Appendix K Table of Contents

1 - Introduction and Background K-1
1.1 – Inventory and Monitoring Program Overview K-1
1.1.1 – Data Management Goals and Objectives K-2
1.1.2 – What is Data? – Data Defined K-2
1.2 – Mediterranean Coast Network K-2
1.2.1 – Scope of the Information Management Plan K-3
1.2.2 – Data Covered by the Information Management Plan K-4
1.2.3 – Prioritization of Data K-4
1.2.4 – Revisions K-5
2 - Data Management Roles and Responsibilities K-6
2.1 – Data Stewardship K-6
2.1.1 – Data Management Roles K-6
2.2 - Data Management Coordination K-10
3 - Data Management Program Overview K-12
3.1 – Project Work Flow K-12
3.1.1 – Planning and approval K-12
3.1.2 – Design and testing K-12
3.1.3 – Implementation K-14
3.1.4 – Product integration K-14
3.1.5 – Evaluation and closure K-14
3.2 – Data Life Cycle K-14
3.3 – Sharing Data Products K-16
3.3.1 – Data Distribution K-16
4 - Data Management Infrastructure K-20
4.1 – Data Infrastructure K-20
4.1.1 – National-level Architecture K-20
4.2 – File Management Strategy K-23
4.3 – Database Formats and Standards K-23
4.3.1 – Centralized Database Components 24
4.3.2 – Data Standards K-25
4.3.3 – National Standards Compatability K-25
5 - Data Acquisition and Processing K-28
5.1 – Program Data Collection K-28
5.1.1 – Data Discovery K-28
5.1.2 – Field Studies K-29
5.1.3 – Changes to Procedures & Protocols K-29
5.2 – Non-Program Data Collection K-29
5.2.1 – Non-Program Data Collection: Internal (NPS) Sources K-29
5.2.2 – Non-Program Data Collection: External Data Sources K-29
6 - Quality Assurance and Quality Control K-31
6.1 – QA/QC Overview K-31
6.1.1 – NPS Mandate for Quality K-31
6.1.2 – Goals and Objectives K-32
6.2 – QA/QC Mechanisms K-32
6.2.1 – Data Collection K-32
6.2.2 – Data Entry K-33
6.2.3 – Verification and Validation Procedures K-33
6.2.4 – Version Control K-34
6.2.5 – Data Quality Review and Communication K-34
7 - Data Documentation K-36
7.1 – Integrated Metadata System Plan and Tools K-36
7.1.1 – Metadata Tools K-36
7.2 – Dataset Documentation Process K-38

7.3 – Project Documentation K-39
8 - Data Analysis and Reporting K-40
8.1 –Trends and Analyses K-40
8.2 – Data Reporting K-40
8.2.1 – Types of Reports K-40
8.2.2 – Automated Reporting and Data Summaries K-41
9 - Data Dissemination K-43
9.1 – Data Ownership K-43
9.2 – Restrictions to Sensitive Data K-43
9.3 - Data Requests K-44
9.4 – Data Distribution K-45
9.4.1 – Websites/Repositories K-45
9.5 – Information Analysis and Feedback K-45
10 - Data Maintenance, Storage and Archiving K-46
10.1 – Digital Data K-46
10.1.1 – Digital Data Maintenance K-46
10.1.2 – Data Storage and Archiving K-47
10.2 – Analog Data K-48
10.2.1 – Park Curatorial Staff K-48
10.2.3 – Analog Data Maintenance K-49
10.2.3 – Analog Data Storage and Archiving K-49
References K-50

1 - Introduction and Background

One of the primary goals of the National Park Service (NPS) is to "improve park management through greater reliance on scientific knowledge." The ability of resource managers to make decisions about the resources within and around park boundaries depends on our knowledge of the surrounding ecosystem. These systems are constantly evolving and our knowledge of them and how they work must also evolve for us to maintain an adequate understanding of their dynamics.

Understanding the ecosystem within our parks often begins with the collection of natural resource data collected as part of stewardship projects within the National Park system. These projects include park planning, inventories of resources, short-term and long-term monitoring, restoration, control of invasive species and other species management, fire management, trail and road maintenance, law enforcement, and the communication of natural resource information to the public (interpretation). The data collected by researchers are analyzed, synthesized and used to model various aspects of the ecosystem. Results and interpretation of the analyzed data becomes information for park managers so they can make informed decisions about the vital natural resources within the park.

1.1 – Inventory and Monitoring Program Overview

The NPS Inventory and Monitoring (I&M) Program represents a long-term commitment by the NPS to assess and document the status and trends of park ecological resources. In 1998, the National Parks Omnibus Management Act established a framework for the I&M Program, integrating natural resource monitoring and other scientific activities into the management processes of the National Park system.

The Omnibus Management Act charges the Secretary of the Interior to "continually improve the ability of the National Park Service to provide state-of-the-art management, protection, and interpretation of and research on the resources of the National Park System," and to "… assure the full and proper utilization of the results of scientific studies for park management decisions." Section 5934 of the Act requires the Secretary of the Interior to develop 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."

