

# **Appendix J. Data and Information Management Plan for the Mojave Desert Network**

Author: Robert E. Truitt

Suggested Citation: Mojave Desert Network. 2008. Data and Information Management Plan for the Mojave Desert Network. Mojave Desert Network – National Park Service. Boulder City, Nevada

\*The current version of this document can be found at:

[http://science.nature.nps.gov/im/units/MOJN/datamgmt/datamgmt\\_main.cfm](http://science.nature.nps.gov/im/units/MOJN/datamgmt/datamgmt_main.cfm)

# Contents

	Page
Contents .....	ii
Figures .....	v
Tables .....	vi
Topic .....	vii
Theme Keywords .....	vii
Placename Keywords .....	vii
Technical Information .....	vii
Acronyms .....	vii
Acknowledgments .....	viii
Executive Summary .....	ix
1.0 Introduction .....	1
1.1 Objectives .....	1
1.2 Scope .....	2
1.3 Purpose .....	3
1.4 Supporting Documentation .....	4
2.0 Infrastructure and System Architecture .....	8
2.1 Objectives .....	8
2.2 Scope .....	8
2.3 Supporting Documentation .....	9
3.0 Project Development and Data Management Workflow .....	10
3.1 Objectives .....	11
3.2 Scope .....	11
3.3 Supporting Documentation .....	13

# Contents

	Page
4.0 Data Management Roles and Responsibilities.....	14
4.1 Objectives .....	14
4.2 Scope.....	14
4.3 Supporting Documentation .....	15
5.0 Databases .....	15
5.1 Objectives .....	16
5.2 Scope.....	16
5.3 Supporting Documentation .....	16
6.0 Acquisition, Processing, and Reporting.....	17
6.1 Objectives .....	17
6.2 Scope.....	17
6.3 Supporting Documentation .....	18
7.0 Quality Assurance and Quality Control.....	19
7.1 Objectives .....	19
7.2 Scope.....	19
7.3 Supporting Documentation .....	21
8.0 Dataset Documentation.....	22
8.1 Objectives .....	23
8.2 Scope.....	23
8.3 Supporting Documentation .....	23
9.0 Data Ownership and Sharing .....	24
9.1 Objectives .....	24
9.2 Scope.....	24

# Contents

	Page
9.3 Supporting Documentation .....	25
10.0 Data Dissemination.....	25
10.1 Objectives .....	26
10.2 Scope.....	26
10.3 Supporting Documentation .....	26
11.0 Records Management and Object Curation .....	27
11.1 Objectives .....	28
11.2 Scope.....	28
11.3 Supporting Documentation .....	28
12.0 Project Tracking and Documentation .....	29
12.1 Objectives .....	29
12.2 Scope.....	29
12.3 Supporting Documentation .....	30
13.0 Implementation .....	30
13.1 Objectives .....	30
13.2 Scope.....	31
13.3 Supporting Documentation .....	31
14.0 References.....	32
15.0 Change History .....	34
16.0 Supporting Documentation .....	34

# Figures

	Page
Figure 1. Understanding data (Bellinger 2004) .....	x
Figure 2. Data and Information Management Plan flowchart .....	xv
Figure 3. Project workflow and data management activities .....	xviii
Figure 4. Simplified data flow diagram for water quality data.....	xix
Figure 5. Parks of the Mojave Desert Network. ....	3
Figure 6. Information entropy - without sound management practices (Michener 2000).....	4
Figure 7. Information entropy - with sound management practices .....	4
Figure 8. Higher-level management plan content.....	7
Figure 9. Functional and operational requirements (modified from Maarten and Burner 2002). ....	9
Figure 10. Conceptual Model of MOJN project management.....	11
Figure 11. MOJN staffing chart, showing existing staff (solid lines) and potential staff (dotted lines). ....	15
Figure 12. Data life cycle from a BLM perspective .....	20
Figure 13. Project workflow and quality management activities .....	21
Figure 14. Quality is important in all aspects .....	22
Figure 15. Guidance assistance decision tree .....	36
Figure 16. SOPs relevant to organizational tasks and activities .....	36
Figure 17. SOPs relevant to tabular data .....	36
Figure 18. SOPs relevant to spatial data. ....	37

