



Northeast Coastal and Barrier Network Inventory and Monitoring Program

Information Management Plan

December 15, 2004
Draft



Developed by:

Sara M. Stevens
&
Gary Entsminger

Mission Statement

This information management plan has been developed to support NPS Inventory and Monitoring Program goals and objectives to ensure that Northeast Coastal and Barrier Network (NCBN) Inventory and Monitoring Program data and information are accessible and useful long beyond the initial field data collection phase.

Executive Summary

As the basic and most important products of scientific research, data and information represent a valuable and an often irreplaceable resource (Michener and Brunt, 2000). Because field experiments and associated data collection are often time consuming and expensive, management of data and information products plays an important role in any scientific program. For long-term ecological monitoring programs, such as the NCBN Inventory and Monitoring Program, retention and documentation of high quality data are the foundation upon which the success of the overall program rests.

To develop and retain high quality data, the Network has developed the following draft information management plan that describes the Network's information management infrastructure (e.g., staffing, hardware, software) and architecture (databases, procedures, archives). This includes procedures to ensure that relevant natural resource data collected by NPS staff, cooperators, researchers, and others are entered, quality-checked, analyzed, reported, archived, documented, cataloged, and made available to others for management decision-making, research, and education.

In addition to the Network's Information Management Plan, NCBN staff are developing specification and guidance documents to share with park, Network, regional and national staff (http://www1.nature.nps.gov/im/units/ncbn/d_guidelines.htm). These guidelines describe methods that will be used by NCBN for managing natural resource information from hard copy reports to digital photos. Standard Operating Procedures (SOP) that describe in detail how to create FGDC compliant metadata or conduct quality control procedures on data collected by the Network are also being developed and included each of the Network monitoring protocols.

In order for the NCBN Information Management Plan to be successful, the Network must:

- provide up-to-date technical guidance for the preparation and management of data (guidance, specification and Standard operating procedure documents)
- maintain efficient standards for data processing, from acquisition to distribution

Success will be based on the resulting:

- production and maintenance of high quality information products that fulfill a wide variety of user needs

The NCBN Information Management plan outlines the approach that the Northeast Coastal and Barrier Network will take to implement and maintain a system that will serve the data and information management needs of the NCBN Inventory and Monitoring Program. The plan reflects the commitment to the establishment, maintenance, description, accessibility, and long-term availability of high-quality data and information.

Roles and Responsibilities

For the NCBN Inventory and Monitoring program to work effectively, everyone within the Network will have stewardship responsibilities in the production, analysis, management, and/or end use of data produced by the program. To meet the new data management goals and standards developed by the National Park Service and its constituents, Network staff must understand how data and information flow, and what their roles and responsibilities are in this process.

There are four main categories of data stewardship roles to be handled by Network personnel. These are:

1. production
2. analysis
3. management
4. end use

Each of these broad categories has principal, or ‘must-do’ responsibilities as well as many potential ancillary tasks. As coordinator of these tasks, the fundamental role of the Network data manager is to understand and determine program and project requirements, to create and maintain data management infrastructure and standards, and to communicate and work with all responsible individuals.

Project Workflow

To better understand the information management needs of the NCBN program, it is useful to understand the general work flow of project development and the information management tasks associated with each stage. There are two main types of projects handled by the Network:

- short-term projects, which may include individual park research projects, inventories, or pilot work done in preparation for long-term monitoring
- long-term projects, mainly the Network vital-signs monitoring projects central to the I&M program, but which may also include multi-year research projects and monitoring performed by other 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 of time (years for vital signs monitoring).

Within this information management plan, both short and long-term projects are divided into four primary stages:

1. initiation and approval
2. planning-design and testing
3. implementation
4. finalization-product integration and evaluation

Each of these four stages are associated with multiple information management tasks. During the *initiation and approval stage*, preliminary decisions are made regarding the scope of the project and its objectives. A scope of work may or may not be written for each project, but a proposal is developed and funding sources, permits, and compliance are addressed. A cooperative agreement or contract is developed and finalized. Although many of these responsibilities rest with the project leader and/or program administrators, data management staff must be involved in identifying project deliverables and assuring that each contract or agreement includes a clear list of these deliverables, with reference to either national, regional or Network information and data standards.

At the *planning-design and testing stage* of a project either an inventory study plan or monitoring protocol is developed that details how data will be acquired, processed, analyzed, and reported. Information management tasks associated with this stage include data design, development and maintenance of guidelines and specifications, and dissemination of this information. This stage is one of the most important as it initiates the development of sound, quality data products.

