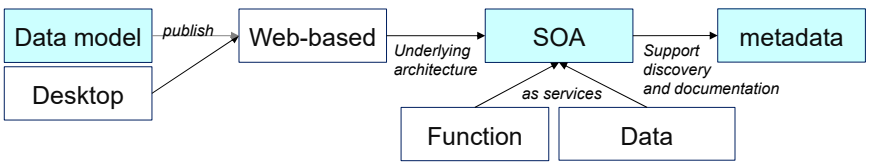
11. Geospatial Metadata

GE3238 GIS Design and Practices
Geography@NUS
Chen-Chieh FENG

1

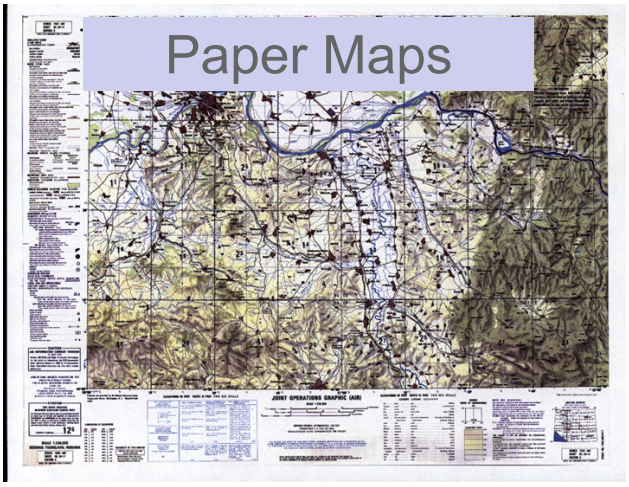
Learning Objectives

- A better appreciation of metadata
  - What is it?
  - Why is it important?
- Metadata content standard
  - Important metadata elements for "data"
- Overview for metadata editing