## Tables

	Page
Table 1. Data that are provided on the MOJN and national I&M websites.....	xiii
Table 2. Roles and responsibilities related to network data management. ....	xvi
Table 3. MOJN Staff and duty station .....	3
Table 4. MOJN management sections and associated SOPs. ....	7
Table 5. MOJN staff duty Stations .....	8
Table 6. Change history table for the MOJN DMP. ....	34
Table 7. MOJN Management Sections and associated SOPs .....	35
Table 8. Table of MOJN SOPs and their development status. ....	37

## Topic

data, information, management, guidelines, specifications

## Theme Keywords

data management, information management, project management, natural resource information, ecological monitoring

## Placename Keywords

Arizona/California/Nevada/Utah State, Mojave/Great Basin deserts, Mojave Desert Network, M

## Technical Information

Version 1.0

Status: Draft

Revision Date: September 30, 2008

Author: Robert E. Truitt

Suggested Citation: Mojave Desert Network. 2008. Data and Information Management Plan for the Mojave Desert Network. Mojave Desert Network – National Park Service. Boulder City, Nevada

\*The current version of this document can be found at:

[http://science.nature.nps.gov/im/units/MOJN/datamgmt/datamgmt\\_main.cfm](http://science.nature.nps.gov/im/units/MOJN/datamgmt/datamgmt_main.cfm)

## Acronyms

DMP - Data and Information Management Plan

DEVA - Death Valley National Park

GRBA - Great Basin National Park

I&M - Inventory & Monitoring

JOTR - Joshua Tree National Park

LAME - Lake Mead National Recreation Area

MANZ - Manzanar National Historic Site

MOJA - Mojave National Preserve

MOJN – Mojave Desert Network

PARA – Grand Canyon-Parashant National Monument

## Acknowledgments

Many people have contributed to the production of this plan. The overall plan outline and components of all chapters were developed by the Inventory and Monitoring Program Data Management Planning Workgroup. The authors wish to thank all subsequent plan workgroups that contributed their thoughts and materials generously and without concern for credit. The collaboration of network data managers has been a highlight of producing this plan, and sets a precedent for cooperation in the data management tasks we all face in the upcoming years.

The authors thank the national-level Inventory and Monitoring Program data management team, who have provided the vision and created the tools that are essential to our work. In particular, Lisa Nelson, Danelle Malget, Chris Dietrich, Willene Hendon, Joe Gregson, Wendy Schumacher, Mark Wotawa, Simon Kingston, and Alison Loar have all provided outstanding technical support, and have been patient and open-minded in addressing our questions and suggestions.

Special thanks must be extended by the authors to Margaret Beer and John Gross for their words of wisdom, knowledgeable council, and copious patience.

We also acknowledge the National Science Foundation's Long Term Ecological Research Program, one of the first to integrate data and information management.

This plan would not have been what it is without the combined collaboration and tireless work of the network's data mining team, the Lake Mead NRA GIS staff, and the effort initiated by Craig Palmer (UNLV), the first MOJN Data Manager.

Finally, we want to thank Steve Fancy for his unwavering commitment to data management in the National Park Service.

## Executive Summary

I shall be telling this with a sigh  
Somewhere ages and ages hence:  
Two roads diverged in a wood,  
and I— I took the one less travelled  
by, And that has made all the  
difference –Robert Frost

*The road not taken* (Frost 1916)



Data management has been adopted traditionally through attrition (Brunt et al. 2002). With the implementation of the Natural Resource Challenge and the subsequent Inventory and Monitoring Program (I&M), the National Park Service (NPS) has instituted a process to mitigate that tradition. A cornerstone of the I&M Program is the strong emphasis placed on data management. All I&M networks, including the Mojave Desert Network (MOJN), expect to invest at least thirty percent of their available resources in data management. The MOJN Data and Information Management Plan (DMP [this document]) is one element in the network’s effort to integrate high-quality data and information management into the I&M program.

