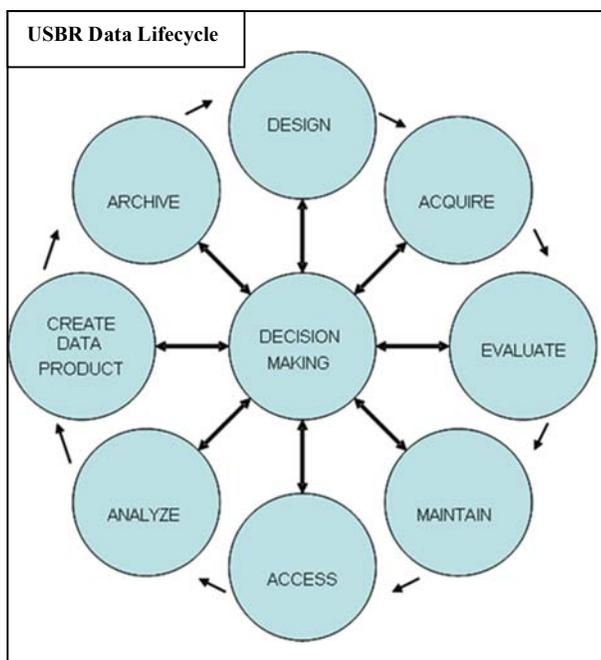


# RECLAMATION

*Managing Water in the West*

## Research Project: Pilot Testing Data Stewardship Processes

Research and Development Office  
Science and Technology Program  
Final Report 2014-01 3760



Douglas Clark, Ph.D., James Nagode, Arthur Coykendall, Barbara Simpson, Julie Eldredge



Research and Development Office  
Bureau of Reclamation  
U.S. Department of the Interior

September 2014

## **Mission Statements**

The U.S. Department of the Interior protects America's natural resources and heritage, honors our cultures and tribal communities, and supplies the energy to power our future.

The mission of the Bureau of Reclamation is to manage, develop, and protect water and related resources in an environmentally and economically sound manner in the interest of the American public.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188		
T1. 09/30/14		T2. REPORT TYPE Research		T3. DATES COVERED 2012-2014	
T4. TITLE AND SUBTITLE Pilot Testing Data Stewardship Processes			5a. CONTRACT NUMBER: 14XR0680A1-RY1541DS201323760 (8)		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER 1541		
6. AUTHOR(S) Douglas Clark, Ph.D., James Nagode, Arthur Coykendall, Barbara Simpson, Julie Eldredge			5d. PROJECT NUMBER 3760		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER 86-68260		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Douglas R. Clark, Ph.D. 86-68260					
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Research and Development Office U.S. Department of the Interior, Bureau of Reclamation, PO Box 25007, Denver CO 80225-0007			10. SPONSOR/MONITOR'S ACRONYM(S) R&D: Research and Development Office BOR/USBR: Bureau of Reclamation DOI: Department of the Interior		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 2014-01 3760		
12. DISTRIBUTION / AVAILABILITY STATEMENT Final report can be downloaded from Reclamation's website: <a href="https://www.usbr.gov/research/">https://www.usbr.gov/research/</a>					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT ( <i>Maximum 200 words</i> ) Reclamation assembled an inter-regional research team to investigate data stewardship processes in the agency with the prime focus on river restoration programs. The team learned that technology and collaboration are the primary predictors of successful data management. Other important factors included upper management support, data management expertise, holistic planning, and adequate funding. In the future, Reclamation will join DOI Open Data efforts and tailor agency policy to Departmental requirements and recommendations. In addition, it will actively support efforts to rationalize management of mission critical data sets such as those pertaining to water, power, and endangered species. Finally, it will make recommendations for data stewardship institutional arrangements within Reclamation.					
15. SUBJECT TERMS data stewardship, data management, data lifecycle					
16. SECURITY CLASSIFICATION OF: U (Unclassified)			17. LIMITATION OF ABSTRACT  U	18. NUMBER OF PAGES  17	19a. NAME OF RESPONSIBLE PERSON Douglas R. Clark
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER 303-445-2125

# PEER REVIEW DOCUMENTATION

## Project and Document Information

Project Name: **Pilot Testing Data Stewardship Processes on River Restoration and Hydrologic Data Sets**

WOID: X3760

Document: Pilot Testing Data Stewardship Processes

Document Author(s): Douglas Clark, Ph.D., James Nagode, Arthur Coykendall, Barbara Simpson, Julie Eldredge

Document date: September 2014

Peer Reviewer: Curtis Brown, Ph.D.

