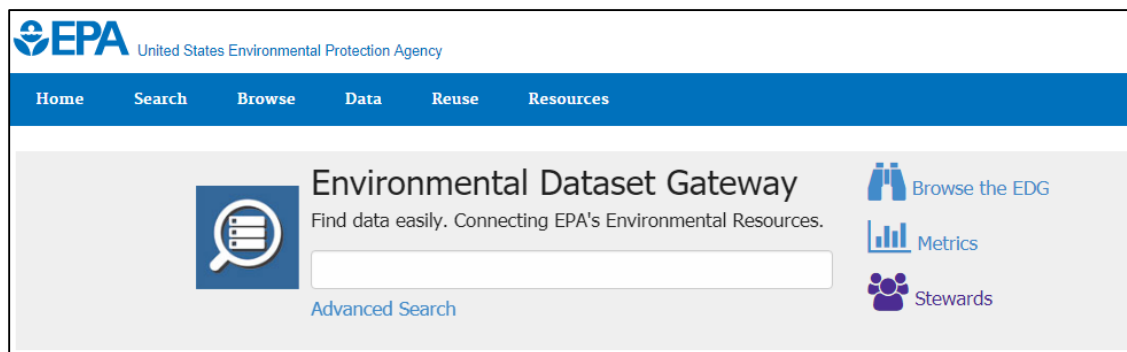


EPA Environmental Dataset Gateway (EDG) Governance Structure and Standard Operating Procedure

Version 4.0

4/25/2018



Contents

1.0 BACKGROUND.....	1
2.0 PURPOSE OF THIS DOCUMENT	2
3.0 INTENDED AUDIENCE.....	3
4.0 ROLES AND RESPONSIBILITIES	3
4.1 EPA Metadata Coordinator	4
4.2 Environmental Dataset Gateway Administrative Group	4
4.3 EPA Data Owner	5
4.4 EPA Data Steward.....	5
4.5 Environmental Dataset Gateway Steward	5
5.0 PROCEDURE FOR CONTRIBUTING CONTENT TO EDG.....	6
5.1 Set up a Program Office, Region, or Research Laboratory for EDG Access.....	7
5.1.1 Designate an EDG Steward	7
5.1.2 Create User Logins and Assign User to Appropriate EDG Groups	8
5.2 Prepare to Contribute to EDG	8
5.2.1 Understand EPA Metadata Compliance Requirements.....	10
5.2.2 Determine Resources Warranting EDG Metadata Contribution	11
5.2.3 Determine Access Level for a Resource.....	11
5.2.4 Determine Method of Access of a Resource	12
5.3 Create Content and Contribute to EDG.....	13
5.3.1 Prepare and Contribute Metadata for a Non-Geospatial Resource	14
5.3.2 Prepare and Contribute Metadata for a Geospatial Resource	14
5.3.3 Post Data to EDG Clip N Ship Application	15
5.3.4 Post Data to EDG Data Download Locations or EPA Data Commons	16
5.3.5 Utilize EDG Reuse Components	17
5.4 Update and Maintain Content in EDG	17
5.4.1 Update Existing EDG Content with New Content	18
5.4.2 Maintain Registration Information for Registered Resources on the Network.....	19
5.4.3 Verify EDG content synchronizes with Data.gov	19
5.4.4 Create and Manage Data.gov Collections.....	19
5.4.5 Create and Manage Compilations.....	20

5.4.5.1 Add and Search Placeholder Records	21
5.4.5.2 Add and Search Members.....	22
6.0 APPENDICES	23
APPENDIX A. KEY EDG CONTACTS.....	23
APPENDIX B. AUTHORITY FOR EDG IMPLEMENTATION AND RELEVANT DOCUMENTS, RELATED PROCEDURES, AND GUIDELINES.....	24
APPENDIX C. WAIVERS.....	25
APPENDIX D. EDG Rules of Behavior.....	26
APPENDIX E. Metadata Steward Contribution Checklist	29

1.0 BACKGROUND

EPA's Environmental Dataset Gateway (EDG) is a group of database, web, and desktop applications used for managing and distributing geospatial and non-geospatial metadata and downloadable data. EDG includes the following components:

- [EDG metadata catalog](#) – the core function of 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 collection of metadata editing applications that allows users to create and edit records to meet EPA metadata requirements.
- [EDG Data Download Locations](#) and [EPA Data Commons FTP Site](#) – locations that allow 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 download geospatial data based on a selected area of interest.
- [EDG Reuse Components](#) – a set of configurable widgets and interfaces that allow application developers to provide access to EDG metadata in their own websites and applications outside the EDG interface.

The purpose of EDG is to: 1) Provide EPA staff with a single point of access to EPA 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, EDG supports external federal initiatives including [Data.gov](#) and Office of Management and Budget's (OMB's) [Project Open Data](#). Figure 1 is a model of EPA's implementation of EDG.

