

Metadata

Metadata Concepts

What is "Metadata"?

Metadata is "data about data".

"Metadata is a summary document providing content, quality, type, creation, and spatial information about a data set. It can be stored in any format such as a text file or database record. Because of its small size compared to the data it describes, metadata is more easily shareable. By creating metadata and sharing it with others, information about existing data becomes readily available to anyone seeking it.

Metadata makes data discovery easier and reduces data duplication."

Metadata and GIS,
ESRI, White Paper, October 2002

- Metadata is a formal data documentation (information about datasets) and it is critical to data discovery
- A metadata record is a set of information fields, usually presented as an XML document, which captures the basic characteristics of data as an information resource
- It represents the **who, what, when, where, why** and **how** of the resource
- **Who** collected and who distributes the data?
- **What** is the subject, projection of the data?
- **When** was the data collected?
- **Where** was the data collected?
- **Why** was the data collected (what is the purpose)?
- **How** was the data collected? **How** should it be used?
- **How** much does it cost?

Metadata - Stored as...

Metadata can be stored as:

- an extra file (e.g an XML)
- within a file (as an attribute)
- as a single extra database documenting summaries

```
<SyncDate>20070525</SyncDate>
<SyncTime>22282600</SyncTime>
<ModDate>20070525</ModDate>
<ModTime>22282600</ModTime>
<MetaID>(58C72CBC-8D7A-4132-8113-67BB929A86FC)</MetaID>
</Esrri>
<idinfo>
  <acritabon>
    <acritabon>
      <conlink>\\halimsvr\SVG_GIS\DATA\STAKEHOLDERS\DFS</conlink>
      <pubdate>30/09/2000</pubdate>
      <title Sync="TRUE">SVG_Landing_Sites_BWI_V45</title>
      <fname Sync="TRUE">SVG_Landing_Sites_BWI_V45</fname>
      <geoform Sync="TRUE">vector digital data</geoform>
    </pubinfo>
    <pubplace>Kingstown, SVG</pubplace>
    <publish>Physical Planning Unit and the Department of Fisheries</publish>
  </pubinfo>
  <congr>Physical Planning Unit and the Department of Fisheries</congr>
</idinfo>
```

ID	Shape*	AREA	PERIMETER	ISLANDS	ISLANDS_ID	TYPE	ACRES	NOTES	TYPES	DIGITIZER	DATE	SOURCE
1366	Polygon	60624.74	20433.454	0	0		189.169		Land	Greemore	10/03/2000	LANDSAT 99
1367	Polygon	42949.397	1094.848	0	0		19.813		Cloud	Greemore	10/03/2000	LANDSAT 99
1368	Polygon	41442.84	1039.764	0	0		19.163		Land	Greemore	10/03/2000	LANDSAT 99
1369	Polygon	47338.562	3022.461	0	0		117.985		Cloud	Greemore	10/03/2000	LANDSAT 99
1370	Polygon	87337.766	1919.899	0	0		21.581		Turbidity	Maguire	12/03/2000	
1371	Polygon	1165468.489	4029.241	0	0		287.905		Turbidity	Maguire	13/03/2000	LANDSAT 99
1372	Polygon	57724.881	3398.563	0	0		142.632		Turbidity	Maguire	13/03/2000	LANDSAT 99
1373	Polygon	58800.251	862.21	0	0		14.53		Cloud	Maguire	13/03/2000	LANDSAT 99
1374	Polygon	14114.993	588.025	0	0		3.503		Shadow	Maguire	13/03/2000	LANDSAT 99
1375	Polygon	70822.246	1280.368	0	0		17.525		Cloud	Maguire	13/03/2000	LANDSAT 99
1376	Polygon	88701.651	2433.45	0	0		21.918		Turbidity	Maguire	13/03/2000	LANDSAT 99
1377	Polygon	3085.548	1202.958	0	0		7.438		Beach	Maguire	13/03/2000	LANDSAT 99

