National Park Service
U.S. Department of the Interior

Natural Resource Program Center

Inventory and Monitoring Program

Data Management Plan for the Chihuahuan Desert Network

Natural Resource Report NPS/CHDN/NRR—2009/096

ON THE COVER
White Sands National Monument
NPS Photo

Chihuahuan Desert Network

Data Management Plan for the Chihuahuan Desert Network

Natural Resource Report NPS/CHDN/NRR—2009/096

James T. Richie
National Park Service
New Mexico State University
MSC PSL/Anderson Hall
Las Cruces, NM 88003

March 2009

U.S. Department of the Interior
National Park Service
Natural Resource Program Center
Fort Collins, Colorado

The Natural Resource Publication series addresses natural resource topics that are of interest and applicability to a broad readership in the National Park Service and to others in the management of natural resources, including the scientific community, the public, and the NPS conservation and environmental constituencies. Manuscripts are peer-reviewed to ensure that the information is scientifically credible, technically accurate, appropriately written for the intended audience, and is designed and published in a professional manner.

Natural Resource Reports are the designated medium for disseminating high priority, current natural resource management information with managerial application. The series targets a general, diverse audience, and may contain NPS policy considerations or address sensitive issues of management applicability. Examples of the diverse array of reports published in this series include vital signs monitoring plans; monitoring protocols; "how to" resource management papers; proceedings of resource management workshops or conferences; annual reports of resource programs or divisions of the Natural Resource Program Center; resource action plans; fact sheets; and regularly-published newsletters.

Printed copies of reports in these series may be produced in a limited quantity and they are only available as long as the supply lasts. This report is also available from the Natural Resource Publications Management Web site (http://www.nature.nps.gov/publications/NRPM) on the Internet.

Please cite this publication as:

Richie, J. T. 2009. Data management plan for the Chihuahuan Desert Network. Natural Resource Report NPS/CHDN/NRR—2009/096. National Park Service, Fort Collins, Colorado.

NPS D-102, March 2009

Change History

Version numbers will be incremented by a whole number (e.g., Version 1.3 to Version 2.0) when a change is made that significantly affects requirements or procedures. Version numbers will be incremented by decimals (e.g., Version 1.6 to Version 1.7) when there are minor modifications that do not affect requirements or procedures included in the plan.

Version #	Date	Revised by	Changes	Justification

Contents

Figures

Tables

Acknowledgements

This document represents a culmination of effort across the Inventory and Monitoring Program and could not have been written without their inspiration and insight. The community of data managers in the I&M Program and parks within the 32 I&M Networks fosters a collaborative environment with the free sharing of information. Credit is due to all who have preceded CHDN in the completion of network data management plans, and for making these plans available to the rest of us.

Several data management plans were used extensively in the development of this plan and associated standard operating procedures and we would like to send our appreciation to the authors of those plans: Southern Plains Network, Sierra Nevada Network, South Florida/Caribbean Network, Rocky Mountain Network, Northern Colorado Plateau Network, Southeast Alaska Network, and Southwest Alaska Network.

In addition, the national-level data management staff have provided the vision and created the tools that have been essential to our work. In particular, thanks go to Margaret Beer, Chris Dietrich, Kathy Dratch, Willene Hendon, Simon Kingston, Alison Loar, and Wendy Schumacher.

We would also like to thank the members of the CHDN Data Management team for their support of, and contributions to, this plan. They include Murugesh Annadurai, Jenny Good, Chris Maher, Missy Powell, Scott Schrader, and Ariel Simpson. Hildy Reiser also provided valuable edits.

Ann Lewis (Physical Science Laboratory, New Mexico State University), provided diligent editing and production of the CHDN Data Management Plan. We are extremely grateful for her efforts.

Acronyms used in this document

ANCS+	Automated National Catalog System
ArcCatalog	ESRI metadata creation tool
ArcGIS	software package
ArcInfo	software program
ASCII	American Standard Code for Information Interchange
CHDN	Chihuahuan Desert Network
CSDGM	Content Standard for Digital Geospatial Metadata
CSDGM	Content Standard for Digital Geospatial Metadata
DBMS	database management system
DLG	digital line graph
DMP	Data Management Plan
DOI	Department of Interior
EPA	Environmental Protection Agency
ESRI	Environmental Systems Research Institute
EXIF	Exchangeable Image File Format
FGDC	Federal Geographic Data Committee
FOIA	Freedom of Information Act
FRC	federal records center
GEOTIFF	Georeferencing Tagged Image File Format
GIS	geographic information system
GPS	global positioning system
GUID	globally unique identifiers
I&M	Inventory and Monitoring
INP	Interior National Park
IPTC	International Press Telecommunications Council
IT	information technology
LAN	local-area network
MRSID	multiresolution seamless image database
MS	Member Server

MS	Microsoft
NAD	North American Datum
NDMP	National Data Management Plan
NIST	National Institute of Standards and Technology
NPS	National Park Service
NPSTORET	National Park STORage and RETrieval (Environmental Protection Agency database)
NRDT	Natural Resources Database Template
NRPC	Natural Resource Program Center
OU	organizational units
PDA	personal data assistants
PDF	portable document format
PEPC	Planning, Environment and Public Comment
PMIS	Project Management Information System
QA/QC	quality assurance/quality control
QMS	quality management system
RAMS	Resource Activity Management System
RPRS	Research Permit and Reporting System
SDLC	System Development Life Cycle
SECN	Southeast Coast Network
SOP	standard operating procedures
SOPN	Southern Plains Network
STORET	STORage and RETrieval (Environmental Protection Agency database)
TI	task instructions
TIFF	Tagged image file format
UTM	Universal Transverse Mercator
VPN	virtual private network
WAN	wide-area network
WASO	Washington Office
XML	extensible markup language

Executive Summary

The central mission of the National Park Service (NPS) Inventory and Monitoring (I&M) Program is to provide timely and usable scientific information about the status and trends of park resources to park managers, planners, cooperators, researchers, and the general public. To meet this challenge, we need an information management system that can effectively produce, maintain, and distribute the products of scientific investigation conducted in our parks.

Good data management is the means by which a thorough understanding of the value of scientific information about our natural resources can become a part of our National Park Service heritage. Data management refers to the framework in which data are acquired, maintained, and made available. Data management is not an end unto itself, but a means of maximizing the quality and utility of our natural resource information. A robust data management system is particularly important for long-term programs when the lifespan of a data set will span the careers of several scientists. Viewed in this way, it becomes obvious that data management is vital to the success of any long-term research program.

The purpose of the Chihuahuan Desert Network Data Management Plan (CHDN DMP) is to provide I&M and other park staff with a conceptual framework for a system that will ensure the production and dissemination of timely and usable scientific information about the status and trends of park resources to park managers. Our strategy for achieving this goal can be summarized as follows: Ensure the quality, interpretability, security, longevity, and availability of our natural resource data. Our objectives include:

- Use of a sound infrastructure that offers easy access to most information and appropriate safeguards for sensitive information
- Implementation of a management strategy that ensures all tasks are completed, responsibilities are clear, resources are available, and overall workflows are synchronized
- A collaborative effort of all parties involved
- Creation and use of databases that are developed as well-functioning vehicles to provide security, flexibility, and ease of use
- Making available standards for acquiring, processing, and reporting data that inspire confidence; provide comprehensive and detailed instruction; ensure quality assurance/quality control; and lead to high quality, vetted results
- Providing metadata to ensure the longevity and usability of all data
- Preparation and use of a project tracking system that provides usability, longevity, and security.

A National Data Management Plan (NDMP) was developed by the I&M program to provide the overarching policies relevant to network data management, to provide national guidance that establishes objectives common to all networks, and to ensure the highest standards for data acquisition and management by all networks (NPS 2008). Much of the structure and content of the CHDN DMP was derived from the NDMP. Many of the topics covered in the NDMP are referenced, but not repeated, in this network plan; therefore, the NDMP should be readily available when reviewing this document.

Specifically, this plan describes how the network will support NPS I&M Program objectives:

- Acquire and process data
- Assure data quality
- Document, analyze, summarize, and disseminate data and information
- Maintain nationally developed data management systems
- Maintain, store, and archive natural resources data and information

Chapter 1. Introduction

In August 25, 1916, President Woodrow Wilson signed the Organic Act that established the mission of the National Park Service. The Act states:

> *… the Service … shall promote and regulate the use of the Federal areas known as national parks, monuments and reservations … by such means and measures as conform to the fundamental purpose of the said parks, monuments and reservations, which purpose is to conserve the scenery and the natural and historic objects and the wild life therein and to provide for the enjoyment of the same in such manner and by such means as will leave them unimpaired for the enjoyment of future generations. (NPS 1916).*

For more than 80 years, park management policies alone were responsible for developing conservation measures to meet this goal. Then in 1998, the National Parks Omnibus Management Act mandated that "The Secretary shall undertake a program of inventory and monitoring of National Park System resources to establish baseline information and to provide information on the long-term trends in the condition of National Park System resources." (NPS 1998). As part of the Service's effort to "improve park management through greater reliance on scientific knowledge," the Natural Resources Challenge Inventory and Monitoring (I&M) Program was established.

The primary purposes of the Chihuahuan Desert Network (CHDN) I&M program are to acquire, develop, organize, archive, and make available relevant natural resource information and data. The CHDN is expected to invest at least one-third of its available resources in data management, analysis, and reporting activities. The I&M program's ability to achieve these tasks will largely determine the program's efficacy and image among critics, peers, and advocates.

1.1 The Chihuahuan Desert Network

The CHDN includes seven national park units, all within the Chihuahuan Desert, one of the most biologically diverse deserts in the Western hemisphere and one of the most diverse arid regions in the world. The park units in the Chihuahuan Desert Network include:

- Amistad National Recreation Area
- Big Bend National Park
- Carlsbad Caverns National Park
- Fort Davis National Historic Site
- Guadalupe Mountains National Park
- Rio Grande Wild & Scenic River
- White Sands National Monument

Collectively, these parks contain 464,544 ha (1,145,444 acres) with elevations ranging from 232 m (760 ft) to the highest peak in Texas at 2,667 m (8,750 ft) (Figure 1.1).

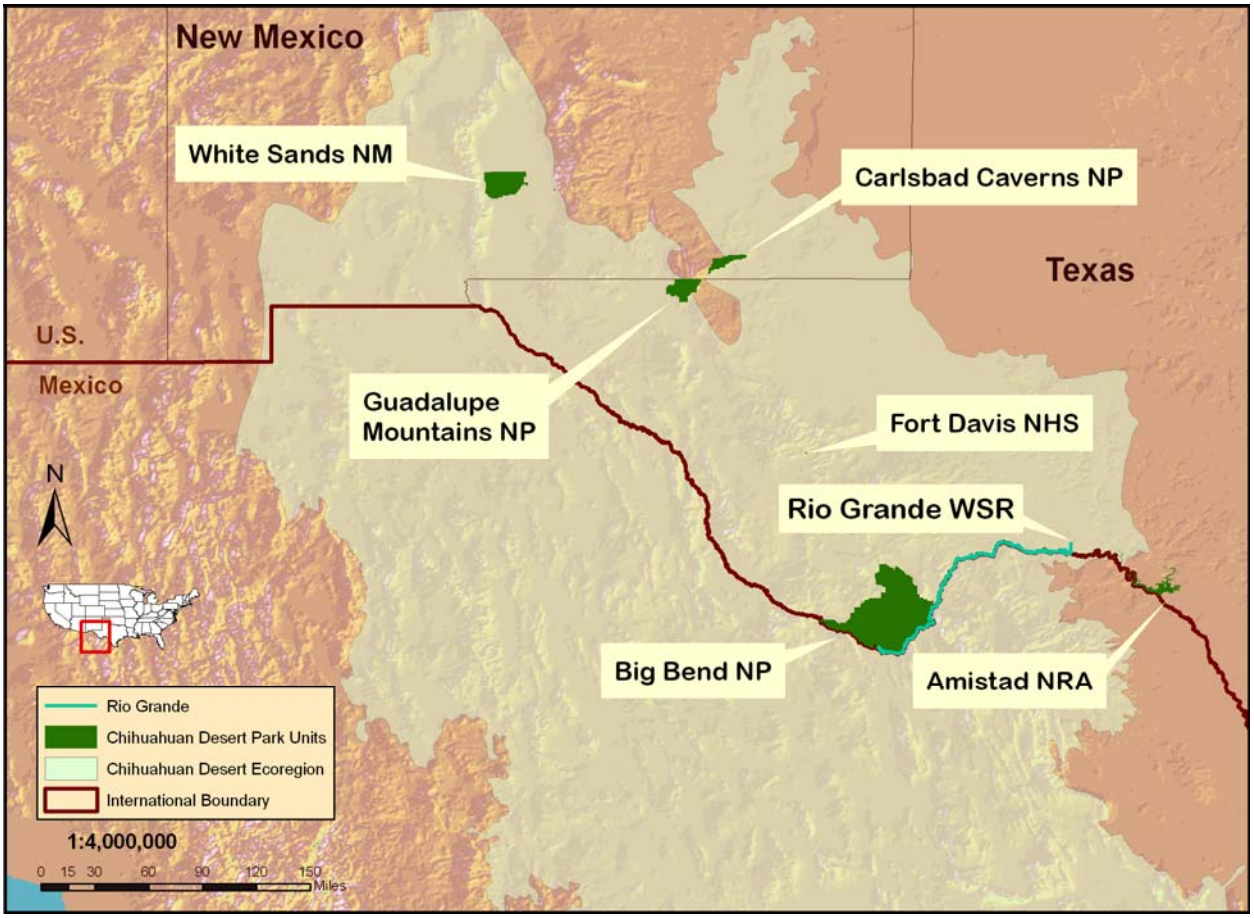

Figure 1.1. Location of CHDN park units.