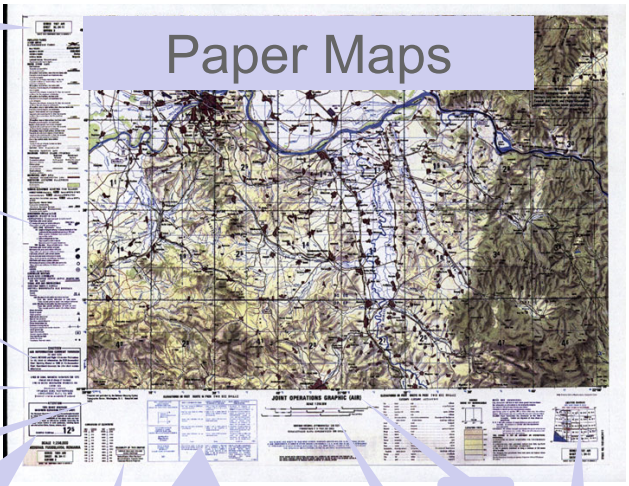
2

Paper Maps



3

Paper Maps



Identification

Symbols

Date

Sources

Producer

Navigation notices

Accuracy

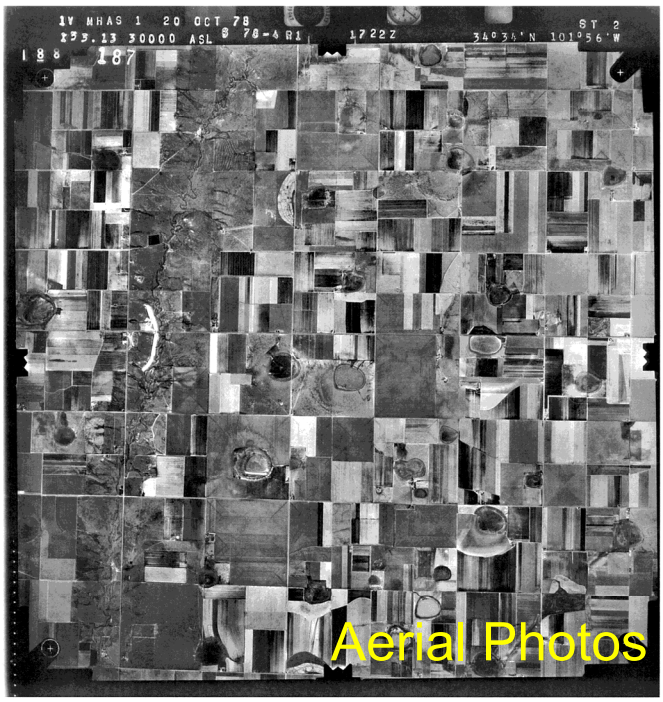
Reference system

Scale

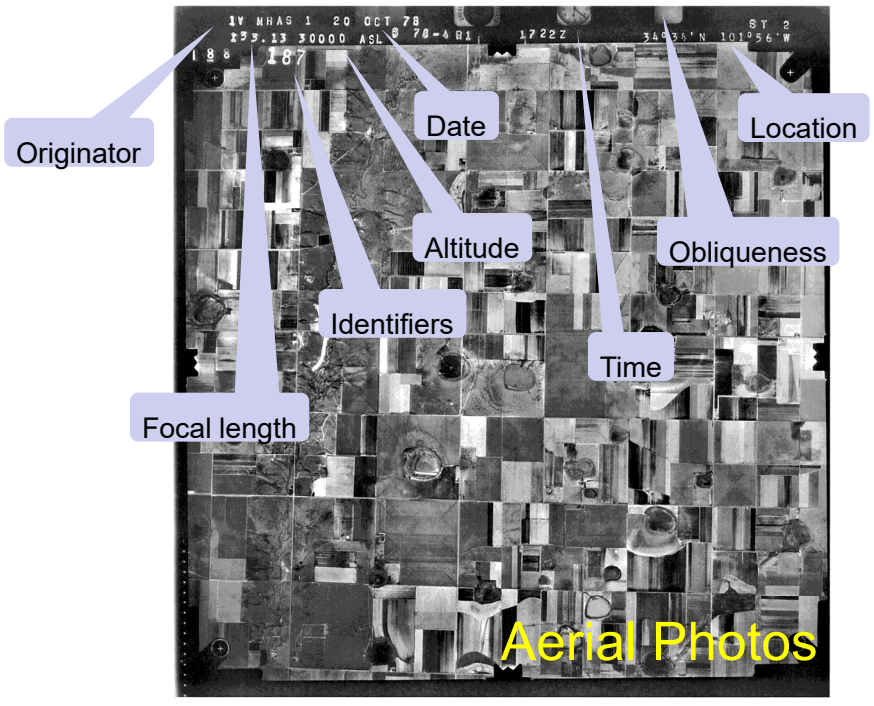
Title

Location

4


Aerial Photos

5


Originator
Date
Location
Altitude
Obliqueness
Identifiers
Time
Focal length
Aerial Photos

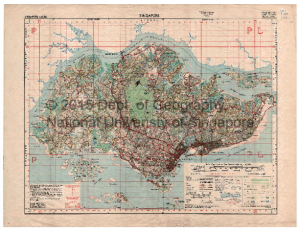
6

```
▼<metadata>
  ▼<idinfo>
    ▼<citation>
      ▼<citeinfo>
        ▶<origin>...</origin>
          <pubdate>19580101</pubdate>
          <title>1958 - SINGAPORE 1:63,360 TOPOGRAPHICAL MAP</title>
          <geoform>remote-sensing image</geoform>
        ▶<onlink>...</onlink>
        </citeinfo>
      </citation>
    ▼<descript>
      ▶<abstract>...</abstract>
        <purpose>Topographic map of Singapore in 1958.</purpose>
      </descript>
    ▶<timeperd>...</timeperd>
    ▶<status>...</status>
    ▼<spdom>
      ▼<bounding>
          <westbc>103.550707</westbc>
          <eastbc>104.072276</eastbc>
          <northbc>1.531221</northbc>
          <southbc>1.183418</southbc>
        </bounding>
      </spdom>
    ▼<keywords>
      ▼<theme>
          <themekt>ISO 19115 Topic Categories</themekt>
          <themekey>boundaries</themekey>
          <themekey>environment</themekey>
          <themekey>imageryBaseMapsEarthCover</themekey>
          <themekey>inlandWaters</themekey>
```

HISTORICAL MAPS, 1951-1964

1958 – SINGAPORE ISLAND MAP

MARCH 6, 2015 KELMAN CHIANG



Summary

Topographic map of Singapore in 1958.

