accompany the data as metadata to provide a legal chain-of-custody for the data. This will provide a level of confidence and protection for CPRA when incorporating data collected by other parties while increasing the quality and relevance of the data provided.

[10] www.access-board.gov/sec508/standards.htm

Data Archiving

A formal, ongoing process, with associated stakeholder input, is needed to determine the archive recommendations for data. A high-level, enterprise-wide approach, driven by user benefits, should ensure that irreplaceable data are preserved, and should explicitly incorporate broad user engagement and coordination with other agencies. For some data sets, archiving requirements may be explicitly spelled out in agreements with other organizations or agencies.
Examples of data archiving include:

Tape Backups—A hardware tape backup system, with a defined procedure for storage of at least one backup copy at a safe, external location (e.g., a safety deposit box). A minimal recommendation would be a weekly full backup and daily differential backups. Differential backups backup what has changed since the last full backup and are preferred because only two media sets are required, the full and the last backup to restore. For the database, daily backups through the backup software would be a minimal recommendation; however, the frequency can be increased depending on the system's reliability and established policies.

Cloud Storage—Some examples of private sector storage resources include the following:

- **Amazon S3.** Requires client software and has no encryption support.[11]
- **S3-based Remote Hard Drive Services.** Some examples include Elephant Drive and Jungle Disk.[12,13]
- **Carbonite.** Free client software that includes 1024-bit Blowfish encryption (an encryption algorithm used for security).[14]

3.3 Capturing Data in Accordance to the Plan

Contractors

Policy could stipulate that contracts that contain a data element as part of a set of deliverables must adhere to the data standards associated with the type of data. All data must be created in an appropriate format, must include metadata describing its collection methods and scope, and stored and backed up securely and appropriately for the type of data collected. Upon completion of the data submittal, the project manager shall review the Data Management Plan in association with the deliverables to ensure that all standards have been met or exceeded.

[11] http://aws.amazon.com/s3/
[12] http://home.elephantdrive.com/
[13] https://www.jungledisk.com/
[14] http://www.carbonite.com/en/v2/index?tp=hpoc

Agency

Policy could stipulate that all CPRA users will be required to comply with the standards and data handling procedures developed during Data Management Plan implementation. The system designers and administrators will engage the various user groups to encourage their input into the architecture and what tools will encourage increased use and adherence to the standards and policies.

Enforcement Policy

Policy could grant the System Administrator the authority to reject data which does not adhere to the Data Management Plan. Rejection of data either through compliance filters or manual review should include a detailed notification to the data author explaining and referencing the standards and/or policies violated. Data manager support contact information should be provided with the notification to provide the author with additional resources as necessary. Policy examples are provided below:

- The National Science Foundation (NSF) provides grant awardees with guidelines and sets out special conditions applicable to grants to implement the Foundation's Sharing Policy. The NSF policy expects investigators to share with other researchers, at no more than incremental cost and within a reasonable time, the data, derived data products, samples, physical collections and other supported materials gathered or created in the course of the research project.[15]

- Louisiana Department of Health and Hospitals provides policy guidance to the Louisiana CAREWare Access Network (LaCAN) data management group detailing the security, data sharing, hardware requirements, and data entry in the referenced document.[16]

- The Digital Curation Centre (DCC) is collecting examples of explicit policies on research data and examples of existing policies amended to encompass research data. For many institutions, effective research data management requires formal policy for support and guidance.[17]

[15] http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/aag_6.jsp#VID4
[16] http://new.dhh.louisiana.gov/assets/oph/Center-PHCH/Center-PH/hiv-aids/2012/HIVAIDSHOME/LaCanProject?LaCANPolicies.pdf
[17] http://www.dcc.ac.uk/resources/policy-and-legal/institutional-data-policies/uk-institutional-data-policies

4. Resources

The operation and maintenance of the data management system will rely on a dedicated group of personnel, adequate financial resources, and system hardware and software. Dedicated staff or other financial resources are the single most important factor to building and maintaining a successful data management system. Technical system resources are driven by the requirements that will be created during the architecture phase of implementation. This section is intended to show the necessary dedicated staffing resources of existing systems to demonstrate the criticality for the overall system success. Utilizing the overall schema presented in this document, an outline of the resources potentially necessary has been presented. Examples of resources necessary to support systems have been presented for planning purposes.

4.1 Personnel

Current Personnel and Level of Effort

Currently, CPRA utilizes resources provided by the Louisiana Department of Natural Resources (LDNR) through an interagency agreement to provide data management services. Current positions include System and Database Administrators, Technical Support, and Application Developers. The System and Database Administrators and Technical Support resources are used for both CPRA and LDNR. Some Application Developers are dedicated to the CPRA and range between three and five personnel.

The CPRA also has contract support from the U.S. Geological Survey (USGS) for GIS database support and maintenance. Three full-time positions are housed at the CPRA offices. These three personnel provide database maintenance and mapping support.

In addition to the staffing resources noted above from the USGS and LDNR, the CPRA has had up to five internal staff dedicated to overall data management. In addition, CPRA Operations Division staff in the regional offices handle second-level QA/QC after it has been submitted by contractors. Data collection is primarily conducted by contractors, and CPRA has developed remote data entry procedures to facilitate efficient data entry. Depending on the time of year and amount of data submitted, QA/QC responsibilities can be time consuming. Current CPRA data management staffing levels are being reevaluated based on internal reorganization. In general, CPRA has assessed that the current personnel resources are inadequate to handle the growing amount of data and information generated internally by the coastal protection and restoration effort and by external data sources that could provide benefits if incorporated.

4.2 Roles

The roles and number of personnel provided below as an example is based on a system of approximately 1,000 users, actively contributing and consuming data. In a typical system that supports an organization that is driven by data, approximately 25 percent of the users are the persons who use the system rigorously for critical functions. This group is expected to have higher support costs as they push the system parameters and request customization. Approximately 50 percent of the users would be considered normal users, who utilize the system for periodic reporting, data requests, calculations, etc. This group is expected to request few customization features or push the system beyond its

intended use. The remaining 25 percent do not use the system very often and would typically make requests such as password resets, general navigation assistance, or infrequent reports.

Some roles require a greater number of personnel initially, but are expected to lessen as the data management system evolves.

System Administrator

The system administrator will be responsible for the day-to-day performance of the data management system. This would involve managing personnel associated with the system operation and maintenance, quality control, legacy data input, training, and user support. Additional role requirements are likely to include budgeting, planning, development, monitoring user survey data, arranging training sessions, and troubleshooting. The System Administrator would be the intermediary to the stakeholder group.

This role could be expected to require at least one full-time person dedicated to the system upkeep and continued development.

Database Administrator

The Database Administrator will be responsible for the maintenance and functionality of the databases supporting the data management system. This position would involve maintaining the standards and policy libraries, conducting data audits for conformance, troubleshooting and repairs, installing and upgrading system tools, monitoring system storage capacity and security applications, monitoring access to the databases, conducting system backups, and archiving data. This position would be responsible for the performance, integrity, and security of the database itself.

This role could be expected to require at least one full-time person dedicated to the integrity and maintenance of the system databases.

Application Developer

Application developers will work under the system administrator to troubleshoot issues, fix operations when needed, and follow procedures outlined by the user and system administrator. The application developers may be called upon to help develop and administer the training program.

This role could be expected to require at least three personnel in the beginning of the project with a potential decrease as the data management system becomes developed.

Data Collection and Input

This position will identify and document data coming into the system, input legacy data, and provide feedback on transforming the data into the system. Current data streams take precedence over legacy data. As they will be familiar with the data sources, they are also responsible for the first QA/QC of the data once it has been input into the system.

Without inventorying the existing data sets, it is expected this role could require at least two full-time personnel; however, this number of personnel could increase with an increased volume of data.

Support / Training

This position will be responsible for performing classroom and one-on-one training, as needed. They will be responsible for user support, user account issues, documenting user feedback, and acting as the user advocate to the rest of the data management team.

This role could be expected to require at least two full-time personnel.

Quality Assurance / Quality Control

This position will be responsible for ensuring the quality of the application and the data imported. A major responsibility in this role is to document and communicate all quality issues to the rest of the data management team. This role could be expected to require at least three full-time personnel before the initial rollout with a decrease in personnel as the application becomes developed.

Personnel and Level of Effort

At the initial rollout, it is expected that at least 12 personnel would be needed to support the data management system. As the data management system expands or contracts, so will the number of personnel needed. The number of personnel discussed should be considered estimates at this time.

- System Administrator – at least 1 person
- Database Administrator – at least 1 person
- Application Developer – at least 3 personnel
- Data Collection and Input – at least 2 personnel
- Support/Training – at least 2 personnel
- QA/QC – at least 3 personnel

4.3 Systems Capability and Recommendations

Hardware

Hardware is dependent upon decisions made by CPRA about hosting requirements. It is expected that two multi-core servers with be utilized with a minimum of 16 GB of RAM. Disk space and backup will be determined for the initial rollout and future growth requirements after an inventory of the current data is completed.

Software

Both servers are expected to operate Windows Operating System (OS). The web server will run IIS7 and .NET 4.0. The database server will run the relational database system selected by CPRA. Any additional software needed to run the system will be specified during implementation.

5. Existing or Anticipated Data

5.1 Data Streams

A profusion of data sources and data sets exist in various databases and repositories that should be a part of the overall Data Management Plan. Also, the plan should remain forward-looking so that data sources and data sets generated in the future can also be integrated into the proposed data management system. Although it is impossible to anticipate every data format that may eventually have to be incorporated, the system should be as open-ended as possible when it comes to allowing new data sets to be added to the collection. The short-term focus will be to incorporate the data streams generated by CPRA and the 2012 Master Plan data. Other data streams of potential use to the CPRA will also be discussed and will be targeted for incorporation as a long-term goal. A preliminary inventory of existing data and data management systems, relevant to coastal Louisiana, is included in Appendix A.

2012 State Master Plan Modeling

The 2012 Coastal Master Plan generated a substantial amount of data and information. One of the most data-heavy components was the predictive modeling. The primary models used included: eco-hydrology, wetland morphology, barrier shoreline morphology, vegetation, storm surge and wave, risk assessment and damage, and 14 upper trophic level / ecosystem service suitability indices. All modeling-related data have been archived and are currently stored on the CPRA's FTP site (ftp://MPM:dnrftp.2010@ftp.dnr.state.la.us/Master-Plan-Modeling).

Appendix B provides a guide to the file nomenclature as they occur on the FTP site (developed by the Master Plan Data Integration Team), a summary of the modeling data, including a guide to the scenarios, data formats and layout of the FTP site.

- **Modeling data: 1.33TB (see Appendix B for more information)**

In addition to the Master Plan modeling data, it should be noted that other datasets related to the development of the 2012 Master Plan are currently archived on the CPRA FTP site. It is possible that these datasets may need to be considered / included in an overall data management effort. These datasets include (but are not limited to):

- **Decision criteria: 1.32GB** (Cultural heritage, Natural processes, Navigation, and Oil and gas
  - Code (.txt, .bat, .sln, .suo, .csv, .F90, .u2d, .vfproj, .exe, .htm, .obj)
  - Spatial (.shp [and associated file-types], .csv, .txt, .ini)
  - Tabular data (.docx, .xlsx)

- **Master Plan document and technical appendices: 3.12GB** (.jpg, .indd, .ai, .png, .img, .rrd, .bmp, .tif, .psd, .gdb, .txt, .docx., .pdf)

- **Miscellaneous GIS files: 113MB** (GIS files related to the 500m reference grid, CPRA spatial viewer, coastal zone, CPRA jurisdiction area, planning units, points of interest, subsidence ranges, urban areas)

- **Outreach and engagement: 2.77GB** (brochure, community meeting agendas, presentations, and summaries, land-water animations, model-related presentations and audio recordings, general presentations, statements of support for the Master Plan)

- **Planning Tool: 1.53GB** (draft and final data sets, in .sql format)

- **Project attributes: 27MB** (shape files for public release, candidate projects, Master Plan projects, influence areas, nonstructural, polders, query list)
    - Spatial (shape files)
    - Tabular (.xlsx)

- **Public comments: 182MB** (.pdf)

- **Targets and constraints: 37MB** (risk reduction targets - census communities and community inundation)
    - Spatial (shape files)
    - Tabular (.xlsx)

Survey Examples

- **Hydrographic (CRMS, USGS, CPRA project-specific, MRHDMS, etc.).** Hydrographic surveys support a variety of activities: nautical charting, port and harbor maintenance (dredging), coastal engineering (beach erosion and replenishment studies), coastal zone management, and offshore resource development. Most surveys are primarily concerned with water level and water quality, including salinity, temperature, PH, DO, etc. Another possible data stream is that the USGS and the state and other agencies operate numerous real time stations throughout the state. Some of them are part of larger data sets, and some are project specific that are housed in sometimes obscure locations.

- **Bathymetry (LASARD, BICM, etc.).** Bathymetric data represents the elevation of the land beneath a body of water. Surveys can also check on the efficiency of dredging operations.

- **Topographic (BICM, etc.).** Topographic surveys are used to identify and map the contours of the ground and existing features on the surface of the earth or slightly above or below the earth's surface (i.e., trees, buildings, streets, walkways, manholes, utility poles, retaining walls, etc.). The survey may include perimeter boundary lines and the lines of easements on or crossing the property being surveyed.

- **Geotechnical (LASARD, BICM, etc.).** The purpose of geotechnical surveys is to determine the quality of sediment (or rock) through physical sampling techniques. Geotechnical investigations pertaining to coastal and wetland issues typically include the collection of borings or surface samples which then undergo laboratory analysis. Additional geotechnical investigations that may be performed include foundation analysis, soil analysis, and soil mechanics.

- **Geophysical (LASARD, BICM, etc.).** This includes seismic, sidescan sonar, and magnetometer data, which are collected to help identify possible borrow sites for marsh creation areas.

- **High-Water/Debris Level Marks.** Surveys of the physical aspects of land-falling storms are used to test the veracity of surge models by comparison with topographically surveyed debris lines that mark the level of marine inundation, along with records from tide gauges, to determine the relative contributions of surge, tide, wave action, and wave run-up to the total marine inundation, and to elucidate the nature and impact of marine inundations on sandy coasts.

- **Biological (CRMS, CPRA project-specific, LDWF, etc.)**. Biodiversity surveys are undertaken to find out what organisms exist in a given area. The data that is gathered from these surveys is used for numerous purposes, such as monitoring endangered populations, evaluating conservation priorities of an area, and bio-prospecting. Museum and herbarium specimens provide a valuable record of the location of organisms but such data are rarely systematic and often subjective. This is why field surveys are so important. Despite their importance in biodiversity research, there are few well-defined rules as to how to perform such surveys. This is due to the vast differences between surveys in terms of the goals of the survey, of available resources and time, the area to be surveyed, the organisms to be found, and a myriad of other factors that change from survey to survey.

- **Emergency Response.** By definition Emergency Response Data is very dynamic, as the data streams, types, and parameters often need to be changed from event to event. It is difficult to define in advance the exact nature of data collected during emergency events. The coastal zone of Louisiana is typically impacted by natural disasters such as hurricanes, flooding, and storm surge type events as well as manmade disasters such as oil spills. Through historical analysis, it is possible to define common data types that are utilized during Response Events. Systems architecture has to incorporate the flexibility to incorporate new data elements for unique situations that arise as a result of man-made or natural disasters. The Data Management Plan should be able to provide data standards for incorporating data services that are published or created as part of Emergency Response efforts.

- **Levees/Flood Protection (CPRA, U.S. Army Corps of Engineers [USACE], Local Levee Boards).** Levee system datasets include topographic survey information, engineering information, and potentially modeling information and outputs. Other information related to operations and maintenance such as floodgate status would also be included.

- **GOHSEP.** The Governor's Office for Homeland Security and Emergency Preparedness has the mandate to lead and coordinate the response efforts in the State of Louisiana. The data streams generated during Emergency Response efforts are in some cases very similar to other data streams already outlined in the plan. However, the nature of Emergency Response dictates that there must be flexibility to account for additional non-traditional elements for each of these data streams. Table 2 lists typical data streams that would be incorporated during Emergency Response Efforts.

Table 2. Data streams typically incorporated during emergency response efforts.

Data Streams	Description and Elements
Biological Data	Data or measurements impacted biological resources in the state of Louisiana.
Flood Data	Data or measurements related to flood levels and impact on the coastal zone.
Fire Damage	Data or measurements of large scale fire damage as a result of natural or man-made disasters.
Hurricane	Data or measurements for the impact of hurricanes to the coastal zone of Louisiana.
HAZMAT	Data or measurements quantifying the impact of Hazardous Materials (HAZMAT) typically chemicals released to the environment during man-made or natural disaster events.
Nuclear/Radiation	Data or measurements quantifying the impact of Nuclear or Radiological events during an emergency response.
Terrorist Response	Data or measurements associated with acts of terror initiated with-in the State of Louisiana.
Tornados	Data or measurements quantifying the impact of Tornados after a storm event.

Additional Datasets

Other data sets that have value to the state's coastal protection and restoration program that may be brought together in a common portal as part of the long-term strategy include the following:

- **LDWF.** The Louisiana Department of Wildlife and Fisheries manages, conserves, and promotes wise utilization of Louisiana's renewable fish and wildlife resources and their supporting habitats through replenishment, protection, enhancement, research, development, and education for the social and economic benefit of current and future generations; to provide opportunities for knowledge of and use and enjoyment of these resources; and to promote a safe and healthy environment for the users of the resources.[18] Available data sets include the Fishery Dependent Monitoring Program and the Louisiana Offshore Oil Port Monitoring Program.

- **LDHH.** The Louisiana Department of Health and Hospitals protects and promotes health and ensures access to medical, preventive, and rehabilitative services for all citizens of the State of Louisiana.[19] Available data sets include the State of Louisiana Safe Drinking Water Information System (SDWIS/State).

