

# EPA Environmental Dataset Gateway (EDG) Governance Structure and Standard Operating Procedure

---

Version 3.1

**DRAFT**  
**11/18/2015**

## Contents

1.0 PURPOSE .....	1
2.0 BACKGROUND .....	1
3.0 SCOPE AND APPLICABILITY .....	2
4.0 INTENDED AUDIENCE .....	3
5.0 ROLES AND RESPONSIBILITIES .....	3
5.1 EPA METADATA COORDINATOR .....	4
5.2 EPA DATA OWNER .....	4
5.3 EPA DATA STEWARD .....	4
5.4 ENVIRONMENTAL DATASET GATEWAY ADMINISTRATIVE GROUP .....	4
5.5 ENVIRONMENTAL DATASET GATEWAY STEWARD .....	5
5.5.1 EDG Stewardship .....	5
6.0 OVERALL PROCEDURE FOR CONTRIBUTING CONTENT TO THE EDG .....	6
6.1 SET UP A PROGRAM OFFICE, REGION, OR RESEARCH LABORATORY FOR EDG ACCESS .....	6
6.1.1 Designate an EDG Steward .....	6
6.1.2 Create User Logins and Assign User to Appropriate EDG Groups .....	7
6.2 PREPARE TO CONTRIBUTE TO THE EDG .....	7
6.2.1 Understand EPA Metadata Compliancy Requirements .....	8
6.2.2 Determine Resources Warranting EDG Metadata Contribution .....	9
6.2.3 Determine Access and Use Constraints for a Resource or Metadata Record .....	9
6.2.4 Determine Method of Access of a Resource .....	10
6.3 CREATE CONTENT AND CONTRIBUTE TO THE EDG .....	12
6.3.1 Prepare and Contribute Metadata for a Non-Geospatial Resource .....	13
6.3.2 Prepare and Contribute Metadata for a Geospatial Resource .....	13
6.3.3 Post Data to the EDG Clip N Ship Application .....	15
6.3.4 Post Data to the EDG Data Download Locations .....	15
6.3.5 Utilize EDG Reuse Components .....	16
6.4 UPDATE AND MAINTAIN CONTENT IN THE EDG .....	16
6.4.1 Update Existing EDG Content with New Content .....	16
6.4.2 Set Access Level for a Metadata Record .....	17
6.4.3 Transfer Ownership of a Metadata Record .....	17

6.4.4 Maintain Registration Information for Registered Resources on the Network.....	18
6.4.5 Verify EDG content syncs with Data.gov.....	18
6.4.6 Create and Manage Collections .....	18
6.4.6.1 Add and Search Placeholder Records.....	21
6.4.6.2 Add and Search Members .....	21
7.0 GLOSSARY.....	21
8.0 APPENDICES .....	26
APPENDIX A. KEY EDG CONTACTS.....	26
APPENDIX B. AUTHORITY FOR EDG IMPLEMENTATION AND RELEVANT DOCUMENTS, RELATED PROCEDURES, AND GUIDELINES.....	27
APPENDIX C. WAIVERS .....	28
APPENDIX D. DETAILED STANDARD OPERATING PROCEDURES .....	29
D.6.1.1 Designate an EDG Steward .....	29
D.6.1.2 Create User Logins and Assign User to Appropriate EDG Groups.....	30
D.6.3.1 Prepare, Validate, and Publish Non-Geospatial Metadata to the EDG Using the EDG Dedicated Editor .....	32
D.6.3.2.1a Download the EPA Metadata Editor (EME).....	34
D.6.3.2.1b Prepare and Validate the Metadata Record Using the EPA Metadata Editor (EME).....	35
D.6.3.2.2a Contribute Metadata Records to the EDG – Manual Upload.....	36
D.6.3.2.2b Contribute Metadata Records to the EDG – Synchronization.....	38
D.6.3.2.2c Contribute Metadata Records to the EDG – ArcSDE Synchronization .....	40
D.6.3.3 Post Data to the EDG Clip N Ship Application .....	42
D.6.3.4 Post Data to the EDG Data Download Locations.....	43
D.6.3.5 Utilize EDG Reuse Components.....	44
D.6.4.1 Set Access Level for a Metadata Record .....	45
D.6.4.2 Transfer Ownership of a Metadata Record .....	46
D.6.4.5 Verify EDG Content Synchronizes with Data.gov .....	47
D.6.4.6 Create and Manage Collections.....	48
APPENDIX E. EPA National Geospatial Data Policy (NGDP) Stewardship Agreement .....	50

## 1.0 PURPOSE

This document identifies the oversight and authority organization of the EPA Environmental Dataset Gateway (EDG), provides an overview of roles and responsibilities for implementing the EDG, and details standard operating procedures for functions within the EDG. Roles and responsibilities outlined within this document align upwardly with the U.S. Environmental Protection Agency's (EPA's) [National Geospatial Data Policy \(NGDP\)](#), procedure, and related documents. This document complements the role structure outlined within the [NGDP Procedure for Metadata Management](#) (NGDP Procedure) by clearly defining the relationship between NGDP Procedure roles and EDG contribution and management responsibilities expected of the individuals fulfilling those roles. Furthermore, this document addresses the interconnectedness between EDG role responsibilities and EPA GeoPlatform role responsibilities. For additional information on the EDG, please visit the [EDG website](#).

## 2.0 BACKGROUND

Historically, the EDG has served as the Agency's metadata repository application. However, while the metadata catalog remains at the core of the EDG, it has evolved to include a collection of applications and web interfaces used for management and delivery of geospatial and non-geospatial data. The EDG includes the following components:

- [EDG metadata catalog](#) – the core function of the EDG. The web-based metadata catalog consolidates metadata files maintained by Data Stewards across the Agency and makes those metadata files available to users via simple, advanced, and targeted searches.
- [EPA Metadata Editor \(EME\)](#) – a geospatial metadata editing application that allows users to create and edit records to meet EPA and Federal Geographic Data Committee (FGDC) metadata requirements.
- [EDG Data Download Locations](#) – a web page that allows users to browse through folders of downloadable data posted by EPA's Regions, Program Offices, and Research Laboratories.
- [EPA's Clip N Ship Site](#) – an interactive web mapping application that allows users to select an area of interest and choose an output format to save the geospatial data locally.
- [EDG Reuse Components](#) – a web page that provides widgets and interfaces that allow application developers the ability to reuse EDG content in other applications. These tools permit direct access to EPA's metadata outside of the EDG interface.

The purpose of the EDG is to: 1) Provide EPA staff with a single access point to EPA's geospatial and non-geospatial resources and metadata records, and 2) Meet legislative and regulatory mandates and Agency requirements for metadata compliance, sharing, and management. As the central location for distributed geospatial and non-geospatial resources and metadata records, the EDG supports external federal initiatives such as Data.gov. Figure 1 is a model of EPA's implementation of the EDG.

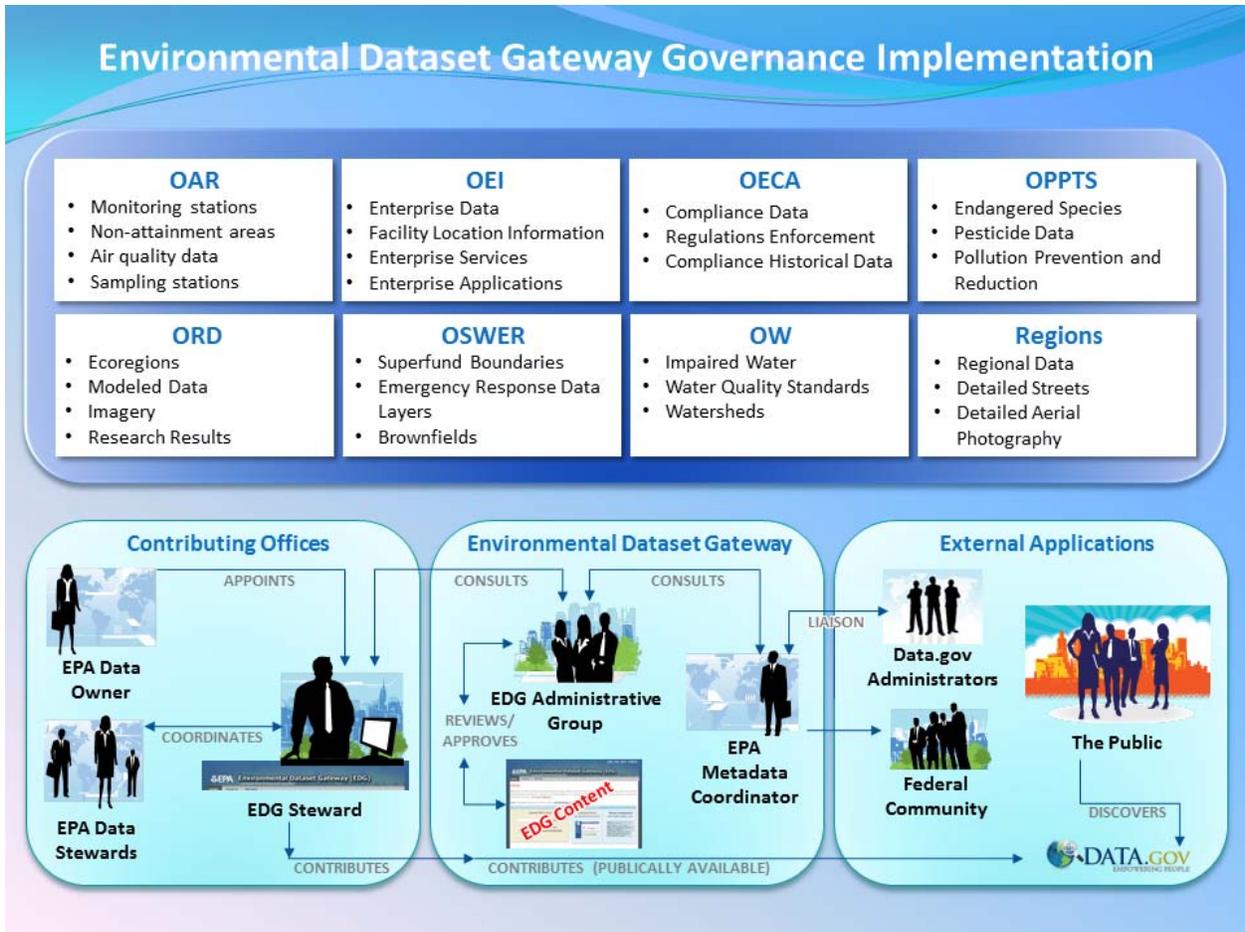


Figure 1: USEPA implementation of the EDG.

### 3.0 SCOPE AND APPLICABILITY

The EDG is a centralized collection of data management applications that consolidates distributed information into a single location. Distributed information is contributed and managed by numerous EDG participants throughout the Agency. As such, a governance structure that details responsibility and authority for EDG contents and standard operating procedures is required and outlined in this document.

Roles and responsibilities that are clearly aligned with EPA’s NGDP and implemented consistently across EPA programs are important for ensuring that information contributed to the EDG is effectively managed and can be traced to owners. Additionally, information presented in this document will enable EDG participants to clearly understand requirements and expectations for contributing to and managing information in the EDG. Agreement by participants to fulfill the specific responsibilities of their role, as detailed in Section 5.0 of this document, continues to make the EDG a useful resource.

## 4.0 INTENDED AUDIENCE

The primary audience for this document includes EPA’s Metadata Coordinator, Data Owners, Data Stewards, the EDG Administrative Group, and EDG Stewards. This document may also be of interest to additional EPA personnel or EPA contractor staff who may find the EDG useful in meeting their resource needs. The secondary audience for this document includes EPA’s owners and stewards of non-geospatial data. The EDG shares EPA metadata that is designated as “unrestricted” with Data.gov. Finally, this document may be considered a resource for individuals with an interest in EPA metadata management.