7

data.gov.sg

Try keywords like: 'school' or 'weather'

Help Feedback Login

Economy

Resident Population by Planning Area/Subzone of Residence, Age Group and Floor Area of Residence (Census of Population 2020)

Updated 17 days ago

SINGSTAT (Singapore Department of Statistics)

Source: SINGAPORE DEPARTMENT OF STATISTICS
Data Last Updated: 18/06/2021
Update Frequency: 10 years
Survey period: Census of Population 2020
Footnotes: Note: Planning areas refer to areas demarcated in the Urban Redevelopment Authority's Master Plan 2019.
Adapted from: https://tablebuilder.singstat.gov.sg/table/CT/17563
See less

Download CSV (237.3 KB) Subscribe

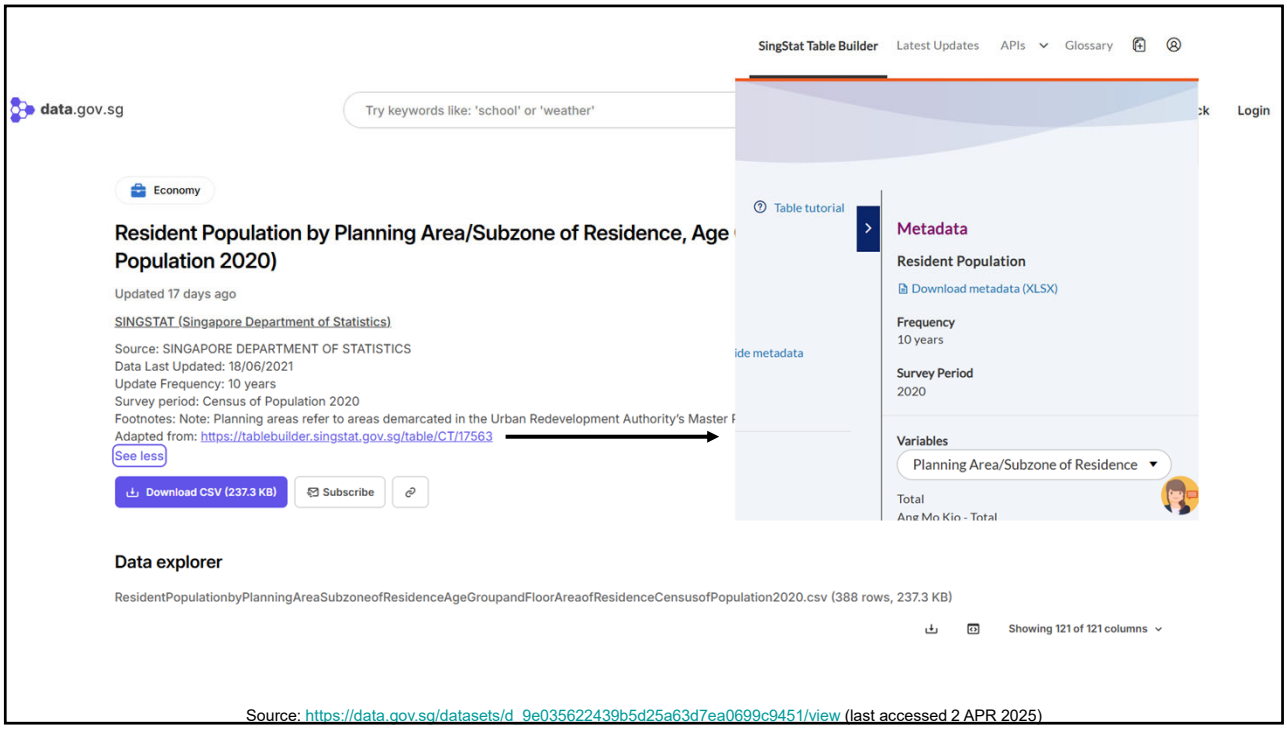
Data explorer

ResidentPopulationbyPlanningAreaSubzoneofResidenceAgeGroupandFloorAreaofResidenceCensusofPopulation2020.csv (388 rows, 237.3 KB)