Once the design and testing and the administrative tasks associated with project information management have been completed, the project *is implemented*. At this stage the technical information management staff is critical to the success of the project. Tasks include acquiring, processing, and documenting data. During this phase, products such as reports, maps, and GIS themes are developed. All raw data undergo QA/QC measures and final manipulated products are reviewed. Although many of these tasks may be completed by Network cooperators, the information management staff must be closely involved in the training, development, and review of all draft and final project products.

Once all products have been developed and gone through extensive review, *product integration and evaluation* occurs. Records are either finalized permanently for short-term projects or finalized for the project year for long-term projects. Records are finalized or closed out for the year in the Network project tracking database to reflect status and deliverables. Information management tasks include the review, dissemination, and archiving of all products.

Although all Network projects vary in terms of the final products they produce, all follow these four basic stages. The differences between projects occur within the stage itself and are dependent

on the category and type of data being collected or compiled. This dictates how data are acquired and processed within each project stage.

Data Acquisition and Processing

There are two categories of data acquired and managed by the Network:

1. **Network-based data**—those data collected by Network staff and/or cooperators working with the Network
2. **Network-integrated data**—those data collected by other entities (parks, universities, other agencies, other NPS programs), but identified as important natural resource data, necessary for the Network to manage

Network-based data are those data originating within the Network or are currently being collected by NCBN staff. These include three of the twelve basic I&M Program biological inventory datasets:

- vegetation maps
- species occurrence inventories
- species distribution inventories

Along with vegetation maps and species inventories, the Network manages long-term monitoring data as part of its Vital Signs program. NCBN is currently developing protocols to monitor:

- salt marsh vegetation
- estuarine nekton
- geomorphologic change
- estuarine nutrient enrichment
- landscape change
- visitor impacts

Network-Integrated data can be divided into two more data type categories:

1. *Current or ongoing datasets*—These datasets are pre-determined for acquisition by the Network and follow very specific acquisition and processing steps. These can either be park-based or from external NPS sources. They are protocol dependent datasets that the Network acquires on a regular basis. These datasets are either used for data analyses and comparison purposes with Network-based monitoring data, or they are baseline datasets essential for the completion of a vital sign monitoring protocol.
2. *Legacy datasets*—are those data found and compiled through the data mining process. These include vertebrate and vascular plant species data, other important natural resource inventory data, specimen or voucher data, bibliographic data, and existing monitoring datasets.

Network-based data and Network-integrated data follow slightly different acquisition and processing steps. These are described in detail in the NCBN Information Management Plan (Chapter 5). Field data (Network-based data), data acquired from external sources and legacy data (Network-integrated data), are handled through a variety of steps to reach their final archiving stage. Although some of the steps differ from one data type to another such as the acquisition and dissemination steps, there are four main steps that all NCBN data undergo once acquired:

1. quality assurance/quality control
2. documentation
3. transcription to master databases
4. archiving

Data and information are stored, maintained and disseminated through Network and nationally based database management systems. Details are available in the Network plan (Chapter 5). NCBN vital signs data are stored in the NCBN monitoring database template. Network water quality data are housed in the national water quality database STORET. Species data are managed in the NPS NPSpecies database, and bibliographic data are compiled by the Network in the NPS NatureBib database.

Quality Assurance, Quality Control

When developing a long term ecological monitoring program, it is imperative that information and data developed as part of the program be of high quality and adequate for its intended use (US EPA, 2001). In order to develop quality products, a plan for quality assurance as well as methods for quality control must be developed at all levels of the program. Network staff and cooperators conducting ecological monitoring must be aware of both the need for and the mechanisms to achieve excellence at all levels of product development. To accomplish this, NCBN is developing a quality management system. The Network's quality management system will include the organizational structure, responsibilities, procedures, processes, and resources for implementing QA/QC for its ecological monitoring program.

NCBN will establish guidelines for the identification and reduction of error at all stages in the data lifecycle, including project planning, data collection, data entry, verification and validation, processing, and archiving. This approach requires that the Network:

- develop a plan for quality assurance that will include the identification of roles and responsibilities of Network, park, and cooperative staff for maintaining quality standards at all levels of the program—from field and laboratory data collection to overall data management procedures
- ensure that the process of achieving quality is not only documented, but maintained through routine review by Network staff
- develop protocols and SOPs to ensure data quality
- evaluate the quality of all data and information based on NPS standards before data are distributed

- perform periodic data audits and quality control checks to monitor and improve the Network's data quality program.

Much QA/QC work involves defining and enforcing standards for electronic formats, locally defined codes, measurement units, and metadata. This process begins with data design and continues through acquisition, entry, metadata development, and archiving. The progression from raw data to verified data to validated data implies increasing confidence in the quality of the data through time. Documentation of the dataset's quality review process will be added to the project metadata.

