

# A Data Management Plan Template for Ecological Restoration and Monitoring

**Brick M. Fevold**  
**Research Scientist Advisor, GDIT**

Judy Schofield<sup>1</sup>, Rob Sutter<sup>1</sup>, Craig Palmer<sup>1</sup>, Elizabeth Benjamin<sup>1</sup>,  
Molly M. Amos<sup>1</sup>, Louis Blume<sup>2</sup>

<sup>1</sup> GDIT, Alexandria, Virginia

<sup>2</sup> U.S. Environmental Protection Agency, Chicago, IL



# Data Management Planning

## Outline:

- **Data Management Policies**
  - Federal Government 'Open Data Policy'
  - EPA Quality System Documentation
  - NSF & NIH Grant Stipulations
- **Data Management System (DMS)**
  - File-Folder Structure vs. Integrated Solutions
- **Data Management Plan (DMP)**
  - Data Management Planning and Implementation
  - DMP Template for Ecological Restoration



# On-Line Resources:

## Federal Government 'Open Data Policy'

- [project-open-data.cio.gov](http://project-open-data.cio.gov)

## EPA Quality System Requirements

- [epa.gov/quality](http://epa.gov/quality)

## National Science Foundation

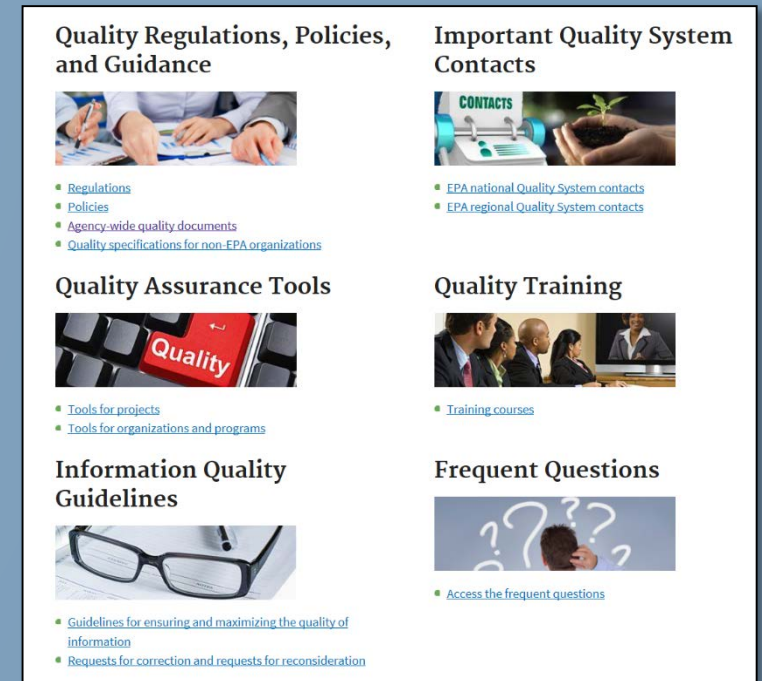
- [nsf.gov](http://nsf.gov)







## National Information Standards Institute

- [niso.org](http://niso.org)

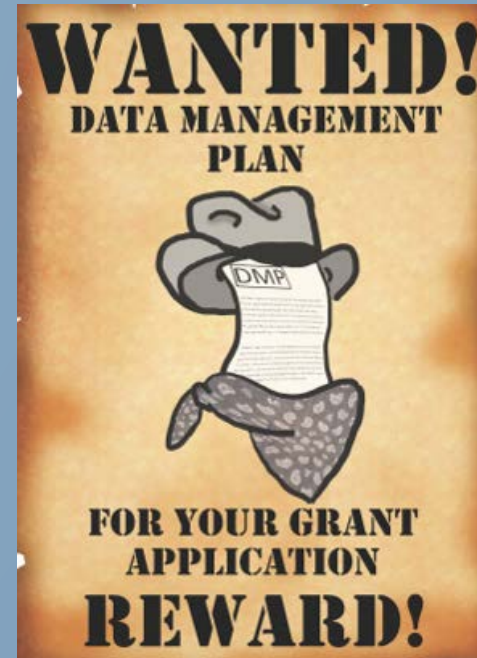
## Developing and Implementing a Data Management Plan

- [dataone.org](http://dataone.org)
- [usgs.gov/datamanagement](http://usgs.gov/datamanagement)
- [lib.umn.edu/datamanagement](http://lib.umn.edu/datamanagement)
- [dcc.ac.uk/resources/how-guides/develop-data-plan](http://dcc.ac.uk/resources/how-guides/develop-data-plan)