## 5.0 ROLES AND RESPONSIBILITIES

Roles and responsibilities outlined in this section reflect duties identified for contributing and managing information posted to the EDG (Figure 2). The roles and responsibilities are derived from NGDP Procedure, with the exception of the EDG Steward which is added to the framework in order to identify this role, which has responsibilities unique to the EDG. The specific responsibilities outlined in this section are a subset of the entire suite of geospatial metadata management responsibilities, as identified in the NGDP Procedure. Responsibilities listed in this document include those that relate specifically to operations involving the EDG, and how those responsibilities overlap with EPA GeoPlatform role responsibilities where appropriate. (For details on EPA GeoPlatform data publishing,

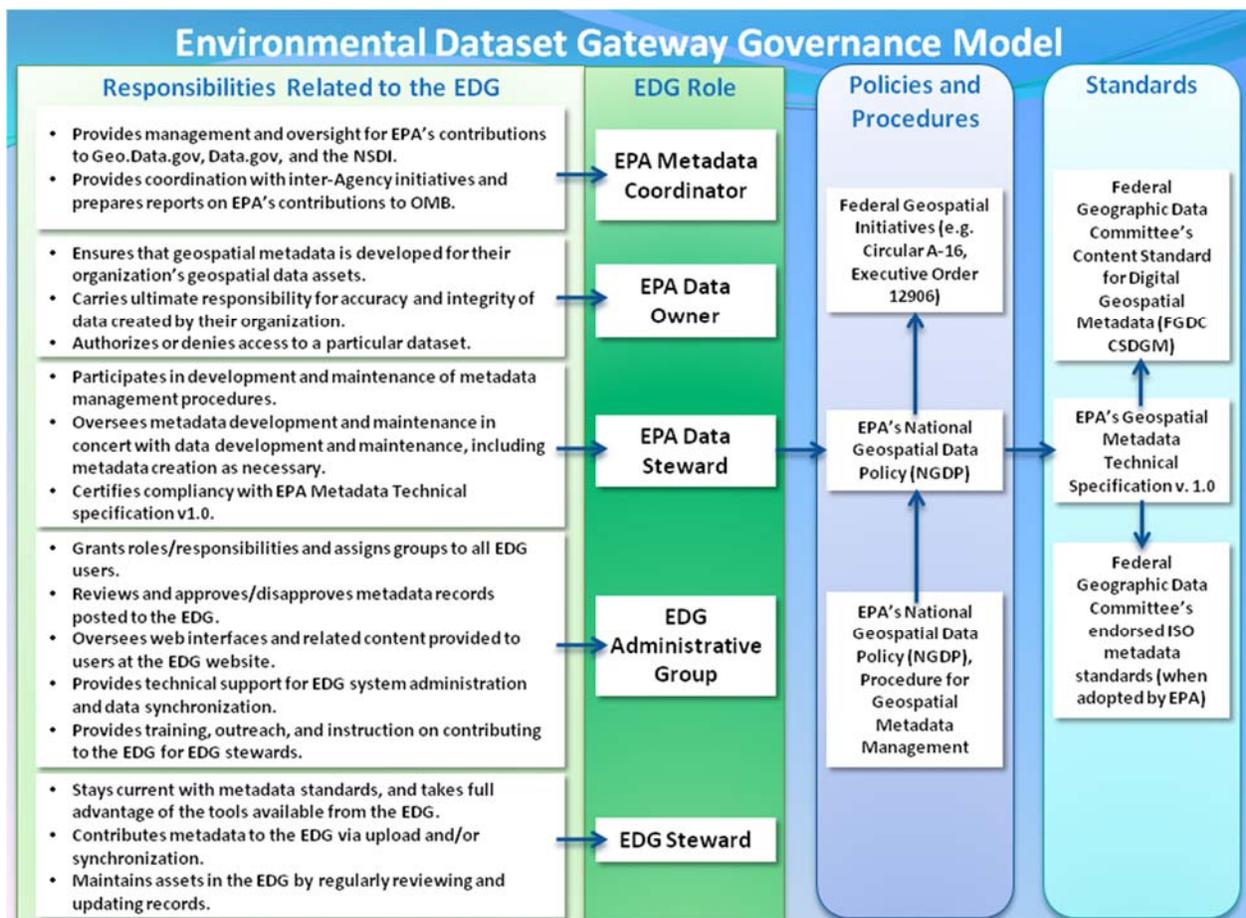


Figure 2: USEPA roles and responsibilities related to the EDG.

please review the [EPA GeoPlatform Data Publishing Standard Operating Procedure](#).) For a full set of geospatial data management responsibilities listed for each EDG role under the NGDP Procedure and for direction on how to appoint these roles within an organization, please refer to the [NGDP Procedure](#).

## 5.1 EPA METADATA COORDINATOR

- **Location:** OEI.
- **Primary Role:** Oversees inter-Agency metadata coordination and reporting.
- **Primary Responsibilities:**
  - Provides oversight for and coordination of metadata contributions to Geo.Data.gov and Data.gov.
  - Prepares and delivers metadata reports to the Office of Management and Budget (OMB).
  - Provides support for cross-Agency metadata coordination efforts.
  - Acts as Project Manager for the EDG, working with the EDG Administrative Group to ensure roles are filled and responsibilities are being followed.

## 5.2 EPA DATA OWNER

- **Location:** EPA Regions, Program Offices, and Research Laboratories.
- **Primary Role:** Manages geospatial and non-geospatial resources for an entire organization.
- **Primary Responsibilities:**
  - Authorizes or denies access to a particular dataset.
  - Carries ultimate responsibility for accuracy and integrity of data created by their organization.
  - Ensures the development of metadata for the data created by their organization.

## 5.3 EPA DATA STEWARD

- **Location:** EPA Regions, Program Offices, and Research Laboratories.
- **Primary Role:** Manages individual datasets, databases, and applications within an organization's data holdings.
- **Primary Responsibilities:**
  - Participates in metadata management for their organization, including creating metadata as appropriate for existing geospatial and non-geospatial resources.
  - Oversees maintenance (review and update) of resources and metadata records in the EDG.
  - Ensures metadata compliancy with *EPA Geospatial Metadata Technical Specification* or the *EPA Implementation of the Data.gov Standard*.
- **Related EPA GeoPlatform Role:** EPA GeoPlatform Data Publishing Steward.
  - **EPA GeoPlatform Responsibilities:**
    - Develops and processes geospatial data.
    - Works with GeoPlatform Data Owner to implement datasets and publish web services and metadata for their organization.

## 5.4 ENVIRONMENTAL DATASET GATEWAY ADMINISTRATIVE GROUP

- **Location:** OEI.
- **Primary Role:** Manages EDG applications and metadata catalog.
- **Primary Responsibilities:**

- Manages and approves information contributed to the EDG metadata catalog.
  - Validates records contributed to the EDG for EPA compliancy.
  - Approves or disapproves records based on validation results.
- Oversees web interfaces and related content provided to users at the EDG website.
- Provides technical support for implementing and maintaining the EDG connection to Geo.Data.gov, Data.gov, and National Spatial Data Infrastructure (NSDI).
- Provides training, outreach, and instruction on contributing to the EDG for EDG Stewards.
- Maintains close contact with EPA Regions, Program Offices, and Research Laboratories to ensure assignment of the EDG Steward role and to address any changes in personnel.

## 5.5 ENVIRONMENTAL DATASET GATEWAY STEWARD

- **Location:** EPA Regions, Program Offices, and Research Laboratories.
- **Primary Role:** Manages and coordinates an organization’s contributions to the EDG.
- **Primary Responsibilities:**
  - Works closely with the EPA’s EDG Administrative Group to plan the organization’s EDG contributions.
  - Stays current with metadata standards, and takes full advantage of the tools available from the EDG.
  - Contributes metadata to the EDG via upload and/or synchronization.
  - Maintains geospatial and non-geospatial resources and metadata records in the EDG by regularly reviewing and updating content.
  - Works within their organization to help others understand how they can reuse EDG resources.
- **Related EPA GeoPlatform Role:** EPA GeoPlatform Data Publishing Owner.
  - **EPA GeoPlatform Responsibilities:**
    - Oversees and is fully responsible for their organization’s content within the EPA GeoPlatform.
    - Ensures organization compliance with requirements for publishing content within the EPA GeoPlatform data publishing environment.
    - Delegates roles and levels of engagement for EPA GeoPlatform data publishing to members of their organization.

### 5.5.1 EDG Stewardship

The EDG Steward role is given additional attention in this section because of its importance in the EDG governance framework. The EDG Steward acts as the central authority for a Region, Program Office, or Research Laboratory’s geospatial and non-geospatial resources and metadata records contributed to the EDG. The EDG Steward also serves as the point-of-contact for other EDG contributors within his or her organization needing assistance with working through processes identified for EDG contribution.

Within EPA programs, the identification of the EDG Steward may vary depending on the nature of the particular organization’s structure. Typically, EPA Data Owners will recommend an EDG Steward for their organization. The EDG Steward identified for an organization may also fulfill one of the other roles listed in this document (e.g., EPA Data Steward or EPA Data Owner). The specific EDG responsibilities listed for the EDG Steward complement other responsibilities identified for the EPA Data Steward and EPA Data Owner roles. If an EDG Steward’s employment is terminated, the EDG Steward will be required to complete an exit interview with his/her supervisor in accordance with EPA standard termination processes.

In general, EDG Stewards will work with EPA Data Stewards in their organizations to:

- Identify and prioritize local geospatial and non-geospatial resources for which metadata should be contributed to the EDG.
- Identify available metadata records for resources identified.
- Create or oversee creation of metadata for resources as appropriate.
- Identify geospatial and non-geospatial resource constraints and characteristics.
- Determine public/private nature of resources.
- Determine resource availability and format (online linkage).

Additionally, the EDG Steward works with the EDG Administrative Group to:

- Understand how to use the EDG.
- Understand non-geospatial metadata requirements according to the Data.gov Standard.
- Understand EPA geospatial metadata requirements according to the *EPA Geospatial Metadata Technical Specification Version 1.0*.
- Contribute metadata to their organization's EDG metadata holdings.
- Maintain geospatial and non-geospatial resources and metadata records on a regular basis, as outlined in the EDG Maintenance Plan.

## **6.0 OVERALL PROCEDURE FOR CONTRIBUTING CONTENT TO THE EDG**

This section provides an overall procedure for contributing to the EDG. The set of process steps described here is intended as guidance and may change based on a particular EPA organization's configuration. The overall procedure is as follows (Figure 3):

1. Set up an EPA organization (Program Office, Region, or Research Laboratory) for EDG access.
2. Prepare to contribute to the EDG.
3. Create content and contribute to the EDG.
4. Maintain and update content in the EDG.

The tasks related to each step in the procedure are provided in this section along with some background information. Detailed procedures, where appropriate, are found in Appendix D.

### **6.1 SET UP A PROGRAM OFFICE, REGION, OR RESEARCH LABORATORY FOR EDG ACCESS**

Preparing an EPA organization for EDG utilization involves designating individuals to fill the local roles as outlined in Section 5 of this document, and providing each individual with access to the necessary tools to perform the functions of his or her role.

#### **6.1.1 Designate an EDG Steward**

The EDG Steward works closely with the EDG Administrative Group to plan the EPA organization's EDG contributions, keep current with metadata standards, and take full advantage of the tools available in the EDG. The EDG Steward is responsible for contributing geospatial and non-geospatial resources and metadata records to the EDG, and maintaining those content items by reviewing records regularly. The EDG Steward acts as a local point-of-contact to assist others in the EPA organization on using EDG resources. Detailed procedures can be found in Appendix D, Section D.6.1.1.

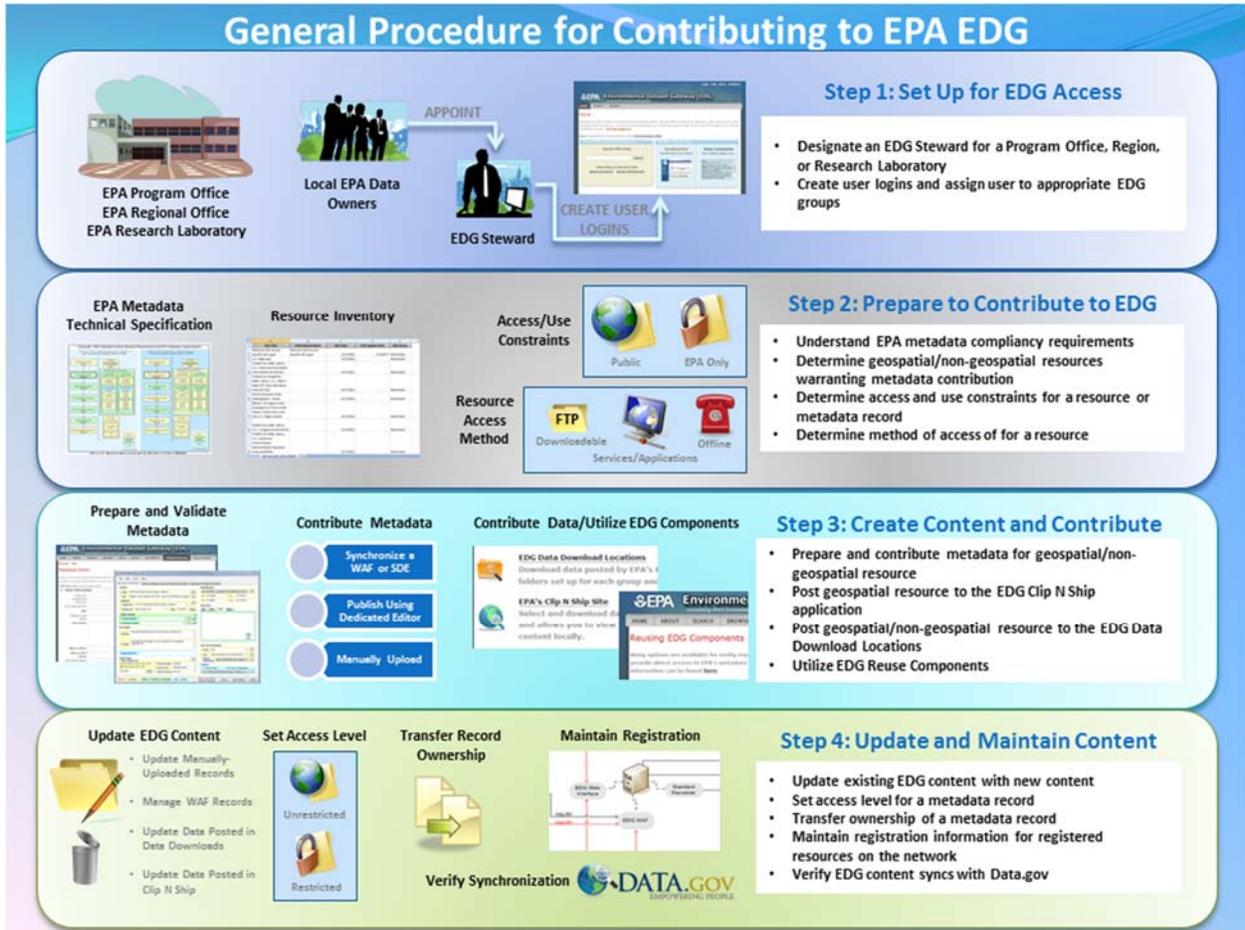


Figure 3: Overall procedure for contributing to the EDG.