This plan is not limited to facts or data contained in the tables, fields, and values that make up a dataset. Its larger purpose is to describe the process for generating, preserving, documenting, and transmitting the context that helps data become information and makes it valuable and interpretable. As such, this plan covers both data—commonly defined as “facts or pieces of information” in scientific or academic literature—and information, defined variously as “knowledge communicated or received concerning a particular fact or circumstance” to “computer data at any stage of processing, as input, output, storage, or transmission” (Merriam-Webster 2006). In other words, this plan is not just concerned with the management of data and facts; it also intends to ensure that facts become information (e.g., interpretation of the data via analyses), which in turn translates into knowledge that we apply to manage the NPS lands that we are entrusted with as stewards. Therefore, it addresses pieces of information, the processing and preservation of those pieces, and the communication of knowledge derived from those pieces (Figure 1).

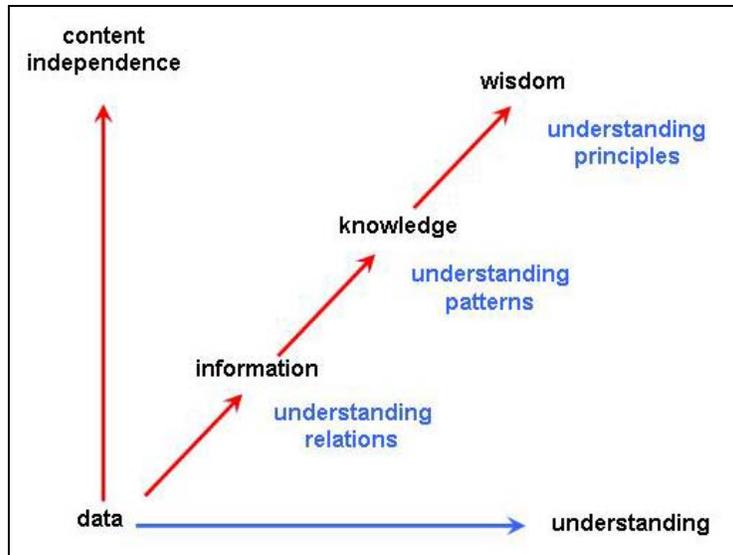


Figure 1. Understanding data (Bellinger 2004)

The central mission of the NPS I&M Program is to provide timely and usable scientific information about the status and trends of park resources to park managers. To meet this challenge, we need a data and information management system that can effectively produce, maintain, and distribute the products (knowledge) of scientific work done in our parks. Information is the common currency among the activities and staff involved in natural resource management in the NPS.

**National Park Service Mission**

*The National Park Service preserves unimpaired the natural and cultural resources and values of the national park system for the enjoyment, education and inspiration of this and future generations...*

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

The overarching goal of MOJN data management systems is to provide timely and usable scientific information about the status and trends of park resources to park managers. The success of our program hinges upon our ability to produce, manage and deliver this information, and the subsequent knowledge derived, to its intended audience. Our strategy for achieving this goal can be summarized as follows: ensure the quality, interpretability, security, longevity, and availability of our natural resource data. In implementing a data and information management system we will strive for the following:

- Confidence in the security and availability of natural resource data and related information
- Easy access to most information, and appropriate safeguards for sensitive information
- Awareness of the intended use and limitations of each dataset
- Infrastructure and documentation that encourages data exploration
- Compatibility of datasets for exploration and analysis at larger scales and across disciplines
- Implementation of standards and procedures that facilitate information management, and that reinforce good habits among staff at all levels of project implementation – project leaders, technicians, and volunteer data collectors
- A proper balance between the standards needed to ensure quality and usability, and the flexibility to meet specific needs and encourage innovation
- A natural resource culture which views data not as a commodity but as the lifeblood of our work

This plan outlines how we intend to implement and maintain a system that will serve the data and information management needs of MOJN. This plan reflects our commitment to establishing and maintaining a robust system for data management to ensure the availability and usability of high-quality natural resource information.