- **LDEQ.** The Louisiana Department of Environmental Quality provides service to the people of Louisiana through comprehensive environmental protection in order to promote and protect

[18] http://www.wlf.louisiana.gov/about-ldwf
[19] http://new.dhh.louisiana.gov/index.cfm/page/2/n/4

health, safety, and welfare while considering sound policies regarding employment and economic development.[20] Available data sets include Emissions Inventory Data Sets and Toxic Pollutants Data Sets.

- **DHS.** The U.S. Department of Homeland Security's available data sets include the Marine Casualty and Pollution Databases and the U.S. Coast Guard Maritime Information Exchange (CGMIX).

- **Demographic (US Census Bureau, etc.).** The United States Census Bureau, as the body charged with being the leading source of quality data about the nation's people and economy, maintains an exhaustive collection of demographic data sets, including population estimates, housing surveys, economic statistics, and geographic information.

- **NOAA.** The National Oceanic and Atmospheric Administration endeavors to understand and predict changes in climate, weather, oceans, and coasts; share that knowledge and information with others; and conserve and manage coastal and marine ecosystems and resources.[21] Climate data sets are available from numerous branches of NOAA, including the Earth System Research Laboratory (ESRL) Physical Sciences Division (PSD), the National Climatic Data Center (NCDC), and the National Environmental Satellite, Data, and Information Service (NESDIS) World Data Center (WDC) for Meteorology, Asheville.

- **NASA.** The National Aeronautics and Space Administration (NASA) Global Change Master Directory (GCMD) holds more than 27,000 Earth science data sets and service descriptions.[22] The project mission is to assist researchers, policy makers, and the public in the discovery of and access to data, related services, and ancillary information (which includes descriptions of instruments and platforms) relevant to global change and Earth science research. Within this mission, the directory also offers online authoring tools to providers of data and services, facilitating the capability to make their products available to the Earth science community. In addition, citation information to properly credit data set contributions is offered, along with direct links to data and services.

- **NWIS.** The U.S. Geological Survey National Water Information System (NWIS) supports the acquisition, processing, storage and dissemination of information about water quantity and quality collected at over 1.5 million sites around the United States. As a long-term database and information delivery system, NWIS provides continual access to data collected over the last 100+ years, as well as real-time data on stream flow, etc.

GIS Data

There are several data formats capable of encoding geographical information. These formats are generally represented in one of two formats: vector data and raster data.

- **Vector Data.** Vector data formats typically contain the following information: shapes usually expressed as points, lines, or polygons that include coordinate system information and sometimes elevation data. Metadata is typically associated with vector data, which explains in

[20] http://www.deq.louisiana.gov/portal/ABOUT/DepartmentProfile.aspx
[21] http://www.noaa.gov/about-noaa.html
[22] http://gcmd.nasa.gov/

plain text the data structure, content, and purposes of the data set. Vector data is typically stored as a single file, such as AutoDesk's DWG and Google's Google Earth Keyhole Markup Language (KML) format, or a collection of files, like in the ESRI shapefile format, or as tables in a relational database, such as Microsoft SQL Server, Oracle, or open-source PostgreSQL databases. Examples of vector data include road centerlines, wetland areas, monitoring well locations, and stream networks.

- **Raster Data.** Raster data formats encode spatial information into a digital, finite series of rasterized grid cells (a.k.a. pixels) consisting of a series of rows and columns with each grid cell containing a single or continuous numerical value. Raster data are typically stored in a standard file based structure, such as a TIFF, JPEG, GRID, or PNG formats. Coordinate system information is often embedded into the raster file or exists outside the data as supplemental files. Metadata is usually included as a supplement file. Examples of raster data include aerial or satellite imagery, digital elevation models, and land use data.

Some datasets include both time and space. Examples of datasets that are considered coverages are model results from hydrodynamic models, bathymetries that change over time and a river with a changing water level. Data is commonly stored in the form of multidimensional arrays (dimensions can include time, one or more spatial dimensions, but also frequencies in the example of wave models). Examples of Coverage Data Formats (Gridded Data) include HDF5, netCDF, etc., and each of the data types has an accompanying service that can be used to stream the data over the internet. These services include Web Map Service (WMS), Web Feature Service (WFS), and Web Coverage Service (WCS). Although the W3C standard for querying coverage data is WCS, the Opendap is the de-facto standard used for querying coverage data.

Most federal, state, and local public entities use WMS to share their maps with the public across different web platforms and clients. Note that WMS does not always provide the data itself, but sometimes merely a visualization of the data (a picture in bitmap form). Examples of consumable WMS include National Oceanic and Atmospheric Administration (NOAA) NowCOAST meteorological data, which includes, but is not limited to, precipitation amounts, wind direction vectors, weather monitoring station locations, and surface water temperatures. Other agencies that provide WMS as well as other file-based GIS data and non-spatial data streams related to water resources include USGS, USACE, FEMA, and the State of Louisiana.

Data Objects

- **Documents.** The data management system should allow uploading and long term storage of file-based documents including a method for associating uploaded documents with particular data sets.

- **Photographs (Aerial, oblique, on-site: CPRA, USGS, NOAA, etc.).** The system should also allow upload and storage of photographs and other graphic files including a method for associating uploaded photographs with particular data sets.

Legacy Data

One of the long-term goals of the planned data management system is to develop methods for incorporating legacy data sets into the repository while requiring a minimum investment of resources.

Since not all legacy data can be determined and analyzed prior to the planning and implementation of the data management architecture, resources will be required to transform the data sets CPRA determines to be candidates for incorporation into the repository. Responsibility for assessing and assigning resources will be determined by CPRA.

Any legacy data incorporated into the data repository should conform to all standards detailed in Section 2. The planning and design of the system should allow for maximum flexibility regarding legacy data integration. However, any data integration tools incorporated into the system, whether in-house built or third-party add-ins, should validate the output data according to the accepted data file and format standards.

5.2 Data Organization

Data Organization is imperative for any large data warehouse in order to maximize the value of data to an agency. Data organization is the strategy to maximize data availability and usability to both users and other systems that need to interact with the data. Data organization is also imperative to the overall architecture of the system to maximize hardware and software resources and ensure operational efficiency.

Users

- **Administration / Decision-makers / Executive.** It is important that the decision-makers are stakeholders in the success of the data management system. They are also primary users of the data and their perspective should be considered during the design and implementation of the data management architecture.

- **Project Managers.** As the stakeholders with the responsibility for the planning, execution, and accomplishment of the stated project objectives, project managers are also primarily users of the data. Generally, they are interested in summary-level information.

- **Operations.** Operations management is responsible for the process that converts inputs, in the forms of materials, labor, and energy, into outputs, in the form of products and/or services. As such, they are interested in consuming data from the perspective of strategic policy, systems analysis, productivity analysis, cost control, and resource planning.

- **Scientific.** The scientific community serves as both users and, in many cases, contributors of data to the repository. The scientific community provides invaluable expertise and contributions through the following activities, among others:

    - Monitoring
    - Research
    - Assessment

- **Public.** Since certain portions of the public will have an interest in at least portions of the data envisioned to be part of the planned repository, the data management system should allow for public consumers of the data. This requires an easy-to-use interface and instructions and guidance (such as a help system) as to the most efficient way to search and visualize the data. The system should also allow some data sets to be “locked” or hidden from public view.

- **Emergency Response Community.** The Emergency Response community will be users and contributors to the overall system. Typically their consumption of the data will be to set a baseline of conditions prior to the Emergency Response Event so impact can be quantified. This will be very important so that proper decisions can be made to minimize or reduce the overall environmental impact of an on-going event. Additionally, data will be collected throughout the response event that could prove valuable to long-term research purposes to help quantify the post event conditions.

Geography

All data sets should be easily searchable by geographic parameters. Some examples of geographic boundaries or features that should be allowed as organizational parameters are:

- Basins
- Systems
- Levees
- Political Districts

Reference

All data sets should also be easily searchable by reference parameters. Examples of reference categories that should be allowed as parameters include:

- Geological
- Modeling
- Biological
- Engineering
- Construction
- Operations and Maintenance
- Mapping

Data File Formats

Formats more likely to be accessible in the future have the following characteristics:

- Are non-proprietary; not protected by trademark, patent, or copyright.
- Utilize open and well-documented standards (open formats).
- Are in common usage by the research community.
- Utilize standard representations, such as American Standard Code for Information and Interchange (ASCII) or Unicode.
- Are Unencrypted.
- Are Uncompressed.

Examples of Preferred Format Choices

- For written documents: PDF/A over Word.
- For spreadsheet or tabular data: ASCII over Excel.
- For video files: MPEG-4 over Quicktime.
- For photographs or image files: TIFF or JPEG2000 over GIF or JPG.

- For data: Extensible Markup Language (XML) or Resource Description Framework (RDF) over Relational Database Management System specific files (RDBMS).
- For GIS: geo-databases or shape files

Recommended File Formats

- **NetCDF.** NetCDF is a set of software libraries and self-describing, machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data. It is widely used because of its compatibility with various platforms and ease of use. NetCDF was developed and is maintained at Unidata. Unidata is part of the University Corporation for Atmospheric Research (UCAR) Community Programs (UCP). Unidata is funded primarily by the National Science Foundation. Unidata supports netCDF Application Programming Interface (API) in C, C++, Fortran 77, Fortran 90, and Java. NetCDF Java implements the Common Data Model (CDM), an abstract data model. The CDM merges the netCDF, OPeNDAP, and HDF5 data models to create a common API for many types of scientific data. Full documentation exists for each API. [23]

- **CSV.** Comma-Separated Values (CSV) files store tabular data (numbers and text) in plain-text form. A CSV file consists of any number of records, separated by line breaks of some kind; each record consists of fields, separated by some other character or string, most commonly a literal comma or tab. Usually, all records have an identical sequence of fields.[24] The CSV file format is often used to exchange data between disparate applications. As is the case with most exchange formats since XML, CSV files have become somewhat of a legacy format. New applications that wish to include an export format will generally use XML today (though there may be exceptions). In legacy systems though (pre-XML), CSV files had indeed become a de facto industry standard. Just as there are still billions of lines of COBOL code in use today that need to be maintained, support for a legacy standard such as CSV is likely to be required long after it has stopped being implemented in new designs.[25]

- **XML.** The Extensible Markup Language defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. It is defined in the XML 1.0 Specification produced by the World Wide Web Consortium (W3C). The design goals of XML emphasize simplicity, generality, and usability over the Internet. It is a textual data format with strong support via Unicode for the languages of the world. Although the design of XML focuses on documents, it is widely used for the representation of arbitrary data structures, for example in web services.[26]

Metadata

Metadata describe the content, quality, condition, and other characteristics of data. Metadata have a wide variety of uses, including organizing and maintaining an organization's investment in data, providing information to data catalogs and clearinghouses, and providing information to aid data transfer. Metadata describe different aspects of data, including:

[23] http://www.unidata.ucar.edu/software/netcdf/
[24] http://en.wikipedia.org/wiki/Comma-separated_values
[25] http://creativyst.com/Doc/Articles/CSV/CSV01.htm
[26] http://en.wikipedia.org/wiki/XML

- **Identification -** What is the name of the data set? Who developed the data set? What geographic area does it cover? What themes of information does it include? How current are the data? Are there restrictions on accessing or using the data?

- **Data Quality -** How good is the data quality? Is information available that allows a user to decide if the data are suitable for his or her purpose? What is the positional and attribute accuracy? Are the data complete? Was the consistency of the data verified? What data and models were used to create the data set, and what processes were applied to these sources?

- **Spatial Data Organization -** What spatial data model was used to encode the spatial data? How many spatial objects are there? Are methods other than coordinates, such as street addresses, used to encode locations?

- **Spatial Reference -** Are coordinate locations encoded using longitude and latitude? Is a map projection or grid system, such as the State Plane Coordinate System, used? What horizontal and vertical datums are used? What parameters should be used to convert the data to another coordinate system?

- **Entity and Attribute Information -** What geographic information (roads, houses, elevation, temperature, etc.) is included? How is this information encoded? Were codes used? What do the codes mean?

- **Distribution -** From whom can I obtain the data? What formats are available? What media are available? Are the data available online? What is the price of the data?

- **Metadata Reference -** When were the metadata compiled? By whom?

The Standards Group will be tasked with selecting the Metadata standard that will be utilized by the data management system. As discussed in the section below, unless other as-yet undefined or undetermined factors affect the planning of the system, an ISO Metadata standard is the recommended path forward.

Metadata Standards

FGDC Endorsed ISO Metadata Standards[27] - In September of 2010, the FGDC Steering Committee endorsed 64 non-federally authored standards that play an important role in enabling geospatial interoperability. Several of these endorsed standards relate to international agreement on geospatial metadata standards that supersede national metadata practices in most countries and are found in the ISO suite of geospatial metadata standards, referred to as the 191** series, which include but are not limited to the following:

- **ISO 19115:2003**. Geographic information – Metadata (corrigendum 1): The base ISO metadata standard for the description of geographic information and services. Expected to be replaced by ISO 19115-1: Geographic Information – Metadata – Part 1: Fundamentals once approved by ISO (currently a Draft International Standard, anticipated to be an International Standard in May 2013).

[27] http://www.fgdc.gov/metadata/geospatial-metadata-standards#fgdcendorsedisostandards

- **ISO 19115-2.** Geographic information - Metadata - Part 2: Extensions for imagery and gridded data. The base ISO 19115 metadata standard plus extended elements for the description of imagery, gridded data and data collected using instruments, e.g. monitoring stations and measurement devices. These extensions also include improved descriptions of lineage and processing information. ISO 19115-2 is expected to be updated to comply with ISO 19115-1 once the final standard is released.

- **North American Profile (NAP) of ISO 19115**. A U.S. and Canada joint profile of ISO 19115:2003 that extends some domains, increases conditionality for some elements, and specifies best practices for populating most elements.

- **ISO 19110.** Geographic information – Methodology for Feature Cataloging: An affiliate standard that supports the detailed description of feature types (roads, rivers, classes, rankings, measurements, etc.) in a manner similar to the Content Standard for Digital Geospatial Metadata (CSDGM) Entity/Attribute Section. The standard can be used in conjunction with ISO 19115 to document geospatial data set feature types or independently to document data models or other feature class representations.

- **ISO 19119.** Geographic information - Services - Amendment 1: Extensions of the service metadata model. An affiliate standard that supports the detailed description of digital geospatial services including geospatial data portals, web mapping applications, data models and online data processing services. The standard can be used in conjunction with ISO 19115 to document services associated with specific data sets or independently to document a service.

- **ISO 19139.** Geographic information – Metadata – XML schema implementation: An XML document that specifies the format and general content of an ISO 19115 the metadata record. Expected to be updated to ISO 19115-1: Geographic Information – Metadata – Part 1: Fundamentals once approved by ISO.

ISO Metadata Standards Utilization. ISO Metadata standards support different agency and collaboration requirements in specific ways:

- **ISO 19115-2** is the preferred standard for organizations actively implementing ISO Metadata as it includes all of the elements of ISO 19115 as well as additional elements that are relevant to many geospatial data sets (raster, imagery, GPS, monitor stations, instruments, etc.).

- **ISO 19139** is an XML schema that specifies the format of the metadata record and is used by application developers to implement the standard.

- Few organizations have implemented NAP of ISO 19115 primarily due to the broader scope of ISO 19115-2. However, the NAP does provide a rich source of best practices as well as strengthened conditionality and expanded domains that are used by many ISO implementers.

- The forthcoming, ISO 19115-1, despite its numbering, was developed after ISO 19115-2 and is an updated version of the base standard (ISO 19115). In addition to changes to the structure of some base elements and the content of some domains, ISO 19115-1 expands upon the former standard by providing more fields to describe geospatial data services (ISO 19119), multi-dimensional gridded datasets, modeling results, etc., and enabling entity/attribute descriptions developed using ISO 19110: Feature Catalog to be associated with or integrated into the metadata record.

Content Standard for Digital Geospatial Metadata (CSDGM) vs. ISO Metadata Standard. Most stakeholders have long utilized the CSDGM, which will continue to have a legacy for many years. International geospatial metadata standards are emerging in the community. FGDC policy states that non-federally authored standards that are endorsed by the FGDC have the same status as FGDC developed standards. Since ISO 19115 and the associated standards are endorsed by the FGDC, federal agencies are encouraged to transition to ISO metadata as their agencies are able to do so. While the selection of appropriate standards is dependent on the nature of metadata collection and publication process, ISO metadata should be considered an option now. It is recognized that the transition to ISO metadata will be occurring over the next few years.

6. User Interface

The Data Management Workgroup tasked with the conception of this Data Management Plan has determined that a key metric of success will be the end user's evaluation. An investment by the system architect team in the time it takes to interview the primary stakeholders, including contributors, users, and database managers, will be recouped through a system that is heavily utilized by the targeted audience.

The primary modes of user interface selected by the workgroup for discussion were wikis and portals. The wiki or portal will act as the main access point to the data management system. Additionally, dashboards should be used within the wiki or portal to provide a very quick, concise summary of the data for a variety of users. The design and look-and-feel of the access point will be determined by the CPRA.

Assembling a Spatial Data Infrastructure (SDI) website involves three key parts: creating catalog registry and search/querying interfaces, incorporating dynamic mapping and data visualization interfaces, and building data modeling and data download tools to acquire data in a useable and well-documented format for use.

6.1 Wiki or Portal Features – Usability

A web portal is a web site that brings in information from diverse sources in a unified way. Usually, each information source gets its dedicated area on the page for displaying information (a portlet); often, the user can configure which ones to display. Portals provide a way for organizations to provide a consistent look and feel with access control and procedures for multiple applications and databases, which otherwise would have been different entities altogether. Some examples of portals are AOL, iGoogle, MSN, and Yahoo!.