## Review Certification

**Peer Reviewer:** I have reviewed the assigned items/sections(s) noted for the above document and believe them to be in accordance with the project requirements, standards of the profession, and Reclamation policy.

Reviewer \_\_\_\_\_

(Signature)

Date reviewed \_\_\_\_\_

9/9/14

## **Acknowledgements**

The Data Stewardship Team would like to acknowledge the cooperation it has received from Reclamation's River Restoration Data Stewards who provided invaluable insights as to the data management challenges they face. We are also indebted to Reclamation researchers for providing us with important information about their data storage and processing requirements.

# Executive Summary

The purpose of this research was to investigate the requirements for implementing data stewardship best practices within Reclamation with river restoration data management efforts as the primary focus.

The researchers investigated various DOI data stewardship policies and procedures in an effort to align Reclamation initiatives with them. They worked with inter-regional teams to learn what data stewardship initiatives were underway within Reclamation river restoration projects, what tools have been developed, what has been achieved, and what challenges remain. They sought to learn what foundational documents should be prepared in advance of the implementation of data stewardship in Reclamation and to the extent possible prepare them. They worked to develop and recommend institutional arrangements for the management of data within Reclamation. Finally, they examined various technologies for storing, documenting, and sharing mission critical data.

An the researchers and the inter-regional teams determined that convincing the Reclamation Leadership Team (RLT) of the need for comprehensive data stewardship policy should be the over-arching goal of this effort. Briefings have been held in pursuit of this goal. Other team accomplishments included:

- Development of a River Restoration Data Stewardship Community of Interest
- Development of a data stewardship SharePoint site
- Development of collaborative partnerships with ongoing data stewardship efforts within DOI bureaus
- A survey was conducted amongst river restoration programs. It found that predictors of successful data stewardship included adoption of relevant technology, strong inter and intra-agency collaboration, management support, consultation with data stewardship and subject matter experts, and holistic planning
- An additional survey was conducted amongst Reclamation researchers. It found that the production of metadata was the exception rather than the rule, issues existed around both the amount of computer storage allowed and existing processing speed, data discoverability was rarely practiced
- The team participated in DOI policy development efforts and in the Water Data Initiative under the direction of LC Regional Director Fulp.

During the course of this work the Data Stewardship Team developed libraries, documents, or templates for documents in the following domains:

- Data Acquisition and Management Plans
- Position Descriptions for Data Stewards
- Metadata Standards
- Data Sharing Agreements
- Data Stewardship Governance Principles (with associated flow charts).
- Briefing papers
- Open Data Policy Documents

## Pilot Testing Data Stewardship Processes

# Contents

<i>Executive Summary</i> .....	<i>1</i>
<i>Contents</i> .....	<i>3</i>
<i>Problem</i> .....	<i>5</i>
<i>The Need for Data Stewardship</i> .....	<i>5</i>
<i>Research Objectives</i> .....	<i>7</i>
<i>Methods</i> .....	<i>7</i>
<i>Results</i> .....	<i>8</i>
<i>Data Information</i> .....	<i>17</i>



## Problem

The purpose of this research was to investigate the requirements implementing data stewardship best practices within Reclamation. At a Reclamation Leadership Team (RLT) meeting in 2011, Dr. Curtis Brown presented the case for data stewardship in Reclamation and offered a draft policy. The Leadership Team directed him and the team he had assembled to conduct a pilot project to determine how data stewardship might be implemented within Reclamation. He chose river restoration programs for this pilot subject.

## The Need for Data Stewardship

Comprehensive data stewardship that encompasses the entire data lifecycle ([Figure 1](#)) continues to be a challenge for river restoration and other data collection activities within Reclamation. The challenges that face data stewards are lack of policy, lack of funding, institutional inertia, hectic staff schedules (that make it difficult to find the time to document data), ignorance of sound data stewardship practices, and the perception that data are the sole property the person(s) collecting and analyzing them. Poor data stewardship practices have resulted in data loss, data destruction, re-collection of the data, and poor management decisions.

Sound science requires careful acquisition, quality control, and management of data. It requires data stewardship. Data must be viewed as assets, whose assured quality is essential for responsible decision-making. Reclamation's mission critical data should be well-documented, discoverable, and accessible. They should be collected and processed using comparable methodologies, maintained using consistent formats, and properly archived.