This plan and supporting documentation (Standard Operating Procedures, SOPs) describes how our network will:

- support I&M Program objectives
- acquire and process data
- assure data and information quality
- document, analyze, summarize, and disseminate data and information
- maintain nationally-developed data management systems
- maintain, store, and archive data

The goal of the MOJN's data management program is to maintain, in perpetuity, the ecological data, information, and knowledge that result from the network's resource I&M work. This plan is to describe the resources and processes required to ensure the following standards for data acquired or managed by MOJN:

- *Accuracy*: The quality of the data collected and managed by the I&M Program 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 I&M Program. To ensure that MOJN produces and maintains data of the highest possible quality, procedures are established to identify and minimize errors at each stage of the data lifecycle.

- *Security*: Digital and hard-copy data must be maintained in environments that protect against loss, either due to electronic failure or to poor storage conditions. MOJN digital data are stored in multiple formats on a secure server, and are part of an integrated backup routine that includes rotation to off-site storage locations. In addition, MOJN is working with NPS 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 datasets have become unusable over time either because the format is outdated (e.g., punchcards), or because metadata is insufficient to determine the data's collection methods, scope and intent, quality assurance procedures, or format. While proper storage conditions, backups, and migration of datasets to current platforms and software standards are basic components of data longevity, comprehensive data documentation is equally important. MOJN uses a suite of metadata tools to ensure that datasets 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 MOJN staff and cooperators are made available for decision-making, research, and education. Providing well-documented data in a timely manner to park managers is especially important to the success of the program. MOJN must ensure that:
  - data can be easily found and obtained
  - data are subjected to full quality control before release
  - data are accompanied by complete metadata
  - sensitive data are identified and protected from unauthorized access and distribution

The MOJN's main mechanism for distribution of the network's I&M data will be the World Wide Web, which will allow data and information to reach a broad community of users. As part of the NPS I&M Program, web-based applications and repositories have been developed to store a variety of park natural resource information (Table 1).

The MOJN's information acquires its real value when it reaches those who can apply it (Figure 1 above). If the web portals listed below (Table 1) do not meet a specific user's requirements, MOJN data management staff will work with users on an individual basis to ensure receipt of the desired information in the requested format.

Table 1. Data that are provided on the MOJN and national I&M websites.

Web Application Name	Data available at site
NPSpecies	Database of plant and animal species known or suspected to occur on NPS park units and as a species keyword search for reference materials ( <a href="#">NPSpecies Home Page</a> ).
NatureBib	Bibliography of park-related natural resource information ( <a href="#">NatureBib Home Page</a> ).
NPS IRMA	Portal to a variety of NPS information sources; will include NPSpecies, NatureBib and NPS Data Store links.
NPS Data Store	Park and network -related metadata and selected datasets (spatial and non-spatial) ( <a href="#">NPS Data Store Home Page</a> ).
NPSstoret	Database for water quality assessment ( <a href="#">NPSstoret home page</a> )
MOJN Websites	Through the use of the network’s inter- and intra-net web sites and the use of MS SharePoint, reports, summaries, outreach materials, as well as other monitoring data and information for MOJN projects and tools for data; data downloads; database templates will be made available ( <a href="#">MOJN Home Page</a> )

### Data Management Plan Model

Network data management plans have been written as an iterative process. Each of the networks has been placed into one of four groups, each group submitting their plans in a subsequent year. As each group of network data managers has submitted a draft plan, the groups have worked to identify and synthesize the salient elements of a complete plan. The first group of network data managers worked collaboratively to develop a plan, with 1-2 data managers working on each identified chapter of the plan. The second and third groups of data managers built off the initial work to fill gaps, revise materials, and build a sound set of chapters, ultimately developing a model that is comprehensive. Unfortunately, the resulting network plans are large (therefore discouraging for others to read and implement) and redundant (each network plan discusses the legal mandates, policies, and general data stewardship guidelines).