Showing 121 of 121 columns

Source: https://data.gov.sg/datasets/d_9e035622439b5d25a63d7ea0699c9451/view (last accessed 2 APR 2025)

8

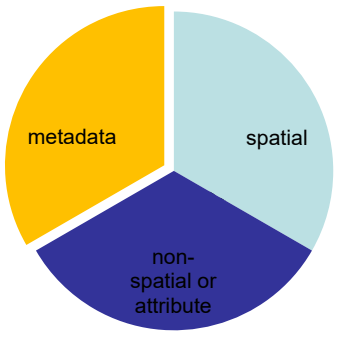


Source: https://data.gov.sg/datasets/d_9e035622439b5d25a63d7ea0699c9451/view (last accessed 2 APR 2025)

9

What is Metadata?

- Data about data; provide context of data
- It describes how geospatial and attribute data and how they were collected and processed
- One of the three components of geospatial data (Source: ESRI Virtual Campus)
  - Description of GIS data is incomplete without it



10

What does this attribute mean?

FID	Shape	AREA	PERIMETER	A10SOILP#	A10SOILP-ID	ARC-ID	LPOLY#	RPOLY#	POLY-ID	SOILCLASS
2	Polygon	205037.40625	2854.017090	2	1	2662	2	2	52005	UpC
3	Polygon	7759044	30731.910156	3	2	2596	3	3	8210	GuD
4	Polygon	98960.21875	1210.482056	4	3	0	4	4	8203	WmE
5	Polygon	4257001	34723.601563	5	4	2357	5	5	8192	VbD
6	Polygon	13112360	123709.296875	6	5	1993	6	6	8155	WmF
7	Polygon	73276.28125	1271.327026	7	6	2653	7	7	8191	WhF
8	Polygon	1802398	14282.209961	8	7	2646	8	8	8188	WhE
9	Polygon	311326.3125	5024.500977	9	8	2648	9	9	8187	WhD
10	Polygon	31646000	130463.296875	10	9	2641	10	10	8554	VbD
11	Polygon	145485.59375	1735.811035	11	10	2643	11	11	8208	WhD
12	Polygon	642108.375	10473.910156	12	11	2582	12	12	9014	EbF
13	Polygon	2296537	24645.380859	13	12	2441	13	13	8576	WhE
14	Polygon	178981.90625	2156.708008	14	13	2636	14	14	9018	VbD
15	Polygon	89592.40625	1754.155029	15	14	2597	15	15	9012	UsD
16	Polygon	2704050	23853.519531	16	15	2631	16	16	8609	GuD
17	Polygon	104809.601563	1589.582031	17	16	2630	17	17	9020	WhE
18	Polygon	73713.78125	1141.522949	18	17	2625	18	18	20666	WhE
19	Polygon	38971	774.400391	19	18	2623	19	19	8202	Ww0
20	Polygon	44667752	193799.90625	20	19	2621	20	20	8604	VbD
21	Polygon	2733101	31330.449219	21	20	0	21	21	9022	WhE
22	Polygon	2743654	32333.269531	22	21	2352	22	22	8605	WmF
23	Polygon	213043.703125	3308.434082	23	22	2613	23	23	9034	VbD

What do these SOILCLASS values mean?

11

Why Metadata

- Maintaining an organization's investment in spatial data
  - Provides an inventory of data assets
  - Helps determine and maintain the value of data
  - Documents legal issues (copyright, liability, privacy)
  - Helps keep data accurate and helps verify accuracy to support good decision making and cost savings
- With metadata support, data producers can publish information about data

https://pro.arcgis.com/en/pro-app/latest/help/metadata/view-and-edit-metadata.htm (last accessed 2 APR 2025)