Data Set	Feature Type	Attributes	Type	Responsible	Location	Observations	Projection	Exam
Basin Layers / Topographic	Line		Baseline	NERO/DFP	Baseline_Data	SVG_basins.shp (or extract from basinset.shp)	UTM20 NAD27 / BVI SV45	YES
All Islands (Coordinates)	Polygon		Baseline	SPS	Geographical	extract coast (ALLISLES.shp)	BVI SV45	YES
Island Area (land) .shp	Polygon		Baseline	SPS	Geographical	ISLAND.shp	BVI SV45	YES
Road (St. Vincent)	Line		Baseline	NERO	Baseline_Data	SVG_roads.shp	UTM20 NAD27	YES
Road (Coordinates)	Line		Baseline	SPS	Geographical	extract roads (ALLRES.shp)	BVI SV45	YES
Rivers (St. Vincent)	Line		Baseline	NERO	Baseline_Data	SVG_rivers.shp	UTM20 NAD27	YES
River (Coordinates)	Line		Baseline	SPS	Geographical	extract rivers (ALLRES.shp)	BVI SV45	YES
Watersheds	Line		Baseline	DOF/NERO	Flood Risk Assessment/Theme	to be provided by DOF / NERO only Hesp.	UTM20 WGS84 / BVI SV45	YES
Hydrography (St. Vincent)	Point		Baseline	NERO	Baseline_Data	SVG_contour.shp	UTM20 NAD27	YES
Hydrography (Coordinates)	Line		Baseline	NERO	Baseline_Data	SVG_division.shp	UTM20 NAD27	NO
Spot Heights (St. Vincent)	Point		Baseline	NERO	Baseline_Data	SVG_division.shp	UTM20 NAD27	YES
Spot Heights (Coordinates)	Point		Baseline	SPS	Geographical	extract features (ALLRES.shp)	BVI SV45	YES
Parishes (Coordinates)	Point		Baseline	NERO	Baseline_Data	SVG_division.shp	UTM20 NAD27	NO
Gazetteer	Point		Baseline	NERO	Baseline_Data	SVG_division.shp	UTM20 NAD27	NO

Metadata Elements - Category Overview:

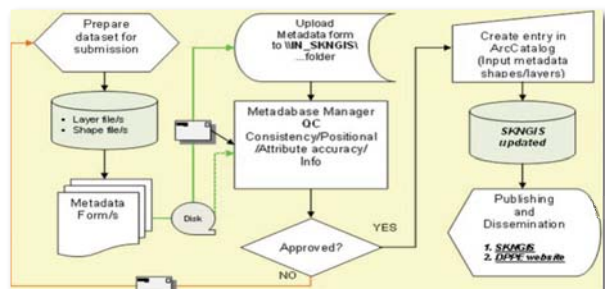
- **Identification** - title, originator, description, area covered, currency, rule for acquiring or using the data, restrictions, liability
- **Data Quality** – positional and attribute accuracy, source of info, method used for producing the data
- **Spatial Data Organization** – method used to represent spatial data (vector, raster, etc.)
- **Spatial Reference** - projection, type, scale, size
- **Entity and Attribute** - entity type, attributes, domain for attribute values
- **Distribution** - contact for distribution, available formats, how to obtain dataset, physical media, fees, etc.
- **Metadata Reference** - currency of the metadata information, responsible party

Metadata production

- Fill in metadata entry form to best of ability

The image shows four metadata entry forms. The first is a 'METADATA ENTRY FORM' with sections for Originator, Contact Information, General, and Descriptions. The second is 'Constraints' with sections for Time-Period, Status, Keywords, and GIS-Metadata. The third is 'Security' with sections for Security and Pricing recommendations. The fourth is 'FIELD INFORMATION' with a table for Field Name, Field Type, and Field Description.