1.2 Purpose and Scope of This Plan

The goal of the CHDN data management program is to maintain, in perpetuity, the ecological data and related analyses that result from the CHDN resource inventory and monitoring work.

This plan establishes the general concepts and procedures the CHDN, its cooperators, partners, and, potentially, individual park units will use to ensure the quality, interpretability, security, longevity, and availability of program data records and related information. Projects initiated by CHDN and the personnel involved with those projects must follow the guidance provided in this plan. This document describes the policies relevant to CHDN data management, establishes objectives for the network, and provides guidance that will ensure the highest standards for data acquired and managed by the network. These standards include:

- *Accuracy*: The quality of the data collected and managed by the CHDN is paramount. Analyses performed to detect ecological trends or patterns require data with minimal error and bias. Inconsistent or poor-quality data can limit the detectability of subtle changes in ecosystem patterns and processes, lead to incorrect interpretations and conclusions, and could greatly compromise the credibility and success of the program. To ensure that CHDN produces and maintains data of the highest possible quality,

procedures are established to identify and minimize errors at each stage of the data life cycle.

- *Security*: Digital and hard-copy data are maintained in environments that protect against loss, either due to electronic failure or to poor storage conditions. CHDN has in place proper storage and backup procedures and a disaster recovery plan, as well as an established records-management process. In addition, CHDN collaborates with the NPS Museum Management Program, enlisting the expertise of museum curators and archivists to ensure that related project materials such as field notes, data forms, specimens, photographs, and reports are properly cataloged, stored, and managed in archival conditions.
- *Longevity*: Countless data sets have become unusable over time either because the format is outdated, or because metadata is insufficient to determine the data's collection methods, scope and intent, quality assurance procedures, or format. Whereas proper storage conditions, backups, and migration of data sets to current platforms and software standards are basic components of data longevity, comprehensive data documentation is equally important. CHDN ensures that data sets are consistently documented and in formats that conform to current federal standards.
- *Usability*: One of the most important responsibilities of the I&M program is to ensure that data collected, developed, or assembled by staff and cooperators are made available for decision-making, research, and education. Providing well-documented data in appropriate formats and in a timely manner to park managers is especially important to the success of the program. CHDN ensures that:
    - data can be easily found and obtained
    - data are subjected to full quality control before release
    - data are accompanied by complete metadata
    - data are provided in formats that are most useful to end users
    - sensitive data are identified and protected from unauthorized access and distribution

1.3 Priorities

Considering the volume of data that has been produced in park units, priorities must be set for CHDN data management efforts. Generally, these are to:

1. Produce and curate high-quality, well-documented data originating with the Inventory and Monitoring Program
2. Assist with data management for current projects, legacy data, and data originating outside the Inventory and Monitoring Program that complement program objectives
3. Help ensure good data management practices for park-based natural resource-related projects.

1.4 Integration with the National Level Data Management Plan

This document is not intended to duplicate national I&M program guidelines, Director's Orders, or legislative requirements. The *Data Management Guidelines for Inventory and Monitoring Networks* (NPS 2008), also referred to as the National Data Management Plan (NDMP), provides broad guidance in terms of the National I&M program, NPS Directives, and other legislative

requirements. This network-level data management plan acts as a supplement to the NDMP and functions as a guide for establishing and maintaining a system that serves the data management needs of the CHDN.

1.5 Quality Management System

W. Edward Deming, known as the father of quality management, recognized the direct relation between productivity and quality, but also realized improving quality involved people and procedures and not just the tools being used (Deming 2000). The detailed components of the CHDN data management program are standard operating procedures (SOPs) and task instructions (TIs) that ensure successful completion of specific tasks (e.g., back-up procedures, data documentation, report and data tracking, GPS (global positioning system device) use). This system of SOPs and TIs is at the core of the CHDN quality management system (QMS).

Some instructions are developed before any monitoring projects are initiated, others are developed as part of the monitoring protocol development process, and some may be developed on an as-needed basis to provide direction to staff and cooperators. Figure 1.2 shows the relationship of the national-level data management plan to network-level data management documentation. Standard operating procedures and task instructions are maintained on the CHDN file server, in hard copy in the CHDN library, and are made available on the CHDN internet Web site (http://science.nature.nps.gov/im/units/chdn/index.cfm).

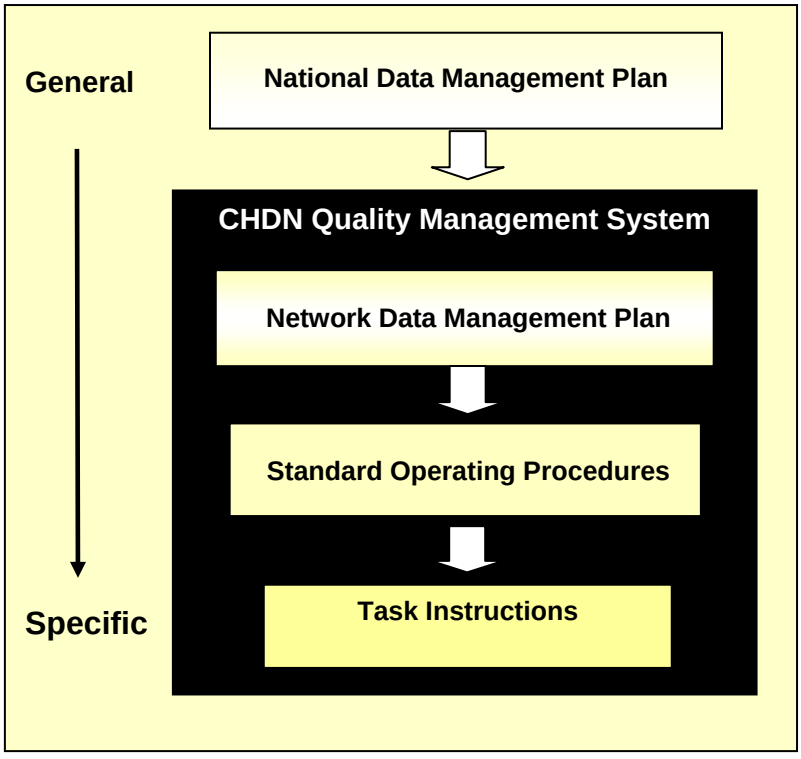

Figure 1.2. Relationship of the national-level data management plan to CHDN data management documents. Adapted from Chapter 1.2 of the NDMP (NPS 2008).

The primary SOP for the CHDN QMS is "Development, Review, Approval, Distribution, and Control of CHDN Standard Operating Procedures and Standard Instructions" (CHDN_SOP-01-0001). This procedure describes the process by which an SOP or TI is established. This tiered system of procedures and instructions allows specific process instruction development without having to repeat the general procedure in every instruction. Additionally, if general procedures are modified, the instructions governed by those procedures may not have to be modified.

> *Example of SOP and TI: The SOP for a pre-field-use inspection of GPS devices could include a general physical condition check, battery and/or power supply check, and accessories check. The SOP would insure that all necessary components and accessories are available and in working condition for field use. The SOP would be generic enough to cover all field GPS units and not any specific make or model. A TI that is specific to a make and model of GPS unit would supplement this SOP. The TI provides the data collection objective, systematic instructions for ensuring the specific GPS unit is configured correctly, the actual data collection process, and the post collection processing and documentation of data. In many cases, in a TI such as this, sections of the operator's manual would be included in the instructions and modified to meet the requirements of a specific data collection process.*

1.6 Plan Revisions

Because this document is based on the NDMP, any significant update to the national plan should trigger a revision of the CHDN plan. Future revised data management plans shall each be issued as a complete unit, not as piecemeal sections. Because CHDN data management has not yet been fully implemented, defects in the plan may not yet have become visible. Deficiencies are expected to surface as network development progresses. Also, rapid evolution of technology regularly provides better ways to accomplish tasks. For these reasons, periodic plan reviews and updates are recommended. A plan that has been static for three years is likely obsolete.

As new processes are honed, their workings will be documented in new and revised standard operating procedures (SOP). The Data Manager will add SOPs as needed. The most up-to-date version of any SOP or task instruction (TI) is posted on the CHDN Internet/Intranet.

1.7 Related Guidance Documents, SOPs and/or TIs (see latest version on CHDN website)

- CHDN_SOP-01-0001; Development, Review, Approval, Distribution and Control of CHDN Standard Operating Procedures and Standard Instructions

Chapter 2. Infrastructure and Systems Architecture

Infrastructure refers to the framework on which a system is built. The Information Technology (IT) infrastructure refers to all the components that make up a system. Processes, such as file storage, printing, and e-mail are independent processes, but depend on a layered infrastructure to function. The layers in a typical IT infrastructure consist of the wires that connect components together, the network hardware that routes communications between components, the workstations, server(s), and printers. System architecture refers to the applications, database systems, repositories, and software tools that make up the framework of an information management system.

2.1 CHDN Program Infrastructure

2.1.1 NPS.DOI.NET

A domain is a group of connected computers that share user account information and a security policy. For administrative and organizational reasons, domains are divided into sub-levels or sub domains in a hierarchal structure. Each level of the hierarchy is related directly to the level above it and below it, if it exists. DOI.NET is the parent domain for the Department of the Interior (DOI) and NPS.DOI.NET is the sub domain for the National Park Service. A domain also can be broken down into organizational units (OU). Organizational units are used as a method of organizing users, groups, and computers and are the smallest unit to which you can assign group policy settings or delegate administrative authority.

The CHDN staff administer its own OU; however, because security policy also follows the hierarchal structure, an administrator at the NPS.DOI.NET level can assist in the CHDN OU administration. The Intermountain Region computer support helpdesk located in Denver, Colorado, is the network's first-line support staff.

2.1.2 Wide Area Network

The CHDN office is located at New Mexico State University, Las Cruces, New Mexico, and not co-located at a park unit. CHDN is not associated with other park units with respect to its IT infrastructure or systems administration. The CHDN infrastructure is both simple and complex: it is a somewhat isolated system, but still part of a much larger network, the NPS Wide Area Network or WAN. The linking together of two or more Local Area Networks or LANs is referred to as a WAN. The CHDN LAN is connected to the larger NPS WAN through a Virtual Private Network (VPN). The VPN allows for secure communications through the public Internet, and provides a remote LAN the same functionality as if both networks were in the same location. One advantage of the VPN is central administration of user account information and security policy. Figure 2.1 illustrates the CHDN's network connectivity.

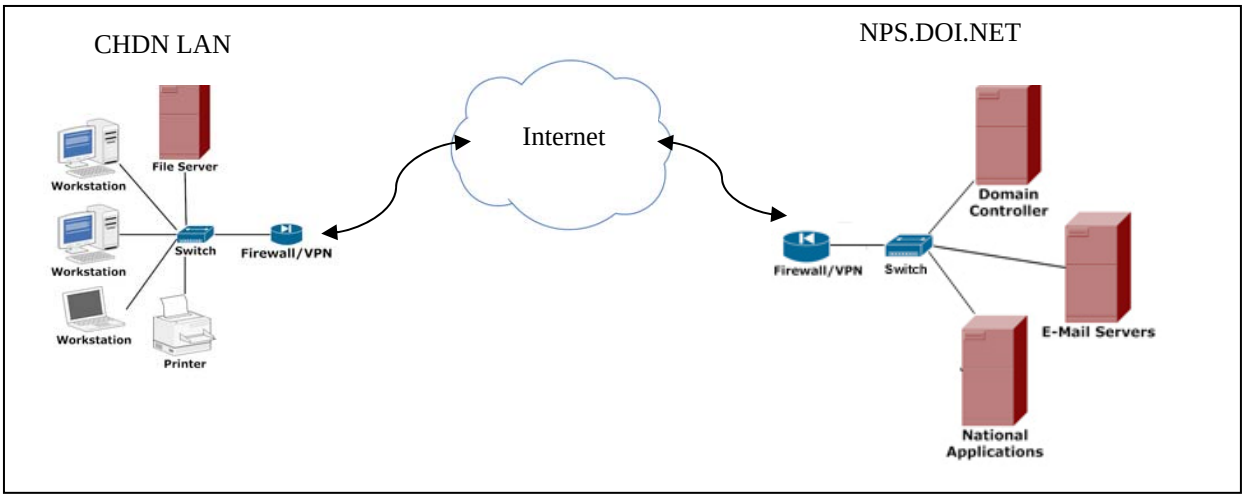

Figure 2.1. CHDN and the wide area network.

2.1.3 CHDN.NPS.DOI.NET

The CHDN.NPS.DOI.NET is the complete or fully qualified domain name of the CHDN local area network. It consists of a single physical server located at the network office, a network-accessible printer, a network-accessible plotter, and several workstations. The physical network is a simple star configuration. Each network jack in the CHDN office connects to a single switch that routes computer communications within the office and is located in a central wiring closet.

All network devices, workstations, server(s), and printers attach to this switch, which is the center of the star topography. Combined, the workstations, server(s), printers, and switches comprise the CHDN LAN. Workstation and servers on the LAN follow the DOI naming structure, which begins with INP for Interior National Park, the park unit acronym CHDN, and the last five digits of the property tag for workstations. CHDN servers are member servers and indentified by MS (Member Server) and a sequential number (e.g., INPCHDNMS01).

2.1.4 Share Drives

The CHDN server, named INPCHDNMS01, acts as the primary file storage location for the network. Users access files on the NPS network through three primary shares mapped logically to drive letters on a user's workstation during the login process. They are the application share, the user's share, and the public share. The Web share is not logically mapped and is restricted to Web managers. Finally, the print server is accessible locally within the network.