<b>Quality Regulations, Policies, and Guidance</b>  <ul style="list-style-type: none"><li>• <a href="#">Regulations</a></li><li>• <a href="#">Policies</a></li><li>• <a href="#">Agency-wide quality documents</a></li><li>• <a href="#">Quality specifications for non-EPA organizations</a></li></ul>	<b>Important Quality System Contacts</b>  <ul style="list-style-type: none"><li>• <a href="#">EPA national Quality System contacts</a></li><li>• <a href="#">EPA regional Quality System contacts</a></li></ul>
<b>Quality Assurance Tools</b>  <ul style="list-style-type: none"><li>• <a href="#">Tools for projects</a></li><li>• <a href="#">Tools for organizations and programs</a></li></ul>	<b>Quality Training</b>  <ul style="list-style-type: none"><li>• <a href="#">Training courses</a></li></ul>
<b>Information Quality Guidelines</b>  <ul style="list-style-type: none"><li>• <a href="#">Guidelines for ensuring and maximizing the quality of information</a></li><li>• <a href="#">Requests for correction and requests for reconsideration</a></li></ul>	<b>Frequent Questions</b>  <ul style="list-style-type: none"><li>• <a href="#">Access the frequent questions</a></li></ul>

# Data Management Policies



**Federal institutions are REQUIRING 1-2 page synopsis of data management planning to be submitted as part of grant requests.**

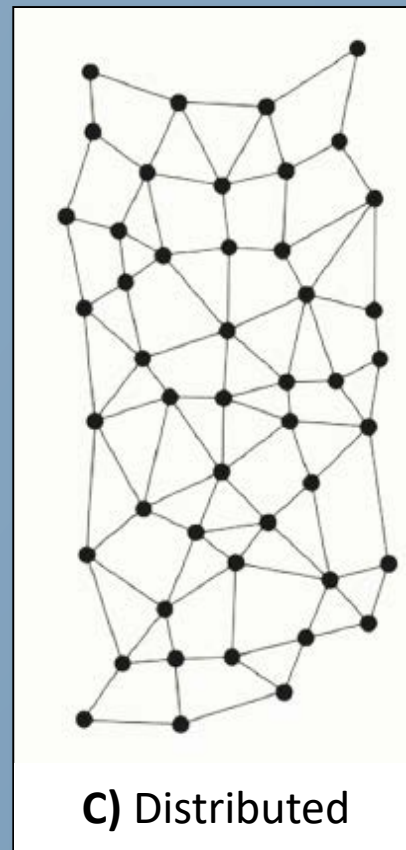
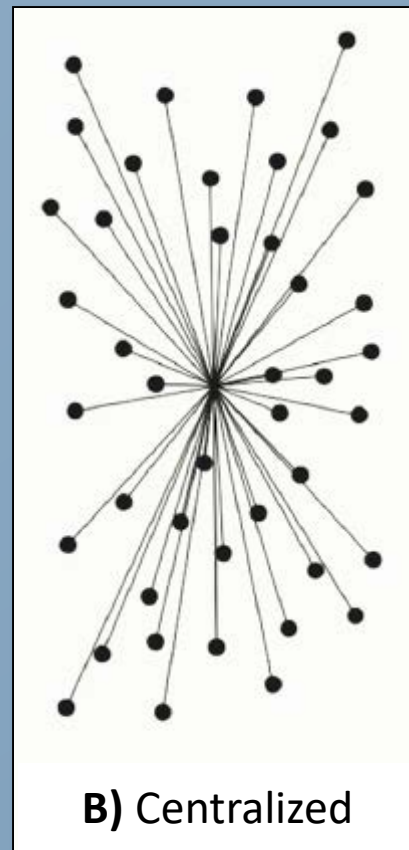
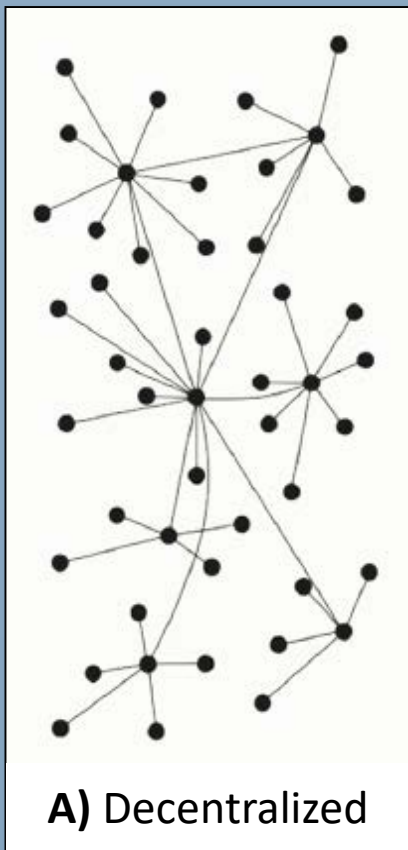
# Data Management Systems (DMS)

Source: Modified from Kolb et al., 2013.

