


Metadata Best Practices

UPDATED! March 2014



3/6/2014

U.S. Environmental Protection Agency

1

Hello, everyone, and thanks for joining us for the third EME training session. Welcome to anyone who is joining us for the first time.



Agenda



- Review
- Workflow for Preparing Metadata
 - Creating and editing templates
 - Naming conventions and keywords
 - QA
- Workflow for Managing Metadata
 - Keeping metadata up-to-date
 - Online links
 - Data organization
 - Stylesheets

3/6/2014

U.S. Environmental Protection Agency

2

In the Advanced Features training session we discussed several ways to customize the EME and troubleshoot problems with synchronization and validation. In this session we're focusing on the overall metadata workflow. We want to provide you with an overview of how to prepare, manage, and share your metadata. Although our focus is on preparing metadata using the EME, you should be able to use many of these best practices regardless of your metadata editor. We will discuss how to use templates, how to manage your metadata efficiently while keeping it up-to-date, and how to share your metadata and link it to its corresponding datasets.



Review

- EME creates FGDC and EPA compliant metadata
 - Includes all mandatory FGDC elements and many optional elements
 - Works with XML, shapefiles, and geodatabases (personal, file, ArcSDE)
 - EPA Technical Specification: implementation of FGDC CSDGM
- MS Access database populates EME defaults

3/6/2014

U.S. Environmental Protection Agency



Let's start with a brief review. In the first session we introduced you to the EPA Metadata Editor. You can run the EME as a standalone application or as an extension of ArcCatalog. You can use the EME with a variety of file formats, including XML records, shapefiles, and various types of geodatabase. The information that populates the EME interface is stored in a Microsoft Access database.



Review

- Synchronization and Validation
- Database Customization
 - Change database location
 - Spell check
 - Compound elements
 - Keywords
- Training Sessions (slides, recordings, Q&A docs)
 - <https://edg.epa.gov/EME/resources.html>



3/6/2014

U.S. Environmental Protection Agency

4

The EME includes the EPA Synchronizer, which is a tool that reads properties of the dataset that the metadata describes and inserts those properties into the metadata.

The EME validator tests your records for compliancy with EPA and FGDC metadata standards. If there are any problems with your metadata, the validator highlights those errors for you.

In the second session we covered several ways to customize the EME's Access database. We showed you how to change the database location, which is helpful if you are running EME on a network, as well as how to customize spell check behavior, compound element behavior, and create your own keyword thesaurus.

Our previous training sessions are available online at the address on this slide. Today's session will also be posted.



Creating and Editing Templates

- Types of metadata templates
 - Customized Access DB
 - C:\Program Files(x86)\Innovate! Inc\EPA Metadata Editor\template
 - Complete XML record
 - Partial XML record



Save time with templates.
Finish faster.

3/6/2014

U.S. Environmental Protection Agency

5

Today's first topic is how to create and edit metadata templates. Templates can dramatically save time and reduce effort by providing default values for your metadata fields. Once you have a template customized for your specific project, that template can be exported and used as the default for metadata values across multiple datasets.

In the previous session we discussed how to customize the EME's Access database to meet your organization's requirements. By editing the EME database, you can modify the EME's default values so that when you click the "D" button in the EME interface, your record is automatically populated with your customized default values. Recall from the previous session that the Access database is stored in a folder called "template." Although you can think of the database as a kind of template for the EME, our focus here will be on XML templates.

An XML template is simply an XML record that contains information that will apply to multiple metadata records. It can be anything from a complete FGDC-compliant XML record with all fields filled in, to a partial record containing entries for just a few fields. When you open your XML template in the EME, all of the completed fields in the template will

automatically be filled in. Then all you need to do is fill in any remaining fields in the EME.

The clear benefit of using a template is that you will save time and reduce the chance that mistakes will be made while manually typing in the information multiple times.