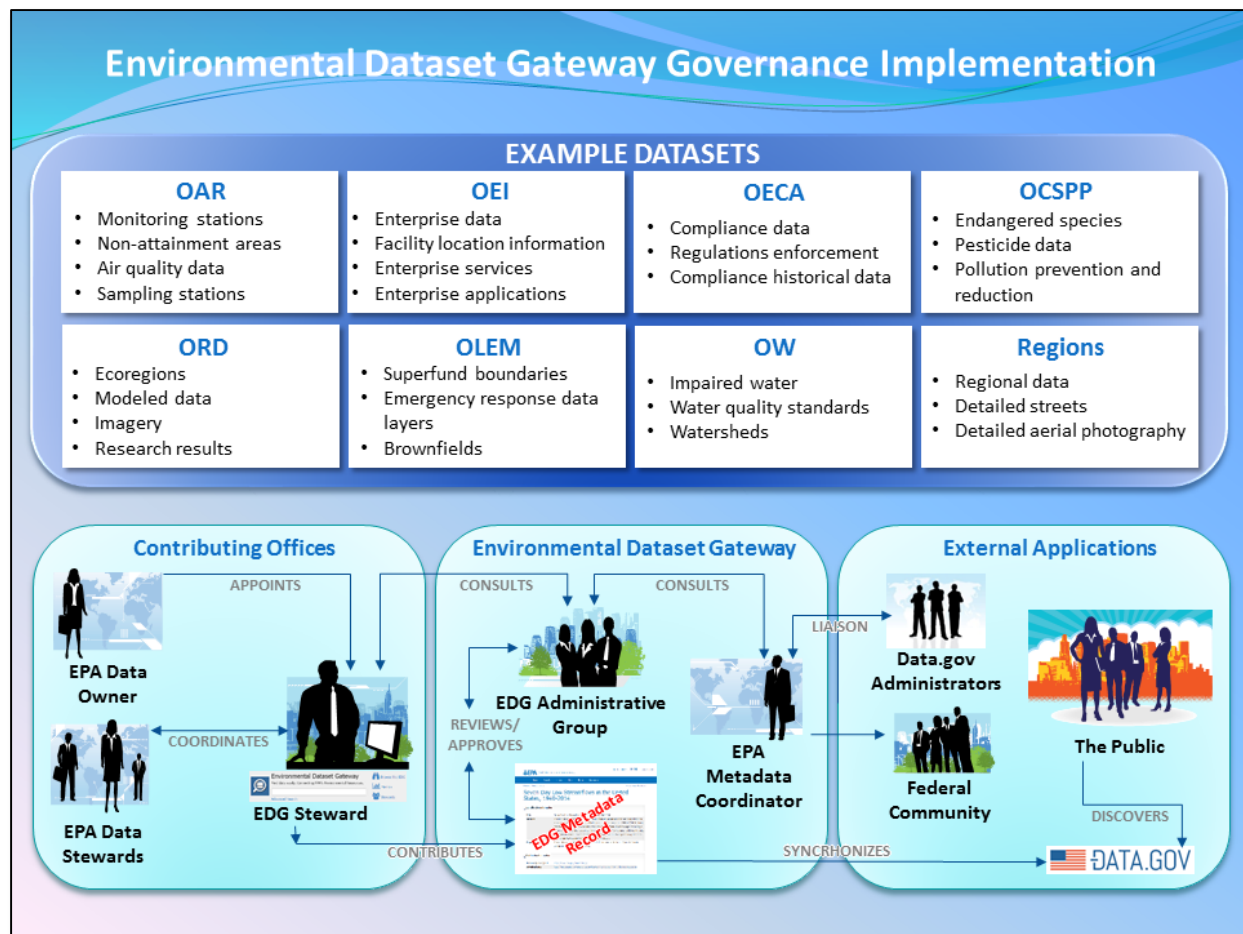


Figure 1. Implementation of EPA Environmental Dataset Gateway.

2.0 PURPOSE OF THIS DOCUMENT

This document has a dual purpose:

1. To identify the governance structure of the EDG, including roles and responsibilities for implementing the EDG ([Section 4.0](#)), and
2. To provide standard operating procedures for functions within EDG ([Section 5.0](#)).

EDG consolidates extensive information holdings, contributed and managed by numerous participants distributed throughout EPA, into a single location. As such, a governance structure that details responsibility and authority for EDG contents is required so that EDG participants clearly understand requirements and expectations for contributing to and managing information in EDG. Roles and responsibilities detailed in this document align with EPA's [National Geospatial Data Policy \(NGDP\)](#) 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.

Clear roles and responsibilities coupled with a set of standard operating procedures that are implemented consistently across EPA programs ensure that information contributed to EDG is effectively managed and can be traced to owners.

Note: Where this document references “metadata” in relation to EDG roles, application functions, and procedures, it applies to both geospatial and non-geospatial metadata.

In several cases this document references policies and initiatives that, in their name and original conception, were written with geospatial metadata in mind. For the purposes of EDG, these references should be construed as applying to all metadata. Any role, function, or procedure that pertains only one type of metadata is explicitly declared as such.

3.0 INTENDED AUDIENCE

The primary audience for this document includes EPA’s Metadata Coordinator, Data Owners, Data Stewards, EDG Administrative Group, and EDG Stewards. This document may also be of interest to additional EPA personnel or EPA contractor staff who may find EDG useful in meeting their geospatial and non-geospatial resource needs. Finally, this document may be considered a resource for individuals with an interest in EPA metadata management.

4.0 ROLES AND RESPONSIBILITIES

Roles and responsibilities outlined in this section reflect duties identified for contributing and managing information posted to EDG (Figure 2). The roles and responsibilities are derived from [NGDP Procedure](#), with the exception of the EDG Steward. The EDG Steward role is unique to EDG and belongs in this document in order to explicitly identify the role within the EDG governance framework.

The specific responsibilities outlined in this section are a subset of the entire suite of metadata management responsibilities identified in the [NGDP Procedure](#). In some cases, EDG role responsibilities overlap with EPA GeoPlatform role responsibilities. This document addresses the nature of the overlap. For details on EPA GeoPlatform data publishing, please review the [EPA GeoPlatform Data Publishing Standard Operating Procedure](#).

