

Network Data Management Best Practices

Overview

This document is part of a series of best practices that will assist in meeting the Inventory and Monitoring Division's (IMD) [Core Expectations for Vital Signs Monitoring](#). The document is a tool to ensure all protocols are following sound data management practices, and that these practices are documented and periodically updated. Networks may use this tool to identify where data management practices are or are not meeting expectations, then develop an approach to bring operations into alignment with expectations.

With the passage of the [OPEN Government Data Act](#) in 2018, all IMD data are expected to be made readily available, discoverable, and usable by stakeholders outside the National Park Service, including the general public. In IMD terminology, this generally means that all datasets be published to an acceptable repository (like the [NPS DataStore](#)) with rich metadata and persistent identifiers. The Central Support Office's Data Science Team is developing [best practices for data publication](#).

While this document may serve as a general orientation to new network data managers, program managers and protocol leads, more detailed guidance and support should be sought from the [National Data Scientist](#), regional program manager, and one or more mentors, such as existing network data managers from the same or different regions.

How to Use This Document

The table below (and available to download at right) includes a list of 10 data management themes that should be explicitly addressed for every protocol implemented by a network. All items in the table should be current and complete prior to publishing a vital sign's associated dataset(s). If all items in the table for a given protocol are not completed and routinely updated, then the network program manager and regional program manager should consider suspending the protocol until they have sufficient time to rectify the shortcomings.

- In **Column 1**, review the "Data Management Themes" during [Operational Reviews](#), by a team consisting of (at minimum) the network data manager, protocol lead, and network program manager.

- In **Column 2** (“Topics”), reflect on each item and either point to existing documentation or determine a plan for creating the necessary documentation. Network staff are not expected to directly answer the questions.
- In **Column 3** (“Where is this documented?”), record where the documentation for each theme/topic exists (e.g., protocol narrative, standard operating procedures, quality assurance plans, data management plans).
- In **Column 4** (“Notes”), determine if the current documentation is adequate or if it needs to be updated, who will do the updating, and by when.

Network: _____ Protocol: _____ Date: _____

Data Management Theme	Topic <i>These prompts are meant to be a helpful guide, but may not include everything that needs to be documented.</i>	Where is this documented? <i>Protocol, SOP, QAP, DMP, etc.</i>	Notes <i>Is existing documentation adequate? Does it need to be updated? By whom and when?</i>
1. Workflow, timeline, and roles	What is the high-level flow of data and expected timeline for each step in the IMD Science Information Life Cycle ?		
	Who is responsible for each step of the workflow? What is the timeline for completion? Be specific.		
2. Required resources	What specialized knowledge, software, tools, licenses, training, IT support, permissions, and positions (FTEs) are required to perform each step in the IMD Science Information Life Cycle ?		
3. Data acquisition, collection, entry, ingestion	What field forms or digital data collection methods are used?		

Data Management Theme	Topic <i>These prompts are meant to be a helpful guide, but may not include everything that needs to be documented.</i>	Where is this documented? <i>Protocol, SOP, QAP, DMP, etc.</i>	Notes <i>Is existing documentation adequate? Does it need to be updated? By whom and when?</i>
	What data entry procedures are used?		
	What lab, logger, or other data must be incorporated and how is that accomplished?		
4. Data storage (Be sure to consider tabular, spatial, photo, video, audio, etc.)	What is the data storage solution, including any relevant documentation or schema?		
	What is the estimated volume of the data collected or generated?		
	Where are the data stored (initially and permanently)?		

Data Management Theme	Topic <i>These prompts are meant to be a helpful guide, but may not include everything that needs to be documented.</i>	Where is this documented? <i>Protocol, SOP, QAP, DMP, etc.</i>	Notes <i>Is existing documentation adequate? Does it need to be updated? By whom and when?</i>
	If paper field sheets are utilized, are they scanned and archived? Where?		
5. Quality Assurance, Quality Control, and Data Completeness	What techniques, steps, scripts, tools, etc. will be used for preventing errors (QA) and detecting errors (QC) and ensuring data completeness?		
	If QC checks have changed or rejected individual data values, do data quality flags, following data quality flagging guidance , document these changes?		
6. Metadata	What metadata standard will be used (e.g., EML, FGDC/CSDGM, ISO)?		

Data Management Theme	Topic <i>These prompts are meant to be a helpful guide, but may not include everything that needs to be documented.</i>	Where is this documented? <i>Protocol, SOP, QAP, DMP, etc.</i>	Notes <i>Is existing documentation adequate? Does it need to be updated? By whom and when?</i>
	Does it follow the metadata best practice guidance ?		
7. Data publication and dissemination	What open data formats will be used?		
	Where will the data and metadata be preserved in the long-term?		
	Who will review the data and metadata?		

Data Management Theme	Topic <i>These prompts are meant to be a helpful guide, but may not include everything that needs to be documented.</i>	Where is this documented? <i>Protocol, SOP, QAP, DMP, etc.</i>	Notes <i>Is existing documentation adequate? Does it need to be updated? By whom and when?</i>
8. Records management, archiving and curation of specimens	How is long-term records management and archiving handled? Are there physical specimens or other samples in need of curation and how is that accomplished?		
9. Access and use restrictions	Is there documentation (in consultation with the park) on what is considered protected information? (e.g., Controlled Unclassified Information, including protected resources and/or resource locations)? How are those data handled (e.g., fuzzing, redacting)?		
10. Taxonomy (when applicable)	What is the taxonomic authority and how frequently is it updated?		

Data Management Theme	Topic <i>These prompts are meant to be a helpful guide, but may not include everything that needs to be documented.</i>	Where is this documented? <i>Protocol, SOP, QAP, DMP, etc.</i>	Notes <i>Is existing documentation adequate? Does it need to be updated? By whom and when?</i>
	What are the data management considerations related to taxonomy updates?		