Consideration or System Feature	File-Folder Structure	Integrated Systems	
		Desktop Relational Database	Enterprise Relational Database
Data Management System	None	Centralized	Decentralized
Example Applications	OS <sup>1</sup> , Excel, Lotus 123, Quattro Pro	Access, Microsoft SQL, Express, SQLite	SQL Server, Oracle, PostgreSQL, MySQL
Technical Capacity	Basic	Intermediate	Advanced
Desktop or Server-based	Both	Both	Server-based
Spatially Enabled	No	Optional	Optional
Security Options	Low	Moderate	High
Multiuser Data Entry	No	No	Yes
Size of Data Set	Unlimited	Limited	Unlimited
Web-based	Optional	No	Yes
Cloud-storage Use	Optional	Optional	No
Cost of Development	Low	Intermediate	High
Level of Programming	Basic	Intermediate	Expert

# Data Management – in Ecological Restoration

## Levels of Organization (Implementation)



### A) Decentralized DM

- Data Collection – implementation of SOPs and QA/QC oversight by collaborating institutions across multiple regions

### B) Centralized DM

- Data Entry – data submitted to a centrally managed DMS by field crews or research laboratories

### C) Distributed DM

- Data Reuse – redistribution and application of shared data by independent institutions or individuals



*The effective management and preservation of project data for primary and secondary uses are, by definition, quality assurance strategies – to preserve is to protect.*

*Data that are preserved are data that can be shared.*

# Data Management Plan Template


## Cover Page - Introduction

## DMP Template - Detailed

## Appendix – Template Headers

*Draft*

**Data Management Plan Template  
for  
Ecological Restoration Projects**



Prepared by General Dynamics Information Technology (GDIT), Alexandria, Virginia, under EPA Contract No. EP-C-17-024, Scientific and Technical Support, through the direction and funding of Louis J. Blume, Quality Assurance Manager, U.S. EPA Great Lakes National Program Office, Chicago, Illinois.

1

GENERAL DYNAMICS Information Technology Great Lakes RESTORATION

**Data Management Plan Template (Draft) - Ecological Restoration Projects**

**DATA MANAGEMENT PLAN TEMPLATE ELEMENTS**

**Data Management Plan (DMP) Cover Page<sup>†</sup>**

- 1) Document title, date, and version number
- 2) List of authors and contributors, including organization affiliation, name and position title; include identification of individual(s) responsible for implementation of the data management plan
- 3) Distribution list of individuals to receive approved DMP

**Project Description, Administration, and Data Management Requirements<sup>†</sup>**

*Project Administration*

- 1) Project title (or phase if part of long-term monitoring program)
- 2) Program or departmental administrating entities including names of the organizations, individuals and other stakeholders that possess intellectual property rights to the data
- 3) Identify of funding institutions and their stipulations relevant to data management (e.g., plan content standards and reporting, long-term archival, policies on making data publically accessible)
- 4) Time period required by data administrators related to first rights to publishing, and intellectual property rights of data authors (e.g., principal investigators, scientists, managers, collaborators, or governing stakeholders)

*Project Description*

- 1) Project goals and objectives, scope
- 2) Roles and responsibilities of key staff or positions involved in data management planning and implementation, including QA oversight, metadata documentation, data backup and security, and training and certification of staff involved in data management
- 3) Description of: anticipated data and/or data-management related products (e.g., certified dataset(s) or database), or other usable end-product resulting from data management planning (e.g., mobile device applications for standardized electronic data-collection)
- 4) Data processing workflow or chart illustrating generalized flow of data management from acquisition/collection through to submittal to an archive or publically accessible repository
- 5) Schedule of data management tasks with task descriptions, role responsibility and timeline

*Data Management Requirements*

- 1) Policies and procedures to guide informed consent and privacy protection or participant confidentiality
- 2) Existing or implied copyright constraints related to copyrighted data formats, proprietary software or programming code, and existing (secondary) data acquired under confidentiality agreements
- 3) Policies regarding specification for hardware and software

**Data Acquisition and Collection<sup>Q</sup>**

- 1) Study and sampling design outlining the hierarchy associated with experimental units, sample and sub-sample units; include sufficient detail necessary to evaluate sampling effort for completeness and representativeness at all levels of hierarchy (e.g., year, sampling period, site, plot, sub-plot)
- 2) Procedures for training and certification of personnel in data management tasks associated with data collection
- 3) Standard data management procedures and checklist (e.g., scheduled daily, weekly, other time period)
- 4) Types and volume of information expected to be generated, including data derived through statistical summary and modeling during data processing and analysis, and all data (and their metadata) acquired from internal and/or external sources
- 5) Reference list of standard operating procedures (SOPs) and data recording forms (electronic and print) created and used to collect and standardize data collection, data dictionaries and reference materials

Q: element is required to be included in U.S. EPA QAPP documentation  
P: element is applicable to agency or departmental program-level policy