Creating and Editing Templates

- Creating template for standalone XML files?
 - Use EME
 - Only complete common fields that you want to include in the template
 - Save file and use as a template for additional records



Depending on the type of data you are working with, you can choose how to create your template XML record. If you are working with a standalone XML metadata record (in other words, an XML file that is not attached to a data set), or if you do not have ArcCatalog installed on your computer, then you can use the standalone EME interface to complete the metadata fields that you want in your template. Once the desired fields are populated, save the file with a name like Metadata_Template.xml. Now you can simply open the file in the EME each time you want to create a new metadata record, fill in that record's individual information, and then save your file with a new name. If you have a metadata record that you find particularly useful, that can serve as a good starting point for creating a template.



Creating and Editing Templates

- Importing template for feature class?
 - Use ArcCatalog
 - ArcCatalog can apply templates to various types of datasets (e.g. file geodatabase, SDE database, service)
 - Save with other documentation of your service



ArcCatalog

However, if you have access to ArcCatalog you might find that using the EME tools from within ArcCatalog provides more flexibility. Using the EME within ArcCatalog, you can perform all the same functions as the standalone version of the EME and some additional functions. For example, you can export metadata from a dataset to create a standalone XML file which you can edit and use as a template. Additionally, you can edit metadata that is attached to a data set such as a geodatabase, shapefile, or Web service. And one of the most helpful functions is that you can import your standalone XML template directly into your data set. After importing the template, click the Edit Metadata button to add remaining values and specifics of the file. If you're working with a Web service, make sure to save your metadata with the rest of the service's documentation.



Creating and Editing Templates



Warning:

By importing a template, you override everything!



After Jared erased important metadata fields, his co-workers took steps to ensure that it would never happen again.

3/6/2014

U.S. Environmental Protection Agency

8

There's one important thing to remember when working with templates: If you import a template, you will override everything in your metadata record. It would be logical to think that by importing a template, you would only change the fields in the template, but that is not the case. If you have an existing metadata record, you will want to be very careful about importing a template.



Naming Conventions – Standalone Metadata



- Use descriptive names
 - In the XML title
 - **Good:** R1_Hydropower.shp.xml
 - **Not So Good:** AlphaXMLRDFSKOS_Web.xml
 - In the record
- Keep names consistent for your office

3/6/2014

U.S. Environmental Protection Agency

9

Metadata that is edited in the EME that is attached to a geodatabase feature class, shapefile, or other data set does not need to be assigned a file name, since it is saved as part of the data set. However, you may want to export your data set metadata to a standalone metadata XML file so that it can be shared in a metadata catalog such as EPA's Environmental Dataset Gateway or EDG. In that case, you will need to provide a name for the file.

It is important to give your record a descriptive file name. This will make it easier for others to work with your record, and it will make multiple records more consistent. It may be helpful for you to include the data set date in the title of your XML record. Whenever possible, you should give your file a name that anyone can understand. Whatever naming convention you use, you might find it useful to keep it consistent for all metadata records originating from your office or program. For example, if you work for EPA Region 1, you may choose to start your filenames with R1, Region-1, EPA_R1, etc.

It is also a good idea to use descriptive text within your metadata record so that individuals searching for data in the EDG, for example, will find

your metadata. When you are editing your metadata in the Metadata Editor, add terms that you think users might search for when looking for your metadata record. This helps users find your record based on common terms that may not be in your Keywords.



Keywords

- Be consistent
 - Acronyms and full terms
 - Include application names
- How will users search for your metadata?
 - Consider common search terms



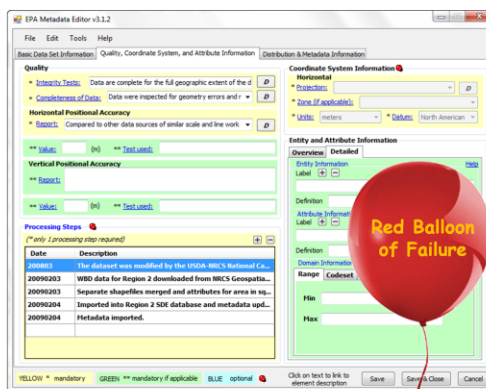
As night fell over the prairie, Kevin continued his long and fruitless search for metadata.