Data Documentation

Another critical step following quality assurance and control is data documentation. Data documentation is another step towards ensuring that datasets are useable for their intended purposes well into the future. Data documentation involves the creation of metadata. Metadata describe how, when, and by whom a particular set of data was collected, and how the data are formatted. Metadata include information about the quality, condition, and characteristics of a dataset. Metadata help us create and maintain a framework for cataloging datasets, which in turn helps make datasets more readily available to a broad range of users.

A significant amount of guidance has become available on proper data documentation (see NCBN Information Plan Appendices). As mandated by the National Park Service, all NCBN metadata associated with geospatial data will conform to Federal Geographic Data Committee (FGDC) standards. There are a variety of software tools available for creating and maintaining FGDC compliant metadata.

For biological datasets, NCBN has adopted the Biological Data Profile Metadata standards developed by the National Biological Information Infrastructure (NBII). All Network-based datasets will be accompanied by the Biological Data Profile when distributed. Northeast Region cooperators have developed helpful guidelines on tools used for creating Biological Data Profile metadata (see Plan Appendices).

NCBN data management staff will provide training and support to project leaders to facilitate metadata development. Upon completion, metadata will be posted and made available and searchable in conjunction with related data and reports via the NCBN website, as well as the national NR-GIS Data Store.

Data Distribution

Access to NCBN data products will be facilitated via a variety of information systems that allow users to browse, search, and acquire Network data and supporting documents. These systems include the NCBN website, and national applications with internet interfaces (NatureBib, NPSpecies, NR-GIS Data Store, etc.). The following table provides a list of repositories and types of data that will or can be housed there (Table A).

Table A. Information management systems that facilitate dissemination of NCBN information

Web Application	Data types available at site	Web Address
NPSpecies	Data on park biodiversity (species information)	http://science.nature.nps.gov/im/apps/npspp/index.htm
NatureBib	Scientific citations related to park resources	http://www.nature.nps.gov/nrbib/index.htm
NR-GIS Metadata and Data Store	Metadata, spatial and non-spatial data products	http://science.nature.nps.gov/nrdata
Biodiversity Data Store	The raw or manipulated data and products associated with I&M data that have been entered into NPSpecies.	http://science.nature.nps.gov/im/inventory/biology/index.htm
NCBN Website	Reports and metadata for all Network projects	http://www1.nature.nps.gov/im/units/ncbn/index.htm

Because Network data will reside in the repositories listed above, these data will automatically be searchable via the integrated metadata and image management system and search gateway called “NPS Focus.” This system is being built with Blue Angel Enterprise software for metadata management and the LizardTech Express Server for image management. Currently, ten NPS and two non-NPS databases have been integrated into the NPS Focus prototype in either full or test bed form for one stop searching. NPS Focus has been released as an Intranet version only (<http://focus.nps.gov/>) – release of a public version is projected in the near future.

Archiving

The final information management step, and one of the most vital to the Network’s Inventory and Monitoring Program, is the long-term maintenance and management of digital and analog information. Technological obsolescence is a significant cause of information loss, and data can quickly become inaccessible to users if they are stored in out-of-date software programs or on outmoded media. Effective maintenance of digital files depends on the proper management of a continuously changing infrastructure of hardware, software, file formats, and storage media. Major changes in hardware can be expected to occur every 1-2 years and in software every 1-5 years (Vogt-O'Connor 2000).

As software and hardware evolve, datasets must be consistently migrated to new platforms, or they must be saved in formats that are independent of specific platforms or software (e.g., ASCII delimited files). Thus, NCBN archiving procedures include saving datasets in both their native format (typically MS-Access or Excel spreadsheet format) and as sets of ASCII text files. As a platform- and software-independent format, ASCII text files ensure future usability of the data in a wide range of applications and platforms. In addition, datasets will periodically be converted to upgraded versions of their native formats.

Chapter 10 of the NCBN Information Management Plan describes procedures for maintaining and managing digital data, documents, and objects that result from Network projects and activities. These procedures will help ensure the continued availability of crucial project information and permit a broad range of users to obtain, share, and properly interpret that information.