1

P: element is applicable to agency or departmental program-level policy

2

P: element is applicable to agency or departmental program-level policy

3

P: element is applicable to agency or departmental program-level policy

4

GENERAL DYNAMICS Information Technology Great Lakes RESTORATION

**Data Management Plan Template (Draft) - Ecological Restoration Projects**

**DATA MANAGEMENT PLAN TEMPLATE ELEMENTS**

- **Data Management Plan (DMP) Cover Page<sup>†</sup>**
  - 1) Document title, date, and version number
  - 2) List of authors and contributors, including organization affiliation, name and position title; include identification of individual(s) responsible for implementation of the data management plan
  - 3) Distribution list of individuals to receive approved DMP
- **Project Description, Administration, and Data Management Requirements<sup>†</sup>**
  - Project Administration*
  - Project Description*
  - Data Management Requirements*
- **Data Acquisition and Collection<sup>Q</sup>**
- **Data Organization, Storage and Security<sup>Q,†</sup>**
  - Organization & Storage*
  - Security*
- **Data Processing and Analysis<sup>Q</sup>**
  - Data Processing*
  - Data Analysis*
- **Data Quality Assurance<sup>Q</sup>**
- **Data Documentation and Metadata Creation<sup>Q</sup>**
  - Project Documentation*
  - Metadata Creation*
- **Data Preservation and Archiving<sup>†</sup>**
  - Data Preservation*
  - Data Archiving*
- **Data Sharing and Reuse<sup>†</sup>**
  - Data Sharing*
  - Data Reuse*

Q: element is required to be included in U.S. EPA QAPP documentation  
P: element is applicable to agency or departmental program-level policy

1



# Data Management Plan (DMP) Implementation

## DMP Template Elements:

- Description & Administration
- Acquisition & Collection
- Organization, Storage & Security
- Processing & Analysis
- Preservation & Archiving
- Sharing & Re-Use

### Co-occurring elements

- Quality Assurance\*
- Documentation & Metadata\*
- Data Backups (and security)



# Data Management Plan Elements

## ▫ Description & Administration

- Acquisition & Collection
- Organization, Storage & Security
- Processing & Analysis
- Preservation & Archiving
- Sharing & Re-Use