Speaking of keywords, they can make a big difference in how easy it will be for users to track down your metadata. When you are brainstorming your list of keywords, it's a good idea to include both acronyms and full terms. For example, your keywords could include both "Toxic Release Inventory" and "TRI." If your metadata is related to any applications, it's a good idea to include the names of those applications in the keywords. Take a moment to put yourself in the shoes of a user searching for your metadata and consider some terms that a user might include in a search query.



QA Practices

- Common validation errors
 - Errors vs. warnings
 - Deleting unnecessary elements



3/6/2014

U.S. Environmental Protection Agency

11

As stated earlier, you may need to upload your metadata records to catalogs like the EPA EDG or Data.gov. After you've completed your metadata and performed a validation, you may find that your validation reveals some errors and warnings. Be aware that it's okay to upload to these catalogs even if your metadata has warnings. Often our users see errors from elements that aren't required to meet EPA and FGDC standards. If that's the case, you can delete those optional elements before uploading to a catalog. Deleting optional elements can solve a lot of your validation problems. The EDG also offers its own validation service, which you can use to double-check your record's compliancy.



Warning:
Removes
any info
entered
using
ArcCatalog
style editor

QA Practices

Review: Removing Esri Tags

1. Import existing XML
2. Validate and fix errors
3. Export from ArcCatalog as FGDC
4. Delete everything with EME Clear All button
5. Re-import FGDC version into EME

Another QA practice is the technique for removing Esri tags. These tags are introduced when metadata is synchronized with the dataset using the ArcGIS synchronizer, and they can introduce problems with validation. This issue was addressed in detail in the Advanced Features session. The steps for removing those tags are summarized on this slide.



Common Metadata Mistakes

<http://www.nj.gov/dep/gis/top10.pdf>

- 10. Defining your data set too finely or too broadly
- 9. Using incorrect State Plane Coordinate System Zone Identifier values
- 8. Confusing 'Currentness Reference' with 'Publication Date'
- 7. Misunderstanding resolution
- 6. Putting too much faith in metadata tools
- 5. Taking the minimalist approach
- 4. Understanding assessments of consistency, accuracy, completeness, and precision
- 3. Glossing over Section 5. Entity and Attributes
- 2. Thinking of metadata as something you do at the end of the data development process
- 1. Not doing it!

Drum roll,
please...



U.S. Environmental Protection Agency

13

To wrap up this discussion of metadata best practices, we've included a link to a helpful document available from the New Jersey Department of Environmental Protection's website. It discusses the top ten mistakes that people make when preparing metadata, which include everything from misunderstanding certain metadata elements to procrastinating until the end of the data development process to create your metadata. All of these mistakes are described in more detail in the PDF, so we encourage you to take a look at it.



Demo 1: Creating & Importing a Metadata Template

3/6/2014

U.S. Environmental Protection Agency

14

With that we'll proceed to a demo of metadata preparation.



Keeping Metadata Up-to-Date

- Automatic updates
- Using EME Synchronizer
 - Use only for metadata attached to a dataset, e.g. file geodatabases
 - For standalone XML records, update manually



3/6/2014

U.S. Environmental Protection Agency

15

The second half of this session focuses on how to manage your metadata. Once you've gone to the trouble of creating and validating complete, well-thought-out, easy-to-find metadata, you might be tempted to save your records in a safe place and never look at them again. Thankfully, the process of keeping your metadata current does not have to be painful. By making regular updates to a few select fields, notably the online links, you can keep your metadata up-to-date.

The EME Synchronizer, which was discussed in the previous sessions, can automatically read some properties from your dataset and apply any changes to the metadata record. However, it's important to remember that this only works with metadata records that are attached to a dataset--for example file geodatabases. If you are working with a standalone XML file, then the EME synchronizer will not provide automatic updates. It won't know where to look for the dataset. If you're working with XML records, you'll need to make your updates manually.