2.1.4.1 Application Share Drive: The application share, *\\INPCHDNMS01\apps*, logically mapped as the N: drive, is intended to store applications installed either on workstations or as the location for program files for a network installation of an application. Because most of the CHDN workstations are laptops, network installation of applications is rarely used. The application share also stores updates or service packs for installed applications. When a service pack is cumulative, previous service packs are purged. The applications stored in the application share are not original media and are stored only as a convenience. Access to the application share is restricted, and software activation codes are not stored in this location.

2.1.4.2 Users Share Drive: The Users share, *\\INPCHDNMS01\Shared\[username]*, is logically mapped as the U: drive. This space is intended for the users' personal work/storage space. Access to a user's space is restricted to the individual user. It is recommended that users back up workstation files, e-mail archives, and browser bookmarks to their user share to ensure safe recovery in case of workstation failure or computer virus infection. Users are responsible for the organization and structure of their share space.

2.1.4.3 Public Share Drive: The public share, *\\INPCHDNMS01\Public*, logically mapped as the P: drive, is a common storage area intended to store working data accessible to all CHDN staff. Because the public share is a common working area, a well documented directory structure or filling system provides ease of filing and efficient access to all users.

Records are grouped together if they relate to a particular subject or function, document a specific kind of transaction, take a particular physical form, or have some other relationship arising out of their creation, receipt, or use (such as restrictions on access). These groupings are called a record series. Directory structure naming reflects the record series and aids in locating records of interest, but naming alone is insufficient. An understanding of the structure's intended content is essential to an efficient filing system.

The NPS system of file codes, Director's Order 19 (NPS 2001a), and other nationally recognized standards, were considered in the development of the public share directory structure. Within the public share, records are grouped either by function or by activity, depending on purpose. For example, record sets not directly related to a project, but essential to day-to-day operations (i.e., forms, policy and guidance documents, vehicles) are grouped by function. Other records, for example data mining, are grouped because those records result from an activity. Records related to a single project or activity are grouped together regardless of their function.

All project electronic files will be stored in an intuitive project directory structure that is clearly understood by all network staff. All digital files associated with a project are filed under a common root directory. Project file names will adhere to the naming conventions established by the CHDN. CHDN_SOP_02-0004, Directory Structure, provides an outline of the public share directory structure with a brief description of each directory's content.

2.1.4.4 Web Share: The Web share, *\\INPCHDNMS01\web*, is not logically mapped to any drive letter. The Web share is a local file storage for Web development and contains Web content used by the local development Web server for the CHDN public and Intranet Web sites. The share access is limited to CHDN staff responsible for Web page development. The files contained in the Web share are duplicates of the files hosted on the national Web servers.

2.1.4.5 Print Server: INPCHDNMS01 acts as the print server for the CHDN office. All CHDN printers and plotters are accessed via INPCHDNMS01 and are restricted to local area network access only.

2.2 National System Architecture

The national I&M program provides several repositories for hosting CHDN information products and applications for summarizing park data at a national level. The applications are available online (with the exception of ANCS+ and NPSTORET) and allow users access to basic natural

resource information for CHDN parks. Specific descriptions of each of these applications are found in Section 2.6 of the NDMP.

- *NatureBib* – master database for natural resource bibliographic references
- *NPSpecies* – master database for species occurrence records and evidence (voucher specimens, references, observations or data sets) at each park
- *NPS Data Store* – master database of metadata for geographic information system (GIS) and natural resource data sets and a repository for that data
- *NPS Focus Digital Library and Research Station* – a decentralized digital imagery and data management system, implemented through a central public Internet portal sponsored by the NPS Office of the Chief Information Officer.
- *NPS Data Clearinghouse* – the central repository for NPS GIS data available to the public. Implemented through the NPS Focus gateway
- *NPS Automated National Catalog System (ANCS+)* – the official curatorial cataloging system of the NPS
- *Biodiversity Data Store* – an Internet-based repository of documents, GIS, and other data sets that contain information of plant and animal species that contribute to knowledge of the biodiversity in national parks
- *NPSTORET (also known as NPS Water Quality Database Templates)* – a NPS database designed to facilitate park-level standardized reporting for STORET, an Internet-based interagency water quality database developed and supported by the Environmental Protection Agency (EPA) to house local, state, and federal water quality data collected in support of managing the nation's water resources under the Clean Water Act.

2.3 CHDN Web Site Development

The CHDN Internet Web site, http://science.nature.nps.gov/im/units/chdn/, and the Intranet Web site, http://www1.nrintra.nps.gov/im/units/chdn/, are developed using Adobe's Dreamweaver. Sites are developed and maintained in the ColdFusion Markup Language, CFML, using INPCHDNMS01 as a development server. Both sites conform to the following guidance:

- NPS Director's Order 70 (NPS 2001b) (http://home.nps.gov/applications/npspolicy/DOrders.cfm)
- NPS Natural Resource Web Manual (NPS 2006c) (http://www1.nrintra.nps.gov/webnotes/NRWebManual/index.cfm)
- NPS I&M Network Data Management Web site guidance (http://www1.nrintra.nps.gov/im/datamgmt/webdev/index.cfm)
- NPS Natural Resource Web Manual (http://inside.nps.gov/helpdesk/manual.htm)

- Section 508 of the Rehabilitation Act which ensures access to electronic information by people with disabilities (http://www.section508.gov/)

The local files used for both Web sites are on the P: Drive under the Web sites folder, and the testing server files are located in the Web share on INPCHDNMS01. When information on either site needs to be updated or other changes are required, the changes are reviewed on the development server, and then posted to the appropriate server following the guidance located at the Data Management Intranet site. See CHDN_SOP_02-0003, Web site management, for additional details.

2.4 Network Security

The CHDN conforms to Department of Interior and National Institute of Standards and Technology security guidelines (NIST 2008). All sensitive electronic files are placed in protected folders with limited read and write access. Electronic file and directory permissions administration will be partially decentralized with file and folder administrative rights controlled by the CHDN Data Manager. A list of users and permissions is maintained in the Active Directory Database. Staff needing read or write access to the directory structure will be enabled through Microsoft Active Directory. There are rare circumstances in which encryption of sensitive electronic files and/or folders may be necessary. As an example, files on removable media or portable devices may need additional protection from unauthorized access or use. Encryption policies and procedures currently are under development for staff that require this assistance.

2.4.1 Disaster Recovery

All too often disaster recovery is centered on hardware failure; however, other concerns must be addressed in order to have a comprehensive disaster recovery plan. Hardware failure is only a small portion of potential problems that can threaten data and impact business.

Two necessary steps must occur before developing a disaster recovery plan. First, evaluate what impact a computer related disaster would have on day-to-day operations and identify what data needs protection. Second, identify threats to your information systems.

An analysis helps to identify system vulnerabilities, which helps target preventative measures. There are three major areas of vulnerability within the CHDN infrastructure:

- INPCHDNMS01 is the only CHDN server on the NPS.DOI.NET domain and is critical to daily operations and data storage. Because of the importance of this server, it has multiple redundancies built in as a precaution against mechanical and electrical failure that could result in a work stoppage. INPCHDNMS01 does comply with the mandatory standard guidelines as established in the DOI IT Security Policy Handbook (DOI 2008).

- The VPN connection also is critical to daily operations. Without the VPN connection, users would not be able to authenticate to the NPS.DOI.NET domain or access resources on INPCHDNMS01. The NPS WAN services group and the Enterprise Services Network manage the VPN, and the CHDN is dependent on their service for this connection.

- CHDN workstations are considered a vulnerability because they are the most likely conduit for malicious software. User awareness and good IT practices are the best prevention against such vulnerabilities. CHDN complies with DOI IT Security Policy Handbook (DOI 2008) regarding workstation configuration and security. A threat analysis identifies potential risks and vulnerabilities to a system that, when combined, can lead to disastrous results. Common threats can be categorized as follows:

- Natural threats (e.g., floods, earthquakes, tornadoes, landslides, avalanches, electrical storms)
- Environmental threats (e.g., power failure, configuration conflict)
- Human threats (intentional or unintentional)

Given the geographic location and facility in which the CHDN office is located, natural threats are minimal. Environmental threats such as power failures do occur, but infrequently, and the facility has been able to supply emergency power for the duration of the outages with only one exception to date. Human threats, intentional or otherwise, are the most probable the CHDN faces. An intentional act of sabotage is highly unlikely, but must be considered as a potential threat. The accidental deletion or overwriting of files is of concern. Malicious software falls under the category of human threats. Although unintentional, the introduction of viruses is the greatest threat to CHDN electronic data.

The threat-vulnerability combination identifies the threats that may cause the greatest impact to normal business operations. The server, INPCHDNMS01, is most vulnerable, and human actions, intentional or otherwise, are the greatest threats. The DOI IT Security Policy Handbook (DOI 2008) outlines many preventative practices. Preventative actions intended to reduce the potential of human threats against both servers and workstations include:

- Access control
- Security awareness training
- Configuration management
- Virus protection
- Security patch management

The ability to recover data and reconstitute the system after preventative measures have failed is essential to normal business operations. Information system backup and recovery it is not complex. The problem with a plan typically lies in follow-through. The implementation of a disaster recovery plan is a last resort and as such, few have ever tested their plan. Even fewer have ever tested their recovery procedures. A disaster recovery plan consists of three parts: the backup, the system recovery and reconstruction, and the contingency plan.

2.4.2 Backups

There are many considerations when planning a backup strategy. One common strategy is to back up everything, including the system state, in hopes of recovering to the original condition if necessary. This is a good strategy in theory, but application recovery is often unreliable, resulting in the need to rebuild the system a second time from scratch. Another strategy is to backup only that data and information that is not restored during a system reconstitution. Both strategies

require identifying data to recover, the backup frequency, and the retention schedule for the backups.

The CHDN uses the latter strategy for backups (Figure 2.2). The backup includes the public and the users' shares located on the E drive for INPCHDNMS01. The contents of the application share can be replaced from CD media or from the original download locations and does not need to be included in the backup strategy. Backups are performed using the backup software that comes with the Windows operating system. The backup frequency is based on a two-week schedule. This backup strategy calls for a full backup of all data every two weeks and a daily differential backup between full backups. Full backups use a series of portable hard drives capable of storing multiple backups. These portable hard drives are transported off site to an alternate storage location so data can be recovered in the event the facility is inaccessible for an extended period. Full backups are retained for a period of one year. Daily differential backups include tape as a backup media and involve backup of all files modified since the last full backup. This backup strategy allows a full data recovery using only the last full backup and the last daily backup. In the event data recovery is necessary due to a computer virus, the most recent backups are likely infected as well and data may have to be recovered from earlier full backups. CHDN_SOP_02-0002-Backups, "General Guidelines for CHDN Backups" covers backup procedures in more detail.

No disaster recovery is complete without a plan for system reconstruction and an alternate work location. To reduce the impact to business during an IT disaster, the system must be recovered in a timely manner or an alternate work location must be available. To reconstruct an IT system in a timely manner all system information must be documented and all necessary software must be available. Copies of all necessary software are kept with the backups at the alternate storage location. Additionally the hardware and personnel resources must be available to reconstruct the system. If it appears that reconstruction will take longer than 48 hours, arrangements have been made with two other park units to host CHDN staff and its data as alternate work locations.

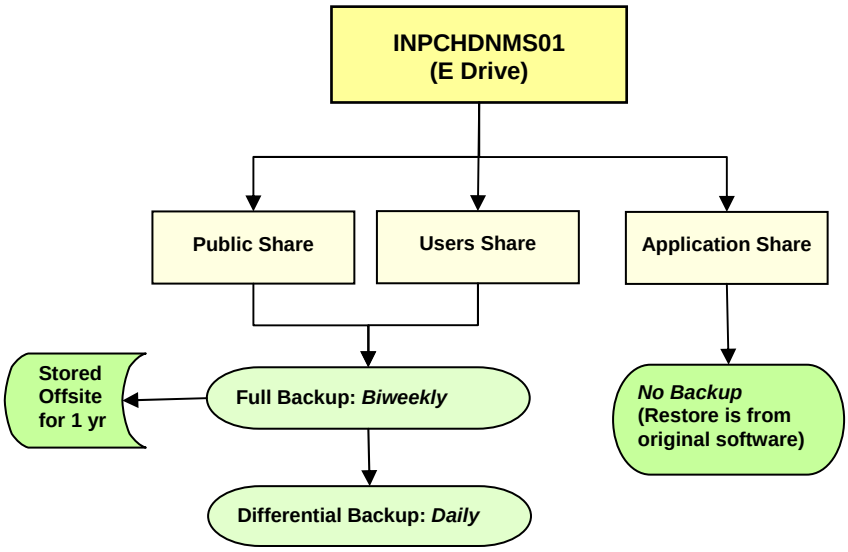

Figure 2.2. CHDN electronic storage backup strategy.

2.5 Geographic Information Systems

The NPS Intermountain Regional GIS office may provide support to CHDN and park staff when needed. The NDMP suggests using NPS Theme Manager to provide GIS layers to staff with limited skills (NPS 2008). CHDN will consider this option for deploying GIS data. CHDN standards for spatial data collection and submission will conform to the service-wide GIS specifications. The service-wide specifications are located at the Natural Resource GIS Program Web site (http://science.nature.nps.gov/nrgis/default.aspx).

2.6 Database Applications Development

Section 2.5.2 of the NDMP calls for desktop versions of network databases to be developed in the latest Microsoft Access, Microsoft SQL Server, or another established industry-standard relational database management system format unless otherwise specified in the project study plan (NPS 2008). For efficiency sake, CHDN will use currently available databases that are easily adapted to CHDN projects. In general, CHDN will use a combination of Microsoft Access and Microsoft SQL Server for database preparation.

2.7 Digital Data Formats

Section 2.5.1 of the NDMP provides guidance and standards for digital data formats (NPS 2008). This includes databases, tabular data, spatial data, and image data. CHDN will adopt the standards listed in the following sections. Additionally, all digital file names will follow specifications found in CHDN_SOP_08-0001, File Naming Conventions.