The last group of data managers have designed and written their plans around a new model for the data management plan. Instead of each network plan containing the same redundant materials (adding to its length) and necessitating that each network update its plan based upon new national guidance and legal mandates, the new model proposed:

- To produce a national-level data management plan guidance document that maintains the overarching documentation (what and why concerning data/information stewardship) and legal mandates relevant to each plan, and can be easily referenced in the development of a network data management plan.
- To produce a new network-level data management plan that is more applicable (how and when concerning data/information stewardship), easily understood, and does not require the lengthy background documentation and legal mandates.

The MOJN plan is written using this new model. For information concerning the national guidance and legal mandates refer to the draft national plan (National Park Service 2008). The network's implementation of that material is contained in this document and the supporting documentation (management sections and SOPs).

This plan is written as both a stand alone document and as a support document for the network's Final Vital Signs Monitoring Plan (Chung-MacCoubrey et al. 2008), for the management of data as well as the subsequently-produced information and knowledge. The plan by itself would be too daunting of a tome to be used or applied (as seen by the plans already developed by the preceding 26 I&M networks) by anyone other than the authors. Hence the plan is only a condensed or abbreviated link between the national data management plan guidance document (NPS 2008) and the more technically oriented and applicable supporting documentation (management sections and SOPs), that are appended to the plan (Figure 2). The supporting documentation are the dynamic guidance that will provide users (park and network staff, cooperators, and others) with the practical know-how to be applied for any particular data and/or information management procedure. The supporting documentation is composed of SOPs that have been arranged into categories of related procedures (i.e. management sections) as illustrated in Figure 2. The national guidance document contains the legal mandates and over-arching justifications, the network plan is the connection between the national guidance and the network level management sections and SOPs.

### **Data Management Roles and Responsibilities**

Data management is collaborative work that involves many persons with a broad range of expertise and abilities. All network staff have a role in data stewardship, and project datasets and products reflect all who have contributed. Table 2 summarizes the roles and responsibilities related to network data management, from field-based data collection, to final distribution and archiving. The fundamental role of the network data manager is to coordinate these tasks.