### Co-occurring Elements:

- Quality Assurance
- Documentation & Metadata
- Backup & Security

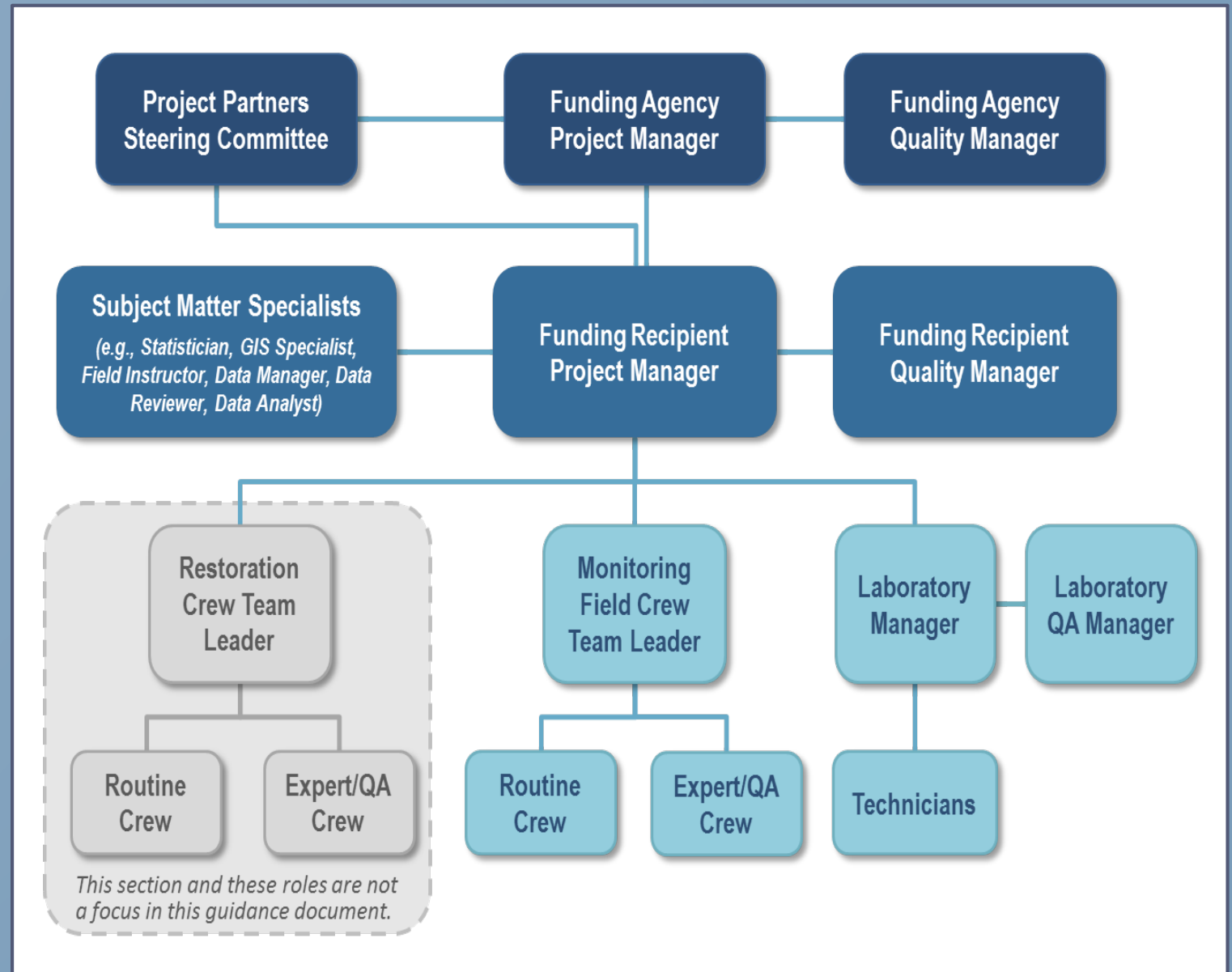


*The who, what, where, when and why of project management.*

- ❑ Managing organization, including roles and responsibilities of key staff involved in data management
- ❑ Summary of project goals and objectives, geographic scale and timeline
- ❑ Funding institutions and key policy stipulations
- ❑ Established 'data-use agreements' and proprietary interests

# Planning and Organizational Structure

Data management roles and responsibilities should be assigned throughout the project's organizational structure



# Data Management Plan Elements

- Description & Administration ✓
  - Acquisition & Collection
  - Organization, Storage & Security
  - Processing & Analysis
  - Preservation & Archiving
  - Sharing & Re-Use
- Co-occurring Elements:
- Quality Assurance
  - Documentation & Metadata
  - Backup & Security



*What are the data, how are they acquired or collected?*

- Type and volume of data expected to be acquired and generated
- Secondary data sources, and logical rules guiding acquisition
- Methods (SOPs) of collection of primary data sources
- Data management training
- Analytical laboratories
- Specialized instrumentation

# Example Process Schema

Schematic diagram illustrating data management and QA/QC (white shaded) procedures

