

Metadata – Frequently Asked Questions

1. What is Metadata?

Metadata is commonly referred to as “data about data”. There are several different types of metadata. These can be broadly defined under the categories of Descriptive, Structural, and Administrative.

- **Descriptive Metadata:** Information that describes the content, quality, and context of a data resource for the purpose of facilitating identification and discovery. It may reference additional information like quality assurance documents and data dictionaries. Through descriptive metadata a user can learn the what, why, when, who, where, and how for a data resource.
- **Structural Metadata:** Information about how the item is put together or arranged such as the table of contents page, individual page numbers, or illustration. It basically describes the structure of an item, such as a book, so that all of the pages of that item can be displayed in the correct order. In the electronic world it facilitates navigation and presentation of electronic resources.
- **Administrative Metadata:** Includes information about resolution, bit depth, type of equipment used to produce the file, storage format, and file name and location. It can also include basic facts on ownership, rights, and reproduction information.

The Environmental Public Health Tracking Network (EPHTN) makes extensive use of Descriptive Metadata. It is considered the backbone of the Network. To view examples of completed Descriptive Metadata entries go to geodata.gov <http://gos2.geodata.gov/wps/portal/gos>.

2. What are the benefits for creating Descriptive Metadata?

As more data are being created and stored, there is a need to document data resources for future use and improve accessibility. The benefits of creating Descriptive Metadata include:

- Helps an organization arrange and maintain its data assets.
- Limits duplication of effort by ensuring that others in the organization are aware of the existence of data resources.
- Assists in both determining and improving the quality of data resources.
- Improves an organization’s ability to comply with rules, regulations, and policies relating to data access.
- Reduces the loss of institutional memory for data resources when key staff move on.
- By providing information about an organization’s data holdings users can locate available resources relevant to an area of interest or study.

- Provides the ability to advertise and promote the availability of data resources via online services.
- Supplies the means to document limitations about the data resource or disclaimers that are important for potential users to be aware of.

3. **Why is Descriptive Metadata important to the EPHTN?**

Descriptive Metadata is important for two key aspects of network functionality. It allows for the discovery of data resources on the Network. Network users can locate resources through a variety of means including keywords, geographic boundaries, and date and time. All of these elements are part of the Descriptive Metadata entry.

It is also important because it provides descriptive information about data resources available on the Network. Through Descriptive Metadata, a Network user can determine the content of the resource, why it was created, how it was created, any limitations, access and use restrictions, data quality, and contact information. It helps a user to decide if a resource is appropriate for the intended use.

4. **Are there standards for the creation and maintenance of Descriptive Metadata?**

Yes, there are several standards available for the creation and maintenance of Descriptive Metadata. The Metadata Subgroup (MDSG) investigated three primary standards for describing data resources on the EPHTN.

- Dublin Core (<http://dublincore.org/>)
- International Organization for Standardization (ISO; <http://www.iso.org/iso/en/ISOOnline.frontpage>)
- Federal Geographic Data Committee (FGDC; <http://www.fgdc.gov/>).

The results of this investigation showed that the FGDC Content Standard for Digital Geospatial Metadata was the most appropriate and adaptable for use within the EPHTN. Developed in 1994 (revised in 1998) to describe geospatial data, the standard is unique in that it contains elements from all three metadata types (Descriptive, Structural, and Administrative). The MDSG determined that the standard provided the most utility for the EPHTN by allowing for the description of both geospatial and non-geospatial data resources.

5. **What is the EPHTN Metadata Profile?**

The EPHTN Metadata Profile outlines the format and content for describing data resources on the Network. The profile complies with the FGDC Content Standard for Digital Geospatial Metadata. It contains 52 of the 195 total data elements available for completion in the FGDC Standard. This represents the minimum description necessary for providing data to the Network. The profile consists of a template and recommendations to assist in the completion of each element. It was developed to make the creation of Descriptive Metadata easier for data stewards.

6. Are there any costs to Descriptive Metadata development?

There are upfront costs for developing Descriptive Metadata. These primarily involve staff time in collecting the basic information about the data resources, physically creating the metadata, undertaking quality assurance of each metadata entry, and training on using the metadata creation tool / editor. However, these initial costs are offset by the benefits.

7. How do data stewards create and edit Descriptive Metadata?

There are several tools for creating and editing Descriptive Metadata. Some types of software include built-in metadata editors (e.g. ArcGIS) to allow the data steward to create associated metadata files and edit them. Others are freeware that require knowledge of a particular metadata standard.

To both simplify and standardize the creation of Descriptive Metadata for the EPHTN, CDC contracted with Northrop-Grumman for the development of a metadata creation tool/editor. The tool is built around the EPHTN Metadata Profile. It guides the data steward through the process of creating FGDC-compliant metadata. The tool is available to stewards providing data to the Network.

8. When should Descriptive Metadata be created, and who should create it?

As a best practice, Descriptive Metadata should be created in parallel (at the same time) as the data resource itself. If the resource already exists, data stewards will have to rely upon existing documentation and institutional knowledge to develop the metadata.

Descriptive metadata is best developed by those that know and understand the contents of the data resource. If someone is designated to create all Descriptive Metadata for an organization, a process will need to be implemented to provide that individual with the necessary information to complete the metadata entry.

9. How often should Descriptive Metadata be updated?

Descriptive Metadata should be updated whenever the data resource is updated. This is especially important for changes that alter the structure, content, quality, conditions, and other characteristics of the resource. Updating metadata can also establish when the resource was altered, and by whom. This is important for determining the resource version and how it was altered.

10. Where does Descriptive Metadata reside once it is created?

Ideally, Descriptive Metadata should be stored as closely to the actual data resource as possible for ease of access. The storage method might be a central repository (database) with different permission levels, or simply a collection of extensible markup language (XML) files stored in a common location. In either case, the goal is to keep the metadata connected to the data in a way that will encourage people to use this important documentation to obtain a better

understanding of the data, and encourage the metadata's maintenance. The EPHTN is creating a central repository for sharing Descriptive Metadata for data resources made available on the Network. Once operational, data stewards will be able to upload their Descriptive Metadata into the EPHTN repository.

11. How do I access Descriptive Metadata on the EPHTN?

Descriptive Metadata for those data resources available on the EPHTN can be accessed through the EPHTN portal at [[Insert Web Address Here](#)]. The portal includes a search function that provides several criteria for locating data resources.

12. Where can I get more information about Descriptive Metadata?

Additional information on the creation, maintenance, and use of Descriptive Metadata is available on the EPHTN portal [[Insert Web Address Here](#)].

Other resources:

- Federal Geographic Data Committee (www.fgdc.gov)
- U.S. Geological Survey (<http://geology.usgs.gov/tools/metadata>)