A wiki is a website that allows its users to add, modify, or delete its content via a web browser usually using a simplified markup language or a rich-text editor. Wikis are powered by wiki software. Most are created collaboratively. Wikis serve many different purposes, such as knowledge management and notetaking. Wikis can be community websites and intranets, for example. Some permit control over different functions (levels of access). For example, editing rights may permit changing, adding, or removing material. Others may permit access without enforcing access control. Other rules may also be imposed for organizing content. The most well-known example of a wiki is Wikipedia (http://en.wikipedia.org).

The main portal or wiki page should be designed with the following usability principles in mind:

- Robustness
- Quick site navigation and full content listing, available at all times
- Hierarchical representation of visual components
- User authentication and authorization
- Simultaneous logging of multiple users
- Different content according to the logged user's privileges
- Sustainable security model

- User profile support
- Storing portal state for each user, to make it available across sessions
- Skins support

6.2 Portal or Wiki Features - Extensibility

In systems architecture, extensibility means the system is designed to include hooks and mechanisms for expanding/enhancing the system with anticipated capabilities without having to make major changes to the system infrastructure. A good architecture provides the design principles to ensure this—a roadmap for that portion of the road yet to be built. Note that this usually means that capabilities and mechanisms must be built into the final delivery that will not be used in that delivery and, indeed, may never be used. These excess capabilities are not frills, but are necessary for maintainability and for avoiding early obsolescence.

Extensibility can also mean that a software system behavior is modifiable at run time, without recompiling or changing the original source code. For example, a software system may have a public API that allows its behavior to be extended or modified by people who do not have access to the original source code. The extra functionality can be provided through either internally or externally coded extensions.

6.3 Extensible Architecture[28]

A registered user (client) opens the portal, and the portal application receives the client request and retrieves the current user's page data from the portal database. The portal application then issues calls to the portlet container for all portlets on the current page. The portlet container, which holds the user's preferences, calls the portlets via the portlet API, requesting the markup fragment from each portlet and returning the fragment to the portal. The portal aggregates all markup fragments together into one page, which the portal finally returns to the user, serving up an integrated, useful interface (see Figure 1).

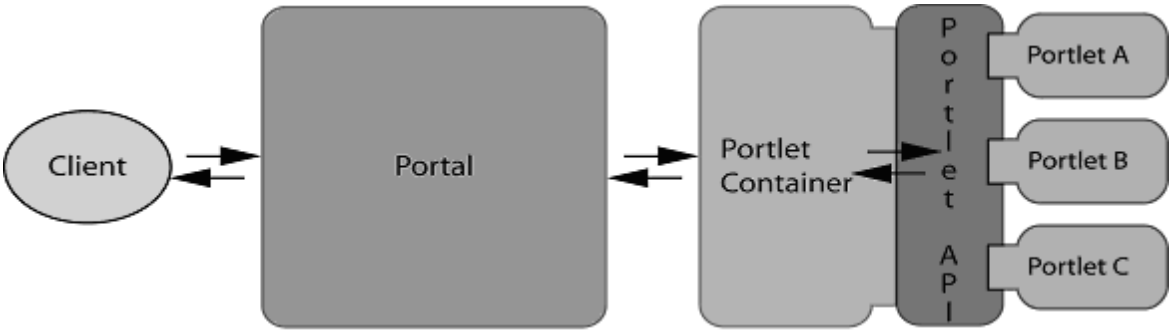

Figure 1. Example diagram of client-to-portlet information exchange.

[28] http://www.editorial.mcpressonline.com/web/mcpdf.nsf/wdocs/5232/$FILE/5232_EXP.pdf

The portal or wiki architecture should be designed with the following extensibility principles in mind:

- Easy deployment of new components
- Strong component modularity
- Interaction between individual components
- Differentiated visual and logical components, ensuring clear division of development efforts
- Simple definition of execution flow and relationships between elements
- Predefined ready-to-use layout and control component templates

One advantage of a wiki over a portal solution is that changes made to wikis by the users are easy to revert back to their prior state. Users are more likely to add and alter content when they know that there is a safety net. The experience of Deltares, which internally runs an Oracle Portal, a Microsoft SharePoint Portal, and a Confluence Wiki, has been that their users prefer to use the wiki due to its better search capabilities and ease-of-use when adding and sharing content.

6.4 Success Metrics

Success metrics should be based upon user input and designed to determine if the data management system is utilized by all groups of stakeholders. Mechanisms for user feedback should be developed to ensure that user input is collected and considered in order to continuously improve the data management system. Further steps that should be taken to accomplish this goal of measuring success include:

- Establishing triggers for conducting reviews of system. Plan and design the system reviews with a focus on flexibility, scalability, and growing the system with forward-thinking process.

- Design a policy to direct resources to update the system as needed. CPRA will be responsible for allocating the resources for system maintenance and architecture and software improvements over time.

6.5 Feedback or Alert Mechanism

The system should have a user feedback or alert mechanism incorporated as part of the user interface. This will alert users if any changes, updates, or corrections made to a dataset that has been previously downloaded. This will require a mechanism that allows people to sign up, and a mechanism to store the list of signed-up users for future reference. The system must also store a history of which users have downloaded which data sets for future cross-reference. It should be made clear to the user that by signing up they will benefit from knowing about these corrections or updates.

6.6 Dashboards

A dashboard is a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance. Generally, dashboards display a set of performance indicators (PIs), key performance indicators (KPIs), and any other relevant information to the user. Dashboard data is often displayed in real-time after retrieval from one or more data sources in an organization. Dashboards are also often interactive, allowing users to drill down into particular aspects of the display or switch between facets or views of the data.

Dashboard Types

On a high level, the two main categories are **operational dashboards** and **analytical dashboards**. Other types of dashboards exist depending upon the type of delivery required or the visualizations desired, but understanding the difference between an operational versus an analytical dashboard is the first step.

- **Operational dashboards** manage intra-daily organizational processes – frequently changing and current performance metrics or KPIs. Examples span almost every department but are particularly common in environments where it is essential to act on opportunities and issues quickly: such as sales, marketing, help-desk, supply-chain etc. Overall, operational dashboards are best suited to departments that require low latency data feeds and a continual view into what is happening within the organizational unit. Although operational dashboards may also help identify trends over time, or provide context around the KPI, the ability to drill through to current information, to get alerts, and to identify potential operational issues as they occur is what sets them apart from their analytical counterpart. These dashboards are often used department or organization wide, as they require less training than analytical dashboards, and eliminate the requirement for report distribution. In short, operational dashboards are meant to help an organization understand if its performance is on or off target, and by how much, in real time.

- **Analytical dashboards** focus on gaining insights from a volume of data collected over time – often the past month or quarter – and use this to understand what happened, why, and what changes should be made in the future. For instance, organizations may want to compare trends over time or identify why certain products are more popular in one region as opposed to another. Analytical dashboards use sophisticated models, what-if analysis and pivots to identify patterns and opportunities, and to help align strategic goals with organizational performance initiatives. These dashboards are often used by business analysts who are responsible for outputting reports for general consumption. In short, analytical dashboards are meant to help an organization establish targets based on insights into historical data.

Dashboard Examples

Along with the standard text and figures, bar graphs, line graphs, pie charts, and scatter plots, some newer types of Data Visualization Tools that may be utilized in modern Dashboards include heat maps, heat matrix visualizations, candlestick graphs, horizon graphs, needle graphs/gauges, spread graphs, stack graphs, and surface plots. This variety of possible visual elements means that dashboards can be designed in an infinite number of styles and to convey all kinds of different information. Figure 2 is an example operational dashboard, and Figures 3 and 4 are example analytical dashboards.

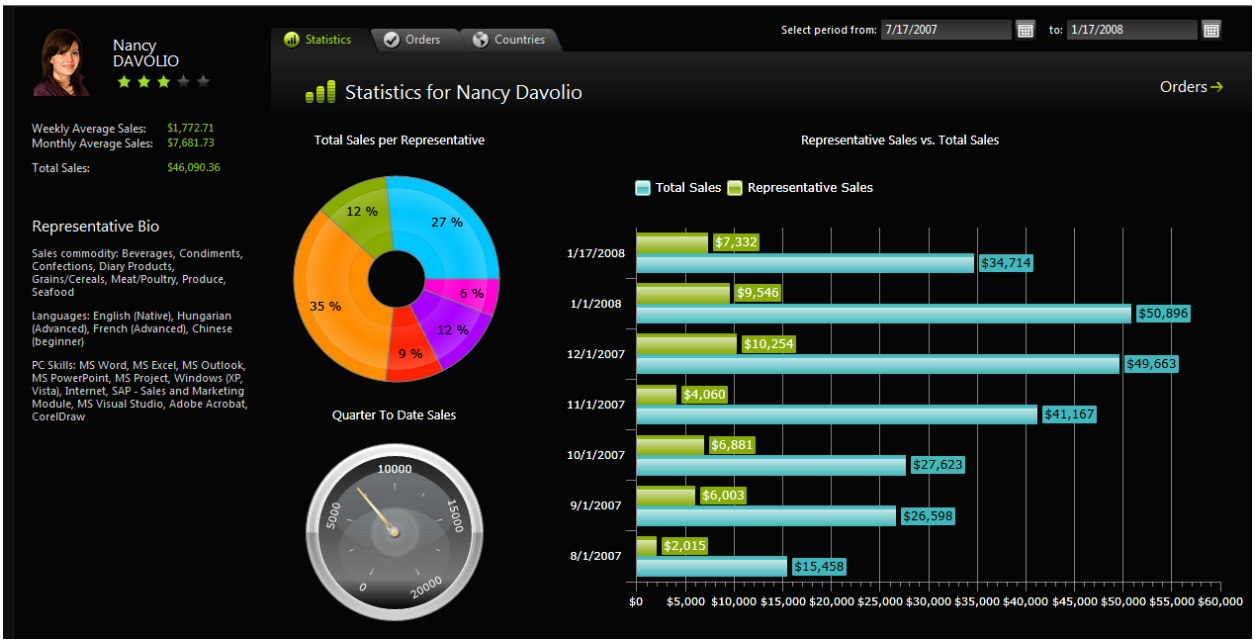

Figure 2. Example operational dashboard.

Figure 3. Example analytical dashboard.

Figure 4. Example analytical dashboard.

Dashboard Properties

- **Key Performance Indicators** - Be aware of similar but different KPIs. Especially in large organizations, a KPI may have been created in one group that is similar to – but not exactly like – a KPI generated by another group. For each KPI it will be important to know how it was generated, how long it has been around, and the exact formula used to generate the KPI. If merging two KPIs into one, understand how the KPIs are similar or different, and what the effect of losing one of these will have on the users.

- **Access** - Access issues are important to consider especially when creating a site for users who are almost always pressed for time. To ensure the dashboard gets used, spend some time during the requirements-gathering phase to understand how users currently get the information that will be newly presented in the dashboard. Understand what they like and dislike about how they get the information.

- **Delivery** - Choose an effective delivery system. A dashboard can have amazing information design, a smoothly functioning back end, and can satisfy every whim and desire of the client, but it still may not be used because the executive or manager travels so much they only use their wireless handheld, or because they're used to reading everything on paper, or because their administrative assistant doesn't know how to find it and therefore doesn't open it up for them. Another option may be developing a mobile application.

Also, the dashboard design should not underestimate the power to produce hard copies. For people who have an enormous amount of information to look at, printed reports is one way of controlling that flow. When ascertaining the users' needs and workflows, ask them about the best way to deliver the dashboard to ensure it gets used by as many users as possible.

- **Plan for Security** - Given the type of information that is being displayed, security will play a big role in the development and design of the dashboards. There are two aspects of security that need to be confronted: who gets to see what, and how the site is protected. Personalization will allow you to control who sees what, but controlling access to the site must also be part of the architecture.

The system should include a Single Sign-On (SSO) solution, if possible. SSO is a mechanism whereby a single action of user authentication and authorization can permit a user to access all computers and systems where he has access permission, without the need to enter multiple passwords. Single sign-on reduces human error, a major component of systems failure and is therefore highly desirable but difficult to implement. Otherwise, one is faced with having to generate a new set of user names and passwords for users to memorize. Given that most of these users are pressed for time and already overwhelmed with systems that are supposed to help them, a new set of user names and passwords can severely hamper the success of a new application. Balancing security concerns and ease of access concerns should be resolved as early as possible in the project lifecycle.

6.7 Citation Tool

The web site interface should include a Citation Tool that allows the researcher submitting data sets to the repository an easy-to-use utility that allows them to attach citation information to the data set.

The requirements for fully usable data citations include: [29]

- The citation itself must be able to uniquely identify the object cited, though different citations might use different methods or schemes to do so.
- It should be able to identify subsets of the data as well as the whole data set.
- It must provide the reader with enough information to access the data set, and, when expressed digitally, it should provide a mechanism for accessing the data set through the web infrastructure.
- It should be usable not only by humans but also by software tools, so that additional services may be built using these citations. In particular, there should be services that use the citations in metrics to support the academic reward system, and services that can generate complete citations.

The information that the Citation Tool should allow the researcher to store may include:

- **Author(s)** - The creator(s) of the data set.
- **Title** - This should include the name of the cited resource itself and may also include the name of a facility and the titles of the top collection and main parent sub-collection (if any) of which the data set is a part.

[29] http://www.dcc.ac.uk/resources/how-guides/cite-datasets

- **Publication Date** - Whichever is the latest: the date the dataset was made available, the date all quality assurance procedures were completed, or the date the embargo period (if applicable) expired (rather than the collection or coverage date).
- **Edition** - The level or stage of processing of the data, indicating how raw or refined the dataset is.
- **Version** - A number that is increased when the data changes, as the result of adding more data points or re-running a derivation process, for example.
- **Feature Name and URI** - The name of an ISO 19101:2002 "feature" (e.g. GridSeries, ProfileSeries) and the URI identifying its standard definition, which is used to pick out a subset of the data.
- **Resource Type** - Examples: "database", "dataset".
- **Publisher** - The organization either hosting the data or performing quality assurance.
- **Unique Numeric Fingerprint (UNF)** - A cryptographic hash of the data, used to ensure no changes have occurred since the citation.
- **Identifier** - An identifier for the data, according to a persistent scheme.
- **Location and Access** - A persistent Uniform Resource Locator (URL) or Digital Object Identifier (DOI) from which the dataset is available. Some identifier schemes provide these via an identifier resolver service (see Figure 5).

The most important of these elements – the ones that should be present in any citation – are the author, title, date, and location. These give due credit, allow the reader to judge the relevance of the data, and permit access to the data, respectively. In theory, they should – between them – uniquely identify the dataset. In practice, however, a formal identifier is often needed. Once a unique formal identifier is generated for the publication, the location of the published information can be permanently mapped by a DOI resolver service so the information can be easily located using any Internet browser. The most efficient solution is to give a location that consists of a resolver service and an identifier.

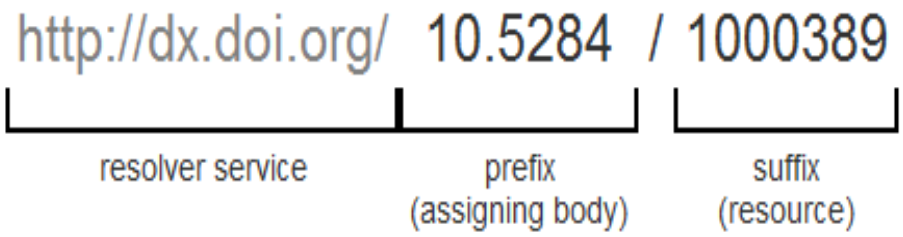

Figure 5. Example of a digital object identifier (DOI).

The DOI resolver service redirects the user from a DOI name to one or more pieces of typed data, such as URLs representing the location of the published data, services such as e-mail, or one or more items of metadata. Some examples of DOI resolver services are The DOI System (dx.doi.org), crossref.org, and the PANGAEA DOI Resolver (doi.pangaea.de).

Citation standards for data are still evolving. A number of groups and consortiums are currently developing consistent practices, including:

The DataCite Metadata Schema Repository (http://schema.datacite.org/) and the Embedded Computer Resources Support Improvement Program (ESIP) Federation Data Citation Guidelines (http://esipfed.org/).

6.8 Project Reports / Summaries

The data management system should include project summary-level reporting. Typical project reports include the following:

- Overview Reports, such as a Project Summary, Top-Level or Critical Tasks Reports, Milestones or Project Scheduling Reports.
- Activity Reports, such as Tasks in Progress or Completed Task Reports, or Aging Reports.
- Cost Reports, such as Budget or Cash Flow Reports, or Financial Resource Reports.
- Assignment Reports, such as To Do Lists, or Personnel or Allocated Resources Reports.
- Workload Reports, such as Task Usage Reports or Resource Usage Reports.
- Custom or "Ad Hoc" Reports.

Reports should include capabilities for printing and exporting (e.g., to Excel, PDF, etc.).

6.9 Query Functionality

Non-Graphical Queries and Data Dumps

CPRA's experience is that many users come to them for "give me everything you have" data requests and have found that large data dump query/delivery mechanisms are useful. For our purposes, Data Dumping refers to the act of copying raw data from one place to another with little or no formatting for readability, usually pulling data from a main database or repository and sending it to a client computer. The system should have a mechanism that allows users to search or query for particular data sets and select and dump entire data sets to the user's machine.

Graphical Queries[30]

The ability to present spatial data graphically is an obvious difference between spatial and conventional information systems. The *display* of query results in graphical form is the most natural form to analyze spatial data. For humans, graphics are much faster to understand than any lexical representation with the same amount of data. Also, by exploiting a second dimension, many spatial relationships – otherwise difficult to perceive – get revealed, such as information about neighborhoods, inclusion, or intersection. Besides rendering the geometry of spatial objects, the graphical display allows for the visualization of issues which are related to spatial objects and represented in the database as non-spatial attributes. Therefore, graphical representation of query results is a key characteristic of a spatial query language.