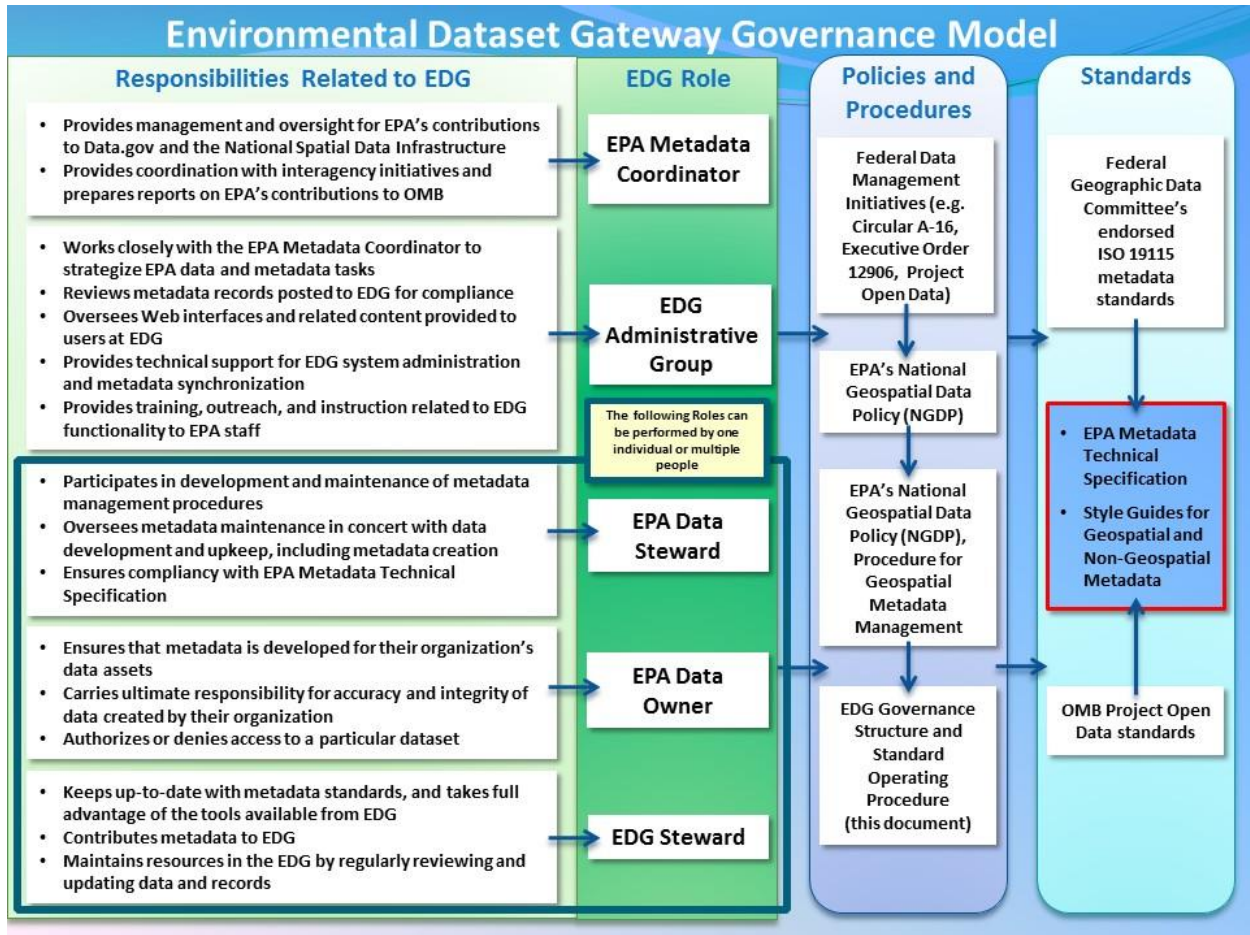


Figure 2. Roles and responsibilities related to EPA Environmental Dataset Gateway.

4.1 EPA METADATA COORDINATOR

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

4.2 ENVIRONMENTAL DATASET GATEWAY ADMINISTRATIVE GROUP

- **Location:** OEI.
- **Primary Role:** Manages EDG applications and metadata catalog.
- **Primary Responsibilities:**
 - Works closely with the EPA Metadata Coordinator to formulate strategies for carrying out Agency-wide initiatives on data and metadata.
 - Monitors resources contributed to the EDG metadata catalog.
 - Reviews metadata records contributed to EDG for EPA compliance.

- Informs EDG Steward of metadata records that do not pass validation.
- Oversees web interfaces and related content provided to users at the EDG website.
- Provides training, outreach, and assistance to EDG Stewards on contributing to EDG.
- 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.
- Provides technical support for implementing and maintaining the EDG connection to Data.gov and National Spatial Data Infrastructure (NSDI).

4.3 EPA DATA OWNER

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

4.4 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.
 - Includes his or her name and contact information in metadata records as the primary metadata contact.
 - Oversees maintenance (review and update) of resources and metadata records in EDG, but does not necessarily interact directly with EDG – may instead work together with an EDG Steward to contribute records to the EDG.
 - Ensures metadata compliance with [EPA's Metadata Technical Specification](#).

4.5 ENVIRONMENTAL DATASET GATEWAY STEWARD

- **Location:** EPA Regions, Program Offices, and Research Laboratories.
- **Primary Role:** Manages and coordinates an organization's contributions to EDG.
- **Primary Responsibilities:**
 - Works closely with the EDG Administrative Group to plan and implement EDG contributions for their organization.
 - Stays current with metadata standards and takes advantage of available EDG tools.
 - Contributes metadata records to EDG via synchronization, reviews harvest reports and works with EDG Administrative Group to resolve any issues.
 - Maintains geospatial and non-geospatial resources and metadata records in EDG by regularly reviewing and updating content.
 - Works within their organization to help others understand how they can reuse EDG resources.

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 contributions to EDG. The EDG Steward also serves as the point-of-contact for other EDG contributors within his or her organization needing assistance with concepts and processes involved in contributing to EDG.

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 ends, the EDG Steward will be required to complete an exit interview with his/her supervisor in accordance with standard EPA termination procedures.

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 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 or restricted 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 EDG.
- Understand EPA geospatial metadata requirements in the [EPA Metadata Technical Specification](#).
- Contribute metadata to their organization's EDG metadata holdings.
- Review geospatial and non-geospatial resources and metadata records on a regular basis and perform maintenance when needed.

5.0 PROCEDURE FOR CONTRIBUTING CONTENT TO EDG

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

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

The tasks related to each step in the procedure are provided in this section along with some background information. Tasks are illustrated in Figure 3.