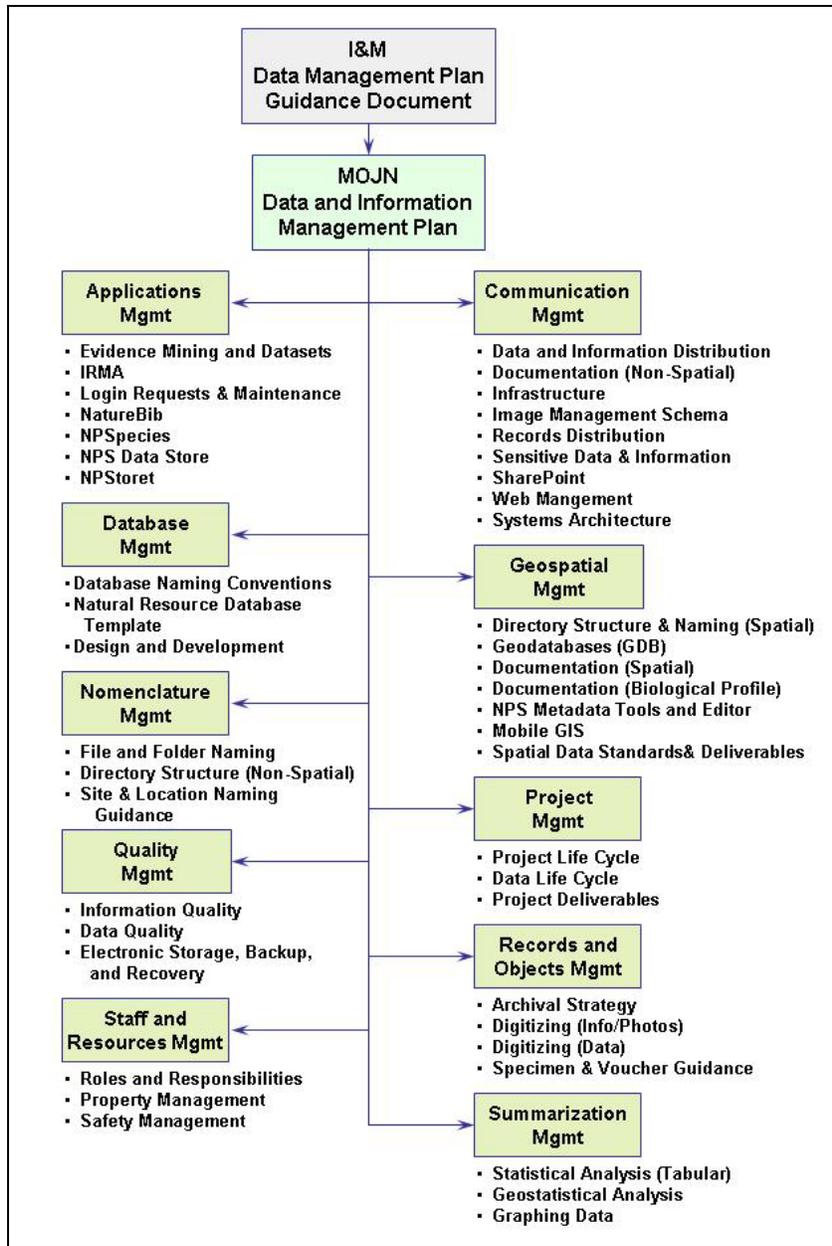


Figure 2. Data and Information Management Plan flowchart

Table 2. Roles and responsibilities related to network data management.

Role	Primary responsibilities related to data management
Project leader	Direct operations, including data management requirements, for network projects
Project crew leader	Supervise crew; communicate regularly with data manager and project leader
Project crew member	Collect, record, perform data entry, verify data; organize field forms, photos, other related materials
Resource specialist	Evaluate validity and utility of project data; document, analyze, publish data and associated information products
GIS specialist	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
Quantitative ecologist	Determine project objectives and sample design; perform and document data analysis and synthesis; prepare reports
Network data manager	Ensure program data and information are organized, useful, compliant, safe, and available
Network coordinator	Coordinate and oversee all network activities
Park or regional curator	Ensure project results (documents, specimens, photographs, etc.) are cataloged and accessioned into 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

### Data Sources and Priorities

There are multiple sources of significant data related to natural resources in the MOJN parks. The types of work that may generate these data include:

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

Because the I&M Program focuses on natural resource inventories and long-term monitoring, MOJN's first priority is the management of data and information that results from these efforts. However,

#### Prioritizing data management efforts in a sea of unmanaged data

- Highest priority is to produce and curate high-quality, well-documented data originating with the Inventory and Monitoring Program
- As time and resources permit, assist with data management for current projects, legacy data, and data originating outside the Inventory and Monitoring Program that complement program objectives
- In addition, help ensure good data management practices for park-based natural resource projects that are just beginning to be developed and implemented

the standards, procedures, and approaches to data management developed by MOJN carry over and are being applied to other natural resource data sources.

For example, all natural resource parks need a basic suite of resource inventory data in order to manage their resources and support a successful monitoring program. The national I&M Program has determined that a minimum of 12 inventory datasets, including both biotic and abiotic components, will be acquired by all parks. MOJN is working with individual parks and national NPS programs to acquire and standardize these basic resource datasets, and make them widely available. The datasets are:

- Natural resource bibliography
- Documented species list of vertebrates and vascular plants
- Species distribution and status of vertebrates and vascular plants
- Vegetation map
- Base cartographic data
- Soils map
- Geology map
- Water body location and classification
- Water quality data
- Location of air quality monitoring stations
- Air quality data
- Weather data

A summary of the status of these datasets for network parks is presented in Appendix J (Status of the 12 natural resource inventories, MOJN).