Selection by Pointing

The formulation of a query is based on either the user's knowledge, or information which is provided from the system. A map on a graphical display, for example, provides a large variety of information which the user can exploit for upcoming queries. Conceptually, this information can be separated into the objects displayed and the information about their spatial distribution across the two-dimensional surface. Using pointing devices *for selection by pointing* promotes the usage of query results as reference in upcoming queries. A second type of information conveyed through a graphical

[30] Query for Geographic Information Systems

representation is the spatial distribution of the objects across the 2-dimensional surface. Upon this surface, a user may perform operations, such as wandering around (*pan*) or viewing a specific part in a more detailed representation (*zoom*). For such iterative approaches, a graphical scene serves as reference to the selection of the following rendering. A typical query with reference to a sub-area is, "show all streets and towns within this region" and the user selects the region from a rendering on the screen. This process can be supported from the dialog using direct-manipulation methods to select the area of concern.

Layered Query Results

A layered approach should be utilized to separate the surface layer, the query layer, and the presentation layer.

Query Result Combinations

The possible combination of one query result with the results of one or more previous queries gives rise to a dynamic and user-friendly interaction. Directing the graphical display to manipulate drawings is essential to overcome the static character of traditional representation or paper maps. It should support the graphical presentation types in an interactive-graphic information system:

- ***New*** refreshes the viewport before drawing the next picture.
- ***Overlay*** adds the result of the current query to the existing picture.
- ***Remove*** erases the result of the current query from the existing picture.
- ***Intersect*** selects all those objects that are both on the display and in the query result.
- **Set Operations.** The semantics of these operations can be derived from conventional set operations. If the current display and the query result are considered as sets, then the following analogies yield:

    - ***New*** yields the ***intersection*** set operation of the query result with the ***empty set.***
    - ***Overlays*** correspond to the ***union*** set operation of combining the set of currently displayed objects with the query result.
    - ***Remove*** is the ***subtraction*** set operation removing all objects of the query result from the set of objects currently displayed.
    - ***Intersect*** corresponds to the ***intersection*** set operation of the set of displayed objects with the query result.

7. Security

All data streams should be evaluated for needed security. Each data stream security level is based on the sensitivity of the data it serves, the security guidelines of the agency or organization serving the data, and the audience it is intended for. For instance, a data service from and agency that serves little to no sensitive data requires minimal security with no role-based permissions. However, a data service deemed to contain sensitive data could serve data restricted to a certain group; this would require user authentication for all consumers and role-based permissions.

The role-based permissions allow for the user groups (whoever is consuming the data) to be classified into security levels. Role-based user permissions are highly-scalable and require minimal maintenance. Another option is service base permissions. This method defines one security level for each data service. What this means is if you have access to this data service, you can view all the data it serves. Both options provide sufficient authentication and security, in slightly different ways. A robust security policy should be established to define the level of security associated with a data stream or service as it is incorporated into the architecture defined by this plan. As the system incorporates information through collaboration agreements with outside entities, the system will need to match standards within the agreement. At a minimum, as data is incorporated, a defined security review should be performed as defined by the security policy.

7.1 Assignment of Authority

The Assignment of Authority will be determined by CPRA. Authorization[31] is the control of access to resources, answering the question: "Does the user have access to this protected resource?" Interactions between users and resources are controlled based on user identity or other information. Roles should be defined to control access to portal resources and administrative tools, such that users can access only the resources and tools that their assigned roles allow.

Software Roles

An Authorization Provider controls the interactions between users and resources to ensure confidentiality. Like a Login Module for an authentication provider, an Access Decision is the component of an authorization provider that determines if access to a resource is allowed. Specifically, an Access Decision determines whether a subject has permission to perform the specified action on a resource, based on the subject's security roles.

Role Mapping Providers control the process by which principals (users or groups) are dynamically mapped to security roles at runtime. A role mapping provider determines which security roles apply to the subject when the subject is attempting to perform an operation on a resource. Because this operation usually involves gaining access to the resource, role mapping providers are typically used with authorization providers.

A Security Role is a privilege granted to users or groups based on specific conditions. Roles are used to determine whether to grant or deny access to resources, and to determine which capabilities on those resources are available to the user. Granting a role to a user or group confers the defined access

[31] http://docs.oracle.com/cd/E13218_01/wlp/docs81/security/security.html

privileges to that user or group, as long as the user or group is granted the role. Any number of users or groups can be granted a single role. Roles are computed and granted to users or groups dynamically, based on role policies, which consist of a role name and a role definition. Role policies are dynamic, and may be based on username, group membership, user profile property values, session and request attributes, and date and time functions.

Security Policies answer the question: "Who has access to a resource?" A security policy is created when you define an association between a resource and one or more users, groups, or roles. Hence, a role policy defines a role and a security policy defines an authorization constraint associated with that role. We recommend basing security policies on roles rather than users or groups. Basing security policies on roles enables you to manage access based on a role that a user or group is granted, which is a more flexible method of management. If a security policy is based on a user or group, the user or group must be defined in the user store for the authentication provider that is configured in the default security realm. If a security policy is based on a role, the role must be defined in the store for the role mapping provider that is configured in the default security realm.

7.2 Portal Security[32] - Best Practices to Secure Portals

Implementation of a multi-tier architecture that isolates the web, application, and servers behind multiple firewalls provides the recommended security measures. The application architecture must consider requirements for intrusion protection.

An appropriate method for providing client accounts must be designed. The system will need a procedure that ensures log-on credentials (i.e., username and password) are delivered to the client in a secure fashion. It should include an efficient method to reset passwords when the client requests. Client credentials should also be coordinated with the master client index to safeguard against confusing access among similarly named clients.

A proactive incident recognition and response program should be implemented. Significant damage to reputation can occur when a security breach is not handled in a timely fashion. A critical aspect of good incident response is proactively monitoring the portal for suspicious events, service interruptions, code errors, and general utilization issues. Timely responses to analyze root causes, correct deficiencies, and communicate with the client population are essential activities.

[32] http://www.coretechrevolution.com/latest-news/139-three-basics-of-secure-client-portal-design

8. System Architecture

System architecture is a set of structuring principles that enables a system to be comprised of a set of simpler systems each with its own local context that is independent of but not inconsistent with the context of the larger system as a whole. A decision needs to be made for each of these systems as to whether to use open- source or closed source solutions.

8.1 Open vs. Proprietary Software: Considerations

There are two competing schools of thought when it comes to selecting software for implementing a new system. These both center around one central issue: whether to utilize open-source or closed-source software.

Open-source is any software, generally created by communities of programmers, that is freely available to the general public and, in most cases, allows the unrestricted access to the source code for the purposes of further development, customization, and redistribution.

Closed-source software is proprietary, licensed to the public usually at a cost, and is legally protected against modification and further distribution. Both concepts offer their own advantages and disadvantages.

General Software Considerations

There are many advantages and disadvantages that need to be considered when choosing between open source and closed source software. Here are a few examples to consider.

General Open Source Appraisal

- Advantages
  - Flexibility. It generally avoids vendor lock-in.
  - Scaling. The environment can scale without restrictions.
  - Adaptable. Source code can be adapted to meet the specific needs of the project.
- Disadvantages
  - Support. Rarely have technical support staff; most problem resolution is done through community discussion and testing, in which response time and assistance is not guaranteed.
  - Synergy. The project community's interests may not necessarily be aligned with the business needs of the organization.

General Closed-Source Appraisal

- Advantages
  - Ease of Use. Generally, the learning curve for using proprietary software is shorter, especially for lay users.
  - Standards. The underlying architecture often uses well-documented and widely-recognized standards.
  - Support. There is generally an on-staff technical support system / team.
  - Security. The security models are generally robust and well-tested.

- Disadvantages
  - Vendor Lock-in. The customer is stuck with the vendor's business decisions.
  - Price. Licensing can be expensive.

8.2 GIS Specific Software Considerations

Both open-source and closed-source web-based GIS systems also offer their set of own advantages and disadvantages. Here are a few examples to consider.

GIS Open-Source Appraisal

- Advantages
  - Maximum Control. Open-source code allows a developer to have full control over software configurations, customization, and bug fixes.
  - Extensive Web Tools. There exists several, tested, and proven web-oriented tools specifically developed for mapping, web authoring, and database integration.
  - Scalability. Open-source software is not regulated to hardware restrictions. In cases of heavy system usage, network and server infrastructure can be added on the fly to expand capacity without incurring additional software licensing costs.
  - Extensibility. Components, such as tilemill, can be easily added and removed.
  - Interoperability. Open data formats are easier and more straightforward than closed-source formats to translate to other data formats, including some closed-source data structures.
- Disadvantages
  - Limited Technical Support. Despite a large online open-source community, there is generally no single point of contact for technical software issues. Answers to issues would usually come from web research or a paid outside consultant familiar with the specific open-source technology used. However, some believe that community support from open-source software (OSS) is often superior to commercial support.

GIS Closed-Source Appraisal

- Advantages
  - Enterprise Specifications. Proprietary software generally tends to come with well-written software specifications, documentation, and service level agreements.
  - Support. Expert level phone, email, web or on-site support is generally available from the software vendor. There also tends to be a larger online knowledge base and training resources for developers to utilize.
  - Acceptance/Market Share. The majority of business enterprises currently rely on Microsoft for their IT infrastructure and desktop computing environments. Leading commercial GIS software products are exclusively developed for the Microsoft computing environment. File sharing and transfer is much easier when multiple people are utilizing the same software.
- Disadvantages
  - Cost Related to Licensing and Maintenance. Initial costs for purchasing and implementing a closed-source, commercial-grade, web-based GIS can be staggering and untenable for most public or grant-funded organizations. Additional maintenance costs are paid yearly in order to receive vendor support and new releases of the software.
  - Cost Related to Scalability. Commercial web-based GIS software tends to be licensed per computer processor and/or network connection basis. Additional licensing and

maintenance costs will occur when computing resources, such as additional web servers, are added to existing IT infrastructure for the purposes of increased performance or to keep up with high traffic volume demand.
- Restrictions. Proprietary software's source code is, for the most part, inaccessible to developers. Certain portions of the code are opened to developers via APIs, but defects in the source code are hard to detect and not resolvable by the developer. Only the vendors are able to make low-level source code changes and updates. Also, individual components cannot be easily replaced as they become legacy.

Web-Based GIS Software Platform Options

There are several legitimate open and closed-source GIS software platforms currently on the market capable of providing access to spatial data and advanced mapping and visualization via the internet. MapServer and GeoServer (both open-source) in addition to MapDotNet and ESRI's ArcGIS Server (both closed-source) are four platforms generally recognized by the GIS community as "best in class" in their respective licensing scheme.

Option 1 – MapServer (Open-Source). MapServer is an open-source platform specializing in publishing spatial data and interactive mapping applications to the web. Originally developed in the mid-1990s at the University of Minnesota, MapServer has a long established history and is considered to be the web-based GIS standard for grant-funded organizations. It is free to use, capable of integrating many data formats, runs on all major operating platforms (Windows, Linux, Mac OS X), and is actively improved upon by developers and users from all over the world.

- Advantages
  - Established software.
  - Cross-platform.
  - Follows OGC standards.
  - Large user community.
  - Stable and fast-performing.
  - Supports PHP, Python, Perl, Ruby, Java, and .NET scripting environments.
  - Google Maps/Bing Maps image tile API support.
- Disadvantages
  - No map authoring Graphical User Interface (GUI); all map configuration and cartography must be coded.
  - Requires extensive programming and configuration.
  - Steep learning curve.
  - Lacks modern REST and GeoWeb 2.0 interfaces.

Option 2 – GeoServer (Open-Source). GeoServer is open-source software written in the Java programming language. Like MapServer, GeoServer is free, cross-platform, widely used by institutions, such as World Bank and the Global Biodiversity Information Facility international organization, and has a large user and developer community. Its strength is the ability to consume other Web Feature Services (WFS) (e.g., WFS of weather data from NOAA) and its capability of sharing geo-data services with other web sites.

- Advantages
  - Fully compliant with current WMS, WCS, WFS standards.

- Mature support for many popular geospatial data formats.
- Friendly GUI for map configurations and authoring.
- Follows OGC standards.
- Commercial support available (fee-based).

- Disadvantages
  - Complicated Java architecture may require additional learning time.
  - In some cases, limited to no scripting language support; known exceptions include Python, Groovy, and Ruby.
  - Learning curve and additional time required to implement a GeoServer solution.
  - Documentation not as complete and developed as MapServer.

Option 3 – MapDotNet (Closed-Source). MapDotNet is proprietary GIS software written specifically to run on Microsoft Windows operating environments. MapDotNet offers a Microsoft .NET software development kit and a suite of design tools for creating high-end, feature rich, interactive mapping applications on Microsoft Silverlight and Windows Presentation Foundation platforms. It strengths are ease of map design, Microsoft Visual Studio integration, ability to consume existing WMS and WFS services, and intuitive end-user map visualizations.

- Advantages
  - Feature-rich cartographic map design GUI.
  - Extensive code samples available for free.
  - Support for SOAP, REST, and WCF endpoints.
  - Bing Maps image tile API support.
  - Web 2.0 and HTML 5 support including mobile support, such as iPhone.
  - Sleek, well designed end-user graphic interfaces and functionally.
  - Upfront and relatively low cost for initial license and year maintenance. $3,800 initial license, plus $760 yearly maintenance. $250 runtime license, plus $50 yearly runtime maintenance. $1,000 yearly developer support.
  - Seamless integration with SQL Server Spatial data types
- Disadvantages
  - Non-existent open community support. Must pay to access MapDotNet user forums.
  - Not a very mature solution when compared to other closed-source GIS software packages, such as ESRI ArcGIS Server.
  - Not cross-platform. Only runs within the Microsoft environment.
  - Additional planning time required for design and development considerations when implementing a MapDotNet solution.

Option 4 – ESRI ArcGIS Server (Closed-Source). ESRI is the current de facto standard for GIS technology in the world. The ArcGIS Server product is an enterprise, web-based GIS software widely accepted by the GIS community as “best in class” commercial grade, closed-source solution. ArcGIS Server represents decades of development and is used by various Federal, State, County, and Local municipalities, as well as grant-funded, international, academic, and commercial organizations.

- Advantages
  - Large open user community and extensive documentation.
  - Mature stable product that is actively maintained.
  - Extensive web mapping and mobile APIs not limited to Microsoft development environments.

- Extensive support for various database types and spatial data formats.
- Supports Microsoft Windows Server and certain Linux Server operating systems.
- Ability to handle large network loads and extensive processing.
- Well-designed intuitive user interfaces.
- All maps designed through industry standard ArcGIS Desktop software and authored to ArcGIS Server with little to no coding.
- Adheres to OGC standards and supports open data standards.
- Wide network of ESRI ArcGIS professionals ranging from analyst to developers.
- Advanced functionality available, such as network routing, geostatistics, back-end geoprocessing, and spatial analysis.
- Cloud-infrastructure support through Amazon EC2 cloud services.

▪ Disadvantages
- Expensive software for entire suite of products depending on needs. ArcGIS Desktop - $1,500 to $7,000 initial license, $500 to $3,000 yearly maintenance. ArcGIS Server - $5,000 to $40,000 initial license, $1,000 to $10,000 yearly maintenance. ArcGIS Development Network support license, $5,000 per year.
- Extensions for advanced functionally are at additional cost.
- Requires additional staff to maintain data and maps.

Stand-Alone GIS Software Platform Options

There are also many stand-alone (not web-based) data viewers that individual users may want to use for more sophisticated data visualization, such as creating animations and dissecting 3D model results and datasets.

Option 1 – USGS EverVIEW Data Viewer[33]. This cross-platform 2D-3D visualization platform was built specifically targeting time-series, grid-based, CERP-standard compliant modeling outputs. It was designed for and is widely used by modeler and resource managers working in the Greater Everglades. USGS lead and defined the CERP netCDF modeling data standard to ensure the viewer would be standards compliant. As such, during the 2012 Master Plan modeling effort, those modeling outputs that were compliant with the standard were easily consumed by the viewer. The viewer facilitates spatial data comparisons (visually) through the use of multiple globes which can be geo-spatially synchronized along with actual data comparisons through the table viewer interface. EverVIEW is free and open to the public, but it is not a true open-source application, as there are currently no resources available to maintain a community presence of software submits/code review/etc.

Option 2 – MapWindow[34] (Open-Source). MapWindow is a free and open source desktop geographic information system (GIS) project that includes an extensible plug-in architecture; a GIS ActiveX control; and C# GIS programmer library.

Option 3 – HydroDesktop[35] (Open-Source). HydroDesktop is a free and open source GIS enabled desktop application that helps you search for, download, visualize, and analyze hydrologic and climate data registered with the CUAHSI Hydrologic Information System.

[33] http://www.jem.gov/Modeling/EverView
[34] http://www.mapwindow.org
[35] http://hydrodesktop.codeplex.com

8.3 Hosting Environments: Considerations

There are many choices for hosting environments available. These include Shared Hosting, Virtual Private Server Hosting, Dedicated or Co-Location Hosting, Clustered Server Hosting (also known as Private Cloud Hosting), and Cloud Hosting.

Shared Hosting

Shared hosting is Web hosting in which the service provider serves pages for multiple Web sites, each having its own Internet domain name, from a single Web server. Most Web hosting companies provide shared hosting. Although shared hosting is a less expensive way for businesses to create a Web presence, it is usually not sufficient for Web sites with high traffic. These sites need a dedicated Web server, either provided by a Web hosting service or maintained in-house.