2.7.1 Databases

Well thought-out database design standards are necessary to promote compatibility among data sets, encourage sound database design, and facilitate interpretability of data sets. CHDN has adopted the Natural Resources Database Template (NRDT) as the standard for database design and recommends this method be used as much as possible (NPS 2007). A description of the NRDT and a working database is presented in Chapter 5 – Database Design.

2.7.2 Tabular data

Tabular data can be stored in any number of applications (i.e., word processing, spreadsheets, databases, and tables) and formats (e.g., .dat, .txt, .xml, .dbf, .xls, .csv). The CHDN standard format for tabular data is Microsoft Excel (.xls) or tab- delimited text files (.txt).The CHDN strongly encourages documentation of attribute and table information. Specific guidance regarding tabular data can be found in CHDN_SOP_05-0001, Database Development (Planned FY09).

2.7.3 Spatial data

The coordinate system standard the National Park Service is *Universal Transverse Mercator with North American Datum 1983* (UTM NAD83) (NPS 2008). Efforts will be made to migrate existing spatial data to this data projection, and it will be used for any new spatial data development. Currently, CHDN and park staff use many different raster and vector data formats for storing and managing spatial data. Vector data formats commonly include ESRI ArcInfo coverages, shapefiles, and personal geodatabases. Raster data formats commonly include MRSID, TIFF, GEOTIFF, and Grid structures. Because ESRI software products historically have demonstrated excellent backward compatibility with older data structures, there is no requirement for CHDN to migrate data formats to more modern data structures and all the above

listed formats will be accepted. Specific guidance regarding spatial data can be found in Specifications for Geospatial and Other Data Deliverables of GIS and Resource Mapping, Inventories, and Studies, http://science.nature.nps.gov/nrgis/standards.aspx.

Section 2.5.1 of the NDMP discusses the file geodatabase format that recently became available with the release of ArcGIS 9.2 (NPS 2008). This data structure does carry numerous advantages such as scalability, ability to handle large data sets, and compatibility across operating systems. Currently, CHDN does not have a large GIS collection and park staff has limited GIS skills; therefore, CHDN will continue to maintain data in shapefile format, but this position will be reconsidered annually.

2.7.4 Image data

Photographic information has become not only an important component of resource management but also a consumer of considerable amounts of electronic storage. Photographic information in the forms of digital images, scanned photographs, and satellite imagery has become a primary concern of both IT and data management staff for two reasons. First, IT staff are looking for simple and easy ways to store the large number of digital images collected by field personnel. Second, data management staff are concerned information may be lost with this large influx of digital imagery and want to encourage personnel to document what they obtain. CHDN has adopted ThumbsPlus from Cerious Software, Inc. for image management due to its ease of use and because image metadata is stored in a Microsoft Access database. This will facilitate exporting image data to other applications. Specific guidance regarding image data is listed in Section 2.9.

2.8 Collaboration Technologies

The CHDN I&M program is a good example of the requirement of collaboration between employees at six national parks spread across a large geographic area, which makes face-to-face meetings difficult. Section 2.4.4 of the NDMP (NPS 2008) provides information regarding the collaboration technologies that can be used by the I&M networks. Several communication options have been approved for use by DOI and NPS IT staff.

- The use of e-mail and telephones to conduct personal communication between individuals at separate locations is ubiquitous. Instant messaging complements these and adds additional functionality. E-mail is comparatively slow and a telephone call limits access to an individual's availability.
- Instant messaging is near real-time and offers additional flexibility for communicating with one or more employees. *Lotus Notes Sametime Instant Messaging* currently is available to all NPS employees with an NPS computer network login profile.
- CHDN likely will adopt the use of Microsoft SharePoint Services in the future. SharePoint is a Web-based collaboration tool that allows multiple users to create, view, or contribute to documents, surveys, wikis, discussion forums, and other types of projects that require the participation of many. Unlike Web sites, all users with appropriate permissions can easily post or contribute content.

2.9 Related Guidance Documents, SOPs and/or TIs (see latest version on CHDN website)

- CHDN_SOP_02-0002; General guidelines for CHDN backups
- CHDN_SOP_02-0003; Web-site management
- CHDN_SOP_02-0004; Directory structure
- CHDN_SOP_08-0001; File Naming Conventions
- CHDN_SOP_08-0002; ThumbsPlus, general procedures for the use of ThumbsPlus (Planned FY09)
- CHDN_TI_08-0001; ThumbsPlus network client installation procedure
- CHDN_TI_08-0002; ThumbsPlus metadata import procedure
- CHDN_TI_08-0003; ThumbsPlus server install instructions
- CHDN_SOP_06-0005; Image management guidelines
- CHDN_SOP_05-0001; Database development (Planned FY09)

Chapter 3. Project Management and the Data Life Cycle

Data management within the CHDN includes coordination of all phases of inventory or monitoring projects, as well as the detailed steps of data acquisition and processing. Establishing a structure, sequence, and steps for these phases is needed to ensure all tasks are completed, responsibilities are clear, resources are available, and overall workflows are synchronized.

There are two main types of projects handled by CHDN natural resources staff and the I&M program:

- Short-term, which may include individual park research projects, inventories, or pilot work done in preparation for long-term monitoring or research.

- Long-term, including network vital-signs monitoring projects central to the I&M program and multi-year research projects and monitoring performed by other park programs, agencies, and cooperators. Long-term projects will often require a higher level of documentation, peer review, and program support.

For information management, the primary difference between short- and long-term projects is an increased need to adhere to and maintain standards for long-term projects. Maintaining standardization from year-to-year will be necessary when comparing data over an extended period (decades for long-term monitoring).

3.1 Stages of Project Development

Projects conducted by CHDN can be divided into a series into five primary stages:

1. *Planning and Approval*. At this stage of initiation, many of the preliminary decisions are made regarding areas such as project scope and objectives. Funding sources, permits, and compliance also are addressed at this time. Primary responsibility rests with project leaders and the network coordinator. To anticipate data management needs, the data manager will need to be informed, even though there are no specific activities during this phase.

2. *Design and Testing.* Planning for acquisition, processing, documentation, analysis, and reporting occurs during this stage. Collaboration between the project leader and the data manager is critical in order to assure data quality and integrity. Developing documentation of databases and formal metadata is worked out during this stage. A team effort is required to develop and document the project methods, data design, data dictionary, and the database itself.

3. *Implementation.* Data are acquired, processed, error-checked, further documented, and products are developed during this stage. The project leader oversees all aspects of this stage, with the data management staff functioning primarily as facilitators to support database applications, GIS, GPS, data verification, summarization, and analysis. Project staff members work to develop and finalize deliverables identified in the project planning documents (i.e., protocol, study plan, contracts, agreements, or permits).

4. *Product Integration and Distribution.* Data are merged from the working database to master databases. Administrative records are delivered to appropriate park and network staff as specified. All project deliverables are distributed according to specifications, which will be stipulated in all protocols, contracts, agreements, and permits. Products that do not meet program requirements should be returned for revision.

5. *Evaluation and Closure.* Project records are updated in the tracking database to reflect the status of the project. After products are catalogued and made available, program administrators, project leaders, and data managers should work together to assess how well the project met its objectives and to determine what might be done to improve various aspects of the methodology, implementation, and formats of the resulting information. For short-term (non-cyclic) projects, this phase represents the completion of the project.

3.2 The Data Life Cycle

Data take on different forms during various phases of a project and are maintained in different places as they are acquired, processed, documented, analyzed, reported, and distributed. These tasks are referred to as the "data life cycle" (Figure 3.1), and further defined as follows:

1. Acquire data – Data are acquired in digital or analog form. Digital data can be recorded on mobile computing devices such as handheld computers and PDAs, tablets, laptop computers, or GPS units. Analog data are entered on field data sheets and should match the protocol database data entry form. The database data entry form will include items to assist in the quality assurance/quality control (QA/QC) of the data being entered. During the protocol development, the range or ranges for the expectable limits of the data collected will be established. Data entry forms will use these data domains to validate data entry forms ensuring the data recorded falls into the expectable limits established for the protocol.

2. Archive raw data – Copies of all raw data files are uploaded to the appropriate project document library. Hard copy forms are scanned and uploaded to the appropriate project document library, and the originals are archived according to the Records Management Handbook (NPS 2005). Completed hard-copy data forms are preserved in a single, safe location until they are archived in the CHDN central files at the end of the field season. Scanning of hard-copy data forms may occur at the end of a season as appropriate.

3. Enter/import data – Analog data are entered manually into the project specific database. Data collected using mobile computing devices is stored in its original format as an archive (CHDN_SOP_02-0004, Directory Structure) and then replicated to the master project database following the established method defined during protocol development.

4. Verify, process, and validate – Accurate transcription of the raw data is verified; data are processed to remove missing values and other flaws; and data are validated through visual inspection and queries to capture missing data, out-of-range values, and logical errors.

5. Data documentation – Develop or update project metadata and certify the data set. Certification is a confirmation by the project leader that the data have passed all quality assurance requirements and are complete and documented. It also means that data and metadata are ready to be posted and delivered.

6. Upload data – Certified data are uploaded from the working database to the master project database. This step may not be required for short-term projects for which there is no need to distinguish working data for the current season from the full set of certified project data.

7. Archive versioned data set - The CHDN Project Data Certification Form is completed (CHDN_SOP_03-0001, Project Data Life Cycle Guidance). Copies of the certified data and metadata are placed in the digital library. This can be accomplished by storing a compressed copy of the working database or by exporting data to a more software-independent format (e.g., ASCII text).

8. Disseminate data and update national databases – Project data, including digital images and metadata, are posted to national repositories (the NPS Data Store, Biodiversity Data Store, NPS Focus) as necessary to make them more broadly available to others.

9. Reporting and analysis – Certified data are used to generate data products, analyses, and reports, including semi-automated annual summary reports for monitoring projects. Depending on project needs, data may be exported for analysis or summarized within the database.

10. Distribute information products – Information products such as reports, maps, and checklists are disseminated to the public through the CHDN Web site, *Learning Center for the American Southwest*, and the appropriate national application as necessary. Only data and/or metadata determined to be non-sensitive will be available for distribution.

11. Share data and information – Reports and other data products resulting from a project are uploaded to the appropriate project document library with all the required document property being completed. This process ensures data are well documented, discoverable, and available for future use. Hard copies are mailed in response to specific requests.

12. Track Changes - All subsequent changes to certified data are documented in an edit log that accompanies project data and metadata upon distribution. Significant edits will trigger reposting of the data and products to national databases and repositories.

3.3 Related Guidance Documents, SOPs and/or TIs (see latest version on CHDN website)

- CHDN_SOP_03-0001; Project Data Life Cycle Guidance
- CHDN_SOP_02-0004; Directory Structure
- CHDN_SOP_06-0006, Library Management Strategy

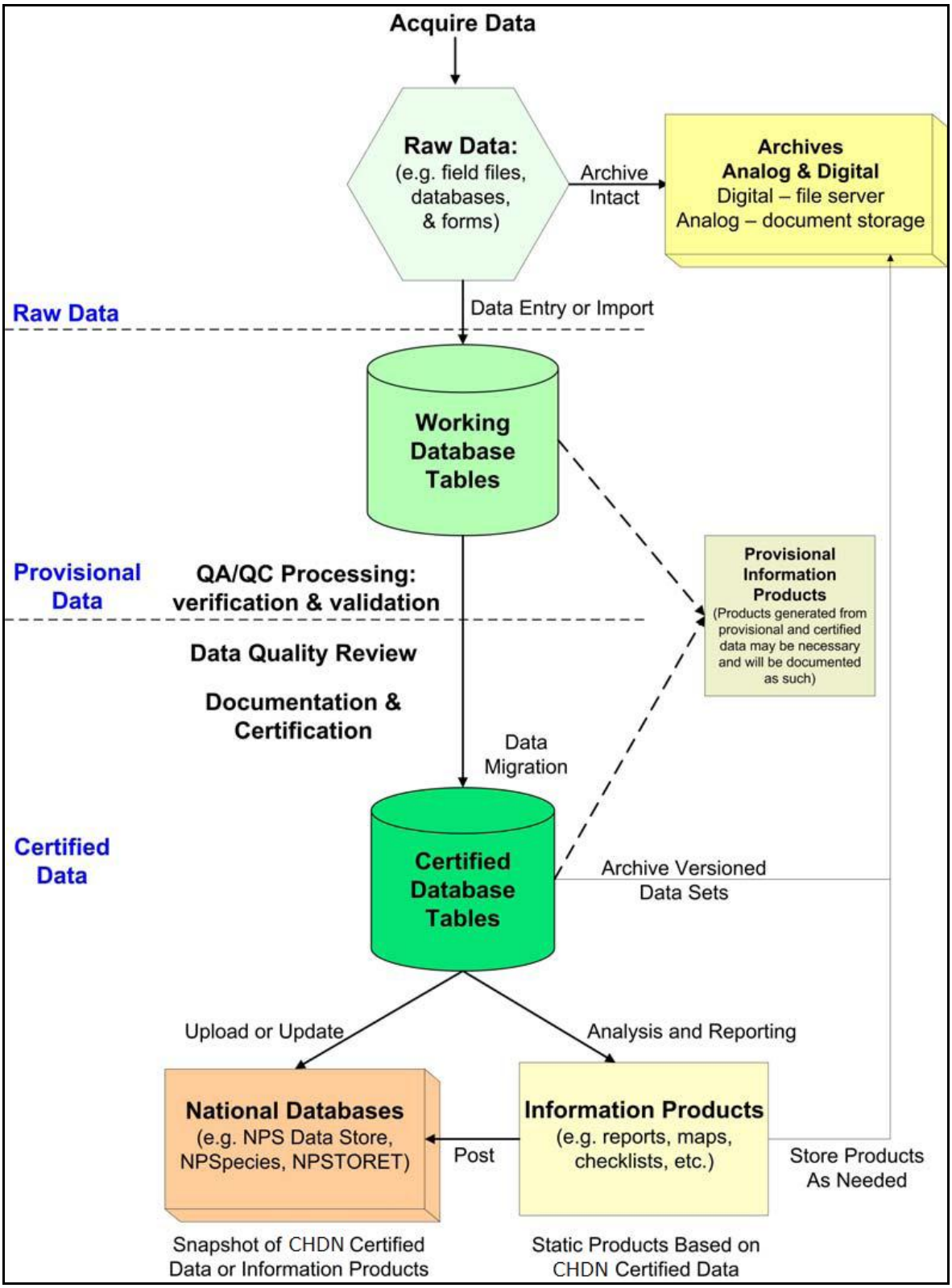

Figure 3.1. Schematic of the migration of data through the typical data life cycle (Adapted from the Southeast Coast Network Data Management Plan)

Chapter 4. Data Management Roles and Responsibilities

During protocol development, standard operating procedures and task instructions will define the data management roles and responsibilities for all phases of data collection.