Keeping Metadata Up-to-Date

- What happens if your data moves?
 - Update your **online links** to reflect new location
 - Do not rely on synchronization to point to accurate location
 - Syncing points to local drive by default



3/6/2014

U.S. Environmental Protection Agency

16

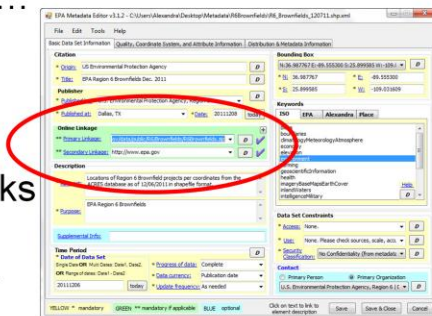
One of the most common problems that we see in metadata records is an out-of-date Online Linkage field. The Online Linkage field is particularly important because it allows anyone who looks at the metadata record to view where the dataset is stored. However, if you move the location of your dataset, the Online Linkage fields in your metadata need to be updated.

If you are working with a metadata record that is attached to a dataset, you can set the EME Synchronizer to automatically update the Online Linkage. However, by default the Synchronizer will automatically point to your local storage location. If you have made your data available to others by placing it in location that allows them to download it, your local storage location may not be that location, or the path may not provide the proper syntax to allow user downloads. Be sure to check the Online Link field and include the proper download address, or addresses of your datasets. Let's spend a few minutes discussing the types of online links you may want to use, and why they matter.



Data Organization

- Online Links to point to...
 - Downloadable data
 - Live services
 - Web pages
- Use multiple Online Links when possible
 - **Primary:** Link to data or service
 - **Secondary:** Link to program website



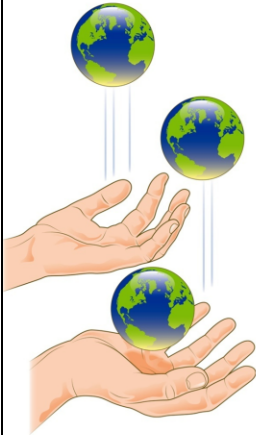
EME Tab 1

If you haven't looked for the Online Linkage field before, you can find it in Tab 1 of the EME. You'll notice that there are two fields: Primary Linkage and Secondary Linkage. The Primary Linkage should be a link to your dataset. Some of the most common and useful types of links are to downloadable data and live data and maps (also known as services).

If you want to include a link to your program's website, or information related to your dataset, you can enter that link in the Secondary Linkage field. You can also use your Primary and Secondary Linkages to link to multiple ways to access the dataset, for example a downloadable data link and a link to a service. Whenever possible, we suggest that you provide multiple online linkages. The more ways to access your datasets and related information, the better.



Data Organization



- Web Accessible Folders (WAFs)
 - Ex: <https://edg.epa.gov/metadata/waf>
- Catalogs/Geoportals
 - Ex: www.data.gov
 - Ex: <http://edg.epa.gov>

3/6/2014

U.S. Environmental Protection Agency

18

There are many ways for you to distribute and share your metadata. You can store it in a Web Accessible Folder, which is an online folder that users can access, or you can contribute it to a larger catalog of metadata such as the EDG. If you share your datasets, they should always have accompanying metadata—preferably metadata that is discoverable through an online catalog.

EPA offers a few different ways to make your data available to others. The following information is EPA-specific, but individuals outside of EPA can still use it as an example of how to link metadata to datasets and distribute data within your user community.



Data Organization

- Environmental Dataset Gateway (EDG) data download options:
 - EDG Data Download Location
 - ZIP files
 - <https://edg.epa.gov/data/>
 - EDG Clip 'N Ship
 - Web services
 - <https://edg.epa.gov/clipship/>



3/6/2014

U.S. Environmental Protection Agency

19

The EPA's metadata catalog is the Environmental Dataset Gateway (EDG). EDG contains thousands of metadata records, which have been contributed by metadata stewards at EPA Regions, Program Offices, and Labs. All of those records are managed by a team of EDG administrators who are responsible for checking and approving contributions. Many of the records are updated automatically on a regular basis through protocols set up within the EDG.

