

Arctic Network Inventory & Monitoring Program

U.S. Department of the Interior

Data Management
Standard Operating Procedure
NPS/ARCN/DMSOP-2007-07



Database Planning Toolkit

Task checklist and guidance for Inventory and Monitoring data application development

Summary

Contents

Database Planning Toolkit.....	1
Task checklist and guidance for Inventory and Monitoring data application development	1
Summary	1
Contents	1
Introduction	1
Database Planning Workflow	2
Database Planning Considerations	2
Database Planning Consultation and Required Materials.....	3
Database Planning Questionnaire.....	3
Stakeholder Requirements.....	3
Conceptual Data Model (Work and Data Collection Flow Chart).....	5
Life span of Data Collection	7
Location of Data Collection	7
Time of Data Collection.....	7
Manner of Data Collection	7
Data Collection Equipment Considerations	7
Power Requirements.....	7
Database Versioning Requirements	8
Data Types.....	8
Data Modeling	9
Entities and Attributes	9
Database Technical Considerations	10
Integration with other software and hardware.....	10
GPS Requirements	10
Milestones.....	11
References	12
About This Standard Operating Procedure	12
Revision History.....	12

Introduction

Long term natural resource monitoring projects not only require a significant investment in data stewardship, but also a highly organized coordination of stakeholders. A good project data management system must accommodate multiple users with varying technical backgrounds and divergent work requirements. Each stakeholder must contribute to the development of the project's data management system in order to ensure

Arctic Network Data Management Standard Operating Procedure

that the requirements are met. This toolkit is intended to serve as an initial consultation tool between data management personnel and researchers. The toolkit will rapidly define the scope of a project's data management needs. By defining the project's needs early in its development stakeholders can more quickly begin data modeling and data application development.

Database Planning Workflow

Constructing data management applications consumes enormous amounts of time. Pre-planning is the only way to minimize development time. All stakeholders must budget a lot of time for participating in the database planning effort. Researchers may not be aware of the various tasks involved in data application development and where they may fit into the process.

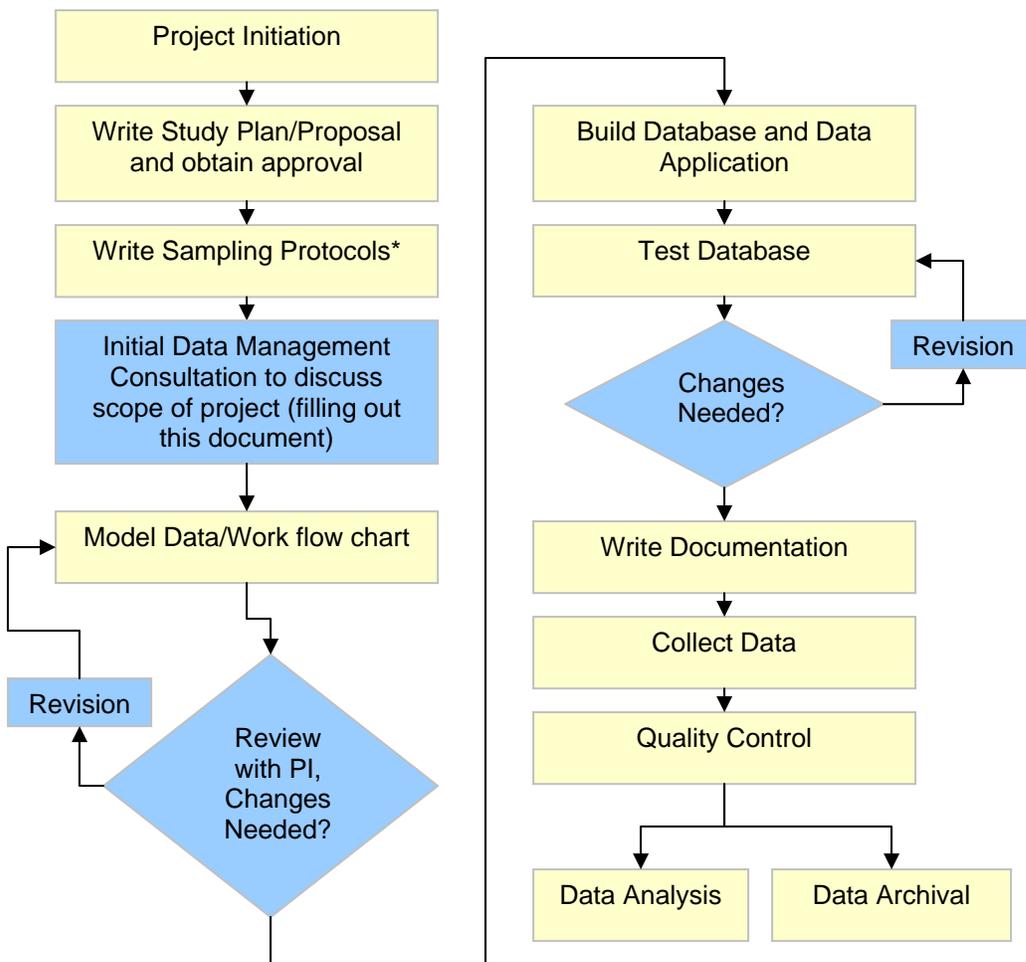


Figure 1. Database development work flow. *Critical element.