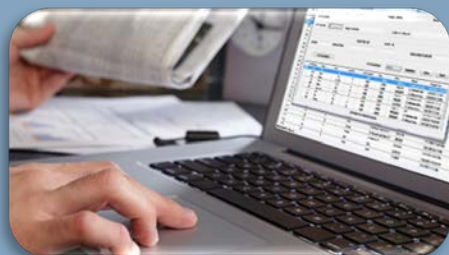


susitna-watanahydro.org

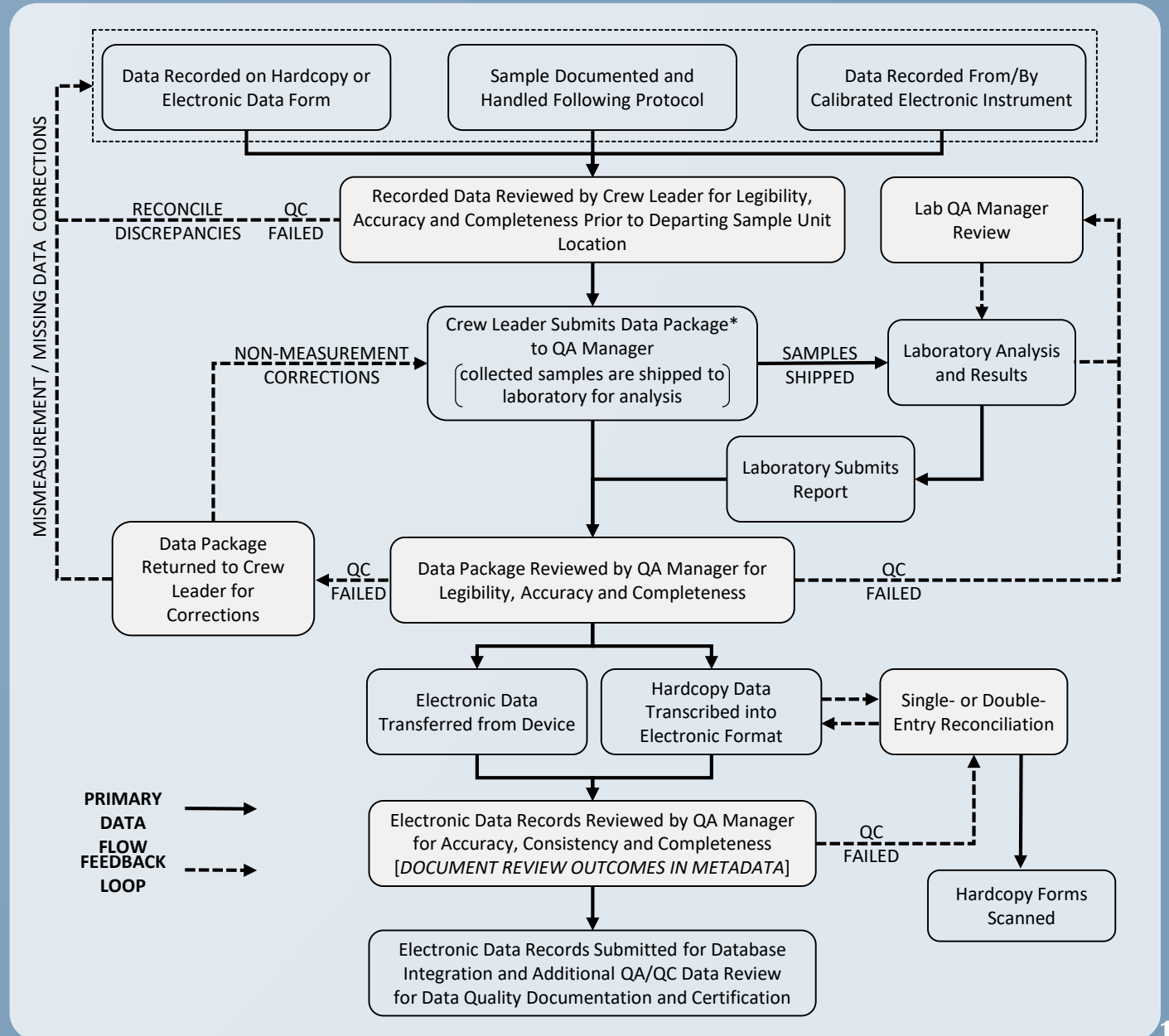
From Field  
to Office



epa.gov



invensis.net



# Data Management Plan Elements

- Description & Administration ✓
  - Acquisition & Collection ✓
  - Organization, Storage & Security
  - Processing & Analysis
  - Preservation & Archiving
  - Sharing & Re-Use
- Co-occurring Elements:
- Quality Assurance
  - Documentation & Metadata
  - Backup & Security



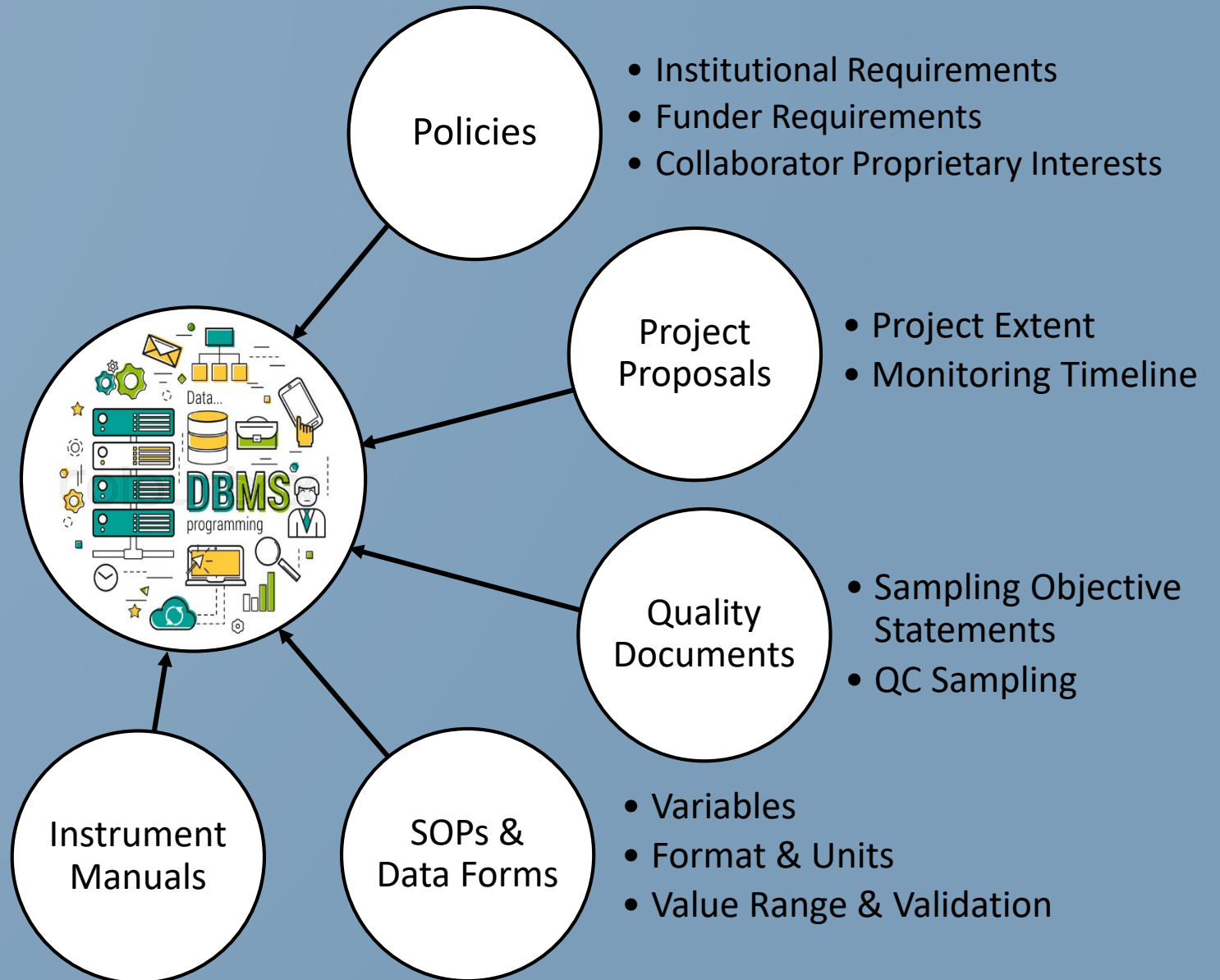
*How, where, and with what are data stored, managed, and secured?*