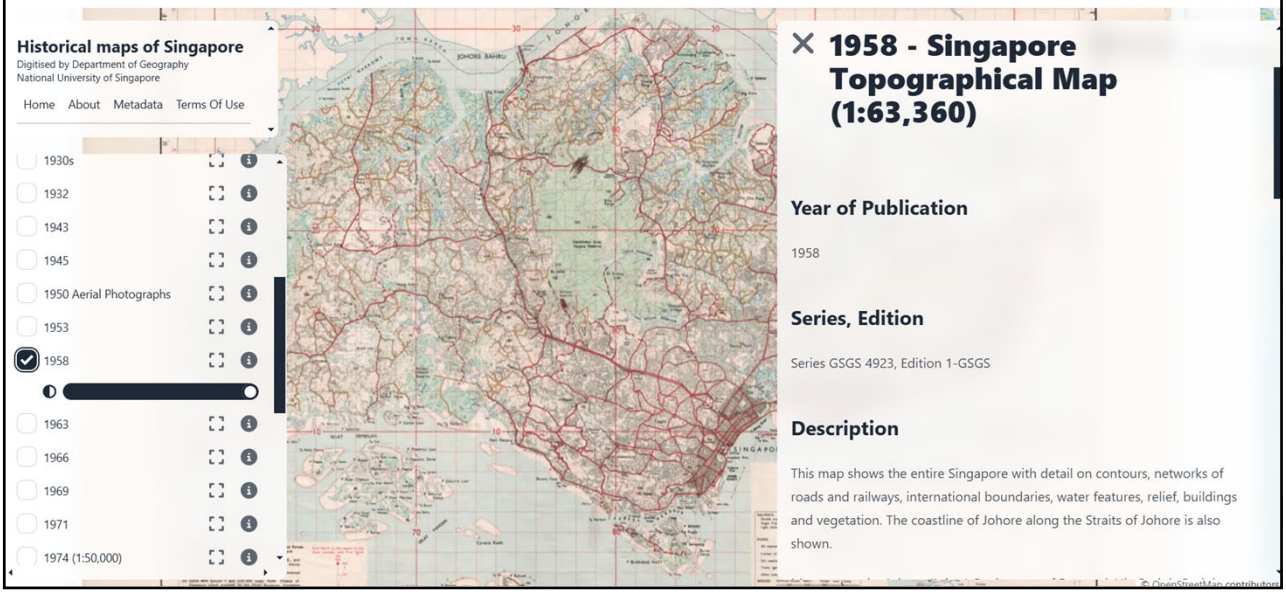
12

Why Metadata (cont.)

- Data users need metadata to locate appropriate data sets
  - Metadata makes data discovery easier and reduces data duplication
- It is important to know if the data will meet **user needs**
  - Helps you determine the reliability and currency of data
  - Data quality of geospatial data? Fitness-for-use
- Metadata not only helps find data, but once data has been found, it also tells how to interpret and use data

13

https://libmaps.nus.edu.sg/maps/1958/



14

Details Review Relationships Preview

1958 - SINGAPORE 1:63,360 TOPOGRAPHICAL

1958 - SINGAPORE 1:63,360 TOPOGRAPHICAL MAP

This map shows the entire Singapore with detail on contours, networks of roads and railways, international boundaries, water features, relief, buildings and vegetation. The coastline of Johore along the Straits of Johore is also shown.Map covers the Johor...

Open Preview Globe (.kml) ArcGIS (.nmf) ArcGIS (.lyr) Add To Map
Details Metadata