A role is a function or position (e.g., Data Manager). A responsibility is a duty or obligation (e.g., review data records). Table 1 summarizes general types of data stewardship activities and the roles typically associated with them.

Table 4-1. Categories of data stewardship involving Chihuahuan Desert Network personnel.

Stewardship Activity	Description of Activities	Principal Roles
Production	Creating data or information from any original or derived source. This includes recording locations, images, measurements, and observations in the field, digitizing source maps, keying in data from a hard-copy source, converting existing data sources, image processing, and preparing and delivering informative products, such as summary tables, maps, charts, and reports.	Project Crew Leader, Project Crew Member, Data/GIS Specialist or Technician
Analysis	Using data to predict, qualify, and quantify ecosystem elements, structure, and function as part of the effort to understand these components, address monitoring objectives, and inform park and ecosystem management.	Ecologists, Resource Specialists
Management	Preparing and executing policies, procedures, and activities that keep data and information resources organized, available, useful, compliant, and safe.	Project Leader, Data Manager, GIS Manager, IT Specialist, Database Manager, National-level Data and Information Managers
End Use	Obtaining and applying available information to develop knowledge that contributes to understanding and managing park resources.	Project Leader, Park Managers and Superintendents, others

Chapter 4 of the NDMP (NPS 2008) defines in detail the roles and responsibilities of specific personnel, but it cannot possibly account for the splitting and lumping that is necessary for any given project and available personnel. As of fiscal year 2009, the CHDN staff consists of two permanent full-time staff and one temporary full-time staff. In addition to NPS personnel, the CHDN also makes use of resources available through a Cooperative Agreement with New Mexico State University.

With respect to data management, the CHDN biological technician is responsible for data mining efforts and ensuring the resulting data is cataloged and stored appropriately. The biological technician also assists the data manger in the development of the DMP. The data manager is responsible for all other duties of managing data, including collateral IT specialist duties.

Because IT specialist responsibilities are generally not a part of a network data manager's position, these corollary duties are described in Chapter 8 of the CHDN Monitoring Plan. As additional positions in the CHDN staffing plan are filled, the roles and responsibilities will be refined accordingly.

Table 2 summarizes the roles and responsibilities of CHDN staff and cooperators with respect to data stewardship. It should be noted that a single person may take on more than one role in a given project. For instance, the CHDN data manager also will take on the role of IT specialist for projects conducted within CHDN.

Table 4-2. Roles of CHDN network staff and cooperators working on monitoring projects.

Role	Data Stewardship Responsibilities
Project crew member	Collect, record, verify data; perform data entry; organize field forms, photos, other related materials
Project crew leader	Supervise crew, communicate regularly with data manager and project leader
Network data manager /GIS specialist	Ensure program data are organized, useful, compliant, safe, and available. Oversee GPS data collection, manage spatial data, prepare maps, perform spatial analyses
IT specialist	Apply database and programming skills to network projects, maintain information systems to support data management
Project leader	Direct operations, including data management requirements, for network projects
Resource specialist	Evaluate validity and utility of project data; document, analyze, publish data and associated information products
Quantitative ecologist	Determine project objectives and sample design; perform and document data analysis and synthesis; prepare reports
Network coordinator	Coordinate and oversee all network activities
Park or regional curator	Ensure project results (documents, specimens, photographs, etc.) are cataloged and stored in NPS or other repositories
I&M data manager (national level)	Provide service-wide database support and services; provide data management coordination among networks
End users (managers, scientists, interpreters, public)	Inform and direct the scope of science information needs; interpret information and use to direct or support decisions

4.1 Standard Operating Procedures

A roles and responsibilities section is to appear in all SOPs created by CHDN to insure that everyone involved in a project clearly understands the tasks assigned to them in all projects.

Chapter 5. Databases

The term database refers to the collection of related records. The software used to develop a database is the database management system (DBMS). Often the term database is used to reference both the data and the software used. The distinction is clear: Any collection of related records is a database, regardless of the format in which it is stored, but not all applications that store a collection of related records are a DBMS.

5.1 Natural Resource Database Template

The Natural Resource Database Template, NRDT (NPS 2007), is a set of standardized relational database tables that make possible high-level integration of data sets across networks or regions. There are four core NRDT tables: locations, events, metadata, and revisions. A database is NRDT compliant if the tables, fields, relations, and naming conventions of mandatory tables are used. Access databases developed by the CHDN will comply with the standards established in the NRDT whenever practicable.

5.2 Database Development

Database development must follow the basic steps of the system development life cycle (SDLC) to increase the likelihood of success. An SDLC is any logical process used by a systems analyst to develop an information system. These steps are similar to project planning phases; rightfully so if you consider database development a project. Also similar to project planning, the number of steps involved in the SDLC model may vary from project to project and some steps may be lumped or split. The development steps involved in most protocol databases are analysis, design and implementation, testing, training, and maintenance.

5.2.1 Analysis

Analysis is perhaps the most important step in database development. It is the analysis that establishes the overall goal and user requirements of the database. During monitoring protocol development, the data manager, the ecologist, and the principal investigator work together to develop the database requirements document. The requirements document serves several purposes, but most importantly, it defines the product and is agreed upon by the members of the development team. The requirements document includes a description of the database, what data is being collected, why it is being collected, and the intended audience of the data. The requirements document also lists all tables, fields, data types, data range if applicable, and field definitions. The requirements document defines the data life cycle, including data collection, retention of raw data, processing of derived data, merging individual data sets into the master database, and final disposition. Any changes to a database or the data life cycle are documented in the database requirements document. See CHDN_SOP_05-0001, Database Development (Planned FY09).

5.2.2 Design and Implementation

Design and implementation involves converting the informational and functional requirements identified in the database requirements document into a functional model. Different aspects of the database are broken down into development sections (e.g., contact information, location information), and an estimated time for completion for each section is established. Most database sections consist of four or five elements (e.g., an input form, a details form, an edit form, a list, and possibly a search form to control the list). The number of sections and the interaction among

sections also is defined in the requirements document. As each section is completed, someone other than the programmer should test it and draft a user manual for that section.

5.2.3 Testing

Thorough testing by potential users is critical to ensuring databases meet protocol and end-user requirements. Although sections of the database are tested during development, when the sections interact the results may not be quite as expected. Testers often identify program defects or features during the testing process that developers overlook. Often testers/end-users do things developers never considered exposing undesirable results. Testing also refines the workflow. What looked good during the design phase may not work well during actual use. Modification to the database resulting from testing may require adjustments to the requirements document. Also during testing, the user manual should be updated as needed.

5.2.4 Training

Data management staff provides training and support for database applications prior to project initiation. Training is important for quality assurance. As the complexity of a database increases so does the need for training. Even the best user's manual is a poor substitute for training, especially when you consider the typical user opens the manual only as a last resort. One important consideration about database development and training is that sometimes the amount of resources required to program to prevent a user's behavior outweighs the effort required to modify their behavior.

5.2.5 Maintenance

Proper maintenance requires making changes to address user requirements and correcting problems, as well as providing regularly scheduled updates and checks. Any changes necessary to address user requirements necessitate the database requirements document be updated and approved by the protocol development team. Versioning and archiving of databases is an important step in maintaining databases and is addressed in CHDN_SOP_05-0001, Database Development (Planned FY09).

5.3 Related Guidance Documents, SOPs and/or TIs (see latest version on CHDN website)

- CHDN_SOP_05-0001, Database Development (Planned FY09)

Chapter 6. Data Acquisition, Processing, and Reporting

Project leaders and data managers are responsible for ensuring data collection, data entry, data verification, validation, storage, and archiving are consistent with I&M network standards. In addition to general SOPs that define network-wide requirements and general operating principles, protocol-specific TIs are developed. TIs contain systematic instructions for performing a single specific task. These SOPs and TIs are established during protocol development and may require modification as monitoring is implemented. Documenting the data collection process and changes to those procedures is critical to producing quality data.

6.1 Data Collection

A wide range of tools is available for field data collection. All methods involve some trade-offs in terms of expense, efficiency, and susceptibility to data entry/transcription errors. Each project's protocols should provide detailed specifications about the data collection method, the use of the equipment, and disaster recovery.

Although the risk of equipment failure must always be compared to the time-saving element, field computers will be used to the greatest extent possible for data collection. The use of devices such as personal data assistants (PDAs), tablet PCs, and GPS units increase data collection efficiency. Data can be downloaded directly to the main database eliminating transcription errors. QA/QC checks can be built directly into the mobile device database, further reducing data entry error and processing time. Although mobile computing devices require a greater upfront cost in equipment and database development, the investment is recovered over time.

Field data forms are currently the most common method of recording field data. Field forms are inexpensive but require neat, legible handwriting. Field forms pose greater opportunities for error during the collection/data entry process compared with mobile computing devices. Field data forms also require data noted on forms be key-entered later into project databases introducing the potential of transcription errors. This method tends to require more data entry time and more rigorous QA/QC.

Automated data loggers are used mainly to collect ambient information such as weather data or water quality and quantity information. These units must be properly calibrated and maintained, which requires proper training of field crews and SOPs to outline these procedures.

Photographs provide an excellent visual record of field visits and are useful for capturing point records of long-term study sites. They also serve well for automated data collection by remotely recording information using Web or trip cameras. Slides, photographs, and digital images should be captured, stored, and archived according to guidelines (see Section 6.6), and must have essential metadata associated with each image for the information to have long-term value.

6.2 Data Processing

Each project will have a database developed in conjunction with project protocols prior to the collection of any data in the field. Built-in procedures for data validation and QA/QC should be part of the database development and user interface. Data processing should occur as soon as possible after data are collected. Original data collected using mobile computing devices will be uploaded to the project directory or uploaded directly to the master database depending on the

protocol procedure. Field data forms will be entered into the project database by the project lead or a delegate and the physical forms sent to the network office to be stored in the CHDN central files.

Photos taken as part of a project's data collection protocol constitute data and need to be organized, documented, and preserved in conjunction with all other project data. Editing of digital photos may be done to improve orientation or correct for lighting and contrast but should never include alterations that change the original content of the photo.

The acquisition and processing of biological specimen vouchers will be guided by the policies outlined by the ANCS+ Guidelines (NPS 2006a) and according to individual park guidance. Project leaders will provide parks with material and data, in MS Excel format or comma-delimited ASCII (American Standard Code for Information Interchange) files. Specimens are owned by the parks in which they were collected and are curated by park staff unless agreements have been made with outside institutions.

6.3 Non-programmatic Data

Non-programmatic data are data that are not a direct result of CHDN efforts (e.g., a plant survey for the Guadalupe Mountains by USDA Forest Service or a lion survey for Brewster County, Texas, done by the Texas Parks and Wildlife Department). This type of data can provide a great deal of information about park natural resources. Non-programmatic data can be obtained from academia, private organizations, and non-profit groups, as well as local, state, and federal government agencies. Non-programmatic data found to be relevant to the CHDN program should be maintained either electronically or in hard-copy format depending on the method of collection, should contain proper metadata as outlined in Chapter 8, Section 8.3, and entered as appropriate into the national databases (see Chapter 10). Non-programmatic data should meet protocol specifications if they are to be used for data analysis.

6.4 Data Analysis and Reporting

Data summary and analysis are essential to providing relevant and useful information for natural resource managers and scientists. Data managers and stewards must provide valid data in formats that support scheduled and ad hoc display, query, analysis, summary, and reporting. Project databases will be designed to provide ad hoc query capabilities with simple display and analysis capabilities and include a means to download raw data in ASCII format for analysis in third-party software. Routine and scheduled data summary, analysis, and reporting requirements and procedures should be identified in project protocols. CHDN will follow the guidelines for analysis and reporting set forth in Section 6.6 of the NDMP (NPS 2008).

6.5 Data Mining

A fundamental step in developing and maintaining the information base for CHDN is locating, evaluating, and documenting, on an ongoing basis, park-related natural resource information. The term "data mining" refers to this process, which requires regular visits to network parks and the establishment of data mining procedures specific to each park. The range of materials that requires documenting is broad, ranging from historical inventories to complex databases to remote-sensing data.

Generally, the goal of network data mining is not to acquire and store copies of all data residing at parks at the network. Instead, cataloging tools are used to describe and document data sources so that potential users can find them. Data that are an exception to this rule are documents or data sets that are of specific need or interest to network staff (e.g., GIS data, monitoring data) or that complement network projects. CHDN also works with parks to identify and obtain electronic copies of documents for full-text posting in NatureBib where they can become widely accessible.