### 6.1.2 Create User Logins and Assign User to Appropriate EDG Groups

All users requiring access to restricted information on the EDG must have a user login. Individuals occupying the EPA Data Owner and EPA Data Steward roles who have a responsibility to manage EDG content for their EPA organization must also be assigned to the appropriate EDG groups by the EDG Administrative Group. Detailed procedures can be found in Appendix D, Section D.6.1.2.

### 6.2 PREPARE TO CONTRIBUTE TO THE EDG

Once user roles are filled for an EPA organization, the organization may begin preparing to contribute to the EDG. Geospatial and non-geospatial resources stored and used by EPA programs may be derived from a number of sources. Some resources are produced and maintained by EPA organizations, while many others are obtained from sources outside of the Agency and are used for performing a variety of EPA business operations. Geospatial and non-geospatial resources obtained from outside of the Agency may be accompanied by metadata of varying quality. Often metadata is absent.

With this in mind, an EPA Region, Program Office, or Research Laboratory wishing to contribute content to the EDG must take the following steps in order to prepare their organization:

1. Understand EPA metadata compliancy requirements.

2. Determine geospatial and non-geospatial resources warranting EDG metadata contribution.
3. Determine access and use constraints for a resource or metadata record.
4. Determine method of access for a resource.

Guidelines to clarify the requirements for contributing geospatial and non-geospatial resources and metadata records to the EDG are summarized in Table 1. These guidelines should be used by EPA organizations to develop metadata, update EDG content, and share EDG content with internal and external parties.

<b>EDG Contribution, Prioritization, and Distribution Criteria</b>			
<b>Resource Type</b>	<b>EPA-Compliant Metadata</b>	<b>Metadata Contribution to EDG</b>	<b>Access Constraints</b>
Resources (data, applications, and services) that EPA creates.	Required.	Required.	Unrestricted and available to the public unless data sensitivity, license agreement, or other EPA restriction on public access exists.
Resources acquired from other agencies but with value-added modifications made by EPA personnel.	Required. Personnel should ensure that all EPA-required fields are present in the metadata, but errors are allowable.	Required.	Unrestricted and available to the public unless data sensitivity, license agreement, or other EPA restriction on public access exists.
Resources acquired from external organizations that are of critical importance or may be difficult to locate by other EPA personnel.	Encouraged, but not required. Personnel should ensure that all EPA-required fields are present in the metadata, but errors are allowable.	No, but highly encouraged to contribute to the EDG.	Generally restricted and not made available to the public.

**Table 1: EDG contribution and distribution criteria.**

### 6.2.1 Understand EPA Metadata Compliancy Requirements

**Non-geospatial metadata.** Non-geospatial resources that originate at EPA must have a metadata record registered in the EDG. Metadata for non-geospatial resources should comply with the Data.gov metadata standard. The EDG’s [dedicated metadata editor](#) is designed to create metadata that complies with the Data.gov metadata standard. Once a metadata record is created in the editor, it may be validated using the validation tool provided within the editor.

**Geospatial metadata.** Geospatial resources that originate at EPA must have a metadata record registered in the EDG. Metadata must meet the [EPA Geospatial Metadata Technical Specification v1.0](#). EPA compliancy means that the metadata record includes all EPA-required fields and that the record passes minimum FGDC compliancy. Questions about this determination may be directed to the EDG

Administrative Group or the EPA Geospatial Metadata Coordinator. EPA-required metadata fields are clearly delineated in the EPA Metadata Editor (EME).

In many cases, high-value geospatial resources used at EPA are acquired from organizations outside of EPA. Sometimes the metadata records accompanying these resources contain FGDC compliancy errors. In these cases, it is permissible to allow records that satisfy the minimum EPA requirements to be contributed to the EDG even though they contain FGDC errors. EDG contributors should review the metadata record to ensure that it contains high-quality information (e.g. contains key information about the resource it references including data source, currency of the data, a description of the data contents, etc.) and should use the EME to complete missing or incomplete metadata fields to bring the record to minimum EPA compliancy.

### **6.2.2 Determine Resources Warranting EDG Metadata Contribution**

EPA requires that metadata for geospatial and non-geospatial resources created or “improved” by EPA personnel and contractor staff be contributed to the EDG. The term “improved” refers to any type of value-add process that alters the resource from its original form.

Frequently, resources are acquired from sources outside of the Agency that do not need to be “improved” but may be of interest to other EPA organizations. In these cases, contribution of metadata records is encouraged but not required. Contributing the metadata record to the EDG exposes the resource in EDG searches and is an excellent way to share geospatial and non-geospatial resources throughout the Agency; particularly resources that are difficult to obtain.

Based on these criteria, the organization will create an EDG Inventory of geospatial and non-geospatial resources requiring metadata contributions to the EDG.

### **6.2.3 Determine Access and Use Constraints for a Resource or Metadata Record**

It is important to determine whether any access or use constraints exist for geospatial and non-geospatial resources prior to contributing the resources and their metadata records to the EDG. EPA organizations have the ability to share geospatial and non-geospatial resources with the general public via the EDG Data Download Locations. Additionally, geospatial resources may be shared with the public via the EDG Clip N Ship application. Metadata records in the EDG may be made shareable with the public by flagging them as “Unrestricted.” These metadata records are then accessible to the public at the EDG, Geo.Data.gov, and Data.gov.

EDG contributors may be uncertain which resources and metadata records are shareable with the public through the EDG. In general, EPA should make EDG content available to the public that [“reflects EPA’s mission areas and business”](#). This includes metadata records that reference geospatial or non-geospatial resources that have been developed by or have had substantial value added by EPA personnel or contractor staff and are not considered “restricted.”

Restricted resources are those that are considered sensitive, have not been published, are restricted by license agreement, have “draft” status, or are otherwise determined by the Agency to be withheld from

public access. Metadata records referencing restricted resources may still be contributed to the EDG but flagged as “Restricted” so that they do not appear in public searches.

The EDG Data Download Locations is provided to EPA Regions, Program Offices, and Research Laboratories as a resource to post downloadable data. There is no requirement to share resources on the EDG Data Download Locations; it is provided as a convenient resource for EPA organizations that want to make specific resources available via download. Restricted geospatial and non-geospatial resources may be posted to the EDG Data Download Locations by flagging the resources as “Restricted” so that they only appear to EDG users with login privileges and are not accessible by the public. Restricted geospatial resources must not be shared on the EDG Clip N Ship application.

#### **6.2.4 Determine Method of Access of a Resource**

Geospatial and non-geospatial resources may be made available to EDG users via the Internet or EPA Intranet, or may be kept offline based on the preference of the organization. The specific method of accessing the resource is recorded in the Data Resource Type metadata element in the resource’s metadata record. It is helpful to determine the desired method by which users will access a resource before creating or editing its metadata record so that the specific data resource type can be included.

The data resource types most commonly used for geospatial resources are “Live Map Services” and “Downloadable Data.” Non-geospatial resources may also be accessed as “Downloadable Data.” Both of these types indicate that the resource is readily accessed via the Internet or Intranet, and a URL to access the resource is provided in the Online Linkage metadata element.

Note: The “Live Map Services” data resource type describes dynamic geospatial data hosted as a service via Internet/Intranet that may be used as cartographic data in mapping applications. The “Geographic Services” data resource type describes geoprocessing tasks, such as address matching or buffering, that are provided as services via the Internet/Intranet.

In cases where data owners do not wish to provide direct online access to their resource, or when the resource is a geospatial dataset contained in a Spatial Database Engine (ArcSDE) database, the owner will record the data resource type as “Offline Data” in the metadata record. In this case, the data owner provides contact information in the metadata record for users to contact them directly to inquire about obtaining the resource.

EDG contributors are highly encouraged to make unrestricted geospatial and non-geospatial resources accessible via the Internet and restricted resources that are shareable within the Agency available via the EPA Intranet. This ensures that users discovering a specific metadata record can directly access the resource that the metadata record describes. Providing online access to geospatial and non-geospatial resources is preferred by users because the resource is immediately available and relieves data owners from handling individual requests from EDG users for access to the dataset.

The Online Linkage metadata element is used to provide a URL for external parties to access the resource. For geospatial resources, information on how to document the Online Linkage within the

metadata record can be found within the [EME](#). For non-geospatial records, the fields at the EDG's [dedicated metadata editor](#) guide the user's input.

It is recommended that EDG contributors provide more than one method of access to geospatial and non-geospatial resources. Providing access through multiple venues increases the options for users to reliably access the resource described in the metadata they have discovered. There are benefits and drawbacks to each method of access, as described below.

#### ***6.2.4.1 Live Map Services***

Benefits (from the EDG perspective):

- Data provided in a web service may directly access a production database, generally providing the user with current data each time the service is accessed.
- Web services are easily accessed using a variety of web-based mapping or client tools, which allows a number of different types of users to access the information (e.g., users may not need to install special software).
- Web services support EPA GeoPlatform Online and may be incorporated directly into web maps.

Drawbacks

- Web services may be more difficult to administer than other simpler data access formats.
- Web services may be more unreliable than other data formats during periods of high network traffic.
- Making web services available typically requires more expensive hardware, software, and system administration than other formats.

#### ***6.2.4.2 Downloadable Data***

Benefits (from the EDG perspective):

- Downloadable data is typically the simplest and most reliable format for different users trying to access data.
- Making downloadable data available does not require as much software, hardware, or other infrastructure robustness as other formats.
- Downloadable data is relatively easy to administer.

Drawbacks

- Data made available for download may require additional maintenance to maintain currentness because it is usually separated from the production database and must be replicated to a Web-accessible folder.
- Downloadable data may be limiting for some users who do not have access to the software required to load or view it.

- Once users have downloaded data, the data owner cannot be certain that users will return to obtain newer, more current versions. If the source data changes frequently, repeating the download process very often may be onerous for a user.

#### **6.2.4.3 Offline Data (Data Stored in ArcSDE)**

Benefits (from the EDG perspective):

- Data provided in ArcSDE may directly access a production database, generally providing the user with current data each time it is accessed.
- Direct connection to a database may require less maintenance as it does not require that owners maintain live web services or generate multiple database copies.

Drawbacks

- Read-only passwords cannot be provided through metadata (as per EPA security restrictions). Therefore, an ArcSDE connection typically requires a user to go through extra steps to access the information.
- Some production databases may not be accessible through Anytime Anyplace Access (AAA). Consequently, connection to an ArcSDE database may be limiting for some EPA contractors or offsite EPA employees.
- Providing a number of different users with direct access to a production database may not be desirable for data owners if usage is extremely high or if they are not willing to share the entire database.
- Opening direct database connections for ArcSDE databases that are on public servers may be problematic in terms of high-volume traffic and usage.

#### **6.2.4.4 Offline Data (Data Not Stored in ArcSDE)**

Benefits (from the EDG perspective):

- Protects data from use by unknown customers.
- Ensures that draft, sensitive, or restricted data are discoverable through the EDG but are not directly accessed by individuals without communication with data owners.