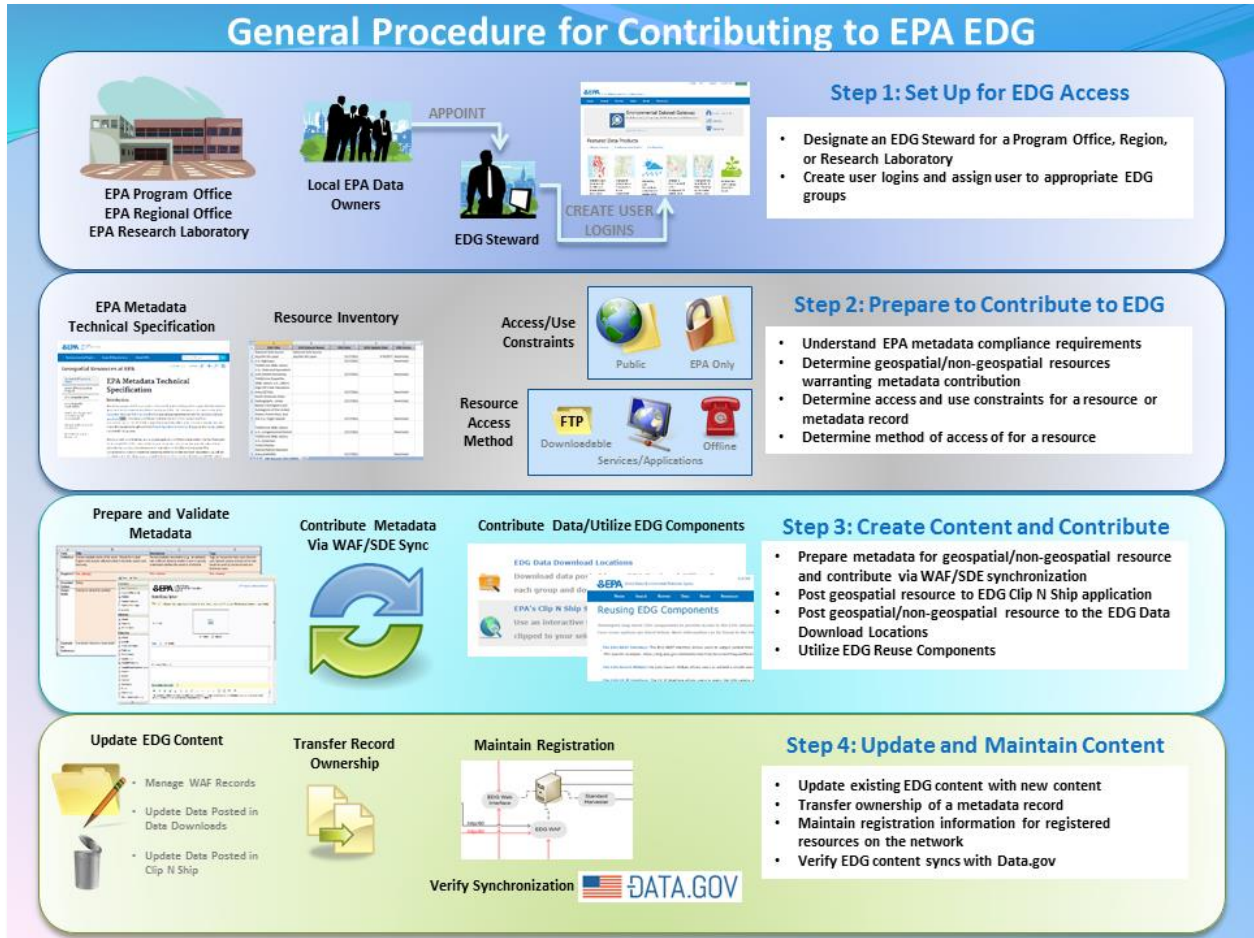


Figure 3. Overall procedure for contributing to EPA Environmental Dataset Gateway.

5.1 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 4 of this document and providing each individual with access to the necessary tools to perform the functions of his or her role.

5.1.1 Designate an EDG Steward

The EDG Steward works closely with the EDG Administrative Group to plan their organization’s EDG contributions, keep current with metadata standards, and take advantage of tools available in EDG. The EDG Steward is responsible for contributing geospatial and non-geospatial resources and metadata records to EDG, reviewing them regularly, and performing maintenance on those resources when needed. The EDG Steward acts as a local point-of-contact to assist others in their organization with using EDG resources.

ACTIONS:

EDG Steward: Contact the EDG Administrative Group (edg@epa.gov) to inform them that you are the EDG Steward.

EDG Administrative Group: Assign the new EDG Steward to the appropriate groups, update their user capabilities, and edit the EDG Stewards webpage (<https://edg.epa.gov/metadata/webhelp/en/qptlv10/inno/Stewards/Stewards.html>) to add the new EDG Steward's name.

5.1.2 Create User Logins and Assign User to Appropriate EDG Groups

All users requiring access to restricted information on EDG must have an EPA LAN account. 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.

ACTIONS:

EDG Steward: Inform the EDG Administrative Group (edg@epa.gov) of individuals in your organization that are responsible for managing EDG content.

EDG Administrative Group: Update the user capabilities of the users and assign them to EDG groups.

5.2 PREPARE TO CONTRIBUTE TO EDG

Once user roles are filled for an EPA organization, the organization may begin preparations to contribute to 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 the Agency and are used for performing EPA business operations. Geospatial and non-geospatial resources obtained from outside the Agency may be accompanied by metadata of varying quality. Often metadata is absent.

CONTENT THAT MAY BE CONTRIBUTED TO EDG INCLUDES:

Data resources for users to download as a complete dataset (EDG Data Download Locations and EPA Data Commons FTP site).

Data resources for users to clip to a geographic extent, then download (Clip N Ship).

Metadata for these data resources and for other data holdings in the organization.

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

1. Understand EPA metadata compliance requirements.
2. Determine geospatial and non-geospatial resources and metadata warranting EDG contribution.
3. Determine access level for a resource.
4. Determine method of access for a resource.

ACTIONS:

EDG Steward:

- *Understand EPA metadata compliance requirements.*
 - *Determine geospatial and non-geospatial resources and metadata warranting EDG contribution (create an inventory).*
 - *Determine access level for each resource.*
 - *Determine method of access for each resource.*
-

Requirements for contributing geospatial and non-geospatial resources and metadata records to EDG vary based on the resource type. Criteria for EDG contribution are summarized in Table 1 by resource type. These criteria should be used by EPA organizations to develop metadata, update EDG content, and share EDG content with internal and external parties. More information about the types of resources is available in [section 5.2.2](#) of this document.

EDG Contribution, Prioritization, and Distribution Criteria			
Resource Type (source)	EPA-Compliant Metadata	Metadata Contribution to EDG	Access Constraints
Resources (data and services) that EPA creates.	Required.	Required. Metadata also published to Data.gov.	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 in non-required fields are allowable.	Required. Metadata also published to Data.gov.	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 in non-required fields are allowable.	Not required, but highly encouraged to contribute to EDG.	Available to the public unless data sensitivity, license agreement, or other EPA restriction on public access exists.

Table 1: EDG contribution and distribution criteria.

5.2.1 Understand EPA Metadata Compliance Requirements

Non-geospatial metadata. Non-geospatial data resources that originate at EPA must have a metadata record published in EDG. Metadata for non-geospatial data should comply with Project Open Data metadata standards as reflected in EPA’s Metadata Technical Specification. EDG provides users with a Microsoft Excel template to assist with creation of non-geospatial metadata.

ACTIONS:

EDG Steward: Review the [EPA Metadata Technical Specification](#) for metadata compliance requirements.

Geospatial metadata. Geospatial resources that originate at EPA must have a metadata record published in EDG. Metadata must meet the [EPA Metadata Technical Specification](#). EPA compliance means that the metadata record includes valid information for all EPA-required fields so that the record meets Project Open Data and either ISO 19115 or FGDC CSDGM standards. EPA-required metadata fields

are clearly identified in the [EPA Metadata Technical Specification](#) as well as all versions of the EPA Metadata Editor (EME).

5.2.2 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 EDG. The term “improved” refers to any type of value-add process that alters the resource from its original form.

ACTIONS:

EDG Steward:

- Review current organization holdings in EDG.
- Inventory current organization data resources, noting those that are created by EPA vs. those created by others and “improved” by EPA.
- Match EDG holdings to current data resources to determine need for new or updated metadata or download availability.

EDG Administrative Group: Assist EDG Steward with list of their organization’s current holdings in EDG.

Important or difficult to obtain external resources. Resources acquired from sources outside the Agency may be important for performing EPA business operations but do not need to be “improved.” Such resources may be of interest to other EPA organizations. In these cases, creation and contribution of a metadata record is encouraged but not required. Contributing the metadata record to EDG exposes the resource in EDG searches and is an excellent way to share geospatial and non-geospatial resources throughout the Agency; particularly for resources that are difficult to obtain. See [Table 1](#) for additional criteria.