But the metadata catalog is only one part of the EDG. The EDG also offers a resource where users can access and download datasets. The EDG makes data available in two ways: the EDG Data Download Location, and the EDG Clip 'N Ship application. Any EPA office that publishes metadata to the EDG can use the Download Locations and the Clip N Ship to store and share their data.




Data Organization

- EDG Data Download Location
 - Provide online access to data
 - Any type of data (large datasets are Zipped)
 - Internal and external locations, organized by region... office
 - Access via FTP
 - EDG metadata record's Online Linkage points to Data Download Location
 - More information:
https://edg.epa.gov/metadata/webhelp/en/gptlv10/inno/EDG_Download_Locations.pdf



The Data Download Location is where any EPA Region, Program Office or Lab can store and share data. Data is loaded in an office-specific folder, frequently as ZIP files if the datasets are large. Data may be made available in a Public folder, which allows any user to access the downloadable data, or in a Restricted folder that requires an EDG login.

If you are interested in storing your data in the EDG Data Download Location, contact the EME/EDG team at edg@epa.gov and we will set you up with access via FTP. Once your datasets are stored in the Download Location, you can link that location in your metadata's Online Linkage field.

 **Environmental Dataset Gateway Download Locations**
Connecting EPA's Environmental Resources

Download EPA's Geospatial Data
Access geospatial data provided by EPA's Program Offices, Regions, and Labs by clicking on the folders below.

- Public
 - EPA Office of Chemical Safety and Pollution Prevention
 - EPA Office of Pesticide Programs
 - EPA Office of Compliance and Enforcement
 - EPA Office of Environmental Information
 - EPA Office of Policy
 - EPA Office of Research and Development
 - EPA Office of Solid Waste and Emergency Response
 - EPA Office of Water
 - EPA Region 1
 - CG_line_r100.zip
 - CG_poly_r100.zip
 - ESA_points_r24.zip
 - ESA_polys_r24.zip
 - Fuelfaci_r_EPA.zip
 - NPL_AIA_Boundaries.zip
 - NPL_NLPAD.zip
 - NPL_Site_Boundaries.zip

<https://edg.epa.gov/data/>

3/6/2014 U.S. Environmental Protection Agency 21

Here's a screenshot of the Data Download Location. As you can see, the Public folder contains datasets and files from various EPA offices, all available for download.



Data Organization

- EDG Clip N Ship
 - Interactive viewer allows users to preview datasets, clip by spatial extent, and download
 - Contact us for more information about contributing data
 - More information:
https://edg.epa.gov/metadata/webhelp/en/gptlv10/inno/EDG_ClipAndShip_procedures.pdf

3/6/2014

U.S. Environmental Protection Agency

22

The other option within the EDG for making your data available via download is the EDG Clip N Ship application. This option lets users access your data via an interactive web map. Users can preview data, clip it by spatial extent, and download it to their local machines. If you're interested in adding your data to the Clip N Ship, our team can help. We will gather some information from you, launch your data as a service, and add it to the Clip N Ship.

EPA
Clip N Ship Application
Latitude: 19.388855 Longitude: -140.179638

Streets Aerial Topo

EPA Clip N Ship

1. Select Area
Choose a state [Clear]

OR Use the draw tool [Rectangle] [Circle] [Polygon] [Line] [Point]

2. Select Data

- Data
 - R1 Brownfields
 - R1 CG Jurisdictional Boundary
 - R1 Environmentally Sensitive Areas - pts
 - R1 Environmentally Sensitive Areas
 - R1 Facility Response Plan
 - R1 Fuel Storage Facilities

3. Select Vector Format
File Geodatabase (.gdb)

4. Select Raster Format (if needed)
ESRI GRID (ESRI GRID - GRID)

5. Select Spatial Reference
Same As Input - Web Mercator Auxiliary Sphere

Download

<https://edg.epa.gov/clipship/>

3/6/2014 U.S. Environmental Protection Agency 23

This is a screenshot of the Clip N Ship viewer. Available data layers are listed in the table of contents. The icon circled in red links to the dataset's metadata record in the EDG. By clicking this button, users see a pop-up preview of the EDG metadata description. That way they can learn a little bit about the dataset before deciding if they want to download it.