Drawbacks

- May be onerous for data providers in responding to multiple inquiries for data.
- Least useful access method for EDG users and most likely to be discarded when returned from a search for information.

### **6.3 CREATE CONTENT AND CONTRIBUTE TO THE EDG**

After evaluating geospatial and non-geospatial resources, an inventory of content to be contributed to the EDG may be created. The content may be metadata records that describe geospatial and non-geospatial resources, or the content may be the resources themselves (GIS databases, reports,

spreadsheets, or other electronic files) chosen to be posted to the EDG Data Download Locations or Clip N Ship application to make available for user download.

Not unlike a library card catalog containing records of books on a shelf, the EDG contains metadata records for resources stored within EPA computer systems. The metadata record, and the resource itself, become discoverable by EDG users performing a search. The metadata record can include information telling users how to locate and obtain the resource.

During the inventory process, for metadata records, it is helpful to make note of which resources currently have metadata (potentially needing updating) and which resources are completely absent metadata. In both cases the process for metadata creation is similar (as described in Sections 6.3.1 and 6.3.2), but extra care should be taken with pre-existing metadata to ensure that important information such as value-add processes and proper contact information are recorded.

### **6.3.1 Prepare and Contribute Metadata for a Non-Geospatial Resource**

On May 9<sup>th</sup>, 2013, the White House issued Memo M-13-13, entitled “Open Data Policy—Managing Information as an Asset” which outlined a set of policy requirements designed to make federal data more open and accessible. One component of this policy was the adoption of a new metadata format based on international standards. Relative to the old data.gov DMS standard for non-spatial metadata, this new “Common Core” standard is both simpler and more reflective of the way data is published and consumed in a modern, service oriented enterprise. Full details on the memo, the other components of the open data policy, and common core metadata are available on the Project Open Data website (<http://project-open-data.github.io/>).

The implementation timeline for this new standard is very aggressive, even as the standard itself is being refined and improved. After first putting into place the infrastructure for storing this new metadata format and publishing it to Data.gov, OEI is now working hard to develop the tools and associated guidance to help metadata stewards produce and maintain metadata. The first of what we anticipate will be several tools for managing common core metadata is a Microsoft Excel template spreadsheet that contains the official list of metadata fields, usage notes for each field, an indication of whether the field is required or optional, and for certain fields the ability to select from a dropdown box of valid values. Please pay careful attention to this guidance – the validation requirements are very strict and failure to comply will result in your metadata being delisted from Data.gov and excluded from the EPA’s enterprise data inventory. Detailed procedures can be found in Appendix D, Section D.6.3.1.

### **6.3.2 Prepare and Contribute Metadata for a Geospatial Resource**

Contributing metadata for geospatial resources is a two-step process.

1. Prepare and validate the metadata record using the EPA Metadata Editor (EME).
2. Contribute the metadata record to the EDG via upload or synchronization.

#### ***6.3.2.1 Prepare and Validate the Metadata Record Using the EPA Metadata Editor (EME)***

The first step is to prepare the metadata. The [EPA Metadata Editor \(EME\)](#) is a simple geospatial metadata editor that allows users to create and edit records that meet EPA and Federal Geographic

Data Committee (FGDC) metadata requirements. Procedures for downloading the EME can be found in Appendix D, Section D.6.3.2.1a. Detailed instructions for installing the EME application and using the EME to prepare a metadata record are included with the download.

Once editing of a geospatial metadata record is complete it must pass validation before being contributed to the EDG. Validation is a QA process that checks the metadata record fields to ensure that the record meets the EPA and FGDC requirements. A validator is included in the EME for this purpose. Instructions for utilizing the EME for preparing and validating metadata records can be found in Appendix D, Section D.6.3.2.1b.

### ***6.3.2.2 Contribute the Metadata Record to the EDG***

Geospatial metadata records may be contributed in the EDG in three ways: upload, synchronization, or ArcSDE synchronization. Before contributing metadata records to the EDG, an organization must determine the method or methods to be used. There are benefits and drawbacks to each contribution method, as described below.

**Upload.** This method involves manually uploading metadata XML files from a local or network drive to the EDG.

#### Benefits

- Ease of use for small number of metadata records.
- Stewards will notice metadata errors immediately.

#### Drawbacks

- Tedious method when dealing with large number of metadata records.
- If the metadata record is moved to a different location on the user's local or network drive without embedding the universally unique identifier (UUID), the metadata record will be duplicated rather than being overwritten when an update is finished and the metadata record is re-uploaded to the EDG.

Detailed procedures for uploading individual metadata XML records to the EDG can be found in Appendix D, Section D.6.3.2.2a.

**Synchronization.** This method requires that an organization set up a local network web-accessible folder (WAF) in which metadata XML files are maintained. The WAF is registered with the EDG and metadata is synchronized automatically on a schedule.

#### Benefits

- Easy to manage a large number of metadata records.
- Resources can be registered at the EDG to automatically synchronize on a regular schedule.
- Updates made to metadata records will automatically be updated in the EDG when the scheduled synchronization occurs.

- Gives the option to receive synchronization results through email.

#### Drawbacks

- Requires somewhat significant local infrastructure configuration.
- EDG Stewards must monitor the synchronization results for their group. If the results are not monitored, a Steward may not immediately notice a metadata record error.

Detailed procedures for creating a WAF and synchronizing associated metadata records with the EDG can be found in Appendix D, Section D.6.3.2.2b.

**Spatial Database Engine (ArcSDE) Synchronization.** This method synchronizes metadata records that are stored in geospatial databases maintained within ArcSDE.

#### Benefits

- Metadata records are not physically separated from the resources they describe.
- Stewards only need to manage one location for all their metadata records.

#### Drawbacks

- Requires that the organization have ArcSDE.
- Requires the EDG Steward send detailed database information to the EDG Administrative Group.

Detailed procedures for synchronizing ArcSDE metadata records with the EDG can be found in Appendix D, Section D.6.3.2.2c.

### 6.3.3 Post Data to the EDG Clip N Ship Application

[EPA's Clip N Ship](#) application allows users to easily view and download EPA's publically-available data, clipped to their area of interest. Clip N Ship users first select a geographic area (either interactively using the Select by Area tool, or by choosing a state) then select data layers from the table of contents and the format in which they would like the data delivered. The layers are then clipped to the geographic area and made available to the user as a compressed file (ZIP format) download.

EDG Stewards may arrange to contribute geospatial data to the EPA Clip N Ship application for their Region, Program Office, or Research Laboratory. The Clip N Ship application is updated quarterly, allowing EDG Stewards an opportunity to prepare and submit their datasets to the EDG Administrative Group for inclusion in the application update. Detailed procedures for posting data to Clip N Ship can be found in Appendix D, Section D.6.3.3.

### 6.3.4 Post Data to the EDG Data Download Locations

The [EDG Data Download Locations](#) gives EPA Regions, Program Offices, and Research Laboratories space to store geospatial and non-geospatial resources and make them available for download. The EDG Data Download Locations provides both "Public" and "Restricted" locations. External public users without EDG login may access data in the "Public" folders while only registered EDG users may access data in the "Restricted" folders after logging in. No login is necessary for users accessing the EDG from within the

EPA Intranet. Detailed procedures for posting data to the EDG Data Download Locations can be found in Appendix D, Section D.6.3.4.

### **6.3.5 Utilize EDG Reuse Components**

[EDG Reuse Components](#) includes three tools that give application developers the ability to reuse EDG content in other applications: the EDG Rest Interface, the EDG Search Widget, and the EDG CS-W Interface. The tools permit direct access to EPA's metadata outside of the EDG interface.

The Rest Interface is used to output catalog content into a new web page. This is accomplished by embedding an EDG REST URL in the web page. EDG users can also access the EDG Search Widget that allows them to embed a simple search tool into a web page. The EDG provides a CS-W interface that allows users to query the EDG catalog using the CS-W Specification. Detailed procedures for utilizing EDG Reuse Components can be found in Appendix D, Section D.6.3.5.

## **6.4 UPDATE AND MAINTAIN CONTENT IN THE EDG**

Content contributed to the EDG, including geospatial and non-geospatial resources and their associated metadata records, must be regularly reviewed, updated, and maintained by the EDG Steward for their Region, Program Office, or Research Laboratory.

Updating of geospatial and non-geospatial metadata records must occur as needed based on a routine review that should be done by the EDG Steward on a regular basis. Procedures for updating metadata records vary based on the method in which the metadata was contributed: EDG Dedicated Editor, manual upload, or synchronization. Updating of geospatial and non-geospatial resources posted to the EDG Data Download Locations and Clip N Ship application must also occur as needed.

Maintenance tasks that the EDG Steward may need to perform include setting the access level of a metadata record, transferring ownership of a metadata record, and maintaining updated registration information for registered resources on a network.

Maintaining EDG content is also a responsibility of the EDG Administrative Group. The Administrative Group is responsible for ensuring that EDG content successfully synchronizes with Data.gov.

### **6.4.1 Update Existing EDG Content with New Content**

To continue to be a relevant and reliable source of data, the EDG must deliver the most current and accurate content available to users. This requires that content posted to the EDG (geospatial and non-geospatial resources and their associated metadata records) be reviewed regularly and updated when appropriate.

#### **6.4.1.1 Update Manually-Uploaded Metadata Records**

Metadata records that have been manually uploaded to the EDG may require updating. In order to prevent record duplication, the new record must have the same UUID tags and file name as the record being replaced. If duplication occurs, the record may be manually deleted from the EDG. Procedures for manually uploading a metadata record can be found in Appendix D, Section D.6.3.2.2a.

#### ***6.4.1.2 Manage Metadata XML Files in a Web Accessible Folder***

The purpose of a web accessible folder (WAF) is to provide a local repository for metadata records scheduled to be synchronized to the EDG. The WAF is hosted at a Region, Program Office, or Research Laboratory location, typically on a web server, and is managed by the organization. The organization's EDG Steward places metadata XML files in the WAF and then registers the WAF at the EDG for metadata synchronization.

Metadata XML files stored in the WAF should be examined regularly for accuracy and currentness. Old metadata records should be removed or updated with current information and new metadata documents should be added after validating they meet EPA requirements. Detailed procedures for managing a WAF can be found in Appendix D, Section D.6.4.1.2.

#### ***6.4.1.3 Update Data Posted to the Clip N Ship Application***

The Clip N Ship application is updated quarterly to maintain current information. EDG Stewards that choose to contribute geospatial resources to the Clip N Ship application should examine their hosted data on the Clip N Ship web page when the quarterly update notice is emailed by the EDG Administrative Group. Additionally, Clip N Ship metadata records should be examined at the same time for accuracy and currentness. Old geospatial resources or metadata records should be removed or updated. Detailed procedures for coordinating Clip N Ship data hosting can be found in Appendix D, Section D.6.3.3.

#### ***6.4.1.4 Update Data Posted to the EDG Data Download Locations***

EDG Stewards should examine the geospatial and non-geospatial resources stored at the EDG Download Locations regularly for accuracy and currentness. If a resource has become out of date, the EDG Steward should remove the resource from the EDG Download Locations and post an updated version, if available. Detailed procedures for posting data to the EDG Data Download Locations can be found in Appendix D, Section D.6.3.4.

### **6.4.2 Set Access Level for a Metadata Record**

EDG Stewards can designate which users see their organization's metadata records. Once logged in, the access controls for metadata records can be found in the EDG Administration tab. There are two types of security configurations in the EDG:

- Unrestricted: All users can search for approved or reviewed EDG records.
- Restricted: Only registered EPA users can access the metadata record.

Note: Restricting user access to a metadata record only restricts which users can see the record itself. It does not determine user access to the actual geospatial or non-geospatial resource that the metadata record references. Detailed procedures for setting access levels for a metadata record can be found in Appendix D, Section D.6.4.2.

### **6.4.3 Transfer Ownership of a Metadata Record**

It may be necessary at times for an EDG Steward to transfer the ownership of a particular metadata record to another EPA Region, Program Office, or Research Laboratory. EDG Stewards can log in to the

EDG and access the controls for transferring ownership under the Administration tab. The current metadata record owner simply selects the metadata record, selects the correct owner, and executes the transfer. Detailed procedures for transferring ownership of metadata records can be found in Appendix D, Section D.6.4.3.

#### **6.4.4 Maintain Registration Information for Registered Resources on the Network**

Each Region, Program Office, or Research Laboratory's EDG Steward must regularly examine registration information to determine the accuracy of their WAF location. If the organization's metadata repository moves to a new location, or the connection information to their network resource changes, the synchronization with the EDG will fail if not updated. It is important for the EDG Steward to maintain correct registration information for their metadata repository at the EDG to ensure consistent synchronization. Detailed procedures for maintaining accurate resource registration can be found in Appendix D, Sections D.6.3.2.2b and D.6.3.2.2c depending on the synchronization method.