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

5.2.3 Determine Access Level for a Resource

EPA organizations have the ability to share geospatial and non-geospatial resources internally or with the public as web services and/or data downloads on EDG servers or on the organization’s own servers. It is important to determine whether any access constraints exist for geospatial and non-geospatial resources prior to contributing the resources and/or their metadata records to EDG. The access level is recorded in the metadata to provide users with information about the dataset’s availability.

ACTIONS:

EDG Steward: Review data resources (web services and downloadable data) to be contributed to EDG to determine suitability for sharing with the public.

Access level for web services. Web services for geospatial or non-geospatial data on Agency servers may be published to allow internal-only access or public access to the resource. The access level is determined at the time of web service publication. Web service publication is beyond the scope of this document. Consult the [EPA GeoPlatform Data Publishing Workflow Standard Operating Procedure](#) for comprehensive instructions on geospatial web service publication.

Metadata for web services that contain EPA data or external data that have been “improved” as outlined in [Table 1](#) must be published to EDG. Evaluate all web services to determine how to record the access level for each web service’s metadata record. Consult the [EPA Metadata Technical Specification](#) for additional information. Web services that do not provide access to data (such as analytical services, validation services, or submission services) do not require EDG metadata, but should be appropriately accounted for in EPA’s Reusable Component Services registry.

Access level for data downloads. Data downloads of geospatial or non-geospatial data may be made available internally or to the public on servers owned by the contributing EPA organization or via one of the two EDG download sites (EDG Data Download Locations or the EPA Data Commons FTP site). The decision of where to host the downloads is at the discretion of the contributor and his or her EDG Steward. Additionally, geospatial resources may be made available to the public as downloads via the EDG Clip N Ship application. Downloads deemed “restricted” must be placed in a location that is not publicly accessible; on EDG, this may be accomplished by uploading the files to a non-public folder in either EDG download site. Restricted geospatial resources must not be shared on the EDG Clip N Ship application.

Metadata for data downloads must be published in EDG, and the metadata must include the access level of the dataset. Consult the [EPA Metadata Technical Specification](#) for additional information.

5.2.4 Determine Method of Access of a Resource

Once the access level of a data resource has been established, the method of access may be determined. “Method of access” refers to ways that users may obtain the data resource. Geospatial and non-geospatial resources may be made available to users via the Internet or EPA Intranet (online data), or may be kept offline based on the preference of the organization. Online data is most often accessed by users as web services and downloadable data files. Non-geospatial resources may also be accessed as downloadable data. Both types indicate that the resource is readily accessed via an Internet or Intranet URL which is provided in the resource’s metadata. In some cases, multiple methods to access the data resource are available (e.g. as both a web service and a data download). In this case, each method is

ACTIONS:

EDG Steward:

- *Determine the ways in which internal customers and/or the public will access the data resource.*
 - *Prepare to record the Distribution URL(s) in the resource’s metadata record.*
-

listed along with a Distribution URL for each. Refer to the [EPA Metadata Technical Specification](#) for additional guidance on recording the Distribution URL.

For geospatial resources, information on how to document the Distribution URL within the metadata record can be found within the [EME](#). For non-geospatial records, the fields in the EDG non-geospatial metadata Excel template guide the user. Users may also refer to the [EPA Metadata Technical Specification](#), specifically the [Distribution URL section](#).

It is recommended that EDG contributors provide more than one method of access to geospatial and non-geospatial resources. Providing access through multiple access points increases the options for users to reliably access the resource described in the metadata they have discovered.

5.3 CREATE CONTENT AND CONTRIBUTE TO EDG

The organization's EDG Steward and EDG contributors, with help from the EDG Administrative Group should have an inventory of geospatial and non-geospatial resources and their associated metadata to contribute to EDG. That inventory includes the following:

1. The resource to be contributed and made discoverable in EDG.
2. Whether metadata for the resource already exists (and whether it is currently in EDG).
3. The level of access to the resource (public or internal-only).
4. The method of access to the resource (download, Clip N Ship, web service, or offline).

During the inventory process, for metadata records, it is important 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 5.3.1 and 5.3.2), 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 maintained or updated if necessary.

ACTIONS:

EDG Steward: *Create an inventory of resources to contribute to EDG. The inventory should include:*

- *Name of the resource to make discoverable in EDG.*
- *Whether metadata already exist for the resource.*
- *Level of access to the resource.*
- *Method of accessing the resource.*

EDG Administrative Group: *Assist EDG Steward with questions about their organization's inventory.*

Not unlike a library card catalog containing records of books on a shelf, EDG contains metadata records for resources stored within EPA computer systems. The metadata record and the resource it describes

become discoverable by EDG users performing a search. The metadata record should always include information telling users how to locate and obtain the resource.

5.3.1 Prepare and Contribute Metadata for a Non-Geospatial Resource

Metadata for non-geospatial resources (documents, reports, or other electronic files in which the primary use is for something other than spatial analysis or visualization in a map) may be contributed to EDG. The EDG's non-geospatial metadata Excel template may be used to prepare and contribute metadata for non-geospatial resources in a single step. The non-geospatial metadata Excel template can be downloaded, [here](#). Refer to the [EPA Non-Geospatial Metadata Style Guide](#) for information on creating ISO/POD-compliant metadata.

ACTIONS:

EDG Steward or contributor:

- Create metadata for non-geospatial resources using the EDG Excel template and following the [EPA Non-Geospatial Metadata Style Guide](#).
- Send the spreadsheet to the EDG Administrative Group.

EDG Administrative Group:

- Assist EDG contributor with their Excel spreadsheet.
 - Convert the spreadsheet to JSON format.
 - Add JSON file to organizational WAF.
 - Synchronize WAF to add non-geospatial records to EDG metadata catalog.
-

5.3.2 Prepare and Contribute Metadata for a Geospatial Resource

Two-step process for preparing and contributing geospatial metadata:

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

5.3.2.1 Prepare and Validate the Metadata Record Using the EPA Metadata Editor (EME)

Prepare the metadata. The [EPA Metadata Editor \(EME\)](#) is a geospatial metadata editor that allows users to create and edit records that meet EPA, OMB Project Open Data, and either ISO-19115 or FGDC CSDGM requirements. Refer to the [EPA Geospatial Metadata Style Guide using EME 5](#) for information on creating ISO/POD-compliant metadata, or the [EPA Geospatial Metadata Style Guide using EME 3](#) for information on creating FGDC-compliant metadata.

Once editing of a geospatial metadata record is complete it must pass validation before being contributed to EDG. Validation is a QA process that checks the metadata record fields to ensure that the record meets the EPA requirements. All versions of EME provide validation methods; refer to the documentation on the [EPA Metadata Editor \(EME\)](#) webpage for more information.