The screenshot shows the EPA Environmental Dataset Gateway (EDG) interface. At the top, there is a navigation bar with links for HOME, ABOUT, SEARCH, BROWSE, DATA, and REUSE. Below this, there are tabs for Details, Review, Relationships, and Preview. The main content area displays the metadata for a dataset titled "Region 9 NPDES Outfalls - Waste Water Treatment Plants". The metadata includes the following information:

- Title:** Region 9 NPDES Outfalls - Waste Water Treatment Plants
- Originator:** US EPA Region 9, GIS Center
- Publication Place:** San Francisco, CA
- Publisher:** US EPA Region 9, GIS Center
- Publication Date:** 20120620
- Website URL:** https://edg.epa.gov/data/Public/R9/R9_Stakeholder_Outreach/NPDES_WWTP_Outfalls.gdb.zip
- Resource(Server) URL:** <https://edg.epa.gov/clipship/>

An abstract and purpose section are also visible, providing context for the dataset. A blue callout box on the right side of the screenshot contains the text: "Online Linkages connect EDG metadata record to Data Download Location and Clip N Ship".

This is an example of an EDG metadata record that is connected via its Online Linkages to both the Data Download Location and the Clip N Ship. A user can go to the EDG, search for a record of interest using search terms like "Region 9," "NPDES," or "Waste Water," and find this metadata record. Then, with a few clicks, the user can either download the entire dataset from the Data Download Location, or preview, clip and download the dataset using the Clip N Ship. You can see how metadata is now much more than just a bunch of words on a page. It is really a tool that allows users to find and access datasets quickly and easily and get detailed information about those datasets without needing to contact your office and ask for help.



Data Organization

- EPA GeoPlatform Online
 - Provide online access to data
 - Use the data to create shareable Web maps



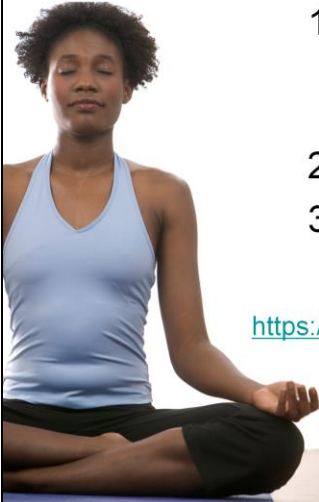
And finally just a brief word about a third option for sharing data -- the EPA GeoPlatform Online. Although not directly linked to the EDG, the GeoPlatform Online allows you to post some datasets and links to data services and share those with the EPA geospatial community. In addition, you can create interactive Web maps with your datasets and share those Web maps with other users.



Here is a screenshot of the GeoPlatform Online homepage. You can find more information about EPA GeoPlatform Online at the webpage, or by going to the EPA Geo Resources intranet page listed here.



Sarah breathed deeply,
embracing the
Circle of Metadata.



Data Organization

1. Contact EDG Team to add your datasets to Download Locations and/or Clip N Ship
2. Enter Online Links in EME
3. Upload metadata records to EDG

<https://edg.epa.gov/metadata/catalog/content/about.page>

J.S. Environmental Protection Agency

27

If you're interested in taking advantage of the data organization options offered within the EDG, you should follow a few steps: First, get in touch with the EDG team to add your datasets to the Download Locations or Clip N Ship. Second, enter the Online Links to those locations in your metadata. Third, upload your metadata records to the EDG.

More details are available at the address on this slide.



Working with Stylesheets



Rachel strived to be as stylish as her metadata.

- The finishing touch
- Change the way metadata is displayed in ArcCatalog

3/6/2014

U.S. Environmental Protection Agency

28

The final topic in this session pertains to working with stylesheets. They put a finishing touch on your metadata, changing the way it is displayed in ArcCatalog.