Stewardship in many domains is a core part of Reclamation's mission "to manage, develop, and protect water and related resources in an environmentally and economically sound manner in the interest of the American public." Reclamation's stewardship ethic expresses itself in many ways: water conservation, river restoration, water reuse, environmental protection, and non-native species control, etc. The broad classes of resources that Reclamation must manage to meet its mission include real property, power resources, water storage and delivery infrastructure, power infrastructure, and threatened and endangered species of plants and animals. To manage these classes of resources, Reclamation must also manage the data and information associated with them. Why? It is because decisions about resources are based upon data and the analysis of data. Such data, then, must be viewed and managed as mission critical assets, as vital resources. This means that they are related to an identified critical resource; difficult, costly, or impossible to recreate, if lost; and widely used by Reclamation offices.

## Overview of the Reclamation Data Lifecycle

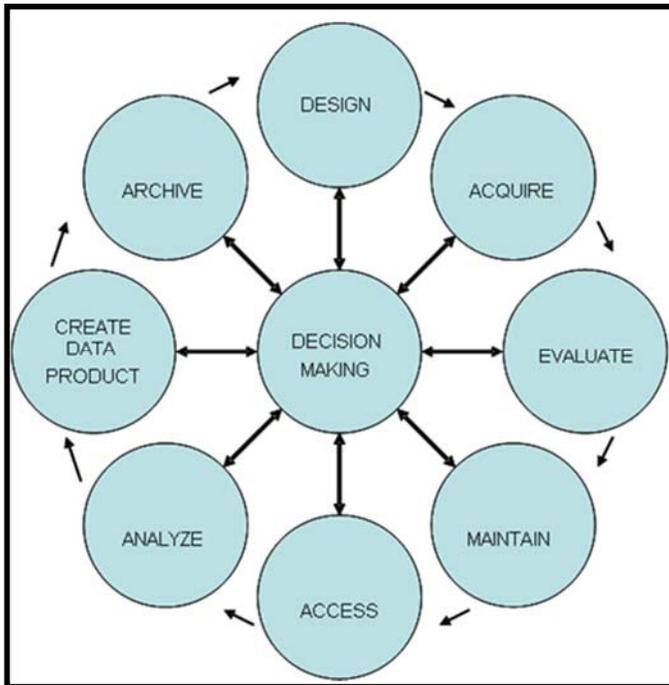


Figure 1: Reclamation Data Lifecycle

needs into variable definitions, sampling plans, collection methods, and recording techniques. It documents specifications, collection protocols, QA/QC procedures, and other requirements for all collected or purchased data.

The **evaluate** stage ensures that the data collected and processed meet the Information Quality Act requirements for objectivity, utility, and integrity. It sets forth the protocols for controlling data quality at each stage of the lifecycle.

The **maintenance** phase of the data lifecycle describes the ongoing processes, procedures and systems for data storage, control, security, and manipulation that must be in place to ensure that the data can provide information to meet the identified business and decision requirements. This phase includes procedures to ensure data and metadata integrity, appropriate updating, and documentation.

During the **access** phase procedures are developed to ensure that the data collection is secure, but also discoverable and retrievable by the designated user community. It goes without saying that successful data sharing implies that metadata associated with the data will be complete and accurate.

In the **analysis** phase the data steward, working with subject matter experts, describes the laboratory, query, modeling, statistical, mapping, graphing, and other investigative techniques that will be used to create the information products needed to meet the original business requirements and the needs of the decision-maker.

Readers new to the concept of data stewardship will want to understand the Reclamation data lifecycle (Figure 1). In the **design** phase of this lifecycle the mission and decision requirements dictated by the business requirement at hand along with data quality metrics and management processes are developed. This phase is concerned with planning the entire data collection and processing effort.

The **acquire** phase speaks to the actual acquisition of the data, whether they come from existing sources, from field work, or from a sensor. A good data acquisition plan translates business/decision

In the **reporting** phase, the information products (reports, maps, charts, etc.) are agreed upon. These products will be delivered to the decision-makers, managers, legislators, or the public.

The **archival** phase describes how the data sets will be retained or retired and the time frames for each. “Archival” refers to the disposition of data after the original business requirements have been met and the need for access will be infrequent. With archival continued discoverability and access will be maintained. Access and maintenance processes for long-term storage are also described in this phase.