ACTIONS:**EDG Steward or contributor:**

- Create metadata for geospatial resources using an EPA Metadata Editor and following the appropriate EPA Geospatial Metadata Style Guide for [EME 5](#) or [EME 3](#).
- Ensure metadata record is valid before contributing to EDG.

EDG Administrative Group:

- Assist contributor with questions about EME and validation.
-

5.3.2.2 Contribute the Metadata Record to EDG

Geospatial metadata records may be contributed to EDG through synchronization. This method requires that an organization set up a web-accessible folder (WAF) in which metadata XML files are maintained. The WAF is registered with EDG and metadata is synchronized automatically on a schedule. The WAF may be set up on a local network or on the EDG server if the local network is not available.

ACTIONS:**EDG contributor (other than EDG Steward):**

- Work with your organization's EDG Steward to load valid metadata XML files into your organization's EDG web-accessible folder (WAF).

EDG Steward:

- Determine location of organization's web-accessible folder (WAF) for metadata. Contact EDG Administrative Group if assistance is needed.
- Load valid metadata XML files into the WAF.
- Contact EDG Administrative Group to inform them of new records.

EDG Administrative Group:

- Assist EDG Steward with loading valid metadata XML files to WAF.
 - Perform a forced synchronization on the WAF to upload new records.
-

5.3.3 Post Data to EDG Clip N Ship Application

[EPA's Clip N Ship](#) application allows users to easily view and download EPA's publicly-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 as needed. EDG

Stewards may coordinate with the EDG Administrative Group to prepare and submit their datasets to the EDG Administrative Group for inclusion in the application update.

ACTIONS:

EDG Steward or contributor:

- *Inform the EDG Administrative Group that you want to make your dataset available via the Clip N Ship application.*
- *Prepare your dataset for publication following the instructions in [Procedures for Publishing Data to the EPA Clip N Ship Application](#).*
- *Review or create metadata for your dataset and contribute it to EDG.*
- *Provide your new content to the EDG Administrative Group for publishing.*

EDG Administrative Group:

- *Ensure that the dataset metadata is compliant and available in EDG.*
 - *Publish the dataset to the Clip N Ship web service.*
 - *Update the Clip N Ship application with the new dataset.*
-

5.3.4 Post Data to EDG Data Download Locations or EPA Data Commons

The EDG Data Download Locations and EPA Data Commons FTP site are provided to EPA Regions, Program Offices, and Research Laboratories as a resource to post downloadable data. EDG Data Download Locations has a file size limit of 250 MB—larger downloads should be posted to EPA Data Commons. Organizations should discuss funding arrangements with the EPA Metadata Coordinator once more than 20GB have been posted to EPA Data Commons. There is no requirement to share resources on either download site; it is provided as a convenient resource for EPA organizations that want to make specific resources available to others internally and/or to the public.

Both the [EDG Data Download Locations](#) and [EPA Data Commons FTP site](#) gives EPA Regions, Program Offices, and Research Laboratories space to store geospatial and non-geospatial resources and make them available for download. These resources provide both Public and Restricted locations. External public users without an EPA LAN account may access data in the Public folders. The Restricted folders may be accessed only individuals with an EPA LAN account. No login is necessary for users accessing either site from within the EPA Intranet.

More information about the procedure for storing datasets in these locations may be found in the [Procedures for Storing Data at the EDG Data Download Locations](#) document. The EDG Administrative Group generally coordinates with the EDG Steward or contributor to obtain login rights to one or both of the download sites to perform their own uploads. The procedure to add new users differs between the two sites.

Grant user access to EDG Download Locations. The EDG Administrative Group must add the new user's AA account to the internal EDG server. Once added, the new user may contribute data to the EDG Data Download Locations using an FTP client tool and connecting to "edg-intranet.epa.gov" as the host.

ACTIONS:**EDG Steward or contributor:**

- Determine accessibility requirements for your dataset (*Restricted or Public*).
- Contact the EDG Administrative Group for assistance in determining the best location to store your downloadable dataset.
- Work with the EDG Administrative group to obtain direct access to EDG Data Download Locations and/or EPA Data Commons.
- Upload your zipped dataset to the appropriate site using an FTP client tool.
- Review or create metadata for your dataset with the online linkage set to the appropriate data download URL or FTP path.
- Contribute your metadata record to EDG.

EDG Administrative Group:

- Grant user access to EDG Data Download Locations and/or EPA Data Commons.
 - Ensure that the dataset metadata is compliant and available in EDG.
-

Grant user access to EPA Data Commons FTP site. The EDG Administrative Group must request that a new user be added to EPA Data Commons. Once the request has been fulfilled, the new user may contribute data to the EPA Data Commons using an FTP client tool and connecting to “newftp.epa.gov” as the host.

5.3.5 Utilize EDG Reuse Components

[EDG Reuse Components](#) includes three tools that give application developers the ability to reuse EDG content in other applications: EDG Rest Interface, EDG Search Widget, and 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. More information about EDG Reuse Components may be found in the [Putting the Environmental Dataset Gateway \(EDG\) to Use](#) document.

5.4 UPDATE AND MAINTAIN CONTENT IN EDG

Content contributed to 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.

Review of EDG resources (geospatial and non-geospatial metadata records, data downloads, and Clip N Ship data) may occur on a regular interval as determined by the EDG Steward. Such a review will

determine whether the resources require maintenance, such as updating or deleting, or whether resources need to be added.

Maintenance tasks that the EDG Steward may need to perform include transferring ownership of a metadata record and maintaining updated registration information for registered resources on a network. Periodically the EDG Steward may receive a maintenance request from the EDG Administrative Group when the Administrative group determines that EDG content needs updating.

Maintaining EDG content is also a responsibility of the EDG Administrative Group. The Administrative Group is responsible for ensuring that EDG content complies with all EPA requirements and that metadata records successfully synchronize with Data.gov. The Administrative Group is in regular contact with the EDG Stewards to inform them of changes in requirements that may require the Stewards to update or revise their EDG content.

The following sections provide additional detail on the review and maintenance routine required of the EDG Steward and EDG Administrative Group.