- Provide new dataset(s) and include symbology for layer file



- Submit for QC of lyr / shp files.
- Include transformation to projection if necessary
- Provide updates to metadata as dataset changes

Concepts and use in search and documentation

Main uses of Metadata

- Organize and maintain an organization's internal investment in spatial data
- Provide information about an organization's data holdings to data catalogs, clearinghouses, and portals (it is the entry point of an SDI and for each SDI node)
- Provide information to process and interpret data received through a transfer from an external source

Benefits of using Metadata

- Facilitates data maintenance
- Facilitate data discovery
- Enables reuse of data
- Informs potential users of inappropriate uses of a dataset

Challenges of using Metadata

- Exchanging records -> Use standards!
- Creating metadata takes time and patience -> Use templates and tools that automatically extract or set metadata (like *GeoNode*). Imperfect metadata is better than none
- Maintaining metadata is essential, especially for Web services and downloads (beware of URL changes!) -> set up processes or tools for automatic checkups

Metadata standards

- Some standards defines the metadata content (fields, values, mandatory v. optional)
 - The most comprehensive is ISO 19115
- From this standard, custom profiles can be defined for specific countries, themes, etc.
- Other standards define appropriate exchange formats
 - The most accepted is ISO 19139, which defines an XML-based format

ISO 19115

The information included in the ISO 19115 standard supports the following use:

- **discovery** - data needed to identify and locate the sets of geographic data that exist for a geographic location;
- **access** - data needed to acquire an identified set of geographic data;
- **fitness for use** - data needed to determine if a set of geographic data meets the user's need and to support the user in applying the geographic information appropriately; and
- **transfer** - data needed to obtain a copy of a set of geographic data.

Metadata Content

Identification: Basic information about the data set. Examples include the title, the geographic area covered, currency, and rules for acquiring or using the data

The screenshot shows a metadata form titled 'Identification info'. It contains the following fields and values:

- Title:** Caribbean Map WMS
- Alternate title:** The most amazing Caribbean map as Web Map Service
- Date:** 2013-01-03T00:00:00
- Date type:** Creation
- Abstract:** This is a WMS service providing access to a map of the Caribbean countries.
- Status:** Completed
- Point of contact:**
 - Individual name: (empty)
 - Organisation name: (empty)
 - Position name: (empty)
 - Role: Point of contact

Information in abstract: data model

- Entity types and their attributes and the domains from which attribute values may be assigned
- The names and definitions of features, attributes, and attribute value

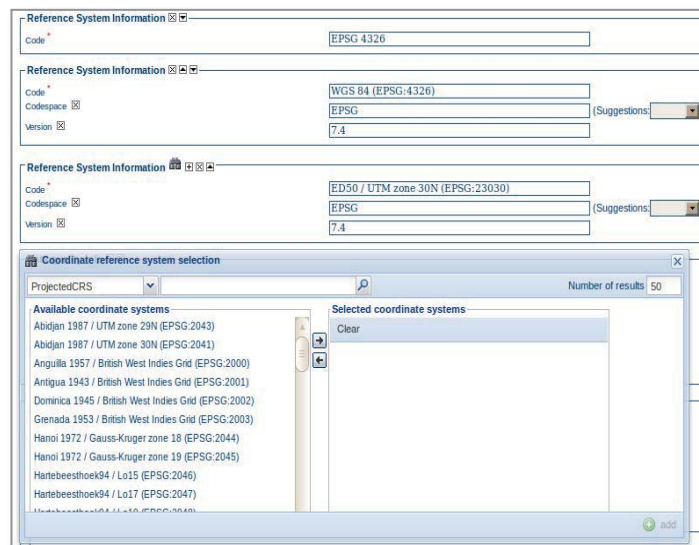
Data quality

- Positional and attribute accuracy
- Completeness and consistency,
- Sources of information
- Methods used to produce the data
- Appropriate uses (scales)

An assessment of the quality of the data set. Examples include the positional and attribute accuracy, completeness, consistency, the sources of information, and methods used to produce the data

Spatial Reference

- Description of the reference frame,
- Means of encoding,
- Coordinates in the data set,



Examples include the name of and parameters for map projections or grid coordinate systems, horizontal and vertical datums, and the coordinate system resolution

Entity and Attribute

- Information in abstract: data model
 - Entity types and their attributes and the domains from which attribute values may be assigned
 - The names and definitions of features, attributes, and attribute value

Information about the content of the data set, including the entities types and their attributes and the domains from which attribute values may be assigned.

Examples include the names and definitions of features, attributes, and attribute values

Distribution information

- How to access the information?
- Multiple transfer options are possible
- Contact for physical media
- Fees

Information about obtaining the data set.