Parks within the network vary widely in how natural resource information is organized and stored. Though some parks have developed efficient and standardized methods of cataloging new information as it arrives, other parks are still working to formalize this approach.

Although the ultimate goal of CHDN is to train individual park staff to become the primary data managers of NatureBib data, CHDN staff will continue to enter data, provide training, and quality control/assurance of network data until individual parks are able to meet the high quality standards for self-management. The online database continues to be updated as new references are added to the network library.

Park visits have and will continue to include searches of each park's library and staff files, staff interviews, and limited searches of local universities, Web-based library catalogs, and other agency files.

6.6 Related Guidance Documents, SOPs and/or TIs (see latest version on CHDN website)

- CHDN_SOP_06-0001, CHDN NatureBib Guidelines
- CHDN_SOP_06-0002, CHDN NatureBib Full-text Document Management
- CHDN_SOP_06-0003, CHDN NPSpecies Guidelines
- CHDN_SOP_06-0004, CHDN Data Mining Strategies
- CHDN_SOP_06-0005, CHDN Image Management
- CHDN_SOP_06-0006, Library Management Strategy
- CHDN_SOP_11-0001, Monitoring in Parks (Planned FY10)

Chapter 7. Quality Assurance and Quality Control

Quality assurance/quality control (QA/QC) can be defined as an integrated system of management activities involving planning, implementation, documentation, assessment, reporting, and quality improvement to ensure that a process, item, or service is of the type and quality needed and expected by the consumer. QA/QC is designed into the protocol during development and is specific to the protocol. During protocol development the project lead and the data manager work together to establish the method by which data is collected, verified, and certified and develop a project specific database for the storage of that data. It is during the database development that QA/QC measures are defined in the database requirements document.

CHDN will adhere to Director's Order #11B: Ensuring Quality of Information Disseminated by the National Park Service (NPS 2002) and the general guidelines of the QA/QC process, referred to in Chapter 7 of the NDMP (NPS 2008).

Chapter 8. Data Documentation

Developing a long-term monitoring program (including elements such as plan development, data collection techniques and analysis, information dissemination, and data storage) requires a considerable investment in time and resources. Because of the long-term nature of the program, data collection will occur over many years and at intervals that may be several years apart. It is essential that data collected retain its usability during the life of the program and beyond. Data documentation is perhaps the single most important step toward ensuring that data sets are usable well into the future. Unfortunately, data developers often perceive the metadata creation as a tedious process, the benefits of which are realized only by those who later utilize or inherit their data. The time invested initially for metadata creation yields benefits in the future and without it, the data collecting efforts may be wasted. The data creator will be primarily responsible for metadata creation. A statement of metadata requirements will be included in each project statement of work.

8.1 Laws and Policies

Even though the importance of metadata is widely accepted within the data management community, the approaches for storing this information and the levels of detail can vary. However, some established metadata strategies apply to NPS data. The NDMP, Section 8.2 (NPS 2008) provides information on the laws and policies related to data documentation followed by the NPS. The CHDN adhere to these laws and policies as it relates to data documentation.

8.2 Metadata Creation

Metadata creation can be as simple as making notes in a text file or on the back of a photo. In fact, many older data sets have proved to be valuable information only because of the availability of simple notes the creator made to jog his or her memory. CHDN requires collection of enough data to meet the needs of the repositories where the data will reside (See Chapter 10). Data created for the I&M program have several repositories to consider for final data product storage. The I&M national applications are the primary data catalogs for CHDN data. See Section 2.6 of the NDMP (NPS 2008).

CHDN uses a series of tools available to facilitate metadata collection. Descriptions of these tools is found in Chapter 8 of the NDMP (NPS 2008). The CHDN data manager, or assignees, will provide training and support in the use of these tools to project leaders and will aid in metadata development where practical. Upon completion, metadata will be posted with project data so they are available and searchable along with their constituent data sets and reports via the CHDN Internet Web site and the NPS Data Store.

8.3 Metadata Content Standards

A metadata record consists of a number of pre-defined elements representing specific attributes of a data set. The number of elements and the order in which the elements are arranged depends on the schema. Many different metadata schema exist, but they all share a common goal, which is to describe the who, what, when, where, why, and how of the data. GIS files and geospatial databases include elements such as geographic extent and projection information. Photographic metadata include data elements about the camera, shutter speed, aperture, and focal length. Biological metadata include elements related to taxonomy, classification system, or authority used. Every metadata schema has its own unique needs and will vary based on those needs.

Regardless of the schema, CHDN adheres to metadata content standards set by the Federal Geographic Data Committee (FGDC) and NPS policy as described in the NDMP, Section 8.2 (NPS 2008).

The CSDGM consists of seven sections (Sections 1-7), all of which contain elements required for spatial data. The NPS Metadata Profile adds another section (Section 0). The Biological and ESRI Profiles are added as elements to Sections 1-7.

Metadata that are *fully compliant* with FGDC and NPS standards have entries in Section 0 and all element fields in Sections 1-7 in which the *Optionality* field contains the term 'mandatory' or 'mandatory if applicable.' A mandatory element must be populated for every data set. A mandatory-if-applicable element must be populated if the data set exhibits the characteristic being documented by the metadata element. For example, the element defining the vertical coordinate system is mandatory if a data set contains elevation data.

Metadata that are *minimally compliant* with FGDC and NPS standards have entries in Section 0 and all 'mandatory' and 'mandatory if applicable' element fields in Sections 1, 6, and 7, and Section 2 for biological data. These include the fields used by the NPS Data Store. Data sets documented to this extent can be distributed via the Data Store's online upload utility.

Different types of data and information require different kinds and levels of documentation. CHDN has outlined standards for documentation of the following data set types:

- ***Spatial Data***. CHDN uses ESRI GIS software, which is the NPS standard software for geospatial data. ESRI GIS includes the metadata creation tool, ArcCatalog, with is well suited for creating FGDC compliant metadata. The NPS Metadata Tools & Editor is a custom application extending the capabilities of ArcCatalog to include the NPS Metadata Profile. Spatial data will contain, at a minimum, all of Section 0 of the NPS Metadata Profile and the required elements of Sections 1-7 of the CSDGM.

- ***Non-Spatial Data*** includes, at a minimum, all elements of Section 0 of the NPS Metadata Profile and the required elements of Sections 1, 6, and 7, and Section 2 for biological data. The minimum requirements for non-spatial data, therefore, meet the requirements for minimum compliance with the FGDC and NPS standards.

- ***Relational Databases*** will be documented according to the standards outlined above (depending on inclusion of spatial components). Complete documentation also will include entity relationship diagrams, business rules, and programming code. The NPS Metadata Profile currently does not support this type of documentation, so it will be stored separately from the formal metadata, in a folder with the database. Relational databases will also utilize internal documentation such as table and field descriptions and will include a table to track modifications.

- ***Legacy Data*** is documented to the extent possible according to the standards outlined above. Metadata that accompany legacy spatial and non-spatial data sets are suitable for upload to the NPS Data Store if they are non-sensitive and include entries sufficient for minimum compliance with FGDC and NPS standards. Priority for documentation will be:

- Data sets needed for current project development
- Data sets used frequently by park staff or cooperators
- Historic data sets archived for possible future use

Any contracts entered into by the CHDN with data miners will stipulate the submission of FGDC and NPS-compliant metadata. The network data manager or project leader will assist with metadata acquisition by providing tools, format protocols, and file transfer services.

- ***External Data*** are the data generated and/or managed outside of CHDN programs but used in analysis with CHDN data or distributed in any manner by the NPS. These data require the same level of documentation produced for CHDN-generated data, including but not limited to, data produced under contract with the NPS. Metadata will be requested from the originating entity by the network data manager or project lead.

Generally, external data will not be posted on network or park local or wide area networks (LAN/WAN) without accompanying metadata. This will include any metadata downloaded with the data, plus additional information regarding date of download and any alterations made to the data by NPS staff. Staff posting data to the CHDN LAN or WAN will make a reasonable effort to make up for any deficiency in the original metadata, but should not create new metadata for data from well-known sources such as U.S. Geological Survey digital line graph (DLG) data.

- ***Sensitive Data.*** Metadata documentation (Section 1 of the CSDGM, Constraints on Access) provides one means of labeling sensitive data in order to ensure their protection and integrity over time. Sensitive data, therefore, are documented according to the standards outlined above, although these data and their metadata will not be uploaded to the NPS Data Store. Instead, they are archived in secure locations on park and/or network servers.

Microsoft applications and Adobe Acrobat have built in metadata capabilities. Document properties are based on the Dublin Core metadata schema (Dublin Core 2009). The real power behind the document properties is the ability to index and search for information based on those document properties. Many applications, in addition to Microsoft Office applications, can take advantage of the indexing.

Document properties are accessed from the Office Button, Prepare, Properties of Microsoft Office 2007 products or File, Document Properties in earlier Microsoft Office versions. Even though the ability to add custom properties is limited, the required document properties for final, business-related CHDN files will include:

- Author
- Title
- Subject
- Keywords
- Status

Additional required *custom* properties are:

- Date
- Park code
- Project
- Sensitivity
- Disposition

Documents converted to Adobe's Portable Document Format (PDF) also have document properties and CHDN documents converted to PDF require the same minimum data fields.

- ***Digital Images***. The NPS Digital Photo Metadata Standards are used to document digital image metadata at CHDN. Additional guidance is available in the CHDN Image Management SOP. The minimum metadata required includes the seven elements considered mandatory as established in the NPS Digital Photo Metadata Standard (NPS 2006b), as well as the author or photographer of the image, as recommended by the NDMP (NPS 2008):

    - Author – who took the photo, with agency affiliation
    - Title – who or what is in the image
    - Location – where the image was taken
    - Date – when the image was taken
    - Park code
    - Access constraints – who may view the image
    - Copyright information – restrictions on using the image
    - Contact information – who to contact for further information

The NPS standard almost entirely overlaps the Exchangeable Image File Format (EXIF) and International Press Telecommunications Council (IPTC) standards. However, the actual metadata for EXIF and IPTC is intended to be embedded in the image file itself. There are many applications available to edit this data. For its photo management software, CHDN uses ThumbsPlus, which is a database driven photo-management software that maintains both IPTC and EXIF metadata and allows for creation and modification of metadata within the file for supported file formats. A spreadsheet containing metadata for multiple images is used by CHDN to import information directly into the ThumbsPlus database, thereby imbedding unique metadata into multiple images in a batch process.

8.4 Metadata Storage

The method(s) for storing metadata at CHDN will vary by file type. Whenever possible, metadata is embedded into the file to maintain the likelihood that the information remains with the data. With a data/metadata file pair, a separate Extensible Markup Language (XML) encoded file accompanies the data throughout its life cycle. This conjoined file pair may become separated unless all handlers pay careful attention during the dissemination of the data. A data/metadata file pair is used when storing data on the NPS Data Store. File/document properties are for files developed with Microsoft applications, PDF documents, and digital images. Using the ThumbsPlus database, images maintain imbedded metadata within each file.

8.5 Related Guidance Documents, SOPs and/or TIs (see latest version on CHDN website)

- CHDN_SOP_08-0002; ThumbsPlus, General Procedures for the use of ThumbsPlus (Planned FY09)
- CHDN_TI_08-0001; ThumbsPlus Network Client Installation Procedure
- CHDN_TI_08-0002; ThumbsPlus Metadata Import procedure
- CHDN_TI_08-0003; ThumbsPlus Server Install Instructions
- CHDN_TI_08-0004; ThumbsPlus Backup and Recovery
- CHDN_SOP_06-0005; CHDN Image Management

Chapter 9. Data Ownership and Sharing

CHDN data and information products are considered property of the NPS. However, the Freedom of Information Act (FOIA) establishes access by any person to federal agency records that are not protected from disclosure by an exemption. The CHDN will comply with all FOIA strictures regarding sensitive data. If the NPS determines that disclosure of information would be harmful, information may be withheld concerning the nature and specific location of the following:

- Endangered, threatened, rare, or commercially valuable National Park System Resources (species and habitats)
- Mineral or paleontological objects
- Objects of cultural patrimony
- Significant caves

Specific laws and guidance for determining whether data should be protected can be found in Chapter 9 of the NDMP (NPS 2008).

9.1 Determining Data Sensitivity

For each project or protocol, the network will identify the network lead. This person, who is typically the project leader, will be responsible for communicating with principle investigators (PIs), cooperators, and park staff. Each project leader, as primary data steward, will determine data sensitivity in light of federal law and will stipulate the conditions for release of the data in the project protocol and metadata. Requests for access to all sensitive information, whether stewarded by the network or the park, will be routed to the respective park contact by the PI for approval. The park will have the ultimate say on how this information will be treated. Network staff will classify sensitive data on a case-by-case or project-by-project basis. The project leader will work closely with investigators and park contacts for each project to ensure potentially sensitive park resources are identified and tracked throughout the project.

Network staff is responsible for identifying all potentially sensitive resources to principal investigator(s) working on a project. Investigators, whether network employees or partners, will develop procedures to flag all potentially sensitive resources in any products that come from the project, including documents, maps, databases, and metadata. When submitting any products or results, investigators should specifically identify all records and other references that contain data pertaining to potentially sensitive resources. Additionally, investigators should not release any information in a public forum before consulting with network staff to ensure the information is not classified as sensitive or protected.

Data fall into one of three categories depending on its sensitivity.

- Public: Approved for general release (Approved for Internet/Public posting). This information is not considered sensitive by national, network, or park-level standards; is of high quality; fully documented; and non-proprietary. This information will be made available to all federal and non-federal entities and the general public.

- Restricted, NPS Only: Only for release to the National Park Service (Approved for Intranet posting). This information is not considered sensitive by national standards, but may reference specific point locations, be of questionable quality, proprietary, or intellectual property. All NPS staff, regardless of position and status, will have read-only access to this information. If information is requested by a federal or non-federal entity, or the public, it will be released only with the approval of the park contact.