Database Planning Considerations

- Databases should be planned for and built prior to any field work takes place.

Arctic Network Data Management Standard Operating Procedure

- Monitoring databases must be built using the Natural Resources Database Template (NRDT, See National Park Service, 2006, Southwold, 2002a and b) and the Microsoft Access and/or SQL Server platform.
- Plan on four to six months database development time at the minimum, depending on developer workload.
- Data collection protocols need to be in substantially finished form at least six months prior to your anticipated field season since these documents will drive the data modeling process.
- Spreadsheets are not acceptable for storing I&M monitoring data.

Database Planning Consultation and Required Materials

Investigator's should plan to have a specific meeting to address database needs. The investigator should be prepared to bring the following to the meeting (* denotes critically important):

- Narrative and SOPs relating to data management.*
- List of stakeholders and end users and their requirements
- Field forms (if using)
- Data collection flow diagram
- Examples (even if hand drawn) of how the information will be used
- List of software or applications that need to be compatible, such as statistical software or export/import formats for other agencies
- Sample or dummy data for testing
- Estimation of number of records per year.

Investigators should be prepared for multiple meetings of an hour or more after the initial consultation as the data model and database are refined.

Database Planning Questionnaire

Stakeholder Requirements

Describe each stakeholder that will interact with the database and what goals the database will accomplish for them. Include non-NPS organizations. Hardware or software interacting with the database should also be listed as stakeholders. An example appears immediately below. Use the blank following table for your requirements.

Stakeholders Inventory (Example)

Actor	Requirements	Is NPS	Database Access Req'd
Investigator	Extract data preformatted for SAS ANOVA Quick trend graphs for reports, website	Y	Unrestricted

Arctic Network Data Management Standard Operating Procedure

Biotech	Field data entry of new data (no edits of existing data needed) Quality control procedures	Y	Seasonally, for keypunching
Investigator (ADF&G)	Export data from NPS database into ADF&G acceptable format	N	Unrestricted
GIS Technician	Extract field locations into network field sites inventory annually	Y	Unrestricted
ADF&G Data Harvesting Server	Harvest updated data weekly via NPS xml web service. (Hardware and software can be stakeholders also).	N	Restricted, through web service

Stakeholders Inventory

Actor	Requirements	Is NPS	Database Access Req'd
--------------	---------------------	---------------	------------------------------

Notes:

Arctic Network Data Management Standard Operating Procedure

Conceptual Data Model (Work and Data Collection Flow Chart)

If you haven't already done so, construct a flow chart of your anticipated data collection work flow. This diagram should correspond closely with your data collection protocols. See Figure 1, p. 6 for an example.