- Advantages
  - Price. Compared to the other environments, it is relatively inexpensive.
  - Administration. No system administration resources are required to maintain the hosting environment.
- Disadvantages
  - Robustness. There are multiple points of failure that could compromise the access to the data.
  - Control. There is a lack of direct control of the hosting environment.
  - Performance. There is a risk of poor or inconsistent performance, depending on the provider.

Virtual Private Servers

A virtual private server (VPS), also called a virtual dedicated server (VDS), is a virtual server that appears to the user as a dedicated server but is actually installed on a computer serving multiple Web sites. A single computer can have several VPSs, each one with its own operating system (OS) that runs the hosting software for a particular user.

The hosting software for each VPS can include a Web server program, a File Transfer Protocol (FTP) program, a mail server program and specialized applications for activities such as e-commerce and blogging.

The VPS alternative is often chosen by small businesses that need a customized Web site but cannot afford a dedicated server. Another useful aspect of VPS technology is the ability of a single subscriber to maintain multiple virtual servers. For example a Web site owner might use one server for the production-level Web site and the other for a "dummy site" that can be used to test planned updates, modifications or new programs.

- Advantages
  - Control. Most providers allow greater access to and more control of the hosting server.
  - Reliability. They are generally more reliable than Shared Hosting environments.
- Disadvantages
  - Administration. They usually require self-management, and therefore deeper system administration knowledge, but this can be outsourced.

- Reliability. Other virtual servers on your server may disrupt your service (e.g. bandwidth).
- Robustness. The hosting environment is not distributed (it is still just one server).

Dedicated or Co-Location Hosting

A dedicated server refers to the rental and exclusive use of a computer that includes a Web server, related software, and connection to the Internet, housed in the Web hosting company's premises. A dedicated server is usually needed for a Web site (or set of related company sites) that may develop a considerable amount of traffic - for example, a site that must handle up to 35 million hits a day. The server can usually be configured and operated remotely from the client company. Web hosting companies claim that the use of a dedicated server on their premises saves router, Internet connection, security system, and network administration costs.

In renting a dedicated server, the client company may be required to use a specified computer system or may be offered a choice of several systems. Some host providers allow a client company to purchase and install its own computer server at the host provider's location, a service known as colocation. Typically, a dedicated server is rented that provides a stated amount of memory, hard disk space, and bandwidth (here meaning the number of gigabytes of data that can be delivered each month). Some hosting companies allow the renter of a dedicated server to do virtual hosting, in turn renting services on the server to third parties for their Web sites. Domain name system, e-mail, and FTP capabilities are typically included and some companies provide an easy-to-use control interface.

- Advantages
  - Control. The customer has complete control of the hosting environment.
  - Reliability. Generally the hosting provider offers guaranteed reliability through Service Level Agreements (SLAs).
  - Security. External data security risks can also be mitigated through SLAs.

- Disadvantages
  - Price. The hosting service can be expensive.
  - Robustness. The hosting environment is not distributed.
  - Administration. System administration skills are essential unless the server is under a managed contract or the administration tasks are outsourced.

Private Cloud

A private cloud is designed to offer the same features and benefits of public cloud systems, but removes a number of objections to the cloud computing model including control over enterprise and customer data, worries about security, and issues connected to regulatory compliance.

Because a private cloud setup is implemented safely within the corporate firewall, it remains under the control of the owner’s IT department. The downside is private cloud ROI (return on investment): The organization implementing the private cloud is responsible for running and managing IT resources instead of passing that responsibility on to a third-party cloud provider.

- Advantages
  - Control. The customer has complete control of the hosting environment.
  - Reliability. Providers guaranteed reliability and external data security through SLAs.

  - Robustness. There is no single point of failure.
- Disadvantages
  - Price. The hardware can be expensive whether locally hosted or leased under a managed contract.
  - Administration. System administration skills are essential unless the cluster is under a managed contract.

Public Cloud

A public cloud is based on the standard cloud computing model, in which a service provider makes resources, such as applications and storage, available to the general public over the Internet. Public cloud services may be free or offered on a pay-per-usage model.

- Advantages
  - Reliability. The hosting environment is fully managed by a third party, meaning better reliability.
  - Performance. There are generally very few performance issues as the environment scales as needed.
  - Administration. No system administration skills are required.
- Disadvantages
  - Control. The service provider may be so large that external factors may affect your site and may take a long time to resolve.
  - Price. The hosting service can be expensive.

8.4 Web Services[36]

Web services provide a framework for fusing computing devices via open networks (the Internet, wireless, and local networks). In Web services, computing nodes have three roles: client, service, and broker. A client is any computer that accesses functions from one or more other computing nodes on the network. Typical clients include desktop computers, Web browsers, Java applets, and mobile devices. A client process makes a request of a computing service and receives results for each request. A service is a computing process that awaits requests, responds to each request, and returns a set of results. A broker is essentially a service metadata portal for registering and discovering services. Any network client can search the portal for an appropriate service. Server and broker technologies are typically used on UNIX, Linux, and Windows platforms.

Web services can support the integration of information and services that are maintained on a distributed network. This is appealing in organizations, such as local governments, that have entities or departments that independently collect and manage spatial data (e.g., roads, pipes, surveys, land records, administrative boundaries). At the same time, many of the functions of a local government require these data sets to be integrated. The use of Web services (a connecting technology) coupled with GIS (an integrating technology) can efficiently support this need. The result is that the various layers of information can be dynamically queried and integrated, while at the same time the custodians of the data can maintain this information in a distributed computing environment.

[36] http://www.esri.com/library/whitepapers/pdfs/spatial-data-standards.pdf

Web services avoid the issues and complications of GIS applications being tied to the spatial schema of a specific RDBMS vendor and allow GIS vendors to manage their own data using the best methods and formats for their tools in whatever database environment they choose. In addition, Web services allow server-to-server sharing of data and services, as opposed to integration only happening at the client level as it does with standards that are focused on the DBMS. Some vendors, such as ESRI, choose to use an RDBMS with schema and methods that perform optimally for their tools. Others use file systems. Web services mean that each GIS vendor can build and manage its own GIS data and readily provide GIS services (data, maps, and geo-processing) to a larger audience in a common environment.

Standards for Web Services

Protocols. The key standards used for Web services are a series of protocols (i.e., XML; Simple Object Access Protocol (SOAP); Web Services Description Language (WSDL); and Universal Description, Discovery, and Integration (UDDI)) that support sophisticated communications between various nodes in a network. They enable smarter communication and collaborative processing among nodes built within any web services compliant architecture.

Discovery and Access. Web services can be accessed with devices such as browsers, mobile devices such as telephones, desktop clients, and other information appliances. To discover these services, a broker is provided. The discovery protocol is referred to as a Universal Description, Discovery, and Integration (UDDI) node. In the GIS context, the UDDI node plays the role of a metadata server of registered Web services. A user can search a UDDI directory and find other distributed service providers or services that exist on a network. Web services interoperate (communicate) through an XML-based protocol known as Simple Object Access Protocol (SOAP). This is an XML API to the functions provided by a web service. Each web service "advertises" its SOAP API using a mechanism called Web Services Description Language (WSDL), allowing easy discovery of any service's capabilities. The Representational State Transfer is another available option for web services.

Application Programming Interfaces (API)[37]

Application Programming Interfaces. An API is a set of code and standards for accessing a Web-based application or service. This allows a web application to interact with other applications programmatically similar to the way human accesses a web application via GUI within web browsers. By exposing their functions or content via API, web applications allow developers to easily enrich their sites with new features.

A Maps API (e.g., Google Maps API, Yahoo! Map Developer API, MapQuest Open API, Map Control of Microsoft Virtual Earth, or ArcGIS API) is a source code interface that grants web developers access to a program library and to request services in generating a map over the Internet. The emergence of the Maps APIs was founded on powerful web map servers that provided extensive spatial data coverage around the globe. The spatial data that comprise the Internet map include the map data (e.g., road network, hydrographic features, or political boundaries) and remotely sensed imagery (both satellite and aerial). In general, the high resolution imagery (with spatial resolutions of 5 meters or less) and street-level map data may only be available in selected metropolitan areas. Thus, a Maps API enables the web developer to request spatial data for a selected geographic region through the Hypertext

[37] http://champs.cecs.ucf.edu/Library/Journal_Articles/pdfs/The%20Potential%20of%20Maps%20APIs%20for%20Internet.pdf

Transfer Protocol (HTTP) and embed the resulting map as an object in any external web site. A Maps API also allows the flexibility to add custom map controls for dynamic navigation by the map users, such as a navigation slide bar for zooming in/out and a toggle button to switch between street, aerial, and hybrid views. From the developers' perspective, the access to such valuable spatial data and dynamic functionalities per request can be regarded as a form of distributed GIS.

API System Architecture. Figure 6 shows the architecture of a simple website that uses the built-in functionalities and data provided by the Maps APIs.

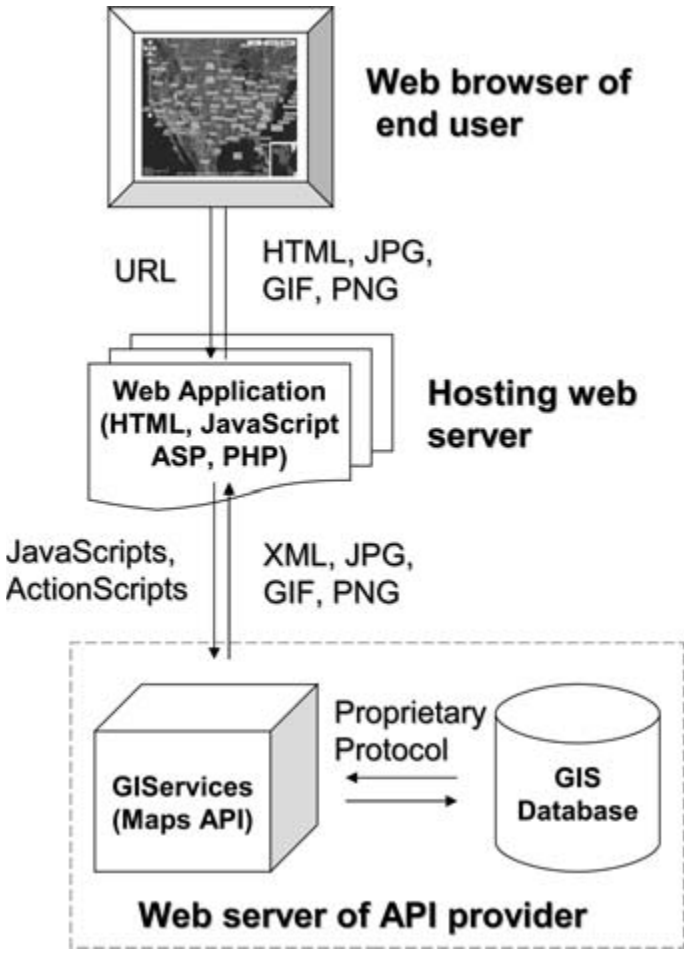

Figure 6. Example architecture of a simple website that uses the built-in functionalities and data provided by the Maps APIs.

The conceptual architecture of a web application that uses the Maps API is quite simple. In general, the web application is hosted in a web server that will return Hypertext Markup Language (HTML) and web-compatible graphics (e.g. JPG, GIF, or PNG) upon the request of a web browser. By using JavaScript to connect to the Maps APIs, the web application has access to the web servers of the API provider in requesting GIS services, such as zoom in/out. Based on the input parameters and values collected by the map interface of the web application, the web server of the API provider will return the spatial data (i.e., map) in the form of web-compatible graphics. Most API providers (with the exception of Virtual Earth) require the registration of a "map key" in accessing the Maps APIs and/or a limit on the number

of page views, queries, and geo-code requests per day for a single registered web directory. Currently, most Maps APIs remain free and do not include advertising.

Limitations. Despite its versatility in dynamic exploration of geographic data online, the existing Maps APIs lack analytical and spatial functionalities compared to other Internet Mapping Services (e.g. ESRIs ArcIMS). For example, spatial operations like buffering, geo-processing or map algebra are not supported in the current version of the Maps APIs. While the web technology of Maps APIs continues to mature, the lack of W3C-endorsed standards, technical support (such as a knowledge base), and literature reviews may present barriers to some.

8.5 Selecting Storage Media for Long-Term Preservation[38]

No computer storage medium can be considered archival, irrespective of its physical longevity: technological obsolescence is inevitable and all media have limited life spans. For the foreseeable future, the need to periodically refresh electronic records onto new media is therefore inescapable. Careful selection of storage media can maximize the periods between refreshment cycles and simplify the refreshment process, in addition to ensuring that data is as secure as possible.

The following criteria should be considered by data creators when selecting removable storage media:

- **Longevity.** The media storage option chosen should have a proven life span of at least 10 years. Longevity greater than this, however, is not necessarily an advantage: over longer timescales, obsolescence of the drive technology used to read the medium will typically be a much more significant factor than physical deterioration of the storage medium itself.

- **Capacity.** The media storage option chosen should provide a storage capacity appropriate for the quantity of data to be stored and the physical size of the storage facilities available. Minimizing the number of actual media to be managed will generally be more efficient and cost effective.

- **Viability.** The media and drives chosen should support robust error-detection methods for both reading and writing data. Provision for testing the integrity of media after writing is also a benefit. Proven data recovery techniques should also be available in case of data loss. Media should be write-once, or have a reliable write-protect mechanism to prevent accidental erasure and maintain the evidential integrity of the data.

- **Obsolescence.** The media and its supporting hardware and software should preferably be based on mature, rather than leading-edge technology. The technology should be well established in the market place and widely available. Media technologies that are based upon open standards for both media and drives are generally preferable to those that are proprietary to a single manufacturer.

- **Cost.** Two elements must be considered when assessing the relative costs of storage media – the cost of the media itself and the total cost of ownership. Valid comparisons of media costs must always be made on a price per Gigabyte (GB) basis. The total cost of ownership will include costs for purchasing and maintaining the necessary hardware and software, and of any

[38] http://www.nationalarchives.gov.uk/documents/selecting-storage-media.pdf

storage equipment required. Support costs and the quoted Mean Time Before Failure (MTBF) of the relevant drive must also be taken into account.

- **Susceptibility.** The media should have low susceptibility to physical damage and be tolerant of a wide range of environmental conditions without data loss. Magnetic media should have a high coercivity value in order to minimize the chances of accidental erasure through exposure to magnetic fields. Any measures required to counter known susceptibilities (such as packaging or storage requirements) should be affordable and achievable.

- **Media Selection Scorecard.** The scorecard approach illustrated in Table 3 uses a simple method for evaluating currently available media against the selection criteria. Each medium should be scored against the criteria on a scale of 1 (does not meet the criterion) to 3 (fully meets the criterion). As a general rule, no medium that scores less than 12 should be considered.

Table 3. Scorecard approach for evaluating currently available media against selection criteria.

Media	CD-R	DVD-R	Hard Disk	Flash Memory Stick and Card	Linear Tape Open (LTO)
Longevity	3	3	2	1	3
Capacity	1	3	3	2	3
Viability	2	2	2	1	3
Obsolescence	1	2	2	2	2
Cost	3	3	1	3	3
Susceptibility	1	1	3	1	3
Total	**11**	**14**	**13**	**10**	**17**

The evaluation process should fully take into account variations within a media type. For example, a number of different dye and metal layer combinations are available for CD-Rs. Recent research suggests that CD-Rs with a gold reflective layer and phthalocyanine-based dyes are the most stable and have the greatest life span.

In situations where multiple copies of data are stored on separate media, it is advantageous to use different media types for each copy, preferably using different base technologies (for example, magnetic and optical). This reduces the overall technology dependence of the stored data. Where the same type of media is used for multiple copies, different brands or batches should be used in each case in order to minimize the risks of data loss due to problems with specific manufacturers or batches.

9. Implementation Activities

As discussed in Section 1, this section expands upon the implementation associated with both the short and long term goals of the data management effort.

9.1 Short-Term Goals

Short-term goals are targeted for completion within six months to two years and focus on gathering requirements, creating the architecture and collecting all of the data inputs. Once all of the data inputs are collected, implementation focuses on categorizing and documenting the data that will enter the system.

Establish Policies and Standards

After discussion with CPRA and major stakeholders, the data management project team will develop a library of policies and standards necessary to maintain consistency and quality of the data. Existing policies and standards will be evaluated against the data inventory to determine the relevance to the current data sets and revised as necessary. New guidance will be developed to address quality issues identified during the data inventory phase. Established policies and standards utilized by federal and state agencies, as well as those employed by academia and industry will be evaluated by the group to determine the applicability and potential use. Priority will be given to 2012 Coastal Master Plan data.

Inventory Data

All existing and anticipated data streams, and the associated meta-data, to be incorporated into the system will be inventoried and documented. Existing data will be evaluated to determine if they are current and complete or if the data have been superseded by a newer data stream. A preliminary inventory of data and data management systems relevant to coastal Louisiana is included as Appendix A. Fragmented or disparate data (data not currently incorporated into a management system or shared) will be assessed for content and categorization. Priority will be assigned to data related to the 2012 Coastal Master Plan. Once all of the data streams have been inventoried, each data stream will be assessed to help define the format moving forward.

Identify and Setup the Data Management System to a Pilot Basin

The pilot study would include several major phases. Phase one would be to identify strategic users in each functional group. The user groups (Executives, Project Managers / Scientists, Public, etc.) will be defined and prioritized. The groups will be prioritized based on which functions will be completed first in the software life cycle plan. Interactions with these users will determine the core of the requirements and workflow mapping phases. Phase two would include gathering requirements for the system from these users. Ideally, in requirements gathering, the proposed users of the system would identify what they would like the system to do and what problems they are trying to solve. Phase three would be workflow mapping. This would involve sitting down with the same users and understanding their workflow. When the workflow is understood, suggestions are made to enhance the workflow in the new tool. Earlier requirements are checked against the workflow to ensure that more complexity than is needed is not introduced and discussion of possible conflicts between stated requirements and

current workflow are addressed. The pilot study would be applied to a basin with a reasonable historical dataset.