To carry out this mission, the NPS initiated a service-wide, natural resource I&M program encompassing 270 parks with significant natural resources. The NPS grouped these parks into 32 networks with each network tasked to:

- document existing park vertebrates and vascular plants
- develop a management based, ecological monitoring program with a written plan and protocols
- create an information management plan that encompasses all aspects of the network program

With these tasks in mind, the I&M program's long-term goals are to:

- establish natural resource inventory and monitoring standards throughout the National Park system that transcend traditional program, activity, and funding boundaries

- inventory the natural resources and park ecosystems under National Park Service stewardship
- monitor park ecosystems to provide reference points for comparisons with other, altered environments
- integrate natural resource inventory and monitoring information into National Park Service planning, management, and decision making

- share National Park Service accomplishments and information with other natural resource organizations and form partnerships for attaining common goals and objectives

A modern information management infrastructure (e.g., staffing, hardware, software) will need to be developed to achieve the last two of these goals. This infrastructure will include procedures to ensure that relevant natural resource data collected by NPS staff, cooperators, researchers, and others will be recorded, quality-checked, analyzed, reported, archived, documented, cataloged, and made available to others for management decision-making, research, and education.

1.1.1 – Data Management Goals and Objectives

As the basic and most important products of scientific research, data and information represent a valuable, and often, irreplaceable resources (Michener and Brunt, 2000). Because field experiments and associated data collection are often time and budget consuming, management of data and information products plays an important role in any scientific program. Data management is an active process with the ultimate goal of ensuring and maintaining the integrity, utility, security, and availability of data. A good data management strategy will ensure that all data:

- meet acceptable accuracy standards
- are accessible to users
- are meaningful
- contain clearly defined relationships to other relevant data
- are protected from corruption, loss, and unauthorized changes
- are maintained with integrity – indefinitely

Good data management can be accomplished through data organization, documentation, and quality control procedures. For example, organization of data in well thought out directory and file structures ensures data are physically accessible to users. Good database design facilitates data exploration and analysis. Documentation of data, or metadata, including description of the extent and purpose of data; why, when, where data are collected; how and by whom data were collected; what was done to the data; and references indicating how data have been used, is critical in maintaining the usefulness and meaning of data through time. Quality control procedures are necessary to ensure data accuracy and quality beginning with data collection and continuing through data storage and archiving. Backup strategies and data access controls will protect from data loss, including corruption and unwanted changes whether inadvertent or malicious.

1.1.2 – What is Data? – Data Defined

Data consists of factual information organized for the purpose of analysis, reason, or decision-making. Scientific data contains observational information that is obtained following a specified method or protocol. It is the intention that the information, once analyzed, will contribute to the knowledge regarding the conditions, processes, and changes occurring within our ecosystem. Such information, though, must be properly collected, verified, and stored in order for it to be used and accepted by both the scientific and non-scientific community. Defining minimum standards and procedures for data managed under the auspices of the I&M program, as well as related data external to the program, will validate their use in scientific analysis and park decision-making.

The term “data” is often used in a broader sense to encompass other products that are generated alongside primary tabular and spatial data. Data and datasets (collections of similar data) can exist in a variety of conditions and states. Generally, data products fall into five categories from raw and derived data, to documents, reports, and administrative records (Table 1.1) and can come in a range of formats including *hard-copy documents* (e.g., reports, field notes, survey forms, maps, references, administrative documents), *electronic documents* (e.g., Word files, email, websites, digital images, databases, spreadsheets, tables, delimited files), *objects* (e.g., specimens, samples, photographs, slides), and *spatial data* (e.g., shapefiles, coverages, remote-sensing data).

1.2 – Mediterranean Coast Network

The Mediterranean Coast Network (MEDN) of National Parks includes Cabrillo National Monument (CABR) in San Diego, Channel Islands National Park (CHIS) off the southern California shoreline, and Santa Monica Mountains National Recreation Area (SAMO) straddling Los Angeles and Ventura Counties along the coast

Table 1.1. Categories of data products and project deliverables.

Data Category	Examples
Raw data	GPS rover files, raw field forms and notebooks, photographs and sound/video recordings, telemetry or remote-sensed data files, biological voucher specimens
Derived (compiled) data	Relational databases, tabular data files, GIS layers, maps, species checklists
Documents	Data collection protocols, data processing/analysis protocols, record of protocol changes, data dictionary, FGDC/NBII metadata, data design documentation, quality assurance report, catalog of specimens/ photographs
Reports	Annual progress report, final report (technical or general audience), periodic trend analysis report, publication
Administrative records	Contracts and agreements, study plan, research permit/application, other critical administrative correspondence

between Santa Monica and Oxnard, California. Together, the parks comprise examples of coastal and island southern California and share a Mediterranean-type ecosystem characterized by hot dry summers, cool wet winters, and evergreen sclerophyll shrub vegetation.