Table of Contents

Chapter 1 Purpose and Scope of the Northeast Coastal and Barrier Network (NCBN) Information Management Plan

- 1.1 Data Management Overview
- 1.2 The NCBN Inventory and Monitoring Program's Long-Term Goals
- 1.3 Overview of the NCBN Information Plan
- 1.4 Sources of Natural Resource Data
- 1.5 NCBN Information Management Plan Principles and Scope
- 1.6 Revisions
- 1.7 Intended Audience and Layout for the NCBN Data Management Plan

Chapter 2 Information Management Staff Coordination of Roles and Responsibilities

- 2.1 Information Management Staff
 - 2.1.1 National Level Staff and Coordination
 - 2.1.2 Northeast Region Data Management Staff
 - 2.1.3 NCBN Information Management Staff
 - 2.1.4 Park Level Data Management Staff
- 2.2 Coordination of Roles and Responsibilities
 - 2.2.1 NCBN Data Management Coordination
 - 2.2.2 Data Management Staff-Roles and Responsibilities
 - 2.2.3 Project Leaders
 - 2.2.4 Project Leaders-Roles and Responsibilities
 - 2.2.5 Sharing Responsibilities
 - 2.2.6 Communication
- 2.3 Acknowledgements

Chapter 3 Information Management Resources: Infrastructure and Systems Architecture

- 3.1 Computer Resources Infrastructure
 - 3.1.1 NCBN Digital File Maintenance and Storage
- 3.2 NCBN Systems Architecture
 - 3.2.1 Project Database Standards
 - 3.2.2 Core tables and fields based on network and national standards
 - 3.2.3 Protocol-specific Field Data Tables
 - 3.2.4 Common Lookup Tables
 - 3.2.5 NCBN NRDT Template User Interface
- 3.3 Project Tracking Application
- 3.4 National Information Management Systems Architecture
- 3.5 Acknowledgements

Chapter 4 Data Management Integration and Project Work Flow

- 4.1 Project Work Flow
- 4.2 Project Initiation and Approval
- 4.3 Planning: Design and Testing
- 4.4 Implementation
- 4.5 Project Finalization
- 4.6 Acknowledgements

Chapter 5 Data Acquisition and Processing

- 5.1 Categories of Network Data
 - 5.1.1 Network-based Data
- 5.2 Acquiring and Processing Field Data (Biological Inventories and Monitoring)
 - 5.2.1 Data Acquisition and Processing, Steps
- 5.3 Network Water Quality Data Integration
- 5.4 Network-Integrated Data
 - 5.4.1 Acquiring and Processing Current or Ongoing Datasets
 - 5.4.2 Acquiring and Processing Legacy Datasets
 - 5.4.2.1 Species Data
 - 5.4.2.2 Bibliographic Data
 - 5.4.2.3 Specimen Collection Data
 - 5.4.2.4 Long-term Monitoring Datasets

Chapter 6 Quality Assurance, Quality Control

- 6.1 National Park Service Mandate for Quality
- 6.2 Quality Assurance (QA) and Quality Control (QC)
 - 6.2.1 Data Verification and Validation
- 6.3 NCBN QA/QC Plan
- 6.4 NCBN QA/QC Plan for Data Collection
 - 6.4.1 QA Measures for Data Collection
 - 6.4.2 NCBN Data Collection Quality Control Measures
- 6.5 NCBN QA/QC Plan for Data Entry
 - 6.5.1 QA Measures for Data Entry
 - 6.5.2 QC Techniques for Data Entry
 - 6.5.3 NCBN Data Verification Methods
 - 6.5.4 NCBN Data Validation Methods
- 6.6 Monitoring Conformance to Plans and Standards
- 6.7 Data Quality Review and Communication
 - 6.7.1 Feedback from QA/QC Measures to Improve Data Quality
 - 6.7.2 NCBN Quality Review of Completed I&M Products
- 6.8 Roles and Responsibilities
- 6.9 Creating a QA/QC Audit Trail

6.10 Acknowledgements

Chapter 7 Data Documentation

- 7.1 Metadata
- 7.2 Sources for Data Documentation Guidance
- 7.3 NPS Integrated Metadata System Plan and Tools
- 7.4 Network Data Documentation Strategy

Chapter 8 Summaries, Analyses, and Reports

- 8.1 Data Users
- 8.2 Data Delivery
- 8.3 Preparation of Data for Exchange and Distribution
- 8.4 Monitoring Data Summaries and Analyses
- 8.5 NCBN Monitoring Project Reports
- 8.6 Species Inventory Summaries, Analyses and Reports
- 8.7 Other Network-generated Reports

Chapter 9 Data Ownership

- 9.1.1 National Park Service Policy on Data Ownership
- 9.1.2 Establishing Data Ownership Guidelines
- 9.2 Data Distribution
 - 9.2.1 Data Distribution Mechanisms
 - 9.2.2 Data Classification: Protected vs. Public
 - 9.2.3 Access Restrictions on Sensitive Data
 - 9.2.4 Public Access to Network Inventory and Monitoring Data
 - 9.2.5 Data Availability
 - 9.2.6 Data Acquisition Policy
- 9.3 Data Feedback Mechanisms
 - 9.3.1 Data Error Feedback Procedures
- 9.4 Acknowledgements