ACTIONS:

EDG Steward:

- *Perform maintenance on EDG content as needed.*
- *Regularly review EDG content to ensure it is up-to-date and relevant.*
- *Comply with requests from the EDG Administrative Group regarding EDG content maintenance.*

EDG Administrative Group:

- *Respond to EPA Metadata Coordinator requests regarding EDG compliance with Data.gov requirements.*
 - *Coordinate with EDG Stewards to maintain EDG content.*
-

5.4.1 Update Existing EDG Content with New Content

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

5.4.1.2 Manage Metadata XML Files in a Web Accessible Folder

The purpose of a web accessible folder (WAF) is to provide an easily accessible repository for metadata records scheduled to be synchronized to EDG. The WAF may be hosted at a Region, Program Office, or Research Laboratory location or it may be hosted at the National Computing Center. Typically the WAF is located on a web server and is managed by the organization. The organization's EDG Steward must

register the WAF location at EDG. Once registered, metadata XML files placed in the WAF are available for synchronization to EDG. A virtual WAF can also be configured to serve metadata documents stored in an Enterprise Geodatabase (Oracle, SQL Server, etc.). Contact the EDG Administrative Group to configure a virtual WAF.

Metadata XML files stored in the WAF should be examined regularly for accuracy and currency. Old metadata records should be removed or updated with current information and new metadata documents should be added after validating they meet [EPA requirements](#).

5.4.1.3 Update Data Posted to the Clip N Ship Application

The Clip N Ship application is updated as needed to maintain current information. EDG Stewards choosing to contribute geospatial resources to the Clip N Ship application should examine their hosted data on the Clip N Ship web page regularly to determine if updates are needed. Additionally, Clip N Ship metadata records should be examined at the same time for accuracy and currency. Old geospatial resources or metadata records should be removed or updated. EDG Stewards should coordinate with the EDG Administrative Group when Clip N Ship updates are needed as detailed in [Section 5.3.3](#) above.

5.4.1.4 Update Data Posted to EDG Data Download Locations and EPA Data Commons

EDG Stewards should examine the geospatial and non-geospatial resources stored at EDG Download Locations and EPA Data Commons regularly for accuracy and currency. If a resource has become out of date, the EDG Steward should remove the resource from the download location and post an updated version, if available. Procedures are detailed in [Section 5.3.4](#) above.

5.4.2 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 EDG will fail if not updated.

5.4.3 Verify EDG content synchronizes 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 publicly-available resources to Data.gov is satisfied through an automatic synchronization with EDG. Records in which the Publishing Organization field indicates the full standardized name of the EPA organization responsible for publishing the dataset are shared with Data.gov. 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.

5.4.4 Create and Manage Data.gov Collections

Project Open Data (<https://project-open-data.cio.gov>) defines a collection as:

Homogeneous series data are all of the same content, share most of the same metadata values, and might only vary in terms of content date and geographic extent. Examples include satellite imagery repositories, or data product series individually available for download. This type of

collection management is not applicable to most heterogeneous collections where every record should be indexed and is unique relative to its peers within the collection.

https://www.geoplatform.gov/sites/default/files/document_library/GeospatialMetadataBestPractices_20141010.pdf

A collection of this type is comprised of a single parent metadata record that describes the collection as a whole, and child records which contain embedded references to the parent. When the parent/child relationship is embedded in the metadata itself, rather than stored as a linkage in a metadata catalog, it allows that relationship to persist as the metadata are harvested and aggregated from catalog to catalog. Examples of homogenous collections of EPA data include Toxic Release Inventory (TRI) data released by year and state, or Re-Powering Alternative Energy data released by EPA Region. Refer to the [EPA Guidance for Creating Data.gov Homogeneous Collections](#) for additional information.

5.4.5 Create and Manage Compilations

EDG contributions, including geospatial and non-geospatial metadata records, can be organized into a Compilation by EDG Stewards and contributors. An EDG Compilation 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

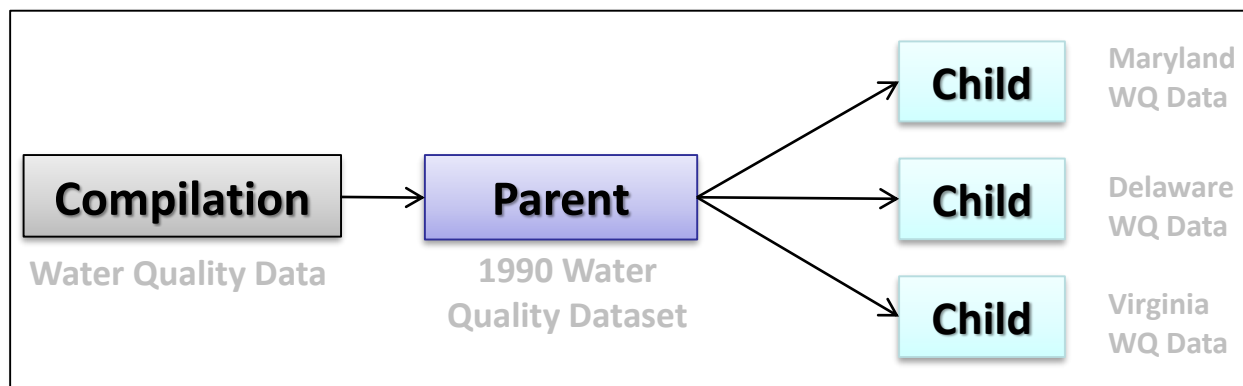


Figure 4: Example of metadata records grouped into a compilation by subject matter.

The creation of Compilations within EDG allows EDG contributors to organize metadata records of interest into meaningful groupings of related data resources. EDG users have the ability to create Compilations, manage Compilation metadata membership, and visualize the Compilations as a diagram or in the EDG Browse Tree (Figures 5 and 6). Compilations 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.

Create/Manage Compilations

Name : Owner : Any ▼

Results 1-50 of 59 record(s) 1 2 > Last













































Name	Owner	Placeholder Record	Approved	Manage	Edit	Delete
Featured Climate Data 	Region 1	<input type="button" value="Add"/>	Y			
Wasted Food Map 	Region 9	<input type="button" value="Add"/>	Y			
Navajo Nation AUM 	Region 9	<input type="button" value="Add"/>	Y			
Navajo Nation 	Region 9	<input type="button" value="Add"/>	N			
Tern Island Metadata Collection 	Region 9	<input type="button" value="Add"/>	Y			
Lower Colorado River 	Region 9	<input type="button" value="Add"/>	N			
Featured FRS Data 	OEI	<input type="button" value="Add"/>	Y			
Data Finder 	OEI	<input type="button" value="Add"/>	Y			
Performance Tracking of Tribal Air Programs 	OAR-OAQPS	<input type="button" value="Add"/>	Y			
National Emission Inventory 	OAR-OAQPS	<input type="button" value="Add"/>	Y			
Air Pollution Permit Statistics 	OAR-OAQPS	<input type="button" value="Add"/>	Y			

Figure 5: User interface for creating and managing Compilations.



Figure 6: Compilation Visualization. This view shows the relationship between the Compilation, Parent Record, and Child Records.