- Data Management System (DMS)
- Logical work flows to guide storage of 'raw' and processed data
- Hard-copy documentation – data forms, log books, custody forms
- Filename conventions, version control and backup and restore plans
- Policies guiding data read/write access and censorship

# Data Management System

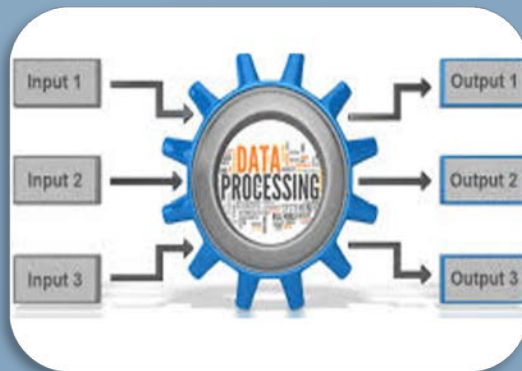
Project documentation provide important details informing DMS development

- Type and volume of data to be collected
- Hierarchy of sampling units
- Spatial and temporal scale/resolution
- Domain ranges and valid values



# Data Management Plan Elements

- Description & Administration ✓
  - Acquisition & Collection ✓
  - Organization, Storage & Security ✓
  - **Processing & Analysis**
  - Preservation & Archiving
  - Sharing & Re-Use
- Co-occurring Elements:
- Quality Assurance
  - Documentation & Metadata
  - Backup & Security



*Policies and procedures involved in data manipulation.*

- ❑ Electronic data-entry and digital file transfer protocols
- ❑ Logical workflows to guide data reduction and metric calculation
- ❑ Design and use of database tables and spreadsheets
- ❑ Computer/software utilities and code used in processing and analysis
- ❑ Statistical models and tests used to validate assumptions



## Electronic Data Conveyance

Develop standard protocols to guide users in the digital transfer of electronic data between devices and the data management system (DMS)

- File-naming conventions
- Data format (e.g., raw, converted)
- Device settings & metadata
- Archive original (raw) files

### Scientific Instruments



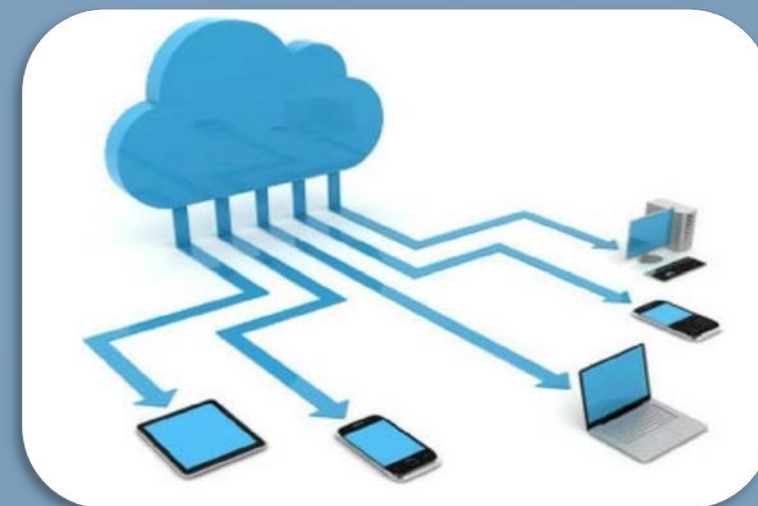
solinst.com

### GPS Instruments



S. Stevens, NPS

### Mobile Devices



# Data Management Plan Elements

- Description & Administration ✓
- Acquisition & Collection ✓
- Organization, Storage & Security ✓
- Processing & Analysis ✓
- **Preservation & Archiving**
- **Sharing & Re-Use**

Co-occurring Elements:

- Quality Assurance
- Documentation & Metadata
- Backup & Security

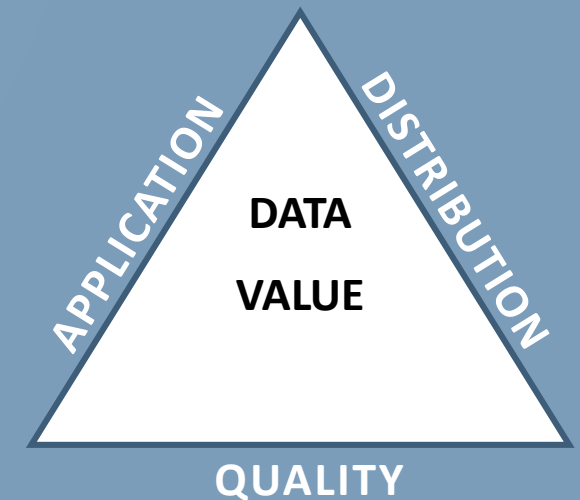


*Enhancing data value by making data available for future and secondary uses.*

- ❑ Policies that guide data sharing for intended primary applications and secondary re-use
- ❑ Identify 'Community of Interest'
- ❑ Data exchanges and repositories
- ❑ Control of access to 'sensitive information' (censorship)
- ❑ Digital Object Identifiers (DOI)