Examples include a contact for the distributor, available formats, information about how to obtain data sets online or on physical media (such as cartridge tape or CD-ROM), and fees for the data

Metadata Reference

- Language, date, update status
- Who is responsible for the metadata?

Information on the currency of the metadata information and the responsible party. Examples include currency and information about the organization that provided the metadata

Access and use constraints

- Use constraint statements can be crafted to express scale, geographic, or temporal limitations to the data
- Liability statements should be written by legal staff to ensure that the legal requirements for use of the data are fully outlined

In general, it is far better to publish your dataset with limitations within your metadata than to later attempt to generate them in response to an inquiry or lawsuit

Access constraints + x Copyright

Use constraints + x Restricted

Other constraints + x

Metadata Tools

As seen in GeoNetwork

Identification info

Title

Date

Date type Publication

Edition

Presentation form mapDigital

Abstract The ISO19115 metadata (metadata standard to use

OnLine resource

URL http://sinpad.indeci.gob.pe:7080/cgi-bin/indexserv.exe

Protocol OGC Web Map Service (ver 1.1.1)

Name consolidado_de_peligros_manzanas

Description http://sinpad.indeci.gob.pe:7080/cgi-bin/indexserv.exe

Access constraints Acceso libre

Use constraints No autorizado para uso comercial

Spatial representation type vector

Equivalent scale

Descriptive keywords

Keyword alerta

Type amenaza

asociación

avenida torrencial

avenida torrencial

capacitación

catastro

centro de reserva

comunitario

cooperación

coordinación

damnificado

Access constraints

Ints

resentation type

As seen in ESRI's ArcCatalog

Contents | Preview | Metadata

SV_elevation_points_BWI_V4
Shapefile

Description | Spatial | Attributes

Keywords

Theme: elevation

Place: St. Vincent, Caribbean

Description

Abstract

This layer consists of 145,393 elevation points on St. Vincent. It is believed that these points were created by the Canadian International Development Agency while doing work in St. Vincent. Presumably these points were digitized using a profiling approach from the 1:2500 topographic maps housed by the St. Vincent and the Grenadines Survey Department. As far as I can tell the 1:2500 set of maps were produced from the same aerial photography used to create the 1:25,000 topographic maps (DOS, 1982).

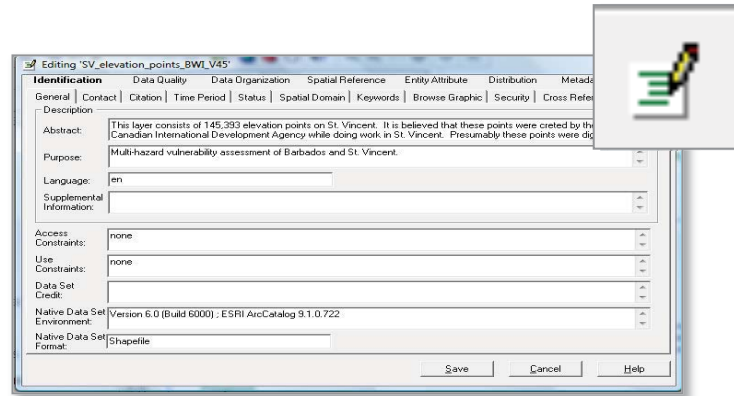
Purpose

Multi-hazard vulnerability assessment of Barbados and St. Vincent.

Status of the data

Complete

Data update frequency: None planned



Conclusions

- **Metadata is important**
 - To know what data exist
 - To know the quality of the data
 - Who is responsible for the metadata
- **You can use it to**
 - Search for suitable data
- **Help the management of the NGIS**
 - Using the metadata for searching and querying
 - Add new metadata using the forms
 - Help the Metadatabase Manager to improve the information about the NGIS

Metadata References

- Metadata Guidelines - Land Information Council of Jamaica - 2006
<http://www.licj.org.jm/licj/documents/metadata.html>
- International Standard ISO/FDIS19115, **Geographic Information - Metadata**
- ESRI White Paper (October 2002), “**Metadata and GIS**” - 4. Vermeij, Bert, “implementing European Metadata Using ArcCatalog”, ESRI ArcUser, July - September
www.esri.com/library/whitepapers/pdfs/metadata-and-gis.pdf