#### **6.4.5 Verify EDG content syncs with Data.gov**

Data.gov provides descriptions of Federal datasets (metadata), information about how to access the datasets, and tools that leverage government datasets. EPA's responsibility for adding publically-available resources to Data.gov is satisfied through an automatic synchronization with the EDG. For resources that should be shared with all users (including the public), EDG Stewards must set the access level of those contributions to "Unrestricted" in the EDG. This facilitates synchronization of those records with Data.gov on a regular basis. The EDG Administration Group and EDG Stewards should ensure that appropriate EPA resources are being synchronized and made discoverable through Data.gov. Detailed procedures to verify EDG content syncs with Data.gov can be found in Appendix D, Section D.6.4.5.

#### **6.4.6 Create and Manage Collections**

EDG contributions, including geospatial and non-geospatial metadata records, can be organized into a Collection by EDG users. An EDG Collection is a group of metadata records that relate to each other in one or more ways. Examples of how metadata records may be related include:

- subject matter
- time frame
- location

Figure 4 illustrates how a Collection groups metadata records by subject matter.

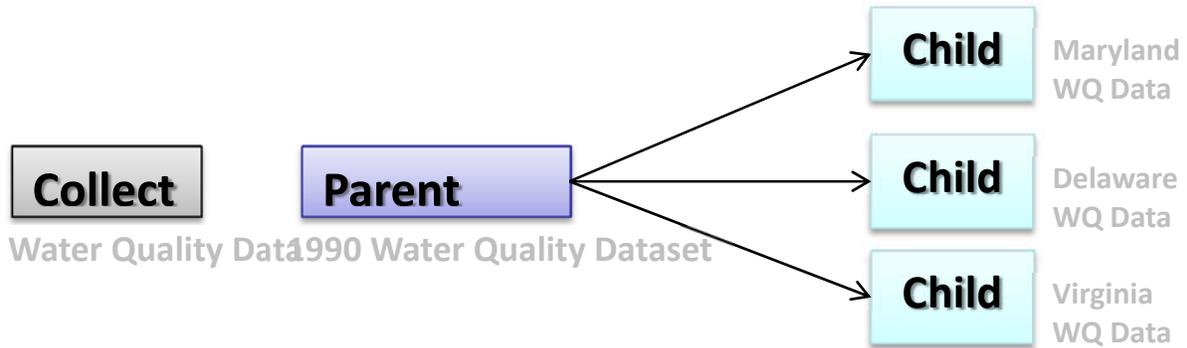


Figure 4: EDG Collections

The creation of Collections within the EDG allows EDG users to organize metadata records of interest into meaningful groupings of related data resources. EDG users have the ability to create Collections, manage Collections metadata membership, and visualize the Collections as a diagram or in the EDG Browse Tree (Figures 5 and 6). The objective of Collections is to provide users with a simple data hierarchy in which a high-level data resource, referred to as a parent record, is comprised of a group of closely-related low-level datasets, or child records. The resulting metadata organization provides a comprehensive, management-level view of an EPA program's data holdings.

[Collections](#) [Members](#) [Placeholder Records](#)

## Create/Manage Collections

Name :  Owner : Any

Name	Owner	Placeholder Record	Approved	Manage	Edit	Delete
Emissions Monitoring Knowledge Base	OAR-OAQPS	<a href="#">Add</a>	Y			
Air Pollutant Emissions Factors	OAR-OAQPS	<a href="#">Add</a>	Y			
RACT/BACT/LAER Clearinghouse	OAR-OAQPS	<a href="#">Add</a>	Y			
Community Multi-Scale Air Quality Model	OAR-OAQPS	<a href="#">Add</a>	Y			
RadNet Fixed Lab Special Analyses	OAR-ORIA	<a href="#">Add</a>	Y			
Radiological Emergency Response	OAR-ORIA	<a href="#">Add</a>	Y			
Radiological Emergency Response for NIMS	OAR-ORIA	<a href="#">Add</a>	Y			
Air Quality (Current)	OAR-OAQPS	<a href="#">Add</a>	Y			
Energy Star	OAR-OAP	<a href="#">Add</a>	Y			
Engine Emissions Compliance	OAR-OTAQ	<a href="#">Add</a>	Y			
Fuel Economy Label and CAFE Data	OAR-OTAQ	<a href="#">Add</a>	Y			
Vehicular Greenhouse Gas Emissions	OAR-OTAQ	<a href="#">Add</a>	Y			
Greenhouse Gas Reporting Program	OAR-OAP	<a href="#">Add</a>	Y			

Figure 5: User Interface for creating or managing Collections

## Browse

Browse metadata records by topic area

- [-] Catalog
  - [-] Content Type
  - [-] Data Categories
  - [-] Owner
    - [-] OAR
      - [-] OAR-OAP
        - Energy Star
        - Greenhouse Gas Reporting Progr
        - Emissions from Power Plants & In
        - Acid Rain Allowance Holdings and
        - Clean Air Status & Trends Networ
        - Atmospheric Deposition Modeling
        - Modeling Results for Policy Analys
        - Control Measures for Electric Gen
        - Monitoring of Acid Lakes & Stream
      - [-] OAR-OAQPS
        - Emissions Monitoring Knowledge f

53 results

Showing 1-10 [1](#) [2](#) [3](#) [4](#) [5](#) [>](#) [Last](#)

**ENERGY STAR Qualified Ventilating Fans**

Ventilating Fans that have earned the ENERGY STAR are more efficient than standard models. ENERGY STAR is the trusted symbol for energy efficiency helping consumers save money and protect the environment through energy-efficient products and practices. M...

[Details](#) [Metadata](#) [Visualize collection](#) [Parent](#)

**ENERGY STAR Qualified New Vending Machines**

New Vending Machines that have earned the ENERGY STAR are more efficient than standard models. ENERGY STAR is the trusted symbol for energy efficiency helping consumers save money and protect the environment through energy-efficient products and practices...

[Details](#) [Metadata](#) [Visualize collection](#) [Parent](#)

**ENERGY STAR Qualified Commercial Clothes Washers**

Commercial Clothes Washers that have earned the ENERGY STAR are more

Figure 6: Collection in Browse Tree Interface. Topic area is selected on left and associated child metadata records are displayed on right

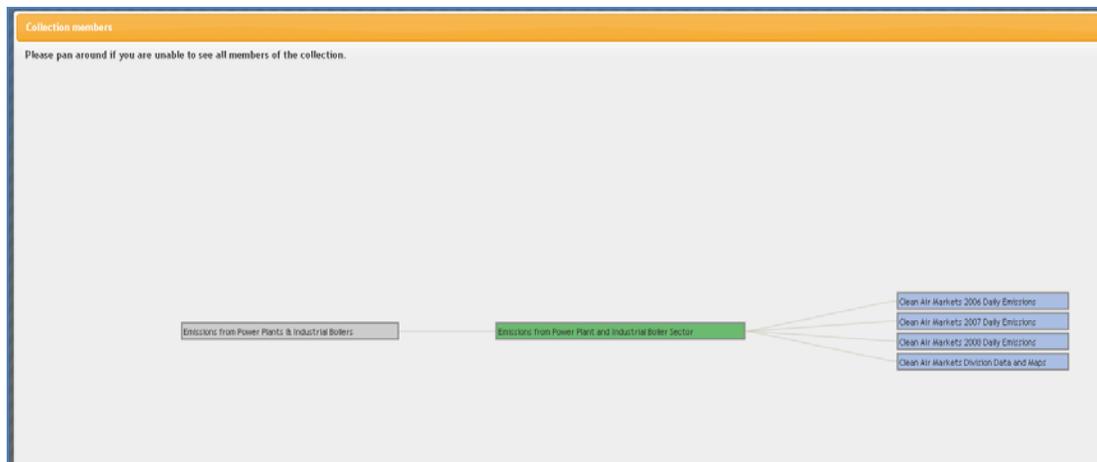


Figure 7: Collection Visualization. This view shows the relationship between the Collection, Parent Records, and Child Records.

#### 6.4.6.1 Add and Search Placeholder Records

EDG Collections allow EDG users to organize their metadata records into high-level parent records that represent a large common theme, subject area, or programmatic interest, and low-level child records that correspond to a specific document, file, or model run pertaining to the larger theme. Once an EDG user creates a Collection, they have the ability to add and search records (referred to as “members” in EDG Collections) to that specific Collection. Detailed procedures for adding and searching Collection members can be found in Appendix D, Section D.6.4.6.

#### 6.4.6.2 Add and Search Members

EDG users have the ability to add and search placeholder records within their Collections. A placeholder record simply contains a title and can be added to an EDG Collection to temporarily retain the site where a complete metadata record will be added once it has been successfully documented, validated, and published to the EDG. A placeholder record simply contains a title. Detailed procedures for creating placeholder records can be found in Appendix D, Section D.6.4.6.

## 7.0 GLOSSARY

**access and use constraints** – Restrictions and legal prerequisites for accessing a geospatial or non-geospatial resource or metadata record. These include any access constraints applied to ensure the protection of privacy or intellectual property, and any special restrictions or limitations on obtaining the resource or metadata record.

**applications** – (EDG Data Resource Type) Custom-built GIS applications that usually have a complete user interface.

**clearinghouses** – (EDG Data Resource Type) Websites that provide links to geographic data.

**Clip N Ship Application** – Interactive website that presents EPA geospatial resources in a map viewer and allows users to view the resources, select geospatial data within an area of interest, and download clipped data and save locally.

**Data.gov** – Website created to *increase public access to high value, machine readable datasets generated by the Executive Branch of the Federal Government*. Data.gov increases the ability of the public to easily discover, download, and use datasets that are generated and held by the Federal Government. Data.gov provides descriptions (metadata) of Federal datasets, information about how to access datasets, and tools that leverage government datasets.

**Data.gov metadata standard** – Metadata standard prescribed for submission of datasets to Data.gov.

**documents** – (EDG Data Resource Type) See *non-geospatial resources*.

**downloadable data** – (EDG Data Resource Type) Data that may be accessed via the Internet and downloaded to a local storage space.

**Environmental Dataset Gateway (EDG)** – EPA's gateway to web-based geospatial information and information services. The EDG enables data consumers to discover, view, and access geospatial resources (e.g., data, web services, or applications) made available by EPA's Program Offices, Regions, and Research Laboratories. Likewise, the EDG enables EPA data contributors to make their geospatial information and web services discoverable, viewable, and accessible by others. Unrestricted information contributed to the EDG is also shared with inter-Agency data sharing portals, including Data.gov and Geo.Data.gov.

**EDG content** – Geospatial and non-geospatial resources and metadata records contributed to the EDG.

**EDG central catalog** – Repository of metadata - data about the geospatial data or services made available to EDG users.

**EDG collections** – An EDG Collection is a group of metadata records that relate to each other in one or more ways.

**EDG Data Download Locations** – Web page in which EPA Regions, Program Offices, and Research Laboratories may post data to make available for download. Data may be made available to the general public or restricted to users with EDG logins.

**EDG Inventory** – Master list of geospatial and non-geospatial resources designated to have metadata in the EDG as identified by the EDG Steward and Data Stewards.

**EDG Reuse Components** – Tools that allow individuals to provide direct access to EPA's metadata outside the EDG interface.

**EPA Metadata Editor (EME)** – Software application that facilitates the documentation of geospatial datasets. The EME allows for creating, editing, and validating metadata records.

**EPA Office of Environmental Information (OEI)** – EPA office that manages the life cycle of information to support EPA's mission of protecting human health and the environment. OEI identifies and

implements innovative information technology and information management solutions that strengthen EPA's ability to achieve its goals. OEI ensures the quality of EPA's information, and the efficiency and reliability of EPA's technology, data collection and exchange efforts, and access services. OEI provides technology services and manages EPA's IT investments.

**Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata (FGDC CSDGM)**

– Metadata content standard established by Office of Management and Budget that provides a common set of terminology and definitions for the documentation of digital geospatial data. The standard establishes the names of data elements and compound elements (groups of data elements) to be used for these purposes, the definitions of these compound elements and data elements, and information about the values that are to be provided for the data elements.

**Geo.Data.gov** – Geospatial-centric version of Data.gov.

**geographic activities** – (EDG Data Resource Type) Events that are geographic in nature, including activities such as land surveys, geographic explorations, and GIS conferences. Information about these events can be presented in many forms, such as a website.

**geographic services** – (EDG Data Resource Type) Internet services that perform geoprocessing tasks such as address geocoding and buffering.

**Geospatial Information Officer (GIO)** – EPA official responsible for management of the Agency's geospatial resources. The GIO provides guidance and coordination in maintaining the [National Geospatial Data Policy](#), setting implementation priorities, and pursuing investments regarding all geospatial activities.