Develop System Architecture for the Pilot Basin

The CPRA has identified three dashboards of particular interest in the short term: executive, operations, and analytical. The overall design of the system architecture will be developed with inputs from a workgroup consisting of a cross section CPRA personnel and stakeholders associated with data production and consumption. The success of the system will depend on the value realized by these groups and their continued input to improve data quality and use.

After gathering requirements from the various user groups, the portal requirements will be assessed and a software roadmap will be planned. Conflicting requirements will be discussed and resolved.

The executive level dashboard will be developed through interactions with the executive team. The most relevant data and functions determined for the executive user will be displayed in an easy to use dashboard application. Non-working dashboard designs will be created to demonstrate the application to the various user groups, and feedback regarding their specific needs will help determine the final design. The basic design will mimic the executive dashboard, but the available data and tools will be different.

Training Programs for the Pilot Basin

The CPRA will conduct system training for the various user groups to increase utility and therefore the return on the resources and funding invested in the data management system. The frequency and content of the trainings will be determined by documenting user requests and monitoring the trends observed by the support personnel. Depending upon the trends, specific workshops can be developed and merged with the existing CPRA training schedule. To create meaningful training sessions, the users will be queried about their needs. Training sessions on the Pilot Basin will be refined based on user feedback and any program changes. It would provide valuable information on the frequency at which trainings should occur, identify the training staff, and would facilitate developing guidelines and tutorials for each of the data streams

Identify Resources

Based upon the pilot study, a resource-requirements report will be developed to document the estimated level of effort needed to design, develop, maintain, and provide user support, as well as the hardware and software necessary to support the broader data management system. Securing these dedicated resources will be necessary to advance the data management system through subsequent phases of this overall effort. During the initial phases of the project, contracted personnel would be expected to be utilized to supplement the CPRA development team. Long term operation, maintenance, and management of the system would be performed by dedicated CPRA personnel. Descriptions of the personnel, responsibilities, and an example of resource requirements for a scenario similar to the proposed CPRA system are provided in Section 4. Phase 1 of the short term implementation activities will determine the necessary resources for the proposed system which may be different then the estimate provided in Section 4.

2012 Coastal Master Plan Modeling Data

The 2012 Coastal Master Plan modeling inputs, outputs, etc. will be gathered and cataloged during the data inventory phase; section 5 contains a summary of the data content currently archived by CPRA. These modeling outputs will facilitate future planning efforts in coastal Louisiana and provide standards by which future modeling data will be organized. The development team will utilize the 2012 Coastal Master Plan modeling data to design an initial component of the data management system. This component will facilitate the re-consumption of the modeling outputs for future planning cycles. The development team will coordinate the design of the data system with the modelers and other data users to determine the tools and organizational structure which would bring the most efficiency to the broader user community.

User Assessment

A protocol to measure the overall user experience will be developed to ensure the system is satisfying or exceeding the goals established by the design team and stakeholders. The system will be truly successful if the users find the system efficiently processes their submittal of data and/or request for information. Initially, a simple survey request to users will provide an overall assessment of the user experience and provide the management team with suggestions on areas of the system which warrant attention. As the system matures, representatives from specific user groups can be identified to provide further refinement of specific applications or requests for the development of additional data management tools.

9.2 Long-Term Goals

Long-term goals are targeted for completion within two to five years, and will continue to be modified as needed. These efforts will, in many cases, be direct follow-ons to efforts undertaken in the near-term plan. Long-term goals are provided below.

Extend the Data Management System to Other Coastal Basins

The data management system established for the Pilot Basin would be extended to other coastal basins. The entire Louisiana Coast would be incorporated into the Data Management System.

2017 Coastal Master Plan

Implementing a method to capture, catalog, and make available final outputs from the 2017 Coastal Master Plan data within the data management system will follow upon the strategy and effort implemented in the near-term plan for the 2012 Coastal Master Plan data. This process should become more streamlined though time, as standards are put into place (early on) for the 2017 Coastal Master Plan effort, based on the strategy adopted for the 2012 data.

Fragmented Databases

In the near-term, an inventory of existing data streams and data management efforts will be completed. It is also important to locate and incorporate fragmented or disparate databases and their data into the broader data management system. It is envisioned that this effort would begin in the near-term, as part of the data inventory task, and incorporation of these data sets will continue into the future.

Legacy Data

Similar to the aforementioned goal of identifying fragmented databases, identifying legacy data and developing the tools needed to capture information in the formats specified by the data management system is also an important task. Because it is often difficult to locate legacy datasets, and a significant amount of resources will be needed to transfer datasets into specific formats, it is envisioned that this task could take several years to complete. It is possible that this task could begin with an electronic query of the coastal technical community, inviting people to share their datasets with the broader community.

New Data

Developing standards to incorporate new data previously not addressed will follow-on to the task initiated in the near-term implementation of the data management effort. Lessons learned from the near-term strategy, utility, and user satisfaction will be used to refine standards that are applied to new data sets to ensure greater functionality. Considering the nature of technology and new data formats of the future, it will be important to maintain flexibility in the system to account for unforeseen changes where possible.

Automated Notifications

Once the data management system is populated and tested, a tool for automatically tracking and communicating data updates to registered users will be developed. In this manner, users who voluntarily sign up to receive such notifications will be alerted anytime a dataset is added or modified. Successful execution and utilization of this feature should help ensure regular users of the datasets are always aware of the most recent versions available.

System Operations and Maintenance

It is envisioned that long term operation, maintenance, and management of the system would be performed by dedicated CPRA personnel. Descriptions of the personnel, responsibilities, and an example of resource requirements for a scenario similar to the proposed CPRA system are provided in Section 4. Once the system is designed and set up, a plan must be put into place to outline system operations and maintenance, including personnel resources. The general operations and maintenance must be flexible and able to account for technological changes as necessary to ensure the system remains effective for the users. Maintenance plans will include considerations for both hardware and software needs.

Resource Projections

As described in the previous item, a long term plan needs to be developed to maintain the hardware, software, and related components of the data management system. In order to sustain operations detailed in the maintenance plan, personnel and equipment resource projections must be done to accommodate the anticipated growth of data and information relevant to the CPRA. An example personnel resource allocation is provided in Section 4.

APPENDIX A Preliminary Inventory of Data Management Systems in Coastal Louisiana

PROGRAMS

Program	Program Description	Responsible Party	Data Types	Spatial Extent	Temporal Collection	Update Frequency	Data Access	Website	Contact / Reference
CRMS (Coastwide Reference Monitoring System)	CRMS data collection effort. *Raw data from CRMS stations are stored on the SONRIS system.*	USGS–NWRC and CPRA	**Data collected at CRMS stations:** • Accretion • DCP and continuous hydrographic data • Discrete hydrographic data • Forested swamp vegetation • Herbaceous marsh vegetation • Soil properties • Surface elevation • Spatial data (land:water analysis)	• Coastwide in Louisiana • Stations in the marsh • 390 monitoring stations total	As early as 2005 for some stations – through present day.	Variable, depending on data set: • Continuous hourly • Monthly • Twice annually • Annually • 3-years	Direct download	http://sonris.com/dataaccess.asp	Dona Weifenbach, CPRA Dona.Weifenbach@la.gov
CRMS (Coastwide Reference Monitoring System)	Synthesized, calculated, or derived data from the CRMS monitoring stations.	USGS–NWRC and CPRA	**Synthesized data available on the CRMS website:** • Hydro (water level range, hydro completeness, salinity, water level, temperature, continuous, hydro index, percent flooded, soil porewater, precipitation) • Vegetation (forested, herbaceous, site floristic quality index, marsh class, basal area, vegetation percent cover) • Soil (percent organic, bulk density, surface elevation/accretion, submergence vulnerability index) • Spatial (percent land)	Same as above	Same as above	Same as above	Same as above	http://www.lacoast.gov/crms2/Home.aspx	Greg Steyer, USGS steyerg@usgs.gov

Program	Program Description	Responsible Party	Data Types	Spatial Extent	Temporal Collection	Update Frequency	Data Access	Website	Contact / Reference
LASARD (**L**ouisian**a** **Sa**nd/Sediment **R**esources **D**atabase)	LASARD is a publically available centralized GIS database of (mostly) sand resources – offshore and river. It is part of Louisiana Sediment Management Plan (LASMP). Data from over 200 sources have been inventoried, reformatted/standardized, archived, and made available.	CPRA	**Data types currently addressed by LASARD:** • Bathymetry/topography • Seismic • Sidescan sonar • Magnetometer • Vibracore/grab • Isopach • Deposit/borrow areas • Shipwrecks • Pipelines and platforms	Louisiana coastal zone from Texas to Mississippi, including the Mississippi River	Variable: as early as 1930 - through 2011.	Variable, based on acquisition of new dataset to inventory	Direct download	http://sonris-www.dnr.state.la.us/gis/agsweb/IE/JSViewer/index.html?TemplateID=21	Ed Haywood, CPRA Ed.Haywood@la.gov MELANY LERANAS' PRESENTATION FROM THE OCTOBER 3-4, 2012, DATA MANAGEMENT WORKSHOP AT THE WATER INSTITUTE
SONRIS (Strategic Online Natural Resource Information System	Hosted by LDNR primarily to house natural resource data and information. Only a section of it is devoted to CPRA's data storage including CRMS, LASARD, and other project-related information/data. CPRA deals with: GIS data layers, monitoring data (CRMS data and otherwise), engineering data (soil borings and GPS elevation benchmarks), documents.	LDNR and CPRA	SONRIS Interactive GIS Map website (GIS data) • CPRA layers (CPRA projects, coastal monitoring stations, soil borings, levees, etc.) • LASARD layers (bathymetry, sand borrow areas, track lines) • Reference layers (primary roads, major inland water bodies, etc.) • Imagery and map layers (coastal photographs, SPOT imagery, Landsat, etc.) SONRIS (Oracle database) - CRMS monitoring data and other coastal data related to LDNR (oil and production, transportation, groundwater, minerals) are also stored here. • CRMS monitoring stations - Accretion - DCP and continuous hydrographic data - Discrete hydrographic data - Forested swamp vegetation - Herbaceous marsh vegetation - Soil properties - Surface elevation - Spatial data (land:water analysis)	Statewide, data from over 17,000 monitoring stations	1987 – present day	Variable	Direct download	SONRIS - http://sonris.com/ SONRIS Interactive GIS Map – http://sonris-www.dnr.state.la.us/gis/agsweb/IE/JSViewer/index.html?TemplateID=21 SONRIS LITE - http://sonris.com/sonlite.asp	Ed Haywood, CPRA Ed.Haywood@la.gov CHRISTOPHER ROBERTSON'S PRESENTATION FROM THE OCTOBER 3-4, 2012, DATA MANAGEMENT WORKSHOP AT THE WATER INSTITUTE

Program	Program Description	Responsible Party	Data Types	Spatial Extent	Temporal Collection	Update Frequency	Data Access	Website	Contact / Reference
National Water Level Observation Network	Long-term, continuously operating tide stations that provide data and information on tidal datum and relative sea level trends, and are capable of producing real-time data for storm surge warning.	NOAA - National Ocean Service (NOS); Center for Operational Oceanographic Products and Services	• Currents • Water level • Meteorological	Coast wide • LAWMA, Alameda Pass • Calcasieu Pass; • East Bank 1, Bayou LaBranche • Freshwater Canal Locks • Grand Isle, • Lake Charles • USCG New Canal Station • Shell Beach • Lake Borgne • Pilots Station SW Pass • Tesoro Marine Terminal	Variable	Real time and predictions	Appears to be viewable only	http://tidesandcurrents.noaa.gov; http://tidesandcurrents.noaa.gov/nowcoast.html	Center for Operational Oceanographic Products and Services (CO-OPS) 1305 East-West Highway Silver Spring, MD 20910-3281 Phone: (301) 713-2815 Fax: (301) 713-4500
Lower Mississippi River PORTS® -	Physical Oceanographic Real-Time System (PORTS®) - responsible for providing real-time oceanographic data and other navigation products to promote safe and efficient navigation within U.S. waters.	NOAA -National Ocean Service (NOS); Center for Operational Oceanographic Products and Services	• Water level • Currents • Air gap data • Meteorological data	In the Mississippi River: • Port Allen • Huey Long Bridge • Carrollton • Crescent City • First Street Wharf • Pilot town • SW Pass	Variable	Real time	Variable	http://tidesandcurrents.noaa.gov/ports/index.shtml?port=lm	Center for Operational Oceanographic Products and Services (CO-OPS) 1305 East-West Highway Silver Spring, MD 20910-3281 Phone: (301) 713-2815 Fax: (301) 713-4500
GCOOS (Gulf of Mexico Coastal Ocean Observing System)	GCOOS Data Portal provides timely information about the environment of the United States portion of the Gulf of Mexico and its estuaries for use by decision-makers, including researchers, government managers, industry, the military, educators, emergency responders, and the general public. Houses stations from other programs across the Gulf of Mexico.	NOAA - National Ocean Service (NOS); Integrated Ocean Observing System (IOOS)	See LUMCON and WAVCIS (for the Louisiana data types)	Gulf of Mexico	Variable	Variable	Direct download	Primary Website http://www.gcoos.org/ Data Access: http://data.gcoos.org/	Felimon Gayanilo Harte Research Institute for Gulf of Mexico Studies; Texas A&M University - Corpus Christi felimon.gayanilo@tamucc.edu

Program	Program Description	Responsible Party	Data Types	Spatial Extent	Temporal Collection	Update Frequency	Data Access	Website	Contact / Reference
WAVCIS	Wave-current information system) provides wave information (sea state) including wave height, period, direction of propagation, water level, surge, near surface current speed and direction and meteorological conditions on a real time basis around the entire Louisiana coast.	LSU - Coastal Studies Institute	• Wave height • Wave period • Current speed • Current direction • Wind speed • Wind direction • Gusts • Air temperature • Barometric pressure • Sea surface temperature • Water depth	Coast wide (near shore)	Variable, as far back as 1998	Real time and forecasting	Data are viewable, but online registration is required for download	http://www.wavcis.lsu.edu/	Chunyan Li, LSU cli@lsu.edu Phone: (225) 578-3619
Digital Coast	Digital Coast is a data synthesis and accessibility hub. It does not appear to host raw data; rather, it provides synthesized data downloads, data use tools, and online data-use training opportunities.	NOAA – Coastal Services Center	**Data accessible on this site:** • Benthic • Hydrography • Marine boundaries • Imagery • Land cover • Elevation • Socioeconomics **SPECIAL FEATURES:** **Tools** - tools to support coastal decision-making by transforming Digital Coast data into information tailored for specific issues. Some are web-based, allowing for direct online analysis and viewing,; others are downloadable for desktop use. • Data analysis • Classification • Change • Spatial visualization • Non spatial visualization	Nationwide – coastal focus	Variable	Variable	Direct download	Primary Website http://www.csc.noaa.gov/digitalcoast/ Data Access http://www.csc.noaa.gov/dataviewer/index.html#app=674c&6ba3-selectedIndex=0	NOAA Coastal Services Center 2234 South Hobson Avenue Charleston, SC 29405 Phone: (843) 740-1200

Program	Program Description	Responsible Party	Data Types	Spatial Extent	Temporal Collection	Update Frequency	Data Access	Website	Contact / Reference
Barrier Island Comprehensive Monitoring (BICM)	Long-term data collection on Louisiana's barrier islands.	CPRA, USGS, and UNO	• Aerial extent of islands • Sub-aqueous extent of islands (depth of closure) • Habitat classifications • Sediment properties / geotechnical • Environmental processes • Vegetation composition	Louisiana barrier shoreline	As early as 2005; as recent as 2008	Single sampling effort; BICM-2 should begin in 2013	Direct download; see Coastal Louisiana Resource Information System	FTP Site: ftp://ftp.dnr.state.la.us/bicm/ Some data available for download here: http://www.ladigitalcoast.uno.edu/data.html All data available via email request: http://sonris-www.dnr.state.la.us/sundown/cart_prod/cart_ocpr_project_summary?pattask_proj_id=5061e122000d82e6d1728b3827f3d7c6&popen_in_attask=N	Darin Lee, CPRA Email: darin.lee@la.gov
Coastal Louisiana Resource Information System	Web-accessible digital clearinghouse; provides data products (e.g., shapefiles, maps, remote sensed images, tabulated text, and point-source data). Hosts the BICM data.	UNO - Pontchartrain Institute for Environmental Sciences (PIES)	• BICM Datasets • Vector shorelines • Shoreline change • Land use / land cover • Bathymetric surveys • LiDar • Aerial surveys	Coastwide	Variable	Variable	Direct download	http://www.ladigitalcoast.uno.edu/data.html	Luis Martinez, UNO lmartinez@uno.edu Phone: (504) 280-4014
National Data Buoy Center	NDBC designs, develops, operates, and maintains a network of data collecting buoys and coastal stations.	NOAA - NWS (National Weather Service)	• Level Press • Wind Speed & Dir • Air Temp • Sea Sfc Temp • Sig Wave Ht. • Dew Point • Tides	Gulf of Mexico	Variable	Varialbe; some real time (hourly)	Direct download	http://www.ndbc.noaa.gov/	National Data Buoy Center Bldg. 3205 Stennis Space Center, MS 39529 Phone: (228) 688-2805