**Decision-making** is at the center of the Reclamation data lifecycle because virtually all mission critical data collections are undertaken to inform and guide it.

The data lifecycle phases are applicable during the period of time for which data have value for their original business requirement. The useful life of the data, however, may extend beyond the original business requirement to serve subsequent requirements of decision-makers (original and future), stakeholders, and/or the public.

Information sharing creates a community of interest around the data sets which may extend their value beyond the original purpose. The determination of the useful life of any data set, and the disposition of those data after the original business requirement has been served, are the joint responsibilities of the decision-maker, project lead, senior Reclamation management, partners, and the business data steward, in consultation with the community of interest.

## Research Objectives

The objectives of this research effort included:

- Investigate Department of the Interior data stewardship policies and procedures and align Reclamation initiatives with them.
- Work with inter-regional teams to learn what data stewardship initiatives are underway within and outside Reclamation river restoration projects, what tools have been developed, what has been achieved, and what challenges remain.
- Learn what foundational documents should be prepared in advance of implementation of data stewardship in Reclamation and to the extent possible prepare them.
- Develop and recommend institutional arrangements for the management of data within Reclamation.
- Examine various technologies for storing, documenting, and sharing mission critical data.

## Methods

There were seven tasks associated with this research and development effort:

- Assembling a data stewardship team.
- Opening internal and external communication.
- Developing collaborative partnerships.
- Developing and implementing surveys.

## **Pilot Testing Data Stewardship Processes**

- Creating documents, document templates, and document libraries.
- Developing recommendations as to the institutional arrangements for the implementation of data stewardship.
- And assessing data stewardship tools and technologies.

# **Results**

## **Data Stewardship Team Formation, Organization, and Meetings**

On November 26, 2012, Michael Gabaldon, Director of Technical Resources, sent out a letter soliciting Regional Offices to contribute team members to the river restoration pilot project. A team of 14 Reclamation staff was assembled to explore data stewardship practices across new and ongoing river restoration programs.

The inter-regional Data Stewardship Team (DST) met in January and May of 2013 to discuss the way forward for the development of related policy, standards, and tools within the agency. The teams determined that the overarching goals of this effort were: (a. to successfully demonstrate to the Reclamation Leadership Team (RLT) the need to adopt a comprehensive policy that would encourage and support the stewardship of mission critical resources data and (b. to propose methods for policy implementation. In undertaking these objectives, the team should make sure that the policy and methods served both current and future data needs-- not only of existing river restoration projects, but other repositories of mission critical data within the agency. In addition, the team saw the need to recommend the requirement to assign mission area data stewards to a selected set of information classes or domains such as water, power, lands, and ecological resources. These stewards would identify and inventory mission critical data sets and develop standards for their management through the data lifecycle.

During these two meetings, a consensus also developed that swift action should be taken to develop a community of interest/practice (COI) amongst river restoration project data stewards and then build bridges between and among them—but then also between them and other established COIs (e.g. Lands, GIS, etc.) to actively foster the implementation of data stewardship standards and practices in the agency. A demonstration Data Acquisition and Management Plan and Template was developed for the new COI review and comment upon.

With respect to specific phases of the Reclamation Data lifecycle, the DST determined that it should identify and use established and recognized standards, protocols, and best practices where possible. For instance, it should identify existing QA/QC project plans to ensure that acquired data met planning and management objectives. It should also inventory existing metadata standards and processes. Useful tools for data access, maintenance, and archival should be identified. In addition, data sharing agreements should be reviewed and evaluated for their current and long term utility.

Finally, the DST also discussed possible institutional arrangements within Reclamation for the management of its data assets. It saw the need to draft definitions for the roles of business data stewards, a data resources manager, and a mission critical data stewardship

coordination team. It was determined that testable model for coordinating data stewardship activities should be developed and evaluated.

### Collaborative Partnerships

A Core Data Stewardship Team (Core Team), a sub-team of the DST, held weekly meetings to steer the pilot project, but frequently solicited input from other team members and from outside agencies. A data stewardship SharePoint site was developed to further foster communication and to host important documents (see: <https://dosp/techResc/TR/RD/data-stewardship/SitePages/Home.aspx>). In addition, the core team initiated a program of speakers in 2014 to communicate the message of data stewardship across the agency.