**geospatial resources** – Information that identifies the geographic location and characteristics of natural or constructed features and boundaries on the earth, typically represented by points, lines, polygons, and/or complex geographic features. This includes original and interpreted geospatial data, such as those derived through remote sensing including, but not limited to, images and raster datasets, aerial photographs, and other forms of geospatial data or datasets in both digitized and non-digitized forms.

**harvesting** – See *synchronization*.

**ISO metadata standard** – Metadata content standard developed by the International Organization for Standardization (ISO). Not currently an Agency standard, it is slated to become the approved standard for EPA datasets in the future.

**live map services** – (EDG Data Resource Type) Dynamic data services, accessible via an Internet/Intranet link (URL), that allow direct interaction with map content in a variety of web-based or desktop mapping applications.

**map files** – (EDG Data Resource Type) Digital maps, such as those created for ArcReader and ArcMap, that can be viewed in a mapping application.

**metadata (FGDC definition)** – Electronic file, usually formatted as a XML document, which captures the basic characteristics of a geospatial or non-geospatial resource. It represents the who, what, when, where, why, and how of the resource. Geospatial metadata are used to document geographic digital

resources such as Geographic Information System (GIS) files, geospatial databases, web services, and earth imagery. A geospatial metadata record includes core library catalog elements such as Title, Abstract, and Publication Data; geographic elements such as Geographic Extent and Projection Information; and database elements such as Attribute Label Definitions and Attribute Domain Values.

**metadata compliancy** – Condition that metadata contains minimum required information to facilitate discovery and reuse of the data and to maintain consistency between datasets. Compliancy is verified through the metadata validation process.

**metadata restrictions** – Availability of a metadata record to particular audiences based on the sensitivity of the data. Unrestricted data may be provided to the Data.gov catalog while restricted data may only be visible to EPA personnel within the EDG. See also *access and use constraints*.

**metadata standard** – Common set of terminology and definitions for the documentation of digital data. The standard establishes the names of data elements and information about the values that are to be provided for the data elements. (<http://www.fgdc.gov/metadata/csdgm/introduction.html>). See also *Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata (FGDC CSDGM)*.

**metadata template** – Metadata model containing fields to be populated with specific information about the resource which it describes to meet the reporting requirements for a particular metadata standard.

**metadata validation** – Process by which a metadata record is evaluated for compliancy with EPA metadata standards.

**non-geospatial resources** – Information including documents, reports, or other electronic files in which the primary use is for something other than spatial analysis or visualization in a map

**offline data** – (EDG Data Resource Type) Data stored on EPA computers and is not accessible via the Internet.

**online linkage** – Internet location at which a geospatial or non-geospatial resource, application, or other asset is stored and accessed.

**other documents** – (EDG Data Resource Type) Geographic information stored in text files, spreadsheets, or other file formats that are not specifically geospatial file formats.

**Quality and Information Council (QIC)** – Body of EPA officials whose primary focus is on addressing and resolving intra-Agency cross-media, cross-program, and interdisciplinary information technology/information management and related policy issues.

**resource undetermined** – (EDG Data Resource Type) Classification of EDG content that does not fit the other content type definitions.

**Senior Information Official (SIO)** – Primary point of accountability for the effective oversight, coordination, and management of information and information technology (IT) within Agency organizations.

**static map images** - (EDG Data Resource Type) Map image files that may be viewed, but are not able to be incorporated into mapping applications (e.g. Adobe PDF maps).

**synchronization** – Process by which EPA offices may contribute metadata to the EDG central catalog in an automated, scheduled manner. Synchronization removes the burden of metadata maintenance from Data Stewards by ensuring that EDG records are automatically refreshed along with local records. EDG stewards who choose to contribute to the EDG using synchronization only need to maintain records locally, and information stored in the EDG catalog is synchronized with the local files on a scheduled basis. Formerly referred to as *harvesting*.

**universally unique identifier (UUID)** – Character string identifier used to uniquely identify an object or entity on the Internet.

**web accessible folder (WAF)** – Local computer network folder maintained at Regions, Program Offices, and Research Laboratories in which metadata files may be placed for synchronization of the metadata records with the EDG.

## 8.0 APPENDICES

### APPENDIX A. KEY EDG CONTACTS

Contacts listed below identify key EDG operational personnel as of April 1, 2013. A list of EDG Stewards for the Agency's Regions, Program Offices, and Research Laboratories may be found at the [EDG Management & Stewardship](#) website.

- EDG Project Manager: Ana Greene  
[green.ana@epa.gov](mailto:green.ana@epa.gov)  
202-566-2132
- EDG Administrative Group  
edg@epa.gov

## APPENDIX B. AUTHORITY FOR EDG IMPLEMENTATION AND RELEVANT DOCUMENTS, RELATED PROCEDURES, AND GUIDELINES

**Authority.** General governance and oversight for the EDG is provided by EPA's Quality and Information Council (QIC) in coordination with the Geospatial Information Officer (GIO). The QIC is responsible for overseeing the implementation of the NGDP, the NGDP Procedure, and related documents among EPA programs. It also acts as the governing authority for questions related to EDG costs, implementation, or structure across programs.

Senior Information Officials (SIOs) are responsible for decisions regarding costs and managing metadata for their individual programs. The SIO has responsibility for ensuring that the roles listed in this document are implemented appropriately within their program. The implementation of roles is expected to vary by organization; procedures for appointing individuals to fulfill roles (excluding the EDG Steward) are listed in the NGDP Procedure. Information regarding the EDG Steward role (including appointing a Steward) is described in Section 7.5.1 of this document. Additional information regarding the roles and responsibilities of the QIC and SIO are available at:

- <http://intranet.epa.gov/oei/imitpolicy/qic/index.htm>
- <http://intranet.epa.gov/oei/imitpolicy/qic/ciopolicy/2101.pdf>

If costs or questions arise that extend beyond the scope of an individual organization, the organization's SIO should consult the QIC. Costs for EDG enterprise resources, EDG training, and application oversight are managed by EPA's Office of Environmental Information (OEI). Program offices may decide the degree to which they would like to be engaged in more specific oversight of the EDG.

### Related Documents.

- Federal Initiatives
  - *E-Government Act of 2002 (H.R. 2458/S. 803)*, December 17, 2002.
  - *Data.gov Metadata Template, version 2.0*.
  - *Executive Order 12906, National Spatial Data Infrastructure*, April 13, 1994, Federal Register, Volume 59, Number 71, pp. 17671 – 17674.
  - *Federal Geographic Data Committee Standard (FGDC-STD) 001-1998, Content Standard for Digital Geospatial Metadata, Federal Geographic Data Committee*, June 1998.
  - *Office of Management and Budget (OMB) Circular A-16, Coordination of Geographic Information and Related Spatial Data Activities*, August 19, 2002.
- EPA Policies, Procedures, Standards, and Guidance
  - *EPA Geospatial Metadata Technical Specification, Version 1.0*, January, 2, 2007.
  - *EPA Guidance, Enterprise Data Policy Implementation Guidance, Office of Environmental Information (In Progress)*.
  - *National Geospatial Data Policy, EPA Order 2121*, August 24, 2005.
  - *National Geospatial Data Policy Procedure for Geospatial Metadata Management*, October 25, 2007.

## APPENDIX C. WAIVERS

**Waiver Process.** The Agency's CIO may grant waivers to selected provisions of this Procedure for sufficient cause. The CIO may re-delegate this authority.

**Applications.** Applications for waivers to specific provisions shall contain the following:

1. Identification of the Procedure provision;
2. A listing of reasons why the Procedure cannot be applied or maintained;
3. An assessment of the impacts resulting from non-compliance; and
4. The signature of the AA, RA, or Laboratory Director responsible for the spatial data collection in question.

**Notification.** The CIO will notify the requesting office in writing of the disposition of the waiver within 60 days of receipt.

## APPENDIX D. DETAILED STANDARD OPERATING PROCEDURES

### D.6.1.1 Designate an EDG Steward

*The EDG Steward works closely with the EPA's EDG Administrative Group to plan the organization's EDG contributions, keep current with metadata standards, and take full advantage of the tools available from the EDG. The EDG Steward will be responsible for contributing to the EDG via upload and/or synchronization. The EDG Steward will maintain geospatial and non-geospatial resources and metadata records in the EDG by regularly reviewing content and may also work within their group to help others understand how EDG resources can be reused.*

#### **Steps:**

1. Understand the role and responsibilities of the EDG Steward as detailed in Section 5.5 of the EPA Environmental Dataset Gateway (EDG) Governance Structure and Standard Operating Procedures Version 3.0.
2. Determine an appropriate person within your Program Office, Region, or Research Laboratory to fulfill the role of the EDG Steward.
3. Notify the EPA Metadata Coordinator and/or the EDG Administrative Group of the person designated as your region or program office EDG Steward. [edg@epa.gov]
4. EDG Administrative Group will add the designated EDG Steward to the EDG Stewards list maintained within the EDG website.
5. EDG Administrative Group will provide an orientation to the new EDG Steward detailing EDG contribution expectations and requirements.

### D.6.1.2 Create User Logins and Assign User to Appropriate EDG Groups

*To access restricted information at the EDG, or to manage content for your EPA organization, you must have an EDG account. You will also need to have your EDG account assigned to the appropriate EDG groups before you can publish, access, and manage your organization's EDG metadata repository.*

#### **FOR EPA EMPLOYEE OR CONTRACTOR WITH A LAN ID AND PASSWORD**

##### **Steps:**

1. If you have a LAN ID and password, you can log into the [EDG](#). Your log in credentials are your LAN ID and password.
2. Send a request to the EDG Administrative Group to have your LAN ID added to the appropriate EDG group. This is accomplished by using the [Feedback Link](#) at the top of the EDG home page.
3. In your request, simply identify the Program Office, Region or Research Laboratory you will be supporting.
4. The EDG Administrative Group will grant you access to the Program Office, Region, or Research Laboratory you request and notify you by email once your request has been completed.
5. Once added to the appropriate Program Office, Region, or Research Laboratory, you can now publish, access, and manage your organization's EDG metadata repository.

#### **FOR NON-EPA EMPLOYEE OR OTHER INDIVIDUALS WITHOUT A LAN ID AND PASSWORD**

##### **Steps:**

1. Send a request to the EDG Administrative Group using the [Feedback Link](#) at the top of the EDG home page.
2. In your message, inform the EDG Administrative Group that you do not have a LAN ID.
3. Go to <https://waa.epa.gov/> and self-register for an account. When registering at the EPA Web Application Access, the user needs to request David Parrish as their EPA Sponsor and request access to the System of Registries community.
4. Once you receive your account information, notify the EDG Administrative Group and also indicate which Program Office, Region, or Research Laboratory you will be supporting.
5. The EDG Administrative Group will grant you access to the Program Office, Region, or Research Laboratory you request and notify you once has been completed.
6. Once added to the appropriate Program Office, Region, or Research Laboratory you can now publish, access, and manage your organization's EDG metadata repository.

#### **EDG ADMINISTRATIVE GROUP STEPS FOR ESTABLISHING NEW USERS:**

1. Go to:  
<https://lumberjack.rtpnc.epa.gov/identity/oblix/apps/groupservcenter/bin/groupservcenter.cgi>.
2. Log in using your LAN username and password.

3. Select the Group Manager tab.
4. Search for an EDG group.
5. Select the appropriate group.
6. Select Manage Group Members.
7. In the search window select Full Name or Login and then type the user's name or LAN ID.
8. Click the Add button next to the correct user after the search.
9. Select Done.
10. Once back on the Group Management page, click the Save button.
  - a. The new user will then be added to the group. A user must be added to the publisher group in addition to their correct Region, Program Office, or Research Laboratory group before they can publish metadata to the EDG.
11. Email the user and have them test logging into the EDG using their LAN ID or username and password.

### **D.6.3.1 Prepare, Validate, and Publish Non-Geospatial Metadata to the EDG Using the EDG Dedicated Editor**

*The purpose of the Environmental Dataset Gateway (EDG) is to provide EPA staff with a single access point to EPA's geospatial and non-geospatial resources from various EPA Program and Regional Offices and Research Laboratories. This is accomplished through the EDG by publishing metadata that provides details about the nature of the resource, its origins and quality, and methods for EPA staff to gain access to the resource. This section addresses publication of non-geospatial metadata to the EDG.*

#### **EDG STEWARD STEPS FOR AUTHORIZING A METADATA RECORD:**

1. Contact the EDG Administrative team and request the Microsoft Excel template spreadsheet. This spreadsheet contains the official list of metadata fields, usage notes for each field, an indication of whether the field is required or optional, and for certain fields the ability to select from a dropdown box of valid values.
2. Conduct a quality check on your metadata and make sure the content is current.
3. Once you have completed the spreadsheet, please email it to [edg@epa.gov](mailto:edg@epa.gov). We will configure an automated harvest of the metadata into the EDG and let you know about any validation errors. Once complete and validated, a member of the EDG Administrative Group will approve your metadata record.
4. Once your metadata record has been approved, it will be registered and discoverable in the EDG.
5. It may take up to 48 hours for the records or any changes or updates to be visible at Data.gov.