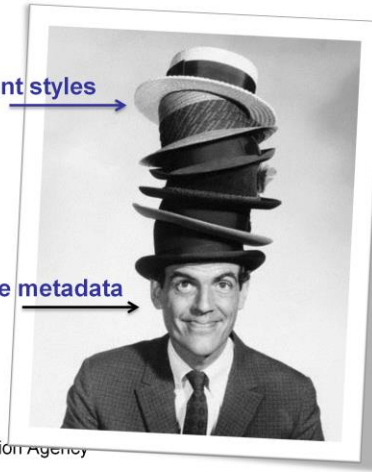
Working with Stylesheets

XLST file – changes appearance of XML doc

“Metadata styles” are designed to make it easier for Esri users to switch between standards

Different styles →

Same metadata →



3/6/2014

U.S. Environmental Protection Agency

When you apply a stylesheet in ArcCatalog, a type of file called an XLST changes the appearance of your XML document.

Metadata styles are designed to help users switch between metadata standards. The idea is that you can switch between standards without having to go back and edit your metadata. Instead, you can simply apply a new stylesheet.



Working with Stylesheets

Esri's default is Description style

Lola's style did not comply with standards.

Simple style designed for users not required to comply with metadata standards

More details:

<http://blogs.esri.com/esri/arcgis/2011/01/12/a-new-approach-for-metadata-with-arcgis-10-part-3/>



3/6/2014

U.S. Environmental Protection Agency

30

The default style that you see in ArcCatalog is called the Description style. This style is designed to be simple and cater to users who are not required to comply with metadata standards—users who just need to create some minimal metadata. When you open a record in Description style, you'll only see a few basic metadata elements.



Working with Stylesheets

- FGDC Style
 - Allows you to export ArcGIS metadata to FGDC format
 - ArcCatalog → Customize → ArcCatalog Options → Metadata Tab
 - Select “FGDC CSDGM Metadata”

The FGDC style allows you to view all FGDC elements and export ArcGIS metadata to FGDC format. With ArcGIS 10.1 and higher, simply navigate to the Metadata Tab in the Customize options within ArcCatalog and select the FGDC CSDGM Metadata style. Folks that are still using ArcGIS 10.0 may need to install the FGDC Style Patch, which is available at Esri’s website, or upgrade to 10.1.



Working with Stylesheets

EPA Metadata Stylesheet

You'll need...

- ArcCatalog 10 or later
- Administrative privileges

Download EPA Metadata Stylesheet and instructions:

<https://edg.epa.gov/EME/Resources.htm>



Download PDF for detailed installation instructions

The EPA also offers its own stylesheet. To download it, go to the EME Resources pages. We have also posted detailed installation instructions. Before you download it, be aware that you will need ArcCatalog 10 or later, along with administrative privileges on your computer.



Demo 2: Working with Metadata from Other Sources

3/6/2014

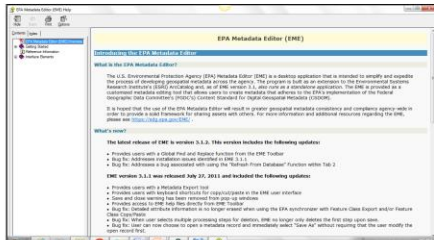
U.S. Environmental Protection Agency

33

With that we'll move on to the second demonstration, which will walk you through how to manage metadata that you receive from other sources.

Getting Help

- EME Help
 - Extensive information available in EME Help
 - Help→Contents



Automatic updates

Templates?

Triple venti caramel machiatto with whip skim milk and cinnamon...

Style patches?

Grumpy Cat!



I mentioned EME help at the beginning of the presentation, but it's worth repeating that a lot of information is available to you in the Help documentation. We've sped through a lot of content today. All of what we've talked about is included in the Help documentation, which you can access from the EME interface.

Getting Help

- EME web resources
 - Fact sheets
 - Training presentations
 - Helpful links



<https://edg.epa.gov/EME/Resources.html>

There are also some useful resources available on the EME website, including fact sheets, training presentations, and links to other metadata resources. This presentation will also be posted at this address.



Getting Help

- Contact Us:
– edg@epa.gov



We're standing by to answer your questions.

If you'd like to get in touch, feel free to contact the EME team. We always appreciate questions and feedback from EME users.