The team developed collaborative partnerships with data stewardship experts in the US Geological Survey, the Bureau of Land Management, the National Park Service, the US Fish and Wildlife Service, and NatureServe. Individuals from these agencies served on a Data Stewardship Advisory Team. The DST also developed partnerships with the following river restoration programs from whom it routinely asked for input:

- Platte River
- Gila River
- Lower Colorado Multi-Species Conservation Program
- Battle Creek Salmon and River Restoration Project
- Central Valley Project Improvement Act
- San Joaquin
- Federal Columbia River Power System: Tributary Habitat Program
- Savage Rapids
- Trinity River Restoration Program
- Glen Canyon Adaptive Management Program
- Middle Rio-Grande
- Upper CO Recovery Program

### River Restoration Survey

The Core Team developed a survey of current data stewardship practices and distributed it to 17 ongoing river-restoration projects to determine what lessons had been learned during the time that their data stewards had managed project data assets.

The survey was based upon the Reclamation Data Lifecycle (Figure 1). Respondents were asked to answer open-ended questions as to the successes and challenges they had experienced for each of the eight phases of the Reclamation data lifecycle. In addition, they were asked to describe the successes and continued challenges related to both data standards and the use of data collected to inform Reclamation decision-making. Finally, each project contact was asked if a written data management plan had been prepared for his/her river restoration project.

The researchers contacted 17 river restoration projects and received 14 completed surveys. *Technology, collaboration, funding, data standards development, expertise, and integrated data management planning* were the most frequently mentioned themes overall. Responses from the various programs indicated that overall they faced similar challenges.

## Pilot Testing Data Stewardship Processes

Like many mission critical data collections in Reclamation, river restoration programs are large and complex enterprises that can generate substantial volumes of data in a variety of formats, often gathered from multiple agencies. These data cannot be managed without the application of *technology*, the most frequently mentioned elements seen as essential to data management success. But, significantly, application of technology alone was insufficient. Without internal and external *collaboration* (the second most-mentioned requirement for success), fully integrated data management *planning*, the implementation of data *standards*, and the use of pertinent data management *expertise*, the data collection and management efforts risked disorganization or failure according to survey respondents.

Without these additional elements, project leads faced administering data fraught with problems. Without them, data sets were acquired using incompatible protocols, naming conventions, formats, sampling designs, and custody practices. And they were processed with incompatible QA/QC processes and analytic procedures. And metadata documentation was inadequate or lacking altogether. Without comprehensive collaboration, planning, standards, and expertise in place, respondents reported trying to manage disorganized and uncoordinated, autonomous data collections undertaken by a variety of cooperators who may or may not be communicating. In short, personnel, including contractors, often took off in different directions, unsure of what they were expected to do, and returned with data products that did not meet project or decision-maker requirements.

Other challenges project data stewards reported facing included:

- determining and implementing data sharing best practices,
- developing successful archives,
- managing multiple databases on multiple computers over multiple years,
- managing legacy data and unconventional data (e.g. specimen collections),
- overcoming field work challenges,
- managing data interpretability issues,
- making certain that contractors delivered quality data,
- serving both technical and policy requirements,
- mission/scope creep,
- lack of funding,
- and managing high volumes of data requests.

Finally, with respect to the question of written data management plans, of the 14 river restoration projects, 9 reported *having no written data management plan*, 3 reported that they were working on a plan or had a draft plan, 1 reported having a partial plan, and only 1 reported having a fully developed plan.

## Reclamation Research Community Survey

The Core Team distributed a second electronic survey to the Reclamation R&D community to better understand the types and amount of the computer resources researchers required to manage and analyze research project data. The survey solicited information about storage needs, access requirements, computer speed requirements, metadata usage, policy guidance, arrangements for the management of sensitive data, and any other data management challenges researchers faced. Several user profile questions were also asked. The response rate was 35%. The vast majority (87%) of respondents

were principal investigators on 2 or fewer research projects. Among the findings were the following:

### Storage and Processing

- The research that occurred using spreadsheets, model runs, GIS layers, imagery, documents, drawings, and databases, primarily took place in the 0-10 gigabyte range.
- In the >10-100 GB range, photos and other imagery predominated.
- In the >100 GB range model runs, imagery, and drawings predominated.
- The primary data storage devices were the Reclamation network drives and personal computers.
- Reclamation network and external hard drives were the back-up systems of choice. Ten (27%) of the researchers had more than 1 back-up system and 1 (2.7%) had 3.