Chapter 10 Data Maintenance, Storage, and Archiving

- 10.1 Digital Data Maintenance
 - 10.1.1 Short-term Datasets
 - 10.1.2 Long-term Monitoring of Datasets
 - 10.1.3 Spatial Data
- 10.2 Storage and Archiving Procedures—Digital Data
 - 10.2.1 Storage Facilities for Electronic Archives
 - 10.2.2 Directory Structure for Electronic Archives
 - 10.2.3 Backup Procedures for Digital Data
 - 10.2.4 Data and Network Security

- 10.3 NCBN Storage and Archiving of Official and Non-official Records
 - 10.3.1 Hard Copy and Digital Reports Produced by the Network
- 10.4 Specimens
- 10.5 Photographs
 - 10.5.1 Aerial Photographs
- 10.6 Role of Curators in Storage and Archiving Procedures
- 10.7 Acknowledgements

NCBN Data Management Guidance Documents and Appendices Associated with the NCBN Information Plan

Appendix: Document Title:

1. Northeast Region GPS Field Collection Standards for Cooperators (*to be developed*)
2. [Northeast Coastal and Barrier Network Standard Operating Procedure \(SOP\) for the Garmin V GPS Unit](#)
3. Northeast Coastal and Barrier Network File Management Plan (*to be developed*)
4. Northeast Region I&M Program Guidelines for developing FGDC compliant metadata for both spatial and biological data (*to be developed*)
5. [Northeast Region Metadata tools used in the creation of the FGDC biological data profile](#)
6. NCBN I&M Program Standards for File Revision Control (*to be developed*)
7. NCBN I&M Program Data Management Standards for Voucher Collection and Cataloging (*to be developed*)
8. Northeast Region I&M Program Database Specifications (*to be developed*)
9. [Northeast Region I&M Program NPSpecies Certifier's Guide](#)
10. [Northeast Region Inventory and Monitoring Program NPSpecies Data Management Plan](#)
11. [Northeast Region Inventory and Monitoring Program NPSpecies Data Entry and Data Management Standards](#)
12. [Northeast Region I&M Program NatureBib Data Management Plan](#)
13. [Northeast Region I&M Program NatureBib Data Management and Data Entry Standards Manual](#)
14. [Northeast Region I&M Program Product Specifications](#)
15. [NCBN Standard Operating Procedure Converting the NCBN Database Template](#)
16. [NCBN Saltmarsh Monitoring Database User's Guide](#)
17. Northeast Region I&M Program Data Mining Specifications (*to be developed*)
18. NCBN I&M Program Project Management Guidelines (based on use of Project KickStart Software) (*to be developed*)
19. NCBN Data Distribution Plan (*to be developed*)

20. Northeast Region I&M Program Web-Page Development and Management Guidelines (*to be developed*)
21. [NCBN GIS Specifications and Guidelines](#)
22. [Cape Cod National Seashore Digital Image Management Guidelines \(Draft\)](#)
23. NCBN Quality Assurance Plan (*to be developed*)
24. [Northeast Coastal and Barrier Network FOIA and Sensitive Data Information](#)
25. [NCBN Data Stewardship-Staff Roles and Responsibilities](#)
26. [Northeast Region I&M Program QA/QC Procedures and Metadata Development for Final Biological Inventory Products](#)
27. [NCBN Data Dictionary Example](#)
28. [NCBN Example Language Describing Deliverables for use in Cooperative Agreements and Contracts](#)
29. Northeast Region Natural and Social Science Study Proposal and Deliverable Guidelines (*download this document at: www.nps.gov/nero/science/*)
30. [Northeast Region GPS Metadata Form](#)
31. [Northeast Region Vegetation Mapping Aerial Photo Specification Example for scale 1:6000](#)
32. [Northeast Region Vegetation Mapping Aerial Photo Specification Example for scale 1:12000](#)
33. NCBN Lidar Data Acquisition and Processing Example (*to be developed*)
34. [Northeast Region Inventory and Monitoring Program Vegetation Map Development Example](#)
35. [Northeast Region Example of Spatial Data Review for Vegetation Mapping Products](#)
36. [Northeast Region Inventory and Monitoring Program Example Vegetation Mapping Deliverables and Mapping Report Guidelines](#)
37. NCBN Metadata Questionnaire (*to be developed*)