Arctic Network Data Management Standard Operating Procedure

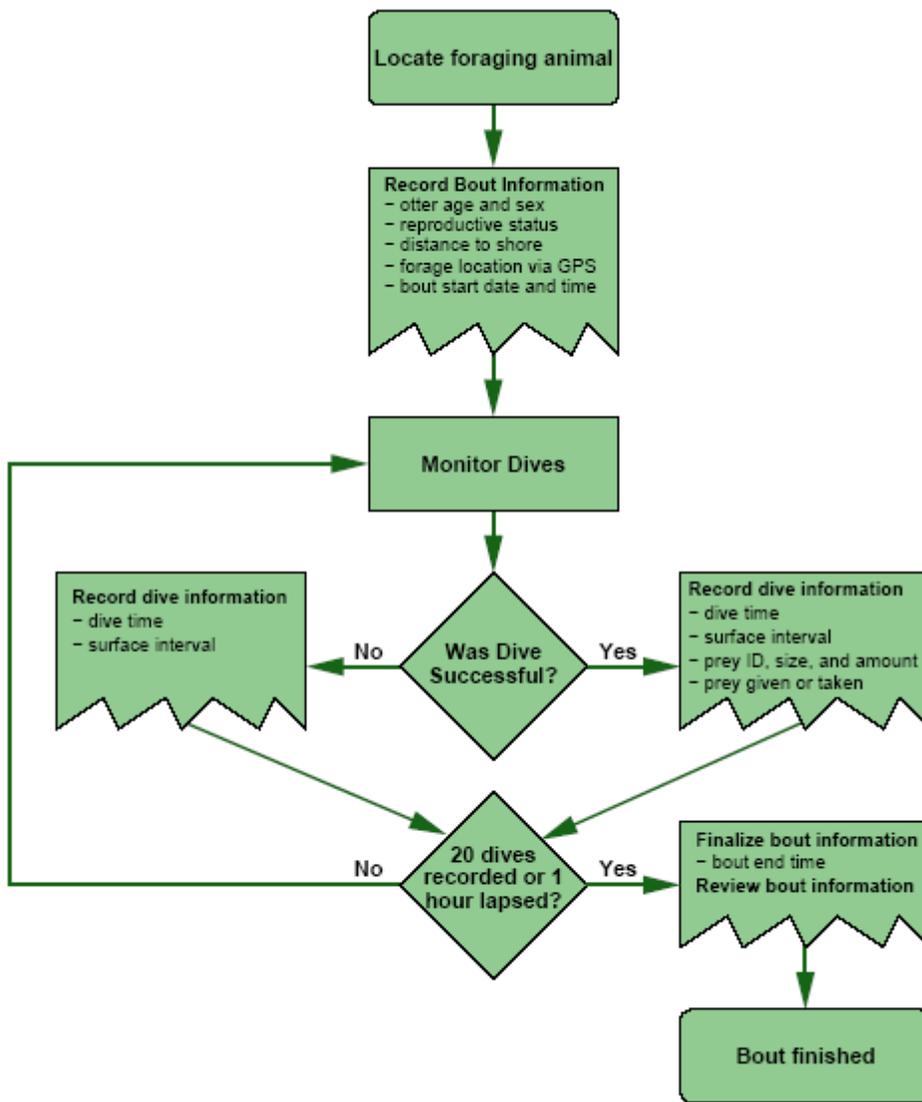


Figure 2. Example conceptual data model showing flow of data collection.

Arctic Network Data Management Standard Operating Procedure

Life span of Data Collection

What is the anticipated life span of the project

Long term monitoring (Indicate sampling interval) Short Term Single Effort

Location of Data Collection

Where will data collection occur? List geographic coordinates, describe the area, or provide a map.

Time of Data Collection

When will field data collection commence? How long will the field season be?

Manner of Data Collection

Describe the manner of data collection (check all that apply).

Aircraft Ground Watercraft Indoors

Other:

Data Collection Equipment Considerations

Describe potential risks to field computing equipment (check all that apply).

Rain/Snow Impact Submersion Dust/Sand/Grit Extreme cold

Extreme heat

Other:

Power Requirements

What will be your power requirements (list equipment and power requirements)?

If you are off the grid, what power options are available (solar chargers, generators, etc.)?

How will the data be collected in the field (check)?

Paper & Pencil Tablet PC Laptop PDA Trimble GPS Unit

Arctic Network Data Management Standard Operating Procedure

Database Versioning Requirements

Will there be more than one field crew deployed at the same time (requires database versioning, disconnected editing)?

Data Types

The data being collected would best be described and handled as:

 Spatial Data (Geodatabase, shapefile or other explicit spatial data format)

 Tabular Data (Non-spatial)

Notes:

Data Modeling

Entities and Attributes

What are the discrete, unique things (entities) from which you are collecting data?

What are the attributes that you are measuring for each of these entities?

Do you have groups of entities (classes) with the same attributes? Draw each entity as a box with the entity name at the top and all attributes below as shown in the example.

Sampling Event (Example)	Otter (Example)
Event ID	Otter ID
Location	Age
Start Date	Sex
Start Time	Pup
	Tags
	Event ID

Database Technical Considerations

Integration with other software and hardware

List all the external hardware and software that your database may need to interact with. List makes and models where possible. Such units may include external GPS units, automated samplers, meters, Geographic Information Systems, etc.

GPS Requirements

Check all that apply. Project requires:

High Accuracy (Sub-meter) Low accuracy (3-10 or more meters) High precision
monumenting

Notes:

References

National Park Service, 2006. Natural Resources Database Template Specifications, Ft. Collins, Colorado. [<http://science.nature.nps.gov/im/apps/template/index.cfm>], November 2006.

Southwold, A. 2002a. Database Specifications for Inventory and Monitoring Studies. National Park Service. Anchorage, Alaska

Southwold, A. 2002b. Recommended Database Strategies including I&M Database Template. National Park Service. Anchorage, Alaska

Washington Support Office, Inventory and Monitoring Program. 2004. Recommended Naming Standards for Inventory and Monitoring Databases. National Park Service. Fort Collins, CO. 14 pg.

About This Standard Operating Procedure

Version: 1

Status: Draft

Publication Date: October 30, 2007

Author(s): Scott D. Miller, Data Manager, Arctic Network Inventory & Monitoring Program.

Abstract: This document describes the general framework and policies for managing all data and information within the Arctic Network Inventory and Monitoring Program.

Suggested Citation: NPS-ARCN (2007). Database Planning Toolkit: Task checklist and guidance for Inventory and Monitoring data application development. Version 1.0, Arctic Network-Inventory and Monitoring Program, National Park Service. Fairbanks, Alaska.

Revision History

Version	Version Date	Revised By	Changes
1.0	20071029	S. Miller and D. Mortenson	Original
1.1	20071102	S. Miller	Added example conceptual data model

This table reflects changes to this document. Version numbers will be incremented by one (e.g., Version 1.3 to Version 2.0) each time there is a significant change in the process and/or changes are made that affect the interpretation of the data. Version numbers will be incremented after the decimal (e.g., Version 1.6 to Version 1.7...1.10....1.21) when there are changes to grammar, spelling, or formatting, or minor modifications in the process that do not affect the interpretation of data.