5.4.5.1 Add and Search Placeholder Records

EDG Compilations 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 Compilation, they have the ability to add and search records (referred to as “members” in EDG Compilations) to that specific Compilation.

5.4.5.2 Add and Search Members

EDG users have the ability to add and search placeholder records within their Compilations. A placeholder record simply contains a title and can be added to an EDG Compilation to temporarily retain the site where a complete metadata record will be added once it has been successfully documented, validated, and published to EDG. A placeholder record simply contains a title.

6.0 APPENDICES

APPENDIX A. KEY EDG CONTACTS

Contacts listed below identify key EDG operational personnel as of August 1, 2017. A list of EDG Stewards for the Agency's Regions, Program Offices, and Research Laboratories may be found at the [EDG Management and Stewardship](#) website.

- EDG Project Manager: Ana Greene
greene.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 EDG is provided by EPA's CIO Strategic Advisory Committee (SAC) in coordination with the Geospatial Information Officer (GIO). The SAC 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. Additional information regarding the roles and responsibilities of the SAC and SIO are available at:

- <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 SAC. 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 EDG.

Related Documents.

- Federal Initiatives
 - [*E-Government Act of 2002 \(H.R. 2458/S. 803\), December 17, 2002.*](#)
 - [*Project Open Data Metadata Schema v.1.1*](#)
 - [*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.*](#)
 - [*OMB Project Open Data policies*](#)
- EPA Policies, Procedures, Standards, and Guidance
 - [*EPA Metadata Technical Specification, 2016.*](#)
 - [*EPA Guidance, Enterprise Data Policy Implementation Guidance, Office of Environmental Information.*](#)
 - [*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. EDG Rules of Behavior

EDG Rules of Behavior (follows OEI/OITO ADC Security Documentation Requirements:
<http://cfint.rtpnc.epa.gov/adc/faq.cfm#adcgeneral55>)

Security Rules of Behavior for EDG:

- **Official Business:** Employees shall use EPA computer systems and information for official business only.
- **Access:** Employees shall access and use only information for which they have official authorization.
- **Accountability:** Employees shall be accountable for their own actions and responsibilities related to information resources entrusted to them.
- **Confidentiality:** Employees shall protect confidentially sensitive information about employees and others from disclosure to unauthorized individuals or groups. Employees shall protect Privacy Act Information (personal information about individuals).
- **Integrity:** Employees shall protect the integrity or quality of information.
- **Availability:** Employees shall protect the availability of information or systems.
- **Passwords and User-IDs:** Employees shall protect information security through effective use of user IDs and passwords.
- **Hardware:** Employees shall protect computer equipment from damage, abuse, and unauthorized use.
- **Software:** Employees shall use software in a safe manner that protects it from damage, abuse, and unauthorized use.
- **Awareness:** Employees shall stay abreast of security policies, requirements, and issues.
- **Reporting:** Employees shall promptly report security violations and vulnerabilities to proper authorities.

EDG General Use Rules:

- Do not leave information on your desk or your screen while away from your desk.
- Do not attempt to view, change or delete data unless you are authorized to do so.
- Do not use your system privileges to obtain information for anyone who is not authorized to do so.
- Do not allow another user to logon using your user ID and password.
- To prevent unauthorized use on your PC, log off whenever you will be away from your PC for an extended period of time
- Use a screen saver with a password.
- Shred all unneeded sensitive documentation.

- Mark all sensitive documentation.

Password Rules:

- Do not share your password with anyone.
- Password must be a minimum 8 characters in length, with at least one numeric and at least one alpha character. Try to create complex password (do not use words).
- Do not use family names, birthdays or other easily solved passwords.
- Password must be changed every ninety days.
- Password should be memorized, not written down.

Rules for Using Confidentially-Sensitive Data:

- Do not leave information on your desk or your screen while away from your desk.
- Ensure that only authorized personnel are allowed to view confidential data on your desk or your computer screen.
- Use a screen saver with a password.
- Position computer screens away from doors, windows and heavy traffic areas.

Rules for Accountability:

- Behave in an ethical and trustworthy manner.
- Do not attempt to perform actions or processing for which you do not have authorization.
- Be alert to threats to EPA applications and data.
- Logout and turn off your computer at the end of the workday.
- Report any potential security violations to your manager immediately.

Consequences of Non-Compliance:

Users must understand the consequences of not complying with principles and rules of behavior. Depending on the severity of the violation, at the discretion of management and through due process of the law, consequences can include:

- Suspension of access privileges
- Reprimand
- Suspension
- Demotion

- Removal
- Criminal and civil penalties, including prison terms and fines

Note: The EDG at time of login displays a system use notification statement as follows: “WARNING NOTICE: This is a United States Environmental Protection Agency (EPA) computer system, which may be accessed and used only for official Government business. Unauthorized access or use of this system may subject violators to criminal, civil, and/or administrative action. All information on this computer system may be monitored, recorded, read, copied, and disclosed by and to authorized personnel for official purposes, including law enforcement. Access or use of this computer system by any person, whether authorized or unauthorized, constitutes consent to these terms. Clicking the Login button constitutes consent to these terms and to the EDG Rules of Behavior.”

APPENDIX E. Metadata Steward Contribution Checklist

- ✓ Contact the EDG Administrative Group (edg@epa.gov) to inform them of the individual(s) in your organization who is responsible for managing EDG content and will act as steward(s).
- ✓ Understand EPA metadata compliance requirements by reviewing the [EPA Metadata Technical Specification](#).
- ✓ Review current organization holdings in EDG.
- ✓ Inventory current organization data resources, noting those that are created by EPA vs. those created by others and “improved” by EPA.
- ✓ Match EDG holdings to current data resources to determine need for new or updated metadata or download availability. Determine access level for each resource.
- ✓ Review data resources (web services and downloadable data) to be contributed to EDG to determine suitability for sharing with the public.
- ✓ Determine the ways in which internal customers and/or the public will access the data resource.
- ✓ Prepare to record the Distribution URL(s) in the resource’s metadata record.
- ✓ Author metadata for resources using an appropriate EPA Metadata Editor and following the EPA Metadata Style Guide.
- ✓ Ensure metadata record is valid before contributing to EDG.
- ✓ Determine location of organization’s web-accessible folder (WAF) for metadata. Contact EDG Administrative Group if assistance is needed.
- ✓ Load valid metadata XML files into the WAF.
- ✓ Contact EDG Administrative Group to inform them of new records.
- ✓ Work with EDG Administrative Group to post data to download and/or Clip and Ship locations.
- ✓ Perform maintenance on EDG content as needed.
- ✓ Regularly review EDG content to ensure it is up-to-date and relevant.
- ✓ Comply with requests from the EDG Administrative Group regarding EDG content maintenance.