#### **EDG ADMINISTRATIVE GROUP STEPS FOR WORKING WITH NON-GEOSPATIAL METADATA STEWARDS:**

1. If the Non-Geospatial Metadata Steward does not already have a copy of the dcat\_inputtemplate.xlsx, send it to them to use for creating records.
2. When the Non-Geospatial Metadata Steward returns a complete inputtemplate.xlsx file, if they don't already have a folder configured, create a new copy of the OCSPP folder to store their xlsx, csv, bat, and output files.
3. Edit the .bat file to reflect the new foldername and input files - instructions are in the .bat file.
4. Review the content in their xlsx file. Make sure required fields are populated, and other content is logical.
5. Copy just the content rows from the xlsx file and past into the template csv file using the "preserve formatting" option (this is necessary to get the date syntax right). Save.
6. Make sure the output folder is empty.
7. Run the .bat file by double-clicking. ISO xml files should appear in the output folder. Review to make sure they look right.
8. Copy the output xml files to the WAF where they will be harvested to the edg - if they don't already have a WAF, Configure one.

#### **EDG ADMINISTRATIVE GROUP STEPS FOR APPROVING A METADATA RECORD:**

1. Login to the EDG application using your LAN ID and password.
2. Select the Administration tab.
3. Perform a search for records that are marked as “posted.”
4. Select a posted metadata record and click on the “View” option.
5. Examine the metadata record for errors (broken onlinks, incorrect resource descriptions, strange titles, etc.).
6. If the record does not have any errors, select “Set as Approved” and execute the action.
7. If the record has errors, select “Set as Incomplete” and execute the action.
  - a. After setting the metadata record as incomplete, contact the EDG Steward for the document and describe the problems with the record.
  - b. Ask the EDG Steward to correct the errors.
  - c. Answer any questions the EDG Steward may have about the error.
  - d. Once the EDG Steward corrects the errors, perform Steps 5 and 6 above.
8. Notify the EDG Steward once their record has been “Set as Approved.”

### D.6.3.2.1a Download the EPA Metadata Editor (EME)

*The EME is a simple geospatial metadata editor that allows users to create and edit records that meet EPA and Federal Geographic Data Committee (FGDC) metadata requirements. The EME is one of the few FGDC Metadata Editors that runs in ArcGIS Version 10 and higher. It can be used as an ArcGIS extension or as a standalone tool and provides ArcGIS 10 users with FGDC synchronization capabilities.*

#### **Steps:**

1. Go to the EME website, located at <https://edg.epa.gov/EME/download.html>
2. Enter your information (fields with [\*] are required).
3. Choose a product to download.
4. Read and agree to the license agreement.
5. For information on installing the EME, read the installation guide provided in the EME download package.

### D.6.3.2.1b Prepare and Validate the Metadata Record Using the EPA Metadata Editor (EME)

*The purpose of the Environmental Dataset Gateway (EDG) is to provide EPA staff with a single access point to EPA's geospatial and non-geospatial resources from various EPA Program and Regional Offices and Research Laboratories. The EDG facilitates this by publishing metadata that provides details about the nature of the resource, origins and quality of the resource as well as ways for EPA staff to gain access to the resource. This document addresses publication of geospatial metadata that is not stored in the Spatial Database Engine (ArcSDE).*

#### **Steps:**

1. Develop or edit metadata for each geospatial layer using the [EPA Metadata Editor \(EME\)](#).
2. Validate metadata using the EPA Metadata Editor
  - a. If the metadata does not validate, correct errors in the EME and re-validate until the metadata is error-free.
  - b. Warnings are acceptable. A metadata record can be published to the EDG with warnings.
3. Export metadata XML files from each geospatial layer using the EME export function and save to a local or network drive.

### D.6.3.2.2a Contribute Metadata Records to the EDG – Manual Upload

*This method involves manually uploading metadata XML files from a local or network drive to the EDG.*

#### **EDG STEWARD STEPS FOR UPLOADING A METADATA RECORD:**

1. [Login to the EDG.](#)
2. Go to [How to Publish Resources](#) page.
3. Click on **2. Upload or validate metadata file from the local disk**. Since you have already validated the metadata, skip to Step 8 in the online instructions to upload the metadata record.
4. Once the metadata records have been uploaded successfully to the EDG, a member of the EDG Administrative Group will approve your new metadata files or contact you if there are questions. After the metadata records have been approved, they are registered and discoverable in the EDG.

#### **BEST PRACTICE: OPTIONAL STEPS FOR EMBEDDING THE METADATA RECORD’S UNIVERSALLY UNIQUE IDENTIFIER (UUID)**

1. Once you have successfully published the metadata record, you need to add the universally unique identifier (UUID) to the metadata record. To locate the metadata record’s UUID, do the following:
  - a. From the EDG Administration tab, search for your record by title.
  - b. Click the **Download** icon associated with your record. This will download the metadata record you just uploaded.
  - c. Save the downloaded metadata record to a local drive. Note that the downloaded XML file has a long character string name (i.e. **{D9F88959-0498-41FA-A328-1351454C5D41}.xml**). The characters surrounded by the curly brackets make up the metadata record’s UUID.
2. The UUID must be inserted into the metadata record. Open your newly-downloaded metadata record in a text editor.
3. The first line of the XML file will usually have `<metadata xml:lang="en">`. Insert the UUID and tags to the metadata record starting on the second line as follows:

```
<Esri>
    <PublishedDocID>{Your new UUID Goes Here}</PublishedDocID>
</Esri>
```

Using the above UUID as an example, the top four lines of your metadata XML file will appear as follows:

```
<metadata xml:lang="en">
    <Esri>
        <PublishedDocID>{D9F88959-0498-41FA-A328-1351454C5D41}</PublishedDocID>
    </Esri>
```

4. Once the UUID has been inserted into the metadata record, you must save the file, maintaining its long character string (UUID) name.
5. Re-upload the record following the instructions starting with Step 2 above. Maintaining the long character string (UUID) name prevents duplication of records.

6. Once successfully re-uploaded, you will receive a notice saying “Your document was **replaced** successfully.” If you receive a notice saying “Your document was **published** successfully.”, you may have created a duplicate record that needs deletion.
7. If your metadata record originated from a GIS feature class, it is good practice to re-import the updated metadata file with the UUID tags into the feature class from which it was generated. You may use the EME Import Metadata functionality to accomplish this.

#### **EDG ADMINISTRATIVE GROUP STEPS FOR APPROVING A METADATA RECORD:**

1. Login to the EDG application using your LAN ID and password.
2. Select the Administration tab.
3. Perform a search for records that are marked as “posted.”
4. Select a posted metadata record and click on the “View” option.
5. Examine the metadata record for errors (broken onlinks, incorrect resource descriptions, strange titles, etc.).
6. If the record does not have any errors, select “Set as approved” and execute the action.
7. If the record has errors, select “Set as Incomplete” and execute the action.
  - a. After setting the metadata record as incomplete, contact the EDG Steward for the document and describe the problems with the record.
  - b. Ask the EDG Steward to correct the errors.
  - c. Answer any questions the EDG Steward may have about the error.
  - d. Once the EDG Steward corrects the errors, perform Steps 5 and 6 above.
8. Notify the EDG Steward once their record has been “Set as approved.”

### D.6.3.2.2b Contribute Metadata Records to the EDG – Synchronization

*This method requires that an organization set up a local network web-accessible folder (WAF) in which metadata XML files are maintained. The WAF is registered with the EDG and metadata is synchronized automatically on a schedule. For uploading multiple metadata records and/or those that may be revised or updated in the future.*

#### **EDG STEWARD STEPS FOR COORDINATING METADATA RECORD SYNCHRONIZATION:**

1. Determine your office's designated WAF location. If your organization does not have a WAF, please go to [How to Publish Resources](#) page, click on **1. Register a resource on the network**, and follow the directions for registering your office's WAF. If you need further assistance, [contact the EDG Administrative Group](#).
2. Place validated metadata XML files in your organization's WAF.
3. The metadata records that pass validation will be registered and discoverable in the EDG on the next scheduled synchronization.
4. You may also force synchronization. Navigate to the [EDG Administration page](#). Search for your WAF by selecting your organization as the Document owner and WAF as the Protocol Type. Once located, click the Synchronize content icon to force synchronization.
5. Once the metadata records have been uploaded successfully to the EDG, a member of the EDG Administrative Group will approve your new metadata files or contact you if there are questions. After the metadata records have been approved, they are registered and discoverable in the EDG.

#### **BEST PRACTICE: OPTIONAL STEPS FOR EMBEDDING THE METADATA RECORD'S UNIVERSALLY UNIQUE IDENTIFIER (UUID)**

1. Once you have successfully published the metadata records, you need to add the universally unique identifier (UUID) to each metadata record. To locate the metadata record's UUID, do the following:
  - a. From the EDG Administration tab, search for your record by title.
  - b. Click the **View** icon associated with your record. This will open a new window with a text block version of the metadata record.
  - c. Note the text string at the beginning of the text block in curly brackets (i.e. **{D9F88959-0498-41FA-A328-1351454C5D41}**). The characters surrounded by the curly brackets make up the metadata record's UUID.
2. The UUID must be inserted into the metadata record. Open the associated metadata record stored in the WAF in a text editor.
3. The first line of the XML file will usually have `<metadata xml:lang="en">`. Insert the UUID and tags to the metadata record starting on the second line as follows:

```
<Esri>  
  <PublishedDocID>{Your new UUID Goes Here}</PublishedDocID>  
</Esri>
```

Using the above UUID as an example, the top four lines of your metadata XML file will appear as follows:

```
<metadata xml:lang="en">
  <Esri>
    <PublishedDocID>{D9F88959-0498-41FA-A328-1351454C5D41}</PublishedDocID>
  </Esri>
```

4. Once the UUID has been inserted into the metadata record save the file, maintaining its original name. This must be performed for each metadata record.
5. The updated metadata record will be synchronized in the EDG on the next scheduled synchronization. Adding the UUID prevents duplication of records in the EDG.
6. If your metadata record originated from a GIS feature class, it is good practice to re-import the updated metadata file with the UUID tags into the feature class from which it was generated. You may use the EME Import Metadata functionality to accomplish this.

#### **EDG ADMINISTRATIVE GROUP STEPS FOR APPROVING A METADATA RECORD:**

1. Login to the EDG application using your LAN ID and password.
2. Select the Administration tab.
3. Perform a search for records that are marked as "posted."
4. Select a posted metadata record and click on the "View" option.
5. Examine the metadata record for errors (broken onlinks, incorrect resource descriptions, strange titles, etc.).
6. If the record does not have any errors, select "Set as approved" and execute the action.
7. If the record has errors, select "Set as Incomplete" and execute the action.
  - a. After setting the metadata record as incomplete, contact the EDG Steward for the document and describe the problems with the record.
  - b. Ask the EDG Steward to correct the errors.
  - c. Answer any questions the EDG Steward may have about the error.
  - d. Once the EDG Steward corrects the errors, perform Steps 5 and 6 above.
8. Notify the EDG Steward once their record has been "Set as approved."

### D.6.3.2.2c Contribute Metadata Records to the EDG – ArcSDE Synchronization

*This method requires that an organization manages their metadata in ArcSDE. This capability provides EDG Stewards with the ability to connect directly to their ArcSDE database to synchronize their metadata at the EDG. The connection made between the organization's SDE database and the EDG is accomplished through an application called WAFer. This application takes metadata managed within an SDE database and moves them to a WAF located on the Intranet EDG server. The EDG Steward registers the URL to this WAF at the EDG for synchronization.*

#### **Steps:**

1. Send a request to the EDG Administrative Group using the [Feedback Link](#) at the top of the EDG home page.
2. In your request, you will need to include:
  - a. Host IP
  - b. Host port
  - c. User
  - d. Password
  - e. Fully qualified name of the metadata table: Usually "sde.dbo.GDB\_USERMETADATA" for SQLServer, "sde.GDB\_USERMETADATA" for Oracle.
3. The EDG Administration Group will set up the connection to the WAFer application and send the URL to be registered at the EDG to the EDG Steward requesting SDE synchronization.
  - a. The URL will look similar to the following:  
`http://edg-intranet.epa.gov/WAFer/NameofGroup/`
4. Register the URL at the EDG for synchronization. (For directions on registering a resource for synchronization, refer to Appendix D Section D.6.3.2.2b).