### **Data Management and the Project Lifecycle**

I&M projects are typically divided into five broad stages: initiation, planning, execution, monitoring and control, and closure (Figure 3). During all stages, data management staff collaborate closely with project leaders and participants.

Specific data management procedures corresponding to these five stages are described in the chapters of this plan. Building upon the data management framework presented in Chapters 1 through 5, Chapter 6 is devoted to data acquisition, processing, and reporting, and Chapter 7 provides a framework for verifying and validating data that are collected and entered into databases. Dataset documentation is the subject of Chapter 8, data ownership and sharing is presented in Chapter 9, and data dissemination, including issues such as compliance with the Freedom of Information Act (FOIA), are addressed in Chapter 10. Chapters 11 and 12 provides a framework for the long-term maintenance, storage, and security of MOJN data.

### **Water Quality Data**

The water quality component of the Natural Resource Challenge requires that networks archive all water quality data collected as part of the monitoring program in a STORET (STORAge and RETrieval, EPA 2006) database maintained by the NPS Water Resources Division (WRD, [NPSTORET home page](#)). MOJN will be developing a MS-Access database that consolidates available water quality data collected in and near the 7 MOJN park units. Associated with this

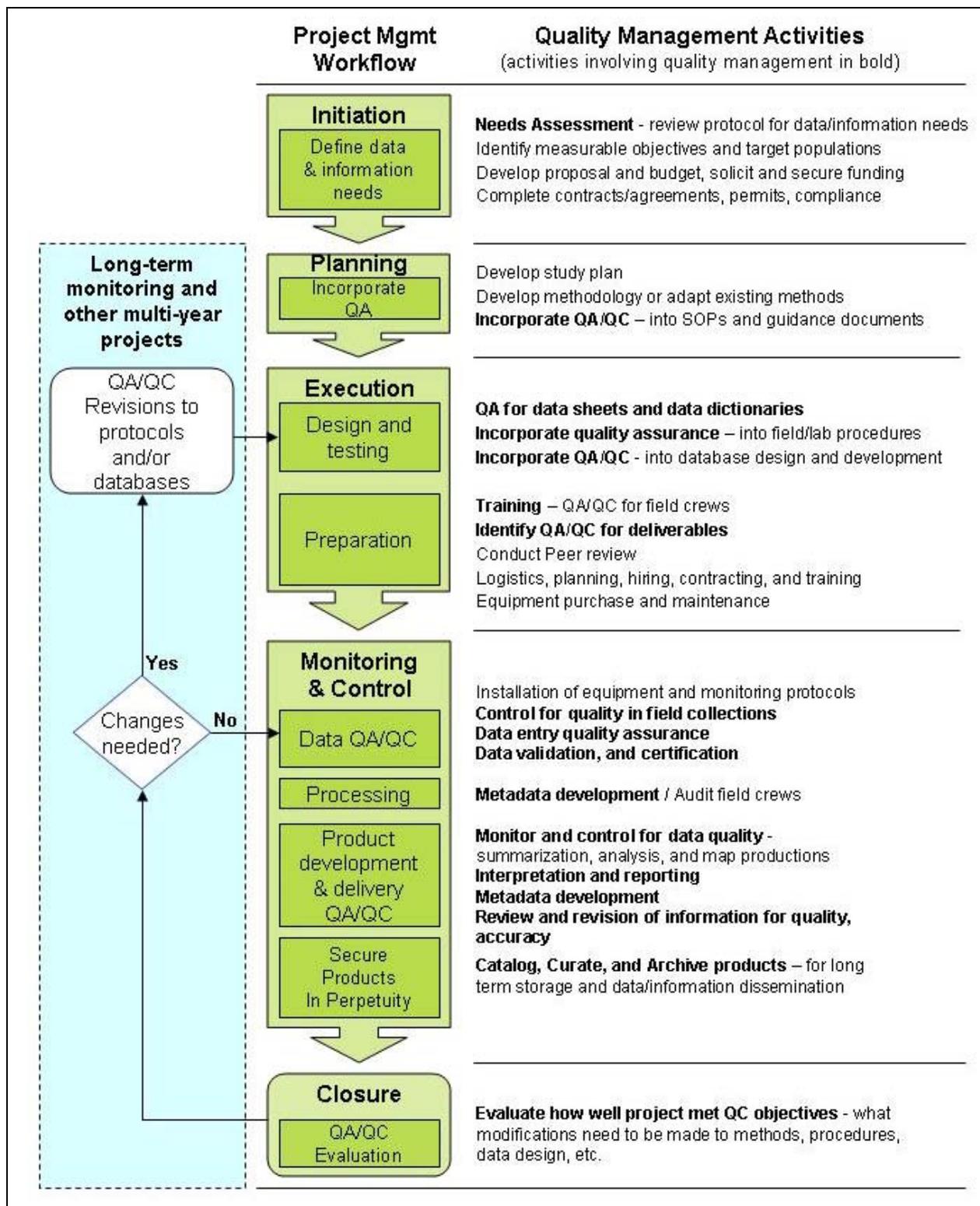


Figure 3. Project workflow and data management activities

database are assessment tools to evaluate water quality that allow comparisons of historical and current data with applicable state standards. MOJN will maintain this database and integrate new data collected so it can serve as an ongoing tool for the network’s long-term water quality monitoring and analysis needs.