- Sensitive, Park Only: Not for release to the public or NPS. This information is considered sensitive by national or park-level standards and may be of questionable quality or be proprietary. This data will not be accessible to a federal or non-federal entity, the general public, or NPS personnel unless approved by the park contact.

Network staff and park contacts will review material for sensitivity. Information is considered sensitive when it (1) describes the specific location, or (2) provides enough information to infer the specific location (e.g., a species is mentioned and there is only one location in the park where it could possibly exist) of the following:

- Wells
- Endangered, threatened, rare, or commercially valuable species
- Minerals
- Paleontological objects
- Objects of cultural patrimony
- Archaeological resources and/or artifacts
- Ethnographic sites
- Caves

Data and information classified as sensitive will be stored in a folder labeled “sensitive” within the appropriate directory as defined in CHDN_SOP_02-0004, Directory Structure, and the appropriate notation made in the file/document properties. This information will be made available only to park and network staff on an individual basis, as approved by the project manager, network coordinator, or designated park point-of-contact. Details related to management of potentially sensitive data that is collected or otherwise needed by non-NPS staff (e.g., cooperators, interns, and student techs.) will be specified in the respective project or protocol. Approval for sharing this information will rest upon the network lead and the park contact.

9.2 Data Ownership

Propriety refers to who holds, possesses, or owns exclusive rights to the data. Any copyrighted materials are assumed to be proprietary. Information is considered non-proprietary if it is publicly owned and distributed by any governmental agency or if written permission by a private party states that it will forego any rights of ownership. The following questions apply when considering whether information is proprietary:

- Is it copyrighted?
- Are there other legal restrictions?

- Did the NPS or network pay a third party for access to the data or does the third party normally sell this information?
- Has the data been collected by a non-NPS Principal Investigator and does this person plan to publish the results in a professional journal?

Answering “Yes” to any of the previous questions suggests the information is likely to be proprietary and/or intellectual property.

The CHDN will not distribute any propriety information outside of the NPS without the consent of the owner. Minimum metadata or better will accompany all information distributed by the CHDN either in the digital file or as a separate metadata file.

Additional information on data ownership and sharing can be found in Chapter 9 of the NDMP (NPS 2008).

9.3 Related Guidance Documents, SOPs and/or TIs (see latest version on CHDN website)

- CHDN_SOP_09-0001; Approving Information for Distribution
- CHDN_SOP_02-0003; CHDN Web site Management
- CHDN_SOP_02-0004; Directory Structure

Chapter 10. Data Dissemination

Providing well-documented data in a timely manner is one of the most important goals of the I&M program and critical to the success of the program. The primary purpose of data and information gathered by the I&M program is to provide information on the long-term trends of the condition of National Park System resources to assist in park management decisions. Park resource staff are the primary recipients of the data. The program is structured using a board of directors and a technical committee comprised of key park personnel that provide a direct avenue for data transfer in both directions. Researchers, educators, and the general public also need access to program data; however, access to program data by other than park staff will be regulated in such a manner as not to disclose sensitive data without proper authorization.

10.1 National Applications

The Natural Resource Program Center (NRPC) has built and maintained a variety of information systems to manage data on species, water resources, GIS layers, reports and publications, and research permits, to name a few. These systems have captured valuable information and provided essential tools for searching and storing natural resource information related to parks.

10.1.1 NPS Data Store

The NPS Data Store is a Web-based system designed to integrate data dissemination and metadata maintenance for natural resource, GIS, and other program data sets, digital documents, and appropriate digital photos. The Data Store utilizes the NPS Metadata Profile for managing GIS and other data for both internal use and publication to the NPS GIS Clearinghouse and NPS Focus Digital Library. The NPS Data Store is accessible to park personnel and the public. All non-sensitive GIS data, final data sets, and associated compliant metadata resulting from CHDN efforts are posted on the NPS Data Store. Detailed guidelines for uploading data are located on the NPS Data Store are found at http://science.nature.nps.gov/nrdata.

10.1.2 NPSpecies

NPSpecies is the NPS database for storing, managing, and disseminating information on all organisms in NPS units. The application was designed to document the presence or absence of a species by documentary evidence, physical evidence, or observations. NPSpecies is not intended to document density and distribution of a species. Evidence confirming the presence of a biological entity resulting from CHDN efforts is entered into NPSpecies if no such evidence exists or if the evidence is more substantial than existing evidence. Specific CHDN requirements and general guidelines are found in CHDN_SOP_06-0003, NPSpecies Guidelines. Additional information about NPSpecies can be found at http://science.nature.nps.gov/im/apps/npspp.

10.1.3 NatureBib

The Natural Resource Bibliography, NatureBib, is the NPS database for citing park and network natural resource-related documents, publications, and references. Full-text documents of cited references are made available in the database. All non-sensitive documents and final reports resulting from CHDN efforts are posted in NatureBib in accordance with the guidance outlined in CHDN_SOP_06-0002, NatureBib Guidelines. Additional information about NatureBib can be found at http://science.nature.nps.gov/im/apps/nrbib.

10.1.4 NPSTORET

NPSTORET is a series of Access-based templates modeled around the Natural Resource Database Template and the EPA's STORET (short for STOrage and RETrieval) for the repository for water quality, biological, and physical data related to water quality and quantity. Use of these templates facilitates archiving NPS data in STORET. The Water Resource Division handles the actual transfer of data to the EPA. NPS Director's Order 77 (NPS n.d.) indicates that the NPS should archive water quality data in STORET, and the NPS Water Resources Division (WRD) requires that any data collected as part of a funded WRD project be archived in STORET. Guidelines for the CHDN use of NPSTORET will be developed during protocol development. Additional information on NPSTORET can be found at http://www.nature.nps.gov/water/infoanddata.

10.2 Network Web sites

The CHDN Web site provides, to the public, general information about the network and its parks and the I&M program, including information on inventories, monitoring, data management, and reports and publications as directed by WASO's general Web page templates. The Web site will be used to serve products such as results of I&M vital signs monitoring (e.g., executive briefs, progress reports, trend reports, etc.). An intranet portal is also in place in which park and network staff can share information and data that are in draft phase and not appropriate for public posting.

The network will enhance its Web site over time to deliver reports and provide supplemental background data and information. In the future, this may include Web-based internet map services and functions for user-controlled queries and summaries of network data. The Web site offers easy access to park managers for up-to-date information generated by the network and other related programs. General guidelines for CHDN Web site management can be found in CHDN_SOP_02-0003, Web site Management.

10.3 Collections

Products (e.g., specimen vouchers, photographic film) collected under park permits during CHDN-specific projects will be available at the park museum or alternative repository. Information associated with these products is located in the ANCS+ database. Information requests related to collections will be coordinated with the park curator or resource manager.

10.4 Communications Plan

The CHDN is also in the development phase of a formal communications plan. The communications plan will enhance awareness and communicate the efforts and findings of the I&M program within the CHDN to both internal and external audiences by:

- Engendering support for the I&M program and vital signs monitoring by promoting its benefits
- Identifying and conceptualizing messages, products, and strategies to facilitate communication between the CHDN and its audiences
- Enhancing communication between and within parks in CHDN
- Assisting the transfer of critical information between I&M scientists and the non-scientific community

The communications plan will have guidelines and instructions specific to topic and target audience as well as the media and method of communication. These guidelines and instructions will become a part of the CHDN quality management system.

10.5 Sensitive Data

Special care must be taken not to post data in any venue that may jeopardize park resources. Throughout its life cycle, data will be evaluated for its sensitivity, and the sensitivity will be documented in the metadata as describe in Chapter 8, Data Documentation. Sensitivity will be checked prior to the release of any data.

Additional information on data dissemination can be found in Chapter 10 of the NDMP (NPS 2008).

10.6 Related Guidance Documents, SOPs and/or TIs (see latest version on CHDN website)

- CHDN_SOP_02-0003; CHDN Web site Management
- CHDN_SOP_06-0001; CHDN NatureBib Guidelines
- CHDN_SOP_06-0002; CHDN NatureBib Full-text Document Management
- CHDN_SOP_06-0003; CHDN NPSpecies Guidelines
- CHDN_SOP_06-0004; CHDN Data Mining Strategies
- CHDN_SOP_09-0001; Approving Information for Distribution
- CHDN_SOP_11-0001, Monitoring in Parks (Planned FY10)

Chapter 11. Records Management and Curation

CHDN is responsible for managing the documents, photographs, digital files, and other products resulting from network projects, administration, and activities. The potential for loss of documents can come from a variety of sources, including catastrophic events (e.g., fire, flood, and earthquake), user error, hardware failure, software failure, data corruption, theft, or intentional acts of vandalism. A fundamental responsibility of network data management is to establish procedures that will help insure CHDN information will be accessible and usable for future generations. The establishment of a useable filing system, some simple filing practices, and knowledgeable trained staff are key factors in managing active records. Effective management of active records also will help to control the records once they become inactive by reducing the need to identify and sort records before they are destroyed or transferred.

11.1 Records

The Federal Records Act (44 U.S.C. 3301) (Federal Records Act 2006) defines records as:

> *"… all books, papers, maps, photographs, machine readable materials, or other documentary materials, made or received by an agency of the United States Government under Federal law or in connection with the transaction of public business and preserved as evidence of the organization, functions, policies, decisions, procedures, operations, or other activities of the Government or because of the informational value of data in them".*

The Records Management Handbook (NPS 2005) defines resource management records as:

> *"Any record regardless of media that documents an NPS cultural, natural, or informational resource."*

The handbook also states that:

> *"Resource management records are valuable, continuously active records that document all cultural, natural, and informational resources that are found within every unit of the National Park Service."*

Resource management records meet the definition of permanent federal records as outlined in the Federal Records Act. Therefore, the same records management and curation policies required for administrative records are required for resource records.

11.2 Records Management

All electronic files will be named according to guidelines and standards outlined in CHDN_SOP_08-0001, File Naming Conventions. These standards apply to all CHDN electronic files created or maintained by staff or cooperators. Electronic files will be managed within a hierarchical set of file directory structures as outlined in CHDN_SOP_02-0004, Directory Structure. These SOPs are part of the quality management system that provides for a means for all users to identify deficiencies or improvements to the system. These corrective action procedures insure the records management system remains current and everyone involved stays informed as the needs change.

CHDN Central Files include administrative files, data, publications, and reports and are organized according to the NPS Records Disposition Schedule (NPS 2003) suggested file names. For example, a CHDN file section and folder labels might include "N14 Records of Animal and Plant Life" and "N1415 Amphibians and Reptiles," respectively, with more specific project file names following this hierarchy. Slight adjustments to this system may be necessary where the I&M Program structure differs from park structure (e.g., CHDN will not have files related to a superintendent, but will have Board of Director files).

11.3 Disposition

The legally required disposition for records no longer needed is the "Records Management Schedule" is Appendix B of NPS-19 (NPS 2001a). Depending on the record series, the required disposition action might be one of the following:

1. Send records to an intermediate records storage facility such as a federal records center (FRC) where they will remain for a defined period of time as indicated in the "Records Management Schedule" after which a review of their status and disposition will occur.

2. Transfer records to an archival storage facility if they have permanent value.

3. Destroy records if they are temporary, are determined by agency staff to have no further value, and they have met the required retention period defined in the Records Management Schedule.

According to the schedule, records will be either temporary, permanent, or their disposition instructions will indicate DISPOSITION SUSPENDED - DO NOT DESTROY. Temporary records have value for only a specified period of time. The National Archives and Records Administration has approved their destruction or deletion. Permanent records have enduring value and must be preserved forever. They have been appraised by the National Archives and Records Administration as having sufficient value to warrant preservation beyond the time they are needed by the agency to meet its administrative, legal, fiscal, or program purposes. These records will be maintained by the National Archives and Records Administration or by agency archival repositories forever. Permanent records cannot simply stay in the office of creation or use for two reasons: these records have mandatory preservation issues that should be dealt with by professionals, and by law, federal records of enduring value must be available to the public for research use. DISPOSITION SUSPENDED is noted in the schedule for record series that should not be destroyed or transferred to the National Archives and Records Administration. These record series have long-term value to the National Park Service for managing its cultural, natural, and informational resources. The CHDN will store permanent records at the network facility as necessary until storage space becomes limited or until files are no longer needed for ongoing management of NPS resources. Permanent records will then be sent to the National Archives and Records Administration as indicated by Appendix B of NPS-19 (NPS 2001a). This transfer of records involves both a physical and legal transfer of the records. Instructions for transferring records to the National Archives are included in Appendix G of the Records Management Handbook (NPS 2005).

Additional information on records management and curation can be found in Chapter 11 of the NDMP (NPS 2008).

11.4 Related Guidance Documents, SOPs and/or TIs (see latest version on CHDN website)

- CHDN_SOP_02-0004; Directory Structure
- CHDN_SOP_08-0001; File Naming Conventions
- CHDN_SOP_11-0001, Monitoring in Parks (Planned FY10)

Chapter 12. Project Tracking

12.1 Tracking

The CHDN will develop and implement a process for tracking I&M projects, including project status, data, and the products of analysis. This process will support program coordination, annual reporting, and improve accountability for network natural resource inventory and monitoring efforts. All projects will be tracked using a database located on the CHDN file server. This will serve as the primary organizational tool for cataloging and searching information for ongoing and completed network projects. This database will be used to:

- Maintain a list of projects, both ongoing and completed
- Provide a method of tracking product deliverables
- Manage project codes used to tie information to other NPS tracking systems (e.g., RPRS, PMIS, PEPC, RAMS)

12.2 Documentation

Projects will be documented by creating project-specific protocol narratives and SOPs. These documents must always accompany the distribution of monitoring data. The network's project-tracking database will trace the project narrative and SOPs by version number and should be updated whenever any narrative or SOP document is modified. The protocol narrative and SOPs will not be distributed without a log of changes from the project-tracking database. Long-term monitoring projects may require additional documentation for items such as algorithms, output files, and analytical products, which may reside in different systems and formats. Data-use, data-request histories, and information on secondary research or publications resulting from long-term monitoring projects should be maintained. All files for any particular project will be stored in a project directory structure as defined in CHDN_SOP_02-0004, Directory Structure. Project file names will adhere to the naming conventions established by CHDN_SOP_08-0001, File Naming Conventions. Individual file metadata will use document properties as described in Chapter 8.