### Data Stewardship Tools and Software

- The most frequently used data format types were spreadsheets, documents, and photos.
- When asked, “Describe any information management systems (e.g. software, services, frameworks, methods, etc.) you use to maintain, organize, analyze and document your data”, surprisingly, 45.9% of researchers indicated that the question was either not applicable to their project or that they made no use of information management or analytic software. For those who did, the most frequently used types were spreadsheets, discipline-specific software, and numerical/statistical packages.

### Data Documentation

- The production and storage of metadata with the project data was far the exception and certainly not the rule.

### Data Access, Discovery, and Sharing

- The most commonly used processes for data sharing were in rank order: email, physical storage devices such as thumb drives or DVDs, and Reclamation network drives.
- 68.6% of respondents said that their data was or would be archived on a Reclamation network drive, a place, significantly, where they would not readily be discoverable to potential users interested in examining or consuming them.
- When researchers were asked where other researchers, managers, or the public could discover their data they reported that they could be discovered in rank order:
  - on a USBR website (26.3%),
  - on PropC/R&D reports (21.1%),
  - by contacting the researcher (21.1%),
  - in a publication (15.8%),
  - on an external website (10.5%),
  - on the Reclamation network (10.5%),
  - by contacting a research partner (10.5%),
  - or by attending a professional meeting (5.3%).

## Pilot Testing Data Stewardship Processes

- Of 23 total responses, 11 (47.8%) were for websites of one type or another, indicating that about half could readily be discovered electronically.

### Data Security and Sensitivity

- Only 10.5% of respondents reported managing sensitive data. Those who did reported using proprietary internet protocols, putting the data under the control of another agency, or strictly restricting access to team members.
- Less than half (45.7%) of responding researchers stored their data in more than one location.

### Data Management Challenges

- 15.8% of respondents reported that sufficient speed in accessing or transferring data was a significant constraint on where they stored their data.
- The most reported overall challenges for data management among researchers were related to data access and sharing.

### Data Stewardship Policy

- 78.2% of respondents wrote that they worked in a unit that did not have any official guidance for the retention, security, and storage of data.

### Other

- When asked what other information, suggestions, and comments they had with respect to their R&D data, the most frequently mentioned remarks suggested that:
  - R&D should have a data port,
  - Reclamation should have a data management program,
  - the Reclamation FTP site should be upgraded,
  - overall, Reclamation should upgrade its data sharing capability.

## Documents

During the course of this work the DST developed libraries, documents, or templates for documents in the following domains:

- Data Acquisition and Management Plans
- Position Descriptions for Data Stewards
- Metadata Standards
- Data Sharing Agreements
- Data Stewardship Governance Principles (with associated flow charts).
- Briefing Papers
- Open Data Policy Documents

## Institutional Arrangements for Data Stewardship Implementation

The President signed Executive Order 13642 in May 2013. This document provided that, going forward, newly generated government data must be freely available in open, machine-readable formats, while safeguarding privacy, confidentiality, and ensuring security. In related Memorandum M-13-13, OMB Director Sylvia Burwell wrote:

*Information is a valuable national resource and a strategic asset to the Federal Government, its partners, and the public. In order to ensure that the Federal Government is taking full advantage of its information resources, executive departments and agencies (hereafter referred to as "agencies") must manage information as an asset throughout its life cycle to promote openness and interoperability, and properly safeguard systems and information. Managing government information as an asset will increase operational efficiencies, reduce costs, improve services, support mission needs, safeguard personal information, and increase public access to valuable government information.*

To repeat, then, Open Data Policy requires agencies to manage their data as assets, release them to the public where practicable, and manage them in ways that make them usable and discoverable. In addition to being accessible, Federal data must be well documented, available without restriction on use, complete, timely, and well managed--with clear points of contact. (See: <http://project-open-data.github.io/principles/>).