## Digital Data Archives and Repositories

- Gulf of Mexico Coastal Ocean Observing System (GCOOS) Data Portal  
**data.gcoos.org**
- Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC)  
**data.gulfresearchinitiative.org**
- NOAA's Data Integration Visualization Exploration and Reporting (DIVER) Explorer  
**diver.orr.noaa.gov**
- EPA's Water Quality eXchange (STORET/WQX) 34 data systems  
**epa.gov/waterdata/water-quality-data-wqx**
- Data.Gov U.S. Government Open Data - **data.gov**
- Knowledge Network for Biodiversity - **knb.ecoinformatics.org**
- Dryad Digital Repository - **datadryad.org**



# Data Management Plan Elements

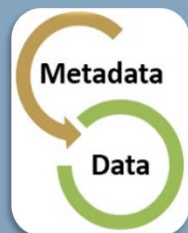
- Description & Administration ✓
- Acquisition & Collection ✓
- Organization, Storage & Security ✓
- Processing & Analysis ✓
- Preservation & Archiving ✓
- Sharing & Re-Use ✓

Co-occurring Elements:

- Quality Assurance ✓
- Documentation & Metadata ✓
- Backup & Security ✓

*On-going activities to maintain data quality, ease of use, and protection.*

- ❑ Ensuring data reliability and logical consistency.
  - QA strategies that maintain data integrity across all DM activities
- ❑ Creating a 'fingerprint' to identify and describe your data.
- ❑ Securing and protecting data for intended use and secondary application (re-use).



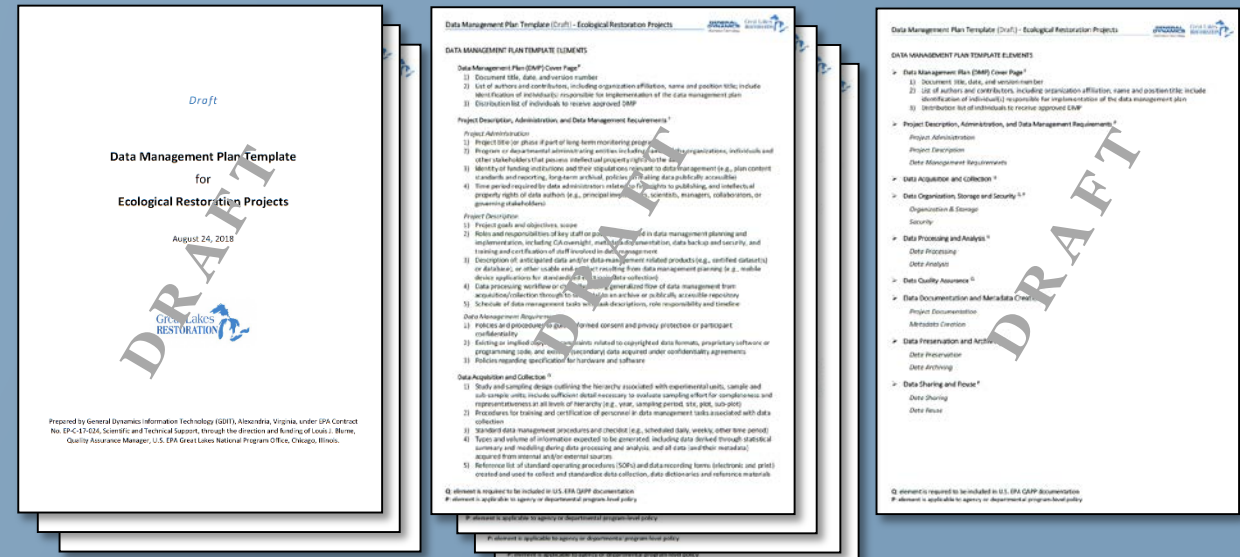
# Summary

- Data management planning should be considered equally important as other best practices conducted in ecological restoration
- Data management planning involves 3 principle components:
  - 1) data management policy,
  - 2) data management system (DMS), and
  - 3) a data management plan.
- Data management planning should be comprehensive and address all data management needs throughout the data life-cycle
- Data stewardship includes preservation and archival to facilitate data sharing and reuse

# Questions ?

Brick Fevold, GDIT

Email: [brick.fevold@gdit.com](mailto:brick.fevold@gdit.com)



• DMP Template Handout\* (near room entrance)

\* Comments welcome through 2018