12.3 Related Guidance Documents, SOPs and/or TIs (see latest version on CHDN website)

- CHDN_SOP_02-0004; Directory Structure
- CHDN_SOP_08-0001; File Naming Conventions
- CHDN_SOP_12-0001, Project Tracking (Planned FY10)

Chapter 13. Implementation

The data management plans for each of the 32 I&M networks are the first comprehensive documents of their kind in the NPS and contain practices that may be new to NPS staff and cooperators. However, almost every requirement stems from federal law and policies, Executive Orders, Director's Orders, or national I&M program guidance. The DMP helps put these requirements into context, and provides operational and functional guidance for achieving them.

13.1 Quality Management System

The main body of the NDMP (NPS 2008) broadly addresses relevant subjects, but the details and implementation for any given network are expanded upon in the individual network-level data management plan and associated SOPs. Although these documents can function independent of each other, the intent is to provide a narrowing prospective beginning from the laws and policies and ending at the systematic instructions for a procedure or task.

The CHDN DMP is a living document. It must evolve as technology changes. The quality management system established by this document, and supported by the SOPs and TIs, establishes a review period for procedures and instructions. This plan is also subject to modifications, as necessary, when deficiencies or improvements are identified by anyone using it.

Implementation will require education and training in order to familiarize network staff, park staff, and cooperators with the quality management system so they understand that they too have a responsibility to identify improvements that will provide the most accurate well documented data that will yield dividends for future generations. These efforts will begin in 2009 and be led, at least initially, by I&M data management staff, with participation by interested parties at all parks actively encouraged.

13.2 Milestone Goals

The network-level plan will be republished in three years (2012), and then every five years afterward. Plan appendices, including SOPs, detailed guidelines, reference manuals, policy statements, etc., will likely require more frequent updates to account for changes and revisions. Goals for the first three years will include:

- Acceptance and understanding by all staff of targeted programs and their cooperators of the fundamentals of data and information management, including
  - File and folder management
  - Network shared drive organization
  - Documentation (e.g., metadata formats and process)
  - Quality assurance and quality control
  - Archive storage
- Improving data management practices by implementing:
  - Accepted database design standards
  - Thorough testing of databases, data collection methods, and their integration prior to field work
  - Quality assurance and control procedures at every stage of project development

- Development of common SOPs and guidance documents for multiple protocols
- Ensure that detailed specifications for data management that is consistent with the national and CHDN DMPs are included in every vital signs monitoring protocol
- Development of procedures and outlets for communication within and among CHDN network parks and with the public

Beyond the first three years, goals should include the development and assessment of:

- Procedures to facilitate the summarization and reporting of monitoring data
- Framework and gateway for integration of monitoring data with other agencies or networks
- Methods for improving file management (e.g., a content management system), database administration and security (e.g., migration to SQL-Server), integration into the network of off-site users, and other needs identified in the DMP

Literature Cited

Deming, W. Edwards. 2000. The New Economics for Industry, Government, Education, Second edition. First MIT Press, Cambridge, MA.

Department of the Interior. 2008. Information Technology Security Policy Handbook: Version 3.1.

Dublin Core. 2009. Dublin Core Metadata Initiative. (http://dublincore.org/). Accessed 23 February 2009.

Federal Records Act. 2006. Disposal of Records, Title 44, Chapter 33, Section 3301. (http://www.archives.gov/about/laws/#statutes). Accessed 23 February 2009.

National Institute of Standards and Technology. 2008. Security assessments: Tools for measuring the effectiveness of security controls. NIST Special Publication 800-53A (http://csrc.nist.gov/publications/nistbul/Aug2008_SP800-53a.pdf). Accessed 27 January 2009.

National Park Service. n.d. Natural resource protection (Under development). Director's Order #77. (http://data2.itc.nps.gov/npspolicy/DOrders.cfm). Accessed 23 February 2009.

National Park Service. 1998. National Parks Omnibus Management Act of 1998. Public Law 105-391. (http://www.nps.gov/legal/laws/105laws.htm). Accessed 27 January 2009.

National Park Service. 2001a. Director's Order 19: Records management. U.S. Department of the Interior, National Park Service, Washington, D.C.

National Park Service. 2001b. Director's Order 70: Internet and Intranet Publishing.

National Park Service. 2002. Director's Order #11B: Ensuring quality of information disseminated by the National Park Service. (http://www.nps.gov/policy/DOrders/11B-final.htm). Accessed 23 February 2009.

National Park Service. 2003. Director's Order 19, Appendix B: Records Disposition Schedule. U.S. Department of the Interior, National Park Service, Washington, D.C. (http://home.nps.gov/applications/npspolicy/DOrders.cfm) Accessed 24 February 2009.

National Park Service. 2005. Records Management Handbook. (http://home.nps.gov/applications/npspolicy/DOrders.cfm) Accessed 24 February 2009.

National Park Service. 2006a. Automated National Catalog System. (http://www.nps.gov/history/museum/publications/ancs.html). Accessed 23 February 2009.

National Park Service. 2006b. NPS Digital Photo Metadata Standard. (http://npsfocus.nps.gov/docs/guide/metadata/MetadataStandards.html). Accessed 20 February 2009.

National Park Service. 2006c. NPS Natural Resource Web Manual, National Park Service. (http://www1.nrintra.nps.gov/webnotes/NRWebManual/index.htm). Accessed 23 February 2009.

National Park Service. 2007. Natural Resource Database Template Version 3.2 documentation. Natural Resource Program Center, Office of Inventory, Monitoring, and Evaluation, Fort Collins, CO. 58 pg. (http://science.nature.nps.gov/im/apps/template/index.cfm). Accessed 27 January 2009.

National Park Service. 2008. Data management guidelines for inventory and monitoring networks. Natural Resource Report NPS/NRPC/NRR—2008/035. National Park Service, Fort Collins, CO.

U.S. Congress. 1916. National Park Service Organic Act. 16USC 1-4; Aug. 25, 1916, ch. 408, 39 Stat. 535.

Glossary

Certified data and metadata Completed data and documentation for short-term projects, or one season of completed data for long-term monitoring projects. Certification is a confirmation by the project leader that data have passed all quality assurance requirements and are complete and ready for distribution. Metadata records include the detailed information about project data needed for proper use and interpretation.

Data Distinct measurements or observations of a variable, usually formatted in a special way. They include symbols or representations of facts or ideas that can be communicated, interpreted, or processes by manual or automated means.

Tabular data are usually organized into logical tables of records and fields, arranged in a matrix of rows and columns. Tabular data can be displayed, manipulated, and stored as simple text files or in applications software (e.g., spreadsheets, relational databases).

Spatial data are any data that reference geographic coordinates. GIS data always contain these references; tabular data that contain spatial references are also considered spatial. The terms *GIS data* and *spatial data* are often used interchangeably.

Raw data are data in their original form, i.e., data that have not been altered, summarized, or grouped into broader categories. Raw data can exist in many forms: as hand-written information on field data forms and in notebooks; as unaltered photographs; sound and video recordings; remote sensing imagery; or Global Positioning System (GPS) files.

Derived data are raw data that have been processed, or converted to another form using some automated or manual process. Raw natural resources data are often processed and packaged for summation, statistical analysis, and graphical display, or the production of maps and other information products.

GIS (geographic information system) data contain information about the location and shape of, and relationships among, features on the surface of the earth and are usually stored as geographic coordinates and topology. Topology is used to compare the geographic locations of features relative to one another (e.g., roads connected to a highway, two vegetation polygons adjacent to one another).

Legacy data are data that are at risk of becoming lost, unusable, or obsolete due to software or metadata limitations.

Non-programmatic data are data obtained from an external source, but that are of value to a program or project (e.g., county land parcel information, weather data managed by the National Weather Service, or taxonomic information managed by the U.S. Department of Agriculture).

Sensitive data are data that through loss, unauthorized access, or modification, could be used in such as way as to adversely affect valuable resources, the national interest, the conduct of federal

programs, or individual privacy. Examples of sensitive natural resources data might include the locations of rare flora or fauna species, caves, or cultural sites.

Data set A grouping of related data, such that the assemblage of the information will be meaningful to prospective users.

Differential backup is a process where all computer files that have been modified since the last full backup of files are saved.

Document Recorded information regardless of physical form or characteristics. Often used interchangeably with the term "record."

Edit log A means of tracking changes to certified data.

Information Created from data as a result of processing, manipulating, synthesizing, or organizing data in a way that provides interpretation or meaning.

Inventories Extensive point-in-time effort to determine location or condition of a resource, including the presence, class, distribution, and status of plants, animals, and abiotic components such as water, soils, landforms, and climate. Inventories contribute to a statement of park resources, which is best described in relation to a standard condition such as the natural or unimpaired state. Inventories may involve both the compilation of existing information and the acquisition of new information. They may be relative either to a particular point in space (synoptic) or time (temporal).

Master database Central repository for project data, used for viewing, summarizing, and analysis. Contains only data that have passed all quality assurance/quality control.

Metadata Information about data. A metadata document contains specific and detailed information about a data set, including who, what, where, when, why and how the data were collected, analyzed, or manipulated. Metadata are considered an essential component of any good data set.

Monitoring The collection and analysis of repeated observations or measurements to evaluate changes in condition and progress toward meeting a management objective. Detection of a change or trend may trigger a management action, or it may generate a new line of inquiry. Monitoring is often done by sampling the same sites over time, and these sites may be a subset of the sites sampled for the initial inventory.

National databases and repositories Applications and repositories maintained at the NPS national level, primarily for the purpose of integration among NPS units and for sharing information with cooperators and the public.

Natural Resource Database Template (NRDT) is a core set of database tables that serves as a foundation for building relational databases for the NPS Inventory and Monitoring Program.

Natural Resource Program Center (NRPC) The core of the NPS Natural Resource Stewardship and Science Directorate. NRPC staff are located in Fort Collins and Lakewood,

Colorado, and in Washington, D.C. NRPC has five divisions: Air Resources; Biological Resources Management; Environmental Quality; Geologic Resources; and Water Resources. NRPC also includes the Office of Inventory, Monitoring and Evaluation; the Office of Education and Outreach; and the Office of Natural Resource Information Systems.

NatureBib The National Park Service bibliographic database, which is used to catalog, search, and manage natural resource-related information sources pertaining to national parks.

NPS Data Clearinghouse – the central repository for NPS GIS data available to the public, implemented through the NPS Focus gateway.

NPS Data Store An online graphical search interface that links dataset metadata to a searchable data server on which datasets are organized by NPS units, offices and programs.

NPS Focus A decentralized digital imagery and data management system, implemented through a central Internet portal sponsored by the NPS Office of the Chief Information Officer. Includes to the NPS Data Clearinghouse.

NPSpecies The master species database for NPS. The database lists the species that occur in or near each park, and the physical or written evidence for the occurrence of the species (e.g., references, vouchers, and observations). NPSpecies is implemented online through secure and public servers.

NPSTORET The NPS version of the EPA (Environmental Protection Agency) STORET database, used for transferring water quality data collected by NPS programs to STORET. (see STORET)

Protocols Highly detailed, formal documents that explain how data are to be collected, managed, analyzed and reported. Protocols are a key component of consistency and quality assurance for natural resource monitoring programs.

Relational Database Management System (RDMS or RDBMS) A system of storing data in related tables, which allows data to be structurally organized for maximum efficiency and a minimum of redundancy.

Research Permit and Reporting System (RPRS) A web-based tool that provides a researcher with applications, procedures, and guidelines for submitting and obtaining a scientific research and collecting permit within NPS units. RPRS also includes Investigator Annual Reports, a report required by all receiving a permit.

Sensitive data Any information which, through loss, unauthorized access, or modification could adversely affect a park resource or program, or the privacy of individuals.

Spatial data Information about the location, shape, and relationships among geographic features, usually stored as coordinates and topology. Can be stored in tabular or GIS file format.

Standard Operating Procedures (SOPs) Detailed step-by-step instructions that outline a formal set of procedures for performing specific tasks.

STORET (STORage and RETrieval) A database application maintained by the U.S. Environmental Protection Agency that contains raw biological, chemical and physical data on surface and ground water quality collected by federal, state and local agencies, Indian Tribes, volunteer groups, academics, and others. (see NPSTORET)

Structured Query Language (SQL) A language used to query and process data in a relational database. All database systems designed for client/server environments support SQL.

Vital Signs Ecological elements or processes chosen to represent the overall health or condition of park ecosystems, known or hypothesized effects of environmental stressors, or elements of value to humans, and are the subject of long-term monitoring by the Inventory & Monitoring Program.

Working database – A project-specific database for entering and processing data for the current season (or other logical period of time). May also constitute the master database for short-term projects.

The Department of the Interior protects and manages the nation's natural resources and cultural heritage; provides scientific and other information about those resources; and honors its special responsibilities to American Indians, Alaska Natives, and affiliated Island Communities.

NPS D-102, March 2009

National Park Service
U.S. Department of the Interior

Natural Resource Program Center
1201 Oakridge Drive, Suite 150
Fort Collins, CO 80525

www.nature.nps.gov