The Open Data initiative has provided additional impetus for the continuation of ongoing data stewardship efforts in Reclamation and it has provided a structural framework. The Core Team is working collaboratively with the DOI team in developing related data management objectives, responsibilities, and policy. The DOI initiative has set forth the following eight goals:

- “The department and its bureaus will have a mature data asset management life cycle following consistent, flexible and effective processes
- DOI employees (in both mission and support roles) will understand their role in ensuring we have an effective data management program that provides benefit to the department, bureaus and the public.
- All employees who produce or procure data adhere to tenets of data stewardship.
- Continuous enhancements in data management processes result in measurable improvements in accomplishing our missions and achieving business value.
- Build a strong data management foundation with a set of clear processes that interpret and implement relevant laws, regulations, and mandates for meeting the missions of the department, bureaus, and offices.
- Departmental and bureau data will be well-documented, of known quality and continually improved over time.
- Data is sharable, discoverable, accessible, reusable, and its value is realized.
- Employees are provided the necessary and appropriate infrastructure, tools, resources and recognition for them to be effective.”

At present, the Core Team is working on ways to incorporate DOI Open Data Policy and related goals and objectives into its own policy recommendations. Overall, the team is working collaboratively with DOI partners to prepare the ground for the implementation of data stewardship in Reclamation on many fronts. One of these is *The Water Data*

## **Pilot Testing Data Stewardship Processes**

*Initiative* and with its state directive to “establish and implement within Reclamation a process to publish water and other mission-related data, in support of Federal open data policy (President’s Executive Order 13642) and making Reclamation resources data more comparable and sharable.” This initiative also:

- Works toward the integration of both spatial and non-spatial natural resources data with a primary initial focus on water data
- Supports the adoption of consensus standards to improve data access, evaluation, and sharing.
- Coordinates efforts to make Reclamation’s water data (e.g. reservoir storage data) more accessible and comparable.
- Builds synergy between Reclamation data stewardship efforts and the efforts of other DOI agencies such as the USGS, NPS, and BLM.

Pursuant to its collaborative efforts with DOI Open Data and Water Data Initiative partners and to its efforts to lay the groundwork for implementation of data stewardship in Reclamation, the Core Team has developed a set of purposes for a new data stewardship function to be called the Mission Critical Data Stewardship Coordination Council (MCDSCC). These include:

### For Communities of Practice

- Understand the full range of data stewardship requirements, activities and practices,
- Identify key data domains within Reclamation that collect and house mission critical data
- Establish data stewardship communities of interest and practice
- Identify and share of useful data management practices and tools

### For Data Stewards

- Support for the creation and maintenance of a body of data stewards
- Development of data stewardship training activities

### In Support of Interoperability and Data Standards

- Facilitate the linking of complementary data, reports, drawings and other content
- Facilitate data standardization and sharing
- Assist in data management document development including: data management plan templates, position description templates, model directives and standards documents, model data sharing agreements, model metadata applications, and the like.
- Facilitate development of bureau-wide data and metadata implementation plans based on DOI policy

### In Support of Policy and Budget

- Conduct data stewardship policy analysis and make policy recommendations.
- Represent mission data needs (Regional and Denver Directorates) by assisting with:

- agency positions
- statements of business needs
- policy and proposal reviews from Agency, Department and Administration
- Serve as point of contact for coordination with the DOI Data Governance Board.
- Coordinate implementation plans for Administration, DOI, Congressional, and Reclamation data stewardship policies.
- Be the outward-facing representative of Reclamation's mission critical data stewardship requirement to DOI, Congress, and the Administration.
- Coordinate strategies to plan and secure the resources necessary to accomplish these core principles.

Significantly, it should be added, the focus of the MCDSCC will be on maintaining the integrity and value of data critical to the support of Reclamation mission requirements and not on selection of IT systems or IT architecture. True enough, the MCDSCC might provide useful information to the IT community about information management requirements, which can help guide IT decisions. But this group is not convened to oversee or otherwise manage systems or architecture.

### Tools Analysis

The Core Team and the DST have also begun to explore a variety of tools and technologies for implementing data stewardship both within Reclamation and in other agencies.

A sampling of these is listed below:

- *USGS ScienceBase*: Provides data cataloging and collaborative data management platform, a central search and discovery application, web services facilitating other applications, research community catalogs.  
<https://www.sciencebase.gov/catalog/>
- *The Exchange Network*: A cooperative effort with representatives from States, Tribes, Territories, and the EPA working together to provide access to environmental data. All partners share in the management and administration of the network. <http://www.exchangenetwork.net/>
- *SOCRATA*: A company that provides data discovery services for open government data. It targets Internet users who want to view and share government, healthcare, energy, education, or environment data.  
[www.socrata.com](http://www.socrata.com)
- *The eMail Enterprise Records and Document Management System (eERDMS)*. This system captures inbound and outbound government emails and assesses them for records content. It also centralizes, standardizes, and automates all internal and external forms.  
[http://www.doi.gov/ofas/support\\_services/upload/JOHN\\_FINAL\\_MDHH-EDIT-06262013\\_eERDMS-EFS-General-Presentation-06212013.pdf](http://www.doi.gov/ofas/support_services/upload/JOHN_FINAL_MDHH-EDIT-06262013_eERDMS-EFS-General-Presentation-06212013.pdf)
- *Pacific Northwest Aquatic Monitoring Partnership (PNAMP)*: a forum to facilitate collaboration around aquatic monitoring, promote best practices for monitoring, and encourage coordination and integration of monitoring activities including data sharing. <http://www.pnamp.org/>
- *The Water Data Exchange (WaDE)*: This capability is designed to better enable states to share important water data with each other, the public and Federal

## Pilot Testing Data Stewardship Processes

agencies, improve the sharing of Federal data with the states, to assist their planning efforts. <http://www.westernstateswater.org/wade/>

- *The Open Water Foundation*: a nonprofit social enterprise that is working to provide an open source software platform to help organizations make better decisions about water. It also supports the Colorado Decision Support Systems. <http://openwaterfoundation.org/about-owf/overview>; <http://cdss.state.co.us/Pages/CDSSHome.aspx>

## Future Actions

Future efforts will be focused on developing institutional arrangements for the implementation of data stewardship within Reclamation. This will involve building on past efforts while coordinating with DOI Open Data efforts and the Water Data Initiative. Specific tasks will include:

- Survey of data stewardship training materials across DOI and assess their utility for our agency.
- Survey and assess metadata and other data management tools, methods, contracting documents, data sharing agreements, data management plans and the like for their utility in Reclamation.
- Survey and assess the data storage and exchange capabilities of other agencies for their utility in Reclamation.
- Develop a release process for Reclamation data.
- Explore Reclamation IT infrastructure for achieving water, hydro power, river restoration, and lands data integration.
- Develop and maintain a data stewardship website and maintain the SharePoint site.
- Work with river restoration, water storage, hydropower, and lands data stewards and subject matter experts to develop a set of data stewardship best practices.
- Develop working partnerships with other DOI bureaus to further the cause of data stewardship. Where appropriate, collaborate on joint projects.
- Charter and stand up a Reclamation wide data stewardship coordination committee.
- In the first year, support and coordinate the effort to make Reclamation's water storage data (and associated metadata) more accessible and comparable across projects and Regions. (We would coordinate this effort with other Federal, state, and local agencies to leverage knowledge and resources.)
- In subsequent years, we will focus on hydropower data, then move to land resources, and thence to data related to species recovery and river restoration. We will focus on building the organizational infrastructure, guidance, and tools that will make it easier for individual data stewardship projects to succeed, and for agency-wide sharing of issues and solutions to occur.
- Host yearly meetings in support of this work.
- Host data stewardship presentations with guest speakers.

## Data Sets that support the final report:

1. Survey of River Restoration Data Stewards
2. Survey of Reclamation Researchers

## Share Drive folder name and path where data are stored:

1. Survey of River Restoration Data Stewards: <https://dosp/techResc/TR/RD/data-stewardship/Shared%20Documents/Forms/AllItems.aspx?RootFolder=%2FtechResc%2FTR%2FRD%2Fdata%2Dstewardship%2FShared%20Documents%2FSurvey%20Results%2FRiver%20Restoration%20Survey>
2. Survey of Reclamation Researchers: <https://dosp/techResc/TR/RD/data-stewardship/Shared%20Documents/Forms/AllItems.aspx?RootFolder=%2FtechResc%2FTR%2FRD%2Fdata%2Dstewardship%2FShared%20Documents%2FSurvey%20Results%2FResearch%20and%20Development%20Survey>

## Point of Contact name, email and phone:

Douglas R. Clark, Ph.D.  
[drclark@usbr.gov](mailto:drclark@usbr.gov)  
303-445-2271

## Short description of the data: (types of information, principal locations collected, general time period of collection, predominant files types, unusual file types.)

Types: survey results

Locations: electronic surveys

Time Period: 2013

File types: Excel spreadsheets

**Keywords:** data, data stewardship, data lifecycle, data management

**Approximate total size of all files:** 100 KB