On an annual basis MOJN will compile and format new water quality data from the MOJN H<sub>2</sub>O MS-Access database into an electronic data deliverable (EDD) that is compatible with WRD-STORET. WRD will ensure that content is transferred to the Environmental Protection Agency’s STORET database (Figure 4).

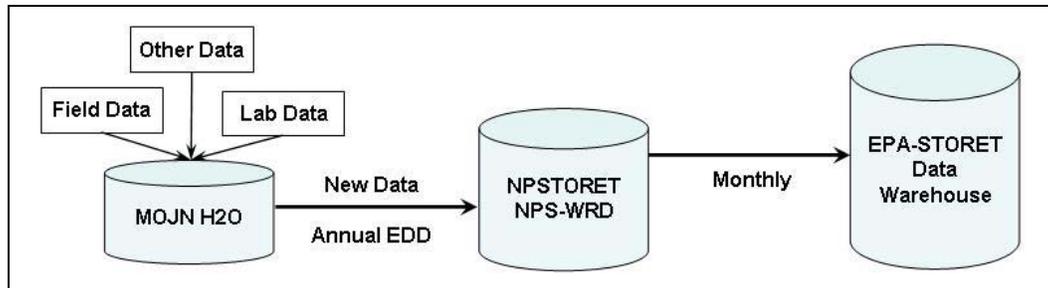


Figure 4. Simplified data flow diagram for water quality data

### Data Management Plan Maintenance

The MOJN approach is to maintain a DMP that is useful to a broad audience, and that can provide guidance on data management practices at a number of different levels. MOJN will keep the plan simple, flexible, and evolving, and include data users in the decision-making process whenever possible.

The document has undergone an initial prescribed review process that included both an internal network review (i.e., by members of the technical committee and network staff), and a service-wide review that involved the regional data/GIS coordinator, data management staff from the WASO I&M Program, and other network data managers.

MOJN will update the plan to ensure that it reflects accurately the network’s current standards and practices. Recommendations for changes can be forwarded to the network data manager by any interested party or user of network I&M data (e.g., park resource managers, project leaders, technicians, superintendents, external users). These recommendations will be discussed by data management and network staff and actions decided upon. Simple changes can be made immediately in the document, while substantive changes will be made during version updates.

*The most current version of the plan is available on the MOJN website (<http://science.nature.nps.gov/im/units/mojn/index.cfm>).*