The network I&M program is organized according to the general guidance received from the national I&M program and includes a permanent network coordinator and data management specialist working in consultation with the network Board of Directors (BOD), consisting of the superintendents of each park with the network coordinator and the Pacific West Region I&M Coordinator participating as ex officio members, as well as an approved charter detailing the goals and mission of the MEDN I&M program. The BOD provides oversight and ensures individual park participation in the I&M planning process. A supporting technical committee comprised of the Chiefs of Resource Science at the three parks is the primary functioning body with responsibility for general program planning and implementation. Other park resource managers may be invited to participate on the technical committee on an ad hoc basis.

As one of the first Long-term Ecological Monitoring (LTEM) prototype parks, vital signs monitoring at CHIS has been on-going since 1991. Development and implementation of an I&M framework encompassing all three parks within the MEDN began in April, 2000. More information on the MEDN I&M program, including organization and monitoring goals, can be found in the MEDN I&M Vital Signs Monitoring Plan.

To facilitate the collection and dissemination of information, the MEDN I&M program must identify, catalog, organize, structure, archive, and provide high-quality natural resource data for park researchers, managers, and administrators. The network must also secure and maintain the quality of these data for over the long-term. Because of these complex data requirements, a process must be employed for the creation, preservation, and integration of data to make them interpretable and valuable. To accomplish this task, the MEDN I&M program has developed an information management plan to address the concerns and issues related to data management.

1.2.1 – Scope of the Information Management Plan

The MEDN I&M program approach to data management is to develop a plan that is “user friendly” to a varied audience from park natural resource managers to data managers. This may include providing guidance on data management practices at a number of different levels. To facilitate this guidance, the MEDN I&M Information Management Plan was developed along three basic principles:

- keep the plan simple, flexible, and evolving
- make it useful to all – from park GIS and data management staff, to regional technical staff, resource management staff, and cooperating scientists
- include the data users in the decision

making process whenever possible

This plan presents a broad strategy to implement and maintain a system that will serve the data and information management needs of the MEDN I&M program. This plan reflects the network's commitment to the establishment, maintenance, description, accessibility, and long-term availability of high-quality data and information. It also identifies the process by which the network will develop more detailed operational guidelines and where those guidelines will be documented. By providing for the preservation of the quality and integrity of resource data collected under the program, as well as "legacy" data currently existing, this plan will ensure that data and information will be available to assist resource managers in daily operations, make informed park management decisions, and facilitate scientific exploration and research. Specifically, the MEDN I&M Information Management Plan describes how the network will:

- support 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 data

This document will detail these and other data management activities, procedures, and resources needed to maintain, over the long-term, the integrity and availability of I&M and related natural resource data for the MEDN.

1.2.2 – Data Covered by the Information Management Plan

There are many potential sources of data and information that concern the condition of natural resources in the parks or network. The types of work that may generate these natural resource data include:

- inventory and monitoring studies
- protocol development pilot studies
- special focus studies done by internal staff, contractors or cooperators
- external research projects
- monitoring or research studies done by other agencies on park or adjacent lands
- resource impact evaluations related to park planning and compliance with regulations
- resource management and restoration work

Each dataset can be grouped into four major data management categories that are coordinated or managed by the MEDN I&M program:

1. **Data managed in service-wide databases**. The MEDN uses three data systems developed by the national-level I&M program (i.e., Nature Bib, NPSpecies, and Dataset Catalog – additional details on these databases can be found in Section 4).
2. **Data developed or acquired directly by the network as a result of inventory, monitoring, or other projects**. This category includes short- and long-term tabular databases as well as project-related protocols, reports, spatial data, and associated materials such as field notes and photographs provided to the MEDN by contractors or by park staff.
3. **Data that, while not developed or maintained by the MEDN, are used as data sources or provide context to other data sets**. Examples of this category include: GIS data developed by parks, other agencies or organizations; national or international taxonomic or other classification systems; and climate or hydrologic data collected by regional or national entities.
4. **Data acquired and maintained by network parks that the MEDN assists in managing**. The MEDN may provide data management assistance for high-priority data sets or those that may benefit from standardized procedures. Examples of this category include: a multi-park database for rare plant data; data sets of legacy natural resource monitoring data; and data on exotic invasive plant species.

1.2.3 – Prioritization of Data

Because I&M data will frequently be analyzed in context with other natural resource data collected outside the program, it is critical that these external datasets meets the same standards, procedures, infrastructure, and attitudes in regards to data management. However, one challenge will be prioritizing and managing workload and other resources to meet the needs of the program.

As the focus of the MEDN I&M program is on long-term monitoring and natural resource inventories, the first priority will be management of the data and information that are derived from these primary efforts. As time and resources permit, the network will work toward raising the level of data management for current projects, legacy data, and data originating outside the I&M Program. Greatest emphasis will be placed on projects that are just beginning to be developed and implemented, because inserting good data management practices into an existing project can be difficult and will generally meet with less success.

1.2.4 – Revisions

The MEDN I&M program Information Management Plan covers program needs based on information systems technology current in 2005. However, it is structured to adapt to changing technology. As changes to technology occur, changes must also be made to the plan and all associated data management documents (e.g., guidelines, SOPs). Revisions to this plan and associated documents will be made as needed to maintain a reliable and efficient data and information handling system.