Program	Program Description	Responsible Party	Data Types	Spatial Extent	Temporal Collection	Update Frequency	Data Access	Website	Contact / Reference
Louisiana Fisheries Mapping Application	Mappling application can be used to view fisheries independent data for selected species found in Louisiana's 7 coastal study areas	Houston Advanced Research Center (HARC)	• Black drum • Blue crab • Brown shrimp • Gafftopsail catfish • Inland silverside • Striped mullet • Salinity • Temperature • Turbidity • Dissolved oxygen	Louisiana's 7 coastal study areas	Seasonal, Yearly	Last updated 2010	ArcGIS online viewer; direct downloadable available for some	http://gulfcoast.harc.edu/CoastalResources/CoastalFisheries/LouisianaCoastalFisheries/tabid/2235/Default.aspx	4800 Research Forest Drive The Woodlands, TX 77381 Phone: (28) 367-1348 Email: webmaster@harc.edu
Atlas	Geospatial data for the state of Louisiana	LSU CADGIS Research Laboratory	GIS data • Census 2000 • Coastwide 2001 imagery • DOQQ 1998, 2004, 2005 • LIDAR • Marsh-Veg Types • NOAA nautical charts • National Wetlands Inventory • Landsat	Coastwide	Variable	Last updated 2009	Direct download; DVD	http://atlas.lsu.edu/	Email: gis@lsu.edu
Multipurpose Marine Cadastre	An integrated marine information system that provides authoritative and regularly updated ocean information.	Bureau of Ocean Energy Management and NOAA Coastal Services Center	• Critical habitat designations • Aids to navigation • Lease blocks • Wind energy areas • Marine mammals	Federal waters of the U.S.	Variable	Variable	Direct download; Online viewer	http://www.marinecadastre.gov/default.aspx	Email: nos.csc.mmc@noaa.gov
Northern Gulf of Mexico (NGOM) Ecosystem Change and Hazard Susceptibility Data Portal	A Google Maps -based web portal for distribution of published lidar, bathymetry, core and imagery data. The second purpose of this tool is for use in coordinating multi-agency surveying and data gathering efforts around the NGOM region in order to minimize reduncy and maximize cooperation.	USGS - Coastal and Marine Geology Program	• Lidar • Bathymetry • Cores • Imagery data	Northern Gulf of Mexico	2006 - 2009	Last updated 1/10/13	Some direct download; some online visualization	http://ngom.usgs.gov/index.html	Project Leader: John Brock - USGS Florida Integrated Science Center, St. Petersburg, FL Email: jbrock@usgs.gov

Program	Program Description	Responsible Party	Data Types	Spatial Extent	Temporal Collection	Update Frequency	Data Access	Website	Contact / Reference
National Weather Service River Forecast	Lower Mississippi River gage readings and weather forecasts.	National Weather Service	• Stage • Flow • Volume • Precipitation	Gage readings in the lower Mississippi River Precipitation (coastwide)	Variable	Real-time	Download, Online visualization	http://www.srh.noaa.gov/lmrfc/	Email: SR-LMRFC.Webmaster@noaa.gov Online contact form: http://www.weather.gov/contact
GulfSource	Program to test seafood, water, and soil across the Gulf coast for compounds.	LA Depts of: • Wildlife and Fisheries • Health and Hospitals • Environmental Quality • Agriculture and Forestry	Laboratory data for • Organic contaminants • Polycyclic aromatic hydrocarbons • Dioctylsulfosuccinate	Louisiana nearshore and inshore in both federal and state waters	May 2010 to December 2012	Variable	Download, Online visualization	http://gulfsource.org/#	Phone: 1.800.442.2511 Email: ASKUS@GULFSOURCE.ORG

ORGANIZATIONS

Organization	Description	Responsible Party	Data Types	Spatial Extent	Temporal Collection	Update Frequency	Data Access	Website	Contact / Reference
Louisiana Universities Marine Consortium (LUMCON)	Marine research facility located in Cocodrie, LA. Manages several weather and hydrographic monitoring stations.	See organization	Weather • Air temperature • Wind direction • Wind speed • Relative humidity • Barometric pressure • Solar radiation • Quantum radiation (PAR) • Precipitation • Hydrographic • Water temperature • Water height • Salinity • Turbidity	Southeast Louisiana • Lake Pontchartrain • Audubon Park (Mississippi River) • Terrebonne Bay • LUMCON (Cocodrie) • Southwest Pass (in-active) • Tambour Bay (in-active)	Variable - as early as 2002 for some stations - through present day	Real-time, continuous (minute)	Weather and Hydrographic = upon request; Archived wind and temperature profile data = direct download	http://weather.lumcon.edu/stationdata.asp?stationid=101	www@lumcon.edu
USACE-MVN (New Orleans District)	Hydrologic monitoring stations managed by the U.S. Army Corps of Engineers, Mississippi River Valley. Daily River Stage and Discharge	See organization	• Daily stage • Discharge	Mississippi River • Vicksburg • Natchez • Knox Landing • Red River Landing • Baton Rouge • Donaldsonville • Reserve • New Orleans • Venice Atchafalaya River • Barbre Landing • Simmesport • Melville • Krotz Sprints • Butte La Rose • Buffalo Cove • Morgan City Old River • Riverside • Channel side • Delta head	Contact MVNHydraulics@usace.army.mil	Daily	N/A – only able to view daily values	http://www.mvn.usace.army.mil/eng/edhd/dailystagedisplay.asp	CEMVN-ED-HD New Orleans, LA Email: MVNHydraulics@usace.army.mil

Organization	Description	Responsible Party	Data Types	Spatial Extent	Temporal Collection	Update Frequency	Data Access	Website	Contact / Reference
USGS - Louisiana Water Science Center	USGS water monitoring in Louisiana; part of the larger USGS National Water Information System (NWIS)	See organization	Real-time data • Streamflow • Groundwater • Water quality • Precipitation Historical data • Streamflow • Groundwater • Water quality • Duration hydrographs	Statewide	Variable	Variable (including some real time)	Direct download	USGS -Louisiana Water Science Center http://la.water.usgs.gov/ National Water Information System http://waterdata.usgs.gov/nwis	Automated text box provided on the website; no direct contact information provided.
LA Dept Environmental Quality	Ambient surface water data at approximately 125 sites across the state each month.	See organization	• Air monitoring • Mercury (fish, vegetation, sediment, water) • Ambient data (water quality monitoring stations and subsegments)	Statewide (rotation of 691 water quality monitoring sites; primarily in rivers / tributaries)	Contact DEQ; monitoring is temporally rotated	Monthly (or less) for some stations	Online registration required	http://www.deq.louisiana.gov/portal/tabid/2742/Default.aspx GIS http://map.deq.louisiana.gov/index2.htm	602 North Fifth Street; Baton Rouge, LA Phone: (225) 219-5337 Email: _DEQ-CustomerServiceCenter@la.gov
LA Dept Wildlife and Fisheries	Wildlife and fisheries sampling in Louisiana	See organization	Fisheries-independent (biological): • Trawl (6 and 16 ft) • Finfish (50-ft seins and gillnets) Alligator and other wildlife data Oyster monitoring	Statewide (estuarine / coastal)	Trawl - as far back as 1966 Finfish - as far back as 1982 Oyster - as far back as 1980	Trawl - weekly to bi-weekly Finfish - monthly to bi-monthly Oyster - annual	Fisheries data in Breton Sound (1988-2009) and data, including oyster, in Barataria Basin (1998-2010) available via SONRIS-Lite. Other biological monitoring require written request to Mike Harden at LDWF.	http://www.wlf.louisiana.gov	Mike Harden Harden_MM@wlf.state.la.us Phone (225) 765-2382
NOAA - National Marine Fisheries Service	Southeast Fisheries Science Center - Field agents serve as the principle data collection agent for marine fisheries throughout the Southeast United States (NC-TX)	NOAA - Southeast Fisheries Science Center	• Recreational and commercial fisheries harvest data (coastal, pelagic, and shrimp) • Landings / economic data • Fisheries monitoring (size, fecundity, sex, age) • Sea turtles and reef fish data	Variable	Variable, some shrimp data date back to 1956, fisheries logbooks date back to 1986	Variable	Some downloads available from the website	http://www.sefsc.noaa.gov/	Southeast Fisheries Science Center 75 Virginia Beach Dr. Miami, FL 33149 Phone (305) 361-4200

Organization	Description	Responsible Party	Data Types	Spatial Extent	Temporal Collection	Update Frequency	Data Access	Website	Contact / Reference
LA Department of Health and Hospitals	Monitors and publicizes resources related to environmental and human health.	See organization - Center for Environmental Health	Beach Monitoring Program tests water at 26 beach sites Molluscan Shellfish Program	Coastwide	Contact LDHH	Beach monitoring (weekly from May 1st through October 31st.) "Regular" monitoring of molluscan shellfish	Visualization (mapping) and via advisories	http://new.dhh.louisiana.gov/index.cfm/page/452/n/279	628 N. 4th Street | Baton Rouge, LA 70802 Phone: (225) 342-9500
LA Governor's Office of Homeland Security & Emergency Preparedness (GOHSEP)	Virtual Louisiana is a GoogleEarth-based project to consolidate all geospatial data layers that would be useful to LA state agencies in an emergency. State-Government access only. Louisiana Earth is a GoogleEarth-based project that makes useful Louisiana geospatial data layers available to public in case of an emergency.	See organization	GIS layers considered useful in an emergency: • Evacuation zones • Evacuation routes • Contraflow • Gas stations GIS layers considered useful during non-emergencies • Cultural events (time and route of parades, festivals, etc.) • Outdoor interests (parks, public hunting lands, fishing areas, hiking trails, bicycle routes, etc.)	Statewide	Variable	Variable	Data viewing via GoogleEarth	Virtual Louisiana (State-government access only) http://www.virtualla.la.gov/site/ Louisiana Earth http://gohsep.la.gov/laearth.aspx	7667 Independence Blvd. Baton Rouge, LA Phone (225) 925-7500 Email: gohsep.webadmin@la.gov
LA Oil Spill Coordinator's Office (LOSCO)	A repository for the Louisiana Oil Spill Coordinator's Office's Environmental Baseline Inventory (EBI) data holdings.	See organization	GIS data layers: • Shoreline elevation • Wildlife • Environmental sensitivity indexes for coastal habitats • Public land surveys • Oil and gas platforms • Pipelines • Federal mineral lease blocks • Flood zones • Bear sightings • Bathymetry • Historical cyclone tracks • Aquifers • Airports/airfields	Coastwide	Variable	Variable	Direct download	http://lagic.lsu.edu/loscoweb/	Email: jkent4@lsu.edu

Organization	Description	Responsible Party	Data Types	Spatial Extent	Temporal Collection	Update Frequency	Data Access	Website	Contact / Reference
LA Geographic Information Center	A critical resource and contributing member of the Louisiana geospatial community, providing outreach, data and technical services, and more to local, state, and private entities in Louisiana.	See organization	Louisiana map portal • 2007 LOSCO GIS DVD • USGS/LAGIC Joint Coastal Infrastructure Project • State agency GIS resources • Local GIS stakeholder LAGIC metadata catalog • Louisiana geospatial data and metadata (similar if not the same data from LOSCO)	Statewide	• Historical (1700s-1900s) maps • 1990-2000s GIS data	Louisiana map portal • Last updated 2011 LAGIC metadata catalog • Last updated 2007	Direct download	http://lagic.lsu.edu/default.asp	E313 Howe-Russell Geoscience Complex, LSU Email: webmaster@lagic.lsu.edu
USGS - National Wetlands Research Center	Although the Wetlands Center conducts a variety of research in coastal Louisiana, this reference specifically applies to the historical and projected land change analyses.	See organization	Land change analysis for coastal Louisiana (maps and reports)	Coastwide	Change analysis dating back to 1932	Variable	Online viewing and download of maps	http://www.nwrc.usgs.gov/special/landloss.htm	nwrcinfo@usgs.gov

APPENDIX B Guide to the 2012 Coastal Master Plan Modeling Data

The 2012 Coastal Master Plan generated a substantial amount of data and information. One of the most data-heavy components was the predictive modeling; in total, the modeling component of the 2012 Master Plan generated 1.33TB of data. The primary models used included: eco-hydrology, wetland morphology, barrier shoreline morphology, vegetation, storm surge and wave, risk assessment and damage, and 14 upper trophic level / ecosystem service suitability indices. All modeling-related data have been archived and are currently stored on the CPRA's FTP site (ftp://MPM:dnrftp.2010@ftp.dnr.state.la.us/Master-Plan-Modeling).

This appendix provides a guide to the file nomenclature as they occur on the FTP site (developed by the Master Plan Data Integration Team), a summary of the modeling data, including a guide to the scenarios, data formats, and layout of the FTP site.

- Scenario – S## (a number of scenarios were considered for each model run)
  - S12 - moderate future conditions
  - S13 - less optimistic future conditions
  - S14 – moderate with high sea level rise future conditions
  - S50 – current conditions (used only for current landscape by the storm surge and damage model teams)
- Group identifier – G## (refers to specific groups of projects; groups of particular interest are identified below)
  - G01 – original Future Without Action (landscape / ecological models)
  - G02-G50 - Restoration Projects (individual project effects by project group)
  - G60 – final Future Without Action (all models)
  - G61 – draft Master Plan (all models)
  - G62 – 2012 Master Plan (all models)
  - G90 – original Future Without Action (storm surge and damage models)
  - G91 - G97 - Protection Projects (individual project effects by project group)
- Variance - V## (used by risk assessment)
  - First # - pump scenario (1, 2, 3)
  - Second # - fragility (0, 1, 2, 3)
- 8 character text descriptor
  - Character 1 indicates from where the file was generated
    - H - hydrology model
    - W - wetland morphology model
    - V - vegetation model
    - U - upper trophic models (also called ecosystem services)
    - S - storm surge model
    - B - barrier island morphology model
    - N – nitrogen model
    - D - data integration (this refers to a file that was modified to serve as input to a model)
  - Characters 2 and 3 indicate the particular eco-hydrology region the file covers
    - CP - Chenier Plains

        - AA - Atchafalaya
          PB - Pontchartrain
        - LA - Louisiana (coast wide)
      - Characters 4, 5, and 6 indicate model parameters (see Table B.1)
      - Characters 7 and 8 indicate a 2 digit year representing the first year in a multi-year data file or the year of a single year data file (see Table B.2)
- Uncertainty analysis realization identifier (R##)
- Free text

Three example model output file names are provided below:

- S12_G29_V00_HAASTG07_R00JustMyFreeText.txt
  - Scenario 12
  - Group ID 29
  - No project variance
  - Hydro file, Atchafalaya Area, Stage type, start year 2007(this is the first year in the multi-year file)
  - Free text: "JustMyFreeText"
- S14_G17_V01_HPBNH405_R01morefreetext.txt
  - Scenario 14
  - Group ID 17
  - Project variance V01
  - Hydro file, Pontchartrain Barataria area, Ammonium Nitrogen, start year 2005
  - Uncertainty Analysis Realization #01 followed by free text = "morefreetext"
- S13_G61_V00_UCPALL07_R02.nc
  - Scenario 13
  - Group ID 61
  - No project variance
  - Upper Trophic file, Chenier Plain area, Alligator HSI data, start year 2007
  - Uncertainty Analysis Realization #02 followed by no free text.

Table B.1. Model parameter naming convention abbreviations (per 2012 Master Plan Data Integration Team).

Model Output Parameter	Abbreviation	File Type	File Size
All Models			
Readme File	AAA	.txt	<1 KB
Eco-Hydrology			
Accretion (from inorganic TSS)	ACC	.txt	~50 KB
Residence Times (aka - water age)	AGE	.txt	~600 KB
Phytoplankton as Chlorophyll-a	ALG	.txt	~600 KB
Detritus	DET	.txt	~600 KB
Dissolved Organic Nitrogen	DON	.txt	~600 KB
Ammonium Nitrogen	NH4	.txt	~600 KB
Nitrate + Nitrite Nitrogen	NO3	.txt	~600 KB
Nitrogen Removal (denitrification)	NRM	.txt	~50 KB
Salinity	SAL	.txt	~600 KB
Box Configuration Shapefile	SHP	6 GIS files	<1 MB

Model Output Parameter	Abbreviation	File Type	File Size
Soluble Phosphorus	SPH	.txt	~600 KB
Stage	STG	.txt	~17 MB
Total Kjeldahl Nitrogen	TKN	.txt	~600 KB
Water Temp	TMP	.txt	~600 KB
Total Phosphorus	TPH	.txt	~600 KB
Tidal Range	TRG	.txt	~600 KB
Sediment Retention (Total Suspended Solids)	TSS	.txt	~600 KB
Wetland Morphology			
Percent Edge	EDG	.img	~1 MB
Elevation	ELV	.dbf .img	~13 MB ~2 MB
Land / Water (500m Grid)	LNW	.dbf	~16 MB
Land / Water (30m Grid)	LNW	.img	~4 MB
Percent Land	PCL	.img	~1 MB
Soil Organic Carbon	SOC	.img	~900 KB
Barrier Shoreline Morphology			
Barrier Elevation	DEM	.asc	~60 MB
Inlet Area Change	IAC	.txt	~5 KB
Vegetation			
All Parameters in one file	VEG	.asc+	15-150 MB
Vegetation Data Conversion			
Bare Ground	BGR	.nc	~100 MB
Wax Myrtle (Morella cerifera)	WXM	.nc	~100 MB
Mangrove (Avicennia germinans)	MNG	.nc	~100 MB
Cut-grass (Zizaniopsis milliacea)	CTG	.nc	~100 MB
Maidencane (Panicum hemitomon)	MDN	.nc	~100 MB
Cattail (Typha domingensis)	CTT	.nc	~100 MB
Sawgrass (Cladium jamasience)	SWG	.nc	~100 MB
Bulltongue (Sagittaria lancifolia)	BTN	.nc	~100 MB
Roseau Cane (Phragmites australis)	RSC	.nc	~100 MB
Bullwhip (Schoenoplectus californicus)	BWP	.nc	~100 MB
Wiregrass (Spartina patens)	WIG	.nc	~100 MB
Paspalum (Paspalum vaginatum)	PSP	.nc	~100 MB
Needlegrass (Juncus roemerianus)	NDG	.nc	~100 MB
Saltgrass (Distichlis spicata)	SLG	.nc	~100 MB
Oystergrass (Spartina alterniflora)	OYG	.nc	~100 MB
Water (pen water without SAV)	WTR	.nc	~100 MB
Swamp Forest (composed of Taxodium distichum and Nyssa aquatica)	SFR	.nc	~100 MB
Shrub-scrub (composed of Iva frutescens and Baccharis halimifolia)	SHS	.nc	~100 MB
Delta Splay (composed of Sagittaria latifolia, Schoenoplectus deltarum and Colocasia esculenta)	DSP	.nc	~100 MB
Thin-mat (composed of Eleocharis baldwinii,	TMT	.nc	~100 MB