#### **Steps for the EDG Administration Group:**

1. Login to the EDG Intranet server.
2. Navigate to D:\Public\Server\apps\Tomcat6029\webapps\WAFer\internal\wafconfig.xml.
3. Open the wafconfig.xml.
4. Input the connection information sent in the request from the EDG Steward into the wafconfig.xml file. Copy the following code into the webconfig file and then input the correct parameters.

```
<SOURCE type="GEODB" shortName="R10" longName="U.S. EPA Region 10" metadataTable=""  
jdbcDriver="com.microsoft.jdbc.sqlserver.SQLServerDriver" jdbcUrl="jdbc:microsoft:sqlserver://host  
IP:Port;databaseName=XX" user="" pwd="" xmlReadMethod="string" />
```

5. Save the webconfig.xml.
6. Restart tomcat.
7. Go to the new URL created and determine if the XML files are listed.

- a. If the XML files are listed, then send the URL to the EDG Steward requesting SDE Synchronization.
  - b. If the XML files are not listed, troubleshoot why the connection did not work.
8. Send the URL to the EDG Steward requesting SDE synchronization.

### D.6.3.3 Post Data to the EDG Clip N Ship Application

*The EPA Clip N Ship application is hosted as a component of the EDG. The Clip N Ship application provides EDG Stewards with the ability to make their datasets easily available to other EPA users for download. Clip N Ship datasets are updated quarterly allowing EDG stewards an opportunity to prepare and submit their datasets to the EDG Administrative Group for inclusion in the application update.*

**Steps:**

Complete steps are located at:

[https://edg.epa.gov/metadata/webhelp/en/gptlv10/inno/EDG\\_ClipAndShip\\_procedures.pdf](https://edg.epa.gov/metadata/webhelp/en/gptlv10/inno/EDG_ClipAndShip_procedures.pdf)

#### D.6.3.4 Post Data to the EDG Data Download Locations

*EPA organizations have the option to store geospatial and non-geospatial resources and make them accessible to EDG users through the EDG Data Download Locations. There are two locations that are available: a Public location accessible by any user, and a Restricted location that is only accessible to users who have an EDG login.*

*EDG Stewards may post resources to the EDG Data Download Locations and reference the download location as the Online Linkage in their resources' metadata records. This facilitates data sharing with the public and with internal EPA users.*

**Steps:**

Complete steps are located at:

[https://edg.epa.gov/metadata/webhelp/en/gptlv10/inno/EDG\\_Download\\_Locations.pdf](https://edg.epa.gov/metadata/webhelp/en/gptlv10/inno/EDG_Download_Locations.pdf)

### D.6.3.5 Utilize EDG Reuse Components

*The EDG hosts several components that allow application developers access to reuse EDG content in other applications. This provides direct access to EPA's metadata outside the EDG interface. The EDG Rest Interface is one method used to output catalog content into a new web page. This is accomplished by embedding an EDG REST URL in the web page. Developers can also access the EDG Search Widget which allows a simple search tool to be embedded in a web page. Finally, the EDG provides a [CS-W interface](#) that allows developers to query the EDG catalog using the CS-W Specification.*

**Steps:**

Complete steps are located at:

[https://edg.epa.gov/metadata/webhelp/en/gptlv10/inno/EDG\\_Reuse.pdf](https://edg.epa.gov/metadata/webhelp/en/gptlv10/inno/EDG_Reuse.pdf)

#### D.6.4.1 Set Access Level for a Metadata Record

*EDG Stewards can designate which types of users(internal to EPA or all users) can view their organization's metadata records. There are two types of access controls available to EDG Stewards: unrestricted and restricted. All approved metadata records marked as unrestricted are searchable for all users. Restricted metadata records are only accessible by registered EDG users. EDG Stewards set the appropriate access level for their metadata records once they have been published.*

**Steps:**

1. Login to the EDG using your LAN ID and password.
2. Click on the Administration tab.
3. Search and then select the metadata record needing to have its access level changed.
4. In the dropdown box "For selected records" select, "Set Access Level."
5. Click the radial button next to either "Unrestricted" or "Restricted" depending on the access required for the metadata record.
6. Click the "Execute Action" button.
7. Check the "Access" column for the metadata record to make sure the correct access level is listed.

#### D.6.4.2 Transfer Ownership of a Metadata Record

*EDG Stewards can transfer the ownership of a metadata record published by their organization. This may be necessary if the resource documented in the record changes ownership. It may also be necessary if an EDG Steward publishes the metadata record under their name instead of the appropriate Region, Program Office, or Research Laboratory.*

**Steps:**

1. Login to the EDG using your LAN ID and password.
2. Click on the Administration tab.
3. Search for and then select the metadata record needing to have its ownership transferred.
4. In the dropdown box "For selected records" select, "Transfer Ownership."
5. In the dropdown box "New Owner" select the correct metadata record owner.
6. Click the "Execute Action" button.
7. Check the "Owner" column for the metadata record to ensure the correct owner is listed.

#### D.6.4.5 Verify EDG Content Synchronizes with Data.gov

*To maintain accurate EPA metadata records at Data.gov, it is important for EDG Stewards to verify that their content synchronizes with Data.gov.*

##### **Steps:**

1. Start at <http://catalog.data.gov>. (Do not try to log in to <http://www.data.gov>, this will take you to a dead-end.)
2. Scroll to the very bottom of the page and click on the word “Login”.
3. Log in.
4. Once successfully logged in, you should see an icon like this  at the top of the page. Click it.
5. On the left side of the page you should see the words “U.S. Environmental Protection Agency”. Click it. (note: throughout the site it can be very tough to tell which words are hyperlinks and which are not. Often the best way to tell is to hover over everything with your mouse and see if the cursor changes.)
6. Next, click the blue “Admin” button on the right. This takes you to where you can start configuring the EPA’s presence on data.gov.
7. Click on the “Harvest Sources” tab.
8. Click on the “Environmental Dataset Gateway” link to see the EDG harvest dashboard. Harvests should occur daily. The timestamps are in Greenwich time, so subtract 5 hours for EST or 4 hours for EDT.
9. The “Full Job Report” gives meaningful errors, though it can take some careful parsing to understand exactly what’s tripping it up. Click on the “Jobs tab” to get the list of all past harvest attempts and reports.
10. The “Edit” tab is where harvesting options are set. Currently the EDG harvests are set to run from the daily cache of the EDG dcat.json file
11. A manual harvest can be run any time by clicking the “Reharvest” button.
12. The “Clear” button purges all existing records from data.gov.

#### D.6.4.6 Create and Manage Collections

*EDG Collections are metadata records grouped together by a common theme, location, or EPA-specific subject matter. EDG users have the ability to manage their own Collections. The EDG's Collections interface allows users to create Collections, manage metadata membership, and visualize the Collections as a diagram or in the EDG Browse Tree. The information below contains detailed steps for creating and managing an EDG Collection.*

##### **Steps to Create a Collection:**

1. Login to the EDG using your LAN account.
2. Click on the Collections tab in the menu bar at the top of the EDG homepage.
3. Type a name for your Collection in the "Name:" text box
4. Select an owner for your Collection from the "Owner:" drop down list.
5. Click the Add button under the "Name:" text box.
6. The newly created Collection will now appear in the main collections list. You can now search and edit your Collection.

##### **Steps to Manage Collection Members:**

1. Locate a Collection to add members (metadata records).
2. Click the "Manage Collection Members" option in the menu for your Collection. Hint: 
3. Input either a Document UUID or Document title along with the Document owner.
4. Uncheck the "Only display current members:" check box
5. Click the Search button.
6. The record you searched for will be returned and listed below the search options on the web page.
7. Click the check box next to the record for either "Parent" or "Child" depending on whether you would like this to be the parent record (a higher level record that other records will be categorized under) or child (a more detailed record that is associated with a higher level record).
8. Click the Execute button.
9. Click the "Collection Visualization" button next to the name of your Collection. The new Placeholder record now shows up as a member of the Collection. Hint: 

##### **Steps to Add a Placeholder Record (First Option):**

1. From the Create/Manage Collections page, select the owner for the Placeholder record from the "Owner:" drop down menu at the top of the page.
2. Next to your Collection, click the "Add" button under the "Placeholder record" column.
3. Type the name to be given to the new Placeholder record in the "Resource Name:" text box.
4. If the Placeholder record should be a parent record, check the box next to "Parent Resource". If the resource is not meant to be the parent record, uncheck the box next to "Parent Resource".

5. Click the “Create record” button.
6. A message will appear stating, “Placeholder record has been added to the collection.”
7. Click the “Collection Visualization” button next to the name of your Collection. The new Placeholder record now shows up as a member of the Collection. Hint: 🗃️

**Steps to Add a Placeholder Record (Second Option):**

1. Click the Placeholder Records menu at the top of the web page.
2. In the “Name:” textbox, type the name you want to give the Placeholder record.
3. Select the owner of the Placeholder record from the “Owner:” drop down menu.
4. Click the “Add” button.
5. Click the “Collections” tab at the top of the page to return to the collections page.
6. Select the Owner of your collection from the “Owner:” drop down list.
7. Click search.
8. Click the “Manage Collection Members” button for your Collection. Hint: 👥
9. Input the title of the Placeholder record you just created for the “Document tile”.
10. Select the owner of the Placeholder record from the “Document Owner:” drop down menu.
11. Uncheck the “Only display current members:” check box.
12. Click the “Search” button.
13. Choose “Parent” or “Child” for the new Placeholder record.
14. Click Execute.
15. Click the “Collection Visualization” button next to the name of your Collection. The new Placeholder record now displays as a member of the Collection. Hint: 🗃️

## APPENDIX E. EPA National Geospatial Data Policy (NGDP) Stewardship Agreement

This agreement sets forth an accord between EPA's Office of Environmental Information (OEI) and \_\_\_\_\_ (participating organization's name) to fulfill the roles and responsibilities listed in EPA's NGDP Procedure for Geospatial Metadata Management. By signing this document, each party is agreeing to commit personnel and resources sufficient to perform the responsibilities listed for the roles assigned for their organization.

### Goals

The goal of the NGDP Procedure for Geospatial Metadata Management is to establish effective geospatial metadata lifecycle management by outlining common requirements and responsibilities for maintaining geospatial metadata used by EPA programs. This agreement ensures the fulfillment of roles identified in the NGDP Procedure for Geospatial Metadata Management in order to guarantee that associated responsibilities are carried out consistently across multiple EPA organizations.

### Resources

1. EPA's OEI commits to providing resources in the form of geospatial infrastructure, personnel, training and assistance to support organizations in fulfilling the Geospatial Data Owner, Geospatial Data Steward and EDG Steward Roles.
2. EPA's OEI agrees to provide sufficient personnel and resources to fulfill the roles of Geospatial Metadata Coordinator and EDG Administrator. All responsibilities required for those roles will be satisfied by OEI personnel and/or subcontractors.

### Roles and Responsibilities

I, (Name, Title, and EPA Organization) \_\_\_\_\_, agree to commit the following personnel to fulfill the roles listed in the NGDP Procedure for Geospatial Metadata Management:

Geospatial Data Owner(s)

\_\_\_\_\_

Geospatial Data Steward(s)

\_\_\_\_\_

EDG Steward(s)

\_\_\_\_\_

Additional support personnel (optional)

---

**Authorization**

I understand that by assigning personnel to the roles listed above, they will be required to fulfill the responsibilities as listed in the NGDP Procedure for Geospatial Metadata Management. I agree to provide sufficient support for personnel listed to enable them to fulfill those duties.

Signed: \_\_\_\_\_, Date: \_\_\_\_\_

EPA’s OEI agrees to fulfill the requirements of the Geospatial Metadata Coordinator and EDG Administrator roles, as well as providing geospatial infrastructure, personnel, training and assistance to support organizations in fulfilling the Geospatial Data Owner, Geospatial Data Steward and EDG Steward Roles.

**NGDP Procedure for Geospatial Metadata Management:**

<http://intranet.epa.gov/oei/imitpolicy/qic/ciopolicy/2131.pdf>

**EPA Geospatial Metadata Technical Specification:**

[https://www.epa.gov/geospatial/docs/EPA\\_Geospatial\\_metadata\\_Technical\\_Specification\\_v1\\_11\\_2\\_2007.pdf](https://www.epa.gov/geospatial/docs/EPA_Geospatial_metadata_Technical_Specification_v1_11_2_2007.pdf)

**EDG Governance Structure and Standard Operating Procedure:**

[https://edg.epa.gov/metadata/webhelp/en/gptlv10/inno/EDG\\_GovernanceStructureReport.pdf](https://edg.epa.gov/metadata/webhelp/en/gptlv10/inno/EDG_GovernanceStructureReport.pdf)