Model Output Parameter	Abbreviation	File Type	File Size
Hydrocotyle umbellata, Bidens laevis)			
Brackish Marsh (composed of Spartina patens, Distichlis spicata and Spartina alterniflora)	BMH	.nc	~100 MB
SAV (open water with SAV)	SAV	.nc	~100 MB
Upper Trophic Level			
Alligator	ALL	.nc	~100 MB
Largemouth Bass	BAS	.nc	~100 MB
Brown Shrimp	BSH	.nc	~100 MB
Crawfish	CRY	.nc	~1.2 GB
Gadwall	GAD	.nc	~100 MB
Green-wing Teal	GRE	.nc	~100 MB
Mottled Duck	MOT	.nc	~100 MB
Muskrat	MUS	.nc	~100 MB
Neotropical Migrant Passerine	NEO	.nc	~100 MB
Otter	OTT	.nc	~100 MB
Oyster	OYS	.nc	~100 MB
Roseate Spoonbill	RSP	.nc	~100 MB
Spotted Sea trout	SST	.nc	~100 MB
White Shrimp	WSH	.nc	~100 MB
Ecosystem Services			
Agriculture	AGR	.nc	~21MB
Freshwater Availability	FRW	.nc	~100 MB
Nature Based Tourism	NBT	.nc	~15 MB
Nutrients	NUT	.txt	~1-3MB
Other Characteristic Fauna (Wildlife)	OCF	.nc	~100 MB
Shrimp	SHR	.nc	~100 MB
Surge/Storm Attenuation	SSA	.nc	~15 MB
Waterfowl	WFL	.nc	~100 MB
Nitrogen Uptake			
Vegetation and Benthic	VBM	.txt	~3MB
Storm Surge & Damage/Risk Assessment			
Damage by Census Block	DBB	.txt	~2 MB
Damage by Community	DBC	.txt	~5 KB
Flooding by Census Block	FDB	.txt	~1 MB

Table B.2. Model data output years (as indicated in the output file names on the FTP site).

Calendar Years	Eco-hydrology	Wetland Morphology	Barrier Shoreline Morphology	Vegetation	Upper Trophic and Ecosystem Service / Nitrogen
2011-2015	10	15	5	15	10 / 15
2016-2020	10	20	10	20	10 / 20
2021-2025	10	25	15	25	10 / 25
2026-2030	10	30	20	30	10 / 30
2031-2035	10	35	25	35	10 / 35
2036-2040	35	40	30	40	10 / 40
2041-2045	35	45	35	45	10 / 45
2046-2050	35	50	40	50	10 / 50
2051-2055	35	55	45	55	10 / 55
2056-2060	35	60	50	60	10 / 59

The most recent model outputs, as presented in the 2012 Master Plan are located in the **Alternatives Analysis** folder (see folders that specify 'new precip'). This is located in the main FTP Master Plan Modeling Archive Directory. The nomenclature above can be used to navigate the model outputs. This folder contains the following subdirectories:

- S12
  - S12_G52
  - S12_G60_new_precip
    - S12_G60_Damage
    - Surge_S12G60Y25
    - Surge_S12G60Y50
  - S12_G61_new_precip
    - Surge_S12G61Y50
  - S12_G62_newprecip
    - S12_G62_Damage
    - Surge_S12G62Y25
    - Surge_S12G62Y50
  - S12_Old Precip
    - S12_G60_OLD_PRECIP_MATCHES PROJECT RUNS
    - S12_G61_OLD_PRECIP_MATCHES PROJECT RUNS
  - S12_Uncertainty_Analysis
    - S12_G60_R01_newprecip…
    - S12_G60_R10_newprecip
    - S12_G61_R01_newprecip…
    - S12_G61_R10_newprecip

- S13
  - S13_G60_new_precip
    - S13_G60_Damage
    - Surge_S13G60Y25
    - Surge_S13G60Y50
  - S13_G61_new_precip
    - Surge_S13G61Y50
  - S13_G62_newprecip
    - S13_G62_Damage
    - Surge_S13G62Y25
    - Surge_S13G62Y50
  - S13_Old Precip
    - S13_G61_OLD_PRECIP_MATCHES PROJECT RUNS
    - S13_G61_OLD_PRECIP_MATCHES PROJECT RUNS
- S14
  - S14_G60_new_precip
    - S14_G60_Damage
    - Surge_S14G60Y25
    - Surge_S14G60Y50
  - S14_G61_new_precip
    - Surge_S14G61Y50
  - S14_G62_newprecip
    - S14_G62_Damage
    - Surge_S14G62Y25
    - Surge_S14G62Y50
  - S14_Old Precip
    - S14_G60_OLD PRECIP_MATCHES PROJECT RUNS
    - S14_G61_OLD PRECIP_MATCHES PROJECT RUNS
- Stitched Files

The **Initialization** folder (in the main FTP Master Plan Modeling Archive Directory) contains all files needed to initialize the models. This folder contains the following subdirectories:

- **Barrier_Morphology**
  - Initialization_Data
    - Bathy_2006_2007
      - 2007
      - LandElevationSampleData
    - Final Model Code and Inputs
- **Damage_Risk**
  - Census_2010
  - Input_points
  - Spatial_units
- **Data_Integration**
  - barrierIslandStitcher_1.0.1
    - barrierIslandStitcher_1.0.1
      - configuration

- org.eclipse.core.runtime
- org.eclipse.equinox.app
- org.eclipse.orgi
    - bundles
        - 21

- MSL
- WM_Widgets

- **Eco-Hydrology**
    - InitializationFiles
        - AA
        - CP
        - PB
- **Ecosystem_Services**
- **Storm_Surge**
    - S12_ProjectAnalysis_HydraulicSubunits
    - S13_ProjectAnalysis_HydraulicSubunits
        - archive
    - Stitched Files
        - G62_YR50_LNW_ASCII
- **Upper Trophic Level**
    - data converter tracking forms
        - GROUP
            - Oyster Tracker
    - Oyster
    - Stitched
        - Alternatives
            - G60
                - S12_new precip
                - S13_new precip
                - S14_new precip
            - G61
                - S12_new precip
                - S13_new precip
                - S14_new precip
            - UA
                - G60
                    - G60_R01…
                    - G60_R10
                - G61
                    - G61_R01…
                    - G61_R10
        - Projects
            - S12 and S13
                - S12
                    - S12_G01…
                    - S12_G50

- S13
  - S13_G01…
  - S13_G50

- **Vegetation**
  - CLEARGrid
  - ExampleConfigurationFilesForCoreModel
  - ExampleConfigurationFilesForRWrapper
  - LAShapefiles
    - lacountiesp
  - LULC
  - VegModel_Grid and Init Cond
  - VegTypeParameters
- **Wetland_Morphology**
  - InitializationData
    - Alternatives
      - Yr35
    - BathyTopo
    - Boundaries
    - Grids
    - Projects
    - Toolboxes
      - BU
        - USGSTools01_05_2011
          - AssignLossGain
          - FWOP
            - AssignLossGain
            - LandAreaPerZone
          - LandAreaPerZone
          - SpatialIntegrityIndiex
          - Utilities
          - WetlandMorphology
          - WVA
        - USGSTools01_08_2011
          - AssignLossGain
          - FWOP
            - AssignLossGain
            - LandAreaPerZone
          - LandAreaPerZone
          - SpatialIntegrityIndiex
          - Utilities
          - WetlandMorphology
          - WVA
        - USGSTools04_19_2011
          - AssignLossGain
          - FWOP
            - AssignLossGain
            - LandAreaPerZone
          - LandAreaPerZone

    - SpatialIntegrityIndiex
    - Utilities
    - WetlandMorphology
    - WVA
- USGSTools04_20_2011
    - AssignLossGain
    - FWOP
        - AssignLossGain
        - LandAreaPerZone
    - LandAreaPerZone
    - SpatialIntegrityIndiex
    - Utilities
    - WetlandMorphology
    - WVA
- USGSTools04_25_2011
    - AssignLossGain
    - FWOP
        - AssignLossGain
        - LandAreaPerZone
        - SLR
    - LandAreaPerZone
    - SpatialIntegrityIndiex
    - Utilities
    - WetlandMorphology
    - WVA
- USGSTools05_16_2011
    - AssignLossGain
    - FWOP
        - AssignLossGain
        - LandAreaPerZone
        - RestorationMeasures
        - SLR
    - LandAreaPerZone
    - SpatialIntegrityIndiex
    - Utilities
    - WetlandMorphology
    - WVA
- USGSTools05_19_2011
    - AssignLossGain
    - FWOP
        - AssignLossGain
        - LandAreaPerZone
        - RestorationMeasures
        - SLR
    - LandAreaPerZone
    - SpatialIntegrityIndiex
    - Utilities
    - WetlandMorphology

    - WVA
- USGSTools05_20_2011
    - AssignLossGain
    - FWOP
        - AssignLossGain
        - LandAreaPerZone
        - RestorationMeasures
        - SLR
    - LandAreaPerZone
    - SpatialIntegrityIndiex
    - Utilities
    - WetlandMorphology
    - WVA
- USGSTools05_23_2011
    - AssignLossGain
    - FWOP
        - AssignLossGain
        - LandAreaPerZone
        - OutputPrep
        - RestorationMeasures
        - SLR
    - LandAreaPerZone
    - SpatialIntegrityIndiex
    - Utilities
    - WetlandMorphology
    - WVA
- USGSTools05_23b_2011
    - AssignLossGain
    - FWOP
        - AssignLossGain
        - LandAreaPerZone
        - OutputPrep
        - RestorationMeasures
        - SLR
    - LandAreaPerZone
    - SpatialIntegrityIndiex
    - Utilities
    - WetlandMorphology
    - WVA
- USGSTools05_24_2011_planB
    - FWOP
        - AssignLossGain
        - LandAreaPerZone
        - OutputPrep
        - RestorationMeasures
        - SLR

- USGSToolsOCPR

- Utilities
    - Agg500m
        - Grid500wvat
        - info

- USGSToolsOCPR2
    - Utilities
        - Agg500m
            - Grid500wvat
            - info

The folder named **Model Group Maps** (in the main FTP Master Plan Modeling Archive Directory) contains the PDF files showing "change in percent land" from individual project effects when compared to Future Without Action. This folder contains outputs for S12 and S13, for all project-level analyses (G01 – G50). Files contained in this folder also show the location of the projects modeled in that group. Files contained in the subdirectory do not include project demarcations. The subdirectory is as follows:

- **20111215_PCL**

The folder named **Nitrogen** (in the main FTP Master Plan Modeling Archive Directory) contains all files related to the nitrogen model, including the project level and alternative level outputs. Nitrogen uptake model runs were completed after many of the other model runs were completed, and therefore it was more streamlined to save the files altogether, instead of by scenario, and then by group, like the other model outputs are currently archived. This folder contains the following subdirectories:

- **Application Files**
    - OCPR_1_0_0_1
- **Excel Summary Files**
    - FEB 2012
    - JAN 2012
    - S12 RUN Oct 21 2011
    - S13 RUN Oct 21 2011
- **S12**
    - G01…
    - G60
    - G60_new_precip
        - aa_y10
        - aa_y59
        - asciito_s12_1
        - asciito_s12_3
        - cp_y10
        - cp_y102
        - cp_y59
        - cp_y592
        - info
        - pb_y10
        - pb_y102

        - pb_y59
        - s12_g60_y10
        - s12_g60_y59
    - G61
    - G61_new_precip
        - aa_y10
        - aa_y102
        - aa_y59
        - aa_y592
        - asciito_s12_1
        - asciito_s12_2
        - cp_y10
        - cp_y59
        - cp_y593
        - info
        - pb_y10
        - pb_y102
        - pb_y59
        - pb_y592
        - s12_g61_y10
        - s12_g61_y59
    - G62_newprecip
        - aa_y10
        - aa_y102
        - aa_y59
        - aa_y592
        - cp_59
        - cp_y10
        - cp_y102
        - cp_y592
        - info
        - la_y10
        - la_y59
        - la_y592
        - pb_y10
        - pb_y102
        - pb_y59
        - pb_y592
        - s12_g62_y10
        - s12_g62_y59
- **S13**
    - G01…
    - G60
    - G60_new_precip
        - aa_y10
        - aa_y102
        - aa_y59
        - aa_y592

        - cp_y10
        - cp_y102
        - cp_y59
        - cp_y592
        - info
        - pb_y10
        - pb_y59
        - s13_g60_y10
        - s13_g60_y59
    - G61
    - G61_new_precip
        - aa_y10
        - aa_y102
        - aa_y59
        - aa_y592
        - aa_y593
        - cp_y10
        - cp_y102
        - cp_y59
        - cp_y592
        - info
        - pb_y10
        - pb_y59
        - pb_y592
        - s13_g61_y10
        - s13_g61_y59
        - s13_g61_y592
    - G62_newprecip
        - aa_y10
        - aa_y102
        - aa_y59
        - aa_y592
        - cp_y10
        - cp_y102
        - cp_y59
        - cp_y592
        - info
        - la_y10
        - la_y59
        - pb_y10
        - pb_y59
        - s13_g62_y10
        - s13_g62_y59
    - info
- **S14**
    - G60
    - G60_new_precip
        - aa_y10

        - aa_y59
        - cp_y10
        - cp_y59
        - info
        - pb_y10
        - pb_y59
        - s14_g60_y10
        - s14_g60_y59
    - G61
    - G61_new_precip
        - aa_y10
        - aa_y59
        - cp_y10
        - cp_y59
        - info
        - pb_y10
        - pb_y59
        - s14_g61_y10
        - s14_g61_y59
    - G62_newprecip
        - aa_y10
        - aa_y102
        - aa_y59
        - aa_y592
        - cp_y10
        - cp_y102
        - cp_y59
        - cp_y592
        - info
        - la_y10
        - la_y59
        - pb_y10
        - pb_y59
        - s14_g62_y10
        - s14_g62_y59

The folder named **Post Processing** contains computer codes that were used to summarize model output into formats usable by the Planning Tool. It also contains the spatial buffers that were applied to each individual project to demarcate project influence areas. Output from post-processing vegetation model output is also contained here. The Post Processing folder contains the following subdirectories:

- **Code**
    - Agriculture_Flood_Area_SI
        - Agriculture_Flood_Area_SI
            - Debug
            - Release
    - Cookie_Cut_for_NetCDFXY

        - Cookie_Cut_for_netCDFXY
            - Debug
            - Release
    - ES_Sum_for_netCDFXY
        - ES_Sum_for_for_netCDFXY
            - Debug
            - Release
    - ES_SUM_FWOA_Case
        - ES_SUM_FWOA_Case
            - Debug
            - Release
    - netCDFProcessor_AllCases
        - netCDFProcessor_AllCases
            - Debug
            - Release
    - Nutrient_Processor
        - Nutrient_Processor
            - Debug
            - Release
    - PreProcessor_UTL_SUM
        - PreProcessor_UTL_SUM
            - Debug
            - Release
    - Preprocessor_UTL_SUM_G01_ONLY
        - Preprocessor_UTL_SUM_G01_ONLY
            - Debug
            - Release
- **Spatial**
    - 500M_Model_Grid
    - Project_Buffers
- **Vegetation**
    - Alternative_Analysis
    - S12
    - S13
    - S14

The folders named **Storm Surge** and **Risk Assessment** (in the main FTP Master Plan Modeling Archive Directory) contain files other than the outputs already stored in the Alternatives Analysis, and S12 and S13 folders. These folders contain additional output files, beyond the typical outputs, current conditions files, and extra material that were generated for CPRA review. These folders each contain the following subdirectories:

- **Storm Surge**
    - CURRENT_CONDITIONS
        - LACPR_results
    - POST_PROCESSING
        - MEOW

        - S12
        - S13
    - PROJECT_SCALE_ANALYSIS
    - REVIEW
        - Mesh
            - 2011CC_MeshImages
            - 2011CC_raised_features
            - 2011CC_weir_images
- **Risk Assessment**
    - Flood Model Inputs
    - Maps
    - PostgreSQL DB
    - PostProcessing Inputs
    - PreProcessing Inputs
    - PT Tables
    - S50

Initial analyses of groupings of independent projects (to assess individual project effects) are found in the folders named **S12** and **S13** (in the main Master Plan Modeling Archive directory). The sub-folders located with the S12 folder (e.g., S12_G01_R00) include the uncertainty analysis results for the project-level analysis. These folders contain the following subdirectories:

- **S12**
    - S12_G01
    - S12_G01_R00…
    - S12_G01_R08
    - S12_G02…
    - S12_G51
    - S12_G90…
    - S12_G97
- **S13**
    - S13_G01…
    - S13_G51
    - S13_G90…
    - S13_G97

The folder named **Uncertainty Analysis** (in the main FTP Master Plan Modeling Archive Directory) contains outcomes of the uncertainty analysis that was performed on the 'alternative' level (G60 and G61) model outputs. Due to time constraints, an uncertainty analysis was not performed on G62. This folder contains the following subdirectories.

- **UA**
    - G60
        - S12g60y25r01…
        - S12g60y25r10
    - G61