Identification Information:
Citation:
Citation Information:
Originator: Singapore Survey Department, Singapore Survey Department
Publication Date: 19580101
Title: 1958 - SINGAPORE 1:63,360 TOPOGRAPHICAL MAP
Geospatial Data Presentation Form: remote-sensing image
Online Linkage: http://www.nusgis.com/arcgis/rest/services/Sing_Hist_Maps/1958/MapServer?f=jsapi
Description:
Abstract: This map shows the entire Singapore with detail on contours, networks of roads and railways, international boundaries, water features, relief, buildings and vegetation. The coastline of Johore along the Straits of Johore is also shown.Map covers the Johore (Johor) Strait, parts of Punggol, Ulu Bedok, Bedok, Changi, Pulau Ubin, Pulau Serangoon and other surrounding islands. It shows Gammon Malaya Quarry, other granite quarries and fisheries reserves in Pulau Ubin, Public Works Department Labourers Lines, Handicapped Childrens' Home, Maternity and Children's Welfare Centre and government bungalows in Changi, Ubin Estate, Ong Ting Lye Estate, Wah Hong Estate, Tampines Estate, Bukit Sembawang Estate, Hun Yeang Village, Kampong Noordin, Pasir Ris Village, Kampong Loyang, Yan Kit Village, Changi Village, proposed international airport under construction, residential development areas, cemeteries, industry sites, rural areas, Sungei Serangoon, Sungei Tampines, streams and rivers, existing and proposed community and public utility facilities, roads and other land use.* Central RegionBishan, Bukit Merah, Bukit Timah, Novena, Potong Pasir, Queenstown, Tanglin, Tiong Bahru, Toa Payoh, Southern Islands, Bras Basah, Chinatown, Downtown Core, Marina Bay, Marina Centre,

15

Details Review Relationships P

1958 - SINGAPORE 1:63,360 TOPO

Identification Information:
Citation:
Citation Information:
Originator: Singapore Survey Depart
Publication Date: 19580101
Title: 1958 - SINGAPORE 1:63,360 T
Geospatial Data Presentation Forr
Online Linkage: http://www.nusg er?f=jsapi
Description:
Abstract: This map shows the entire Sir national
boundaries, water features, relief, buildir e is also
shown.Map covers the Johore (Johor) St erangoon
and other surrounding islands. It shows in Pulau
Ubin, Public Works Department Labourer are Centre
and government bungalows in Changi, U Bukit
Sembawang Estate, Hun Yeang Village, K Changi
Village, proposed international airport un y sites, rural
areas, Sungei Serangoon, Sungei Tampir utility
facilities, roads and other land use.* Cen Queenstown,
Tanglin, Tiong Bahru, Toa Payoh, Southe na Centre,

Keywords:
Theme:
Theme Keyword Thesaurus: ISO 19115 Topic Categories
Theme Keyword: boundaries
Theme Keyword: environment
Theme Keyword: imageryBaseMapsEarthCover
Theme Keyword: inlandWaters
Theme Keyword: location
Theme Keyword: society
Theme Keyword: structure
Theme Keyword: transportation
Theme Keyword: farming
Theme Keyword: economy
Theme Keyword: elevation
Theme Keyword: planningCadastre
Theme:
Theme Keyword Thesaurus: None
Theme Keyword: Singapore
Theme Keyword: 1958
Theme Keyword: Singapore Island map
Theme Keyword: historical data
Access Constraints: None
Use Constraints: None

16

Two paradigms of geospatial data production and distribution

Centralized

- A data set may be produced, stored and maintained by a single organization
- Single machine or closed operation

Distributed – SOA – Enterprise

- Local or regional organizations produce and store data that is accessible via services operated under national standards
- Web GIS

17

17

Big Ten's Geospatial Data Portal connecting access to GIS, historical map data

January 23, 2017

UNIVERSITY PARK, Pa. — The Big Ten Academic Alliance (BTAA) Geospatial Data Portal Project, of which Penn State University Libraries is a contributor, has launched an online spatial data discovery tool called the Big Ten Academic Alliance Geoportal. The p
provide discoverability, facilitate access an
Big Ten Academic Alliance to often scattere

"The BTAA Geoportal is managed collectively by a task force of more than 20 librarians and geospatial specialists at 10 member institutions," said Paige Andrew, University Libraries maps cataloging librarian and BTAA task force member. "Thousands of metadata records have been aggregated and edited to connect scholars across the Big Ten to geospatial data resources, including GIS datasets, web services and digitized historical map images from multiple data clearinghouses and library catalogs or websites." Andrew is one of three Penn State representatives on the task force, along with metadata strategist Linda Ballinger and geospatial services librarian Nathan Piekielek.

https://www.psu.edu/news/impact/story/big-tens-geospatial-data-portal-connecting-access-gis-historical-map-data (last accessed 2 APR 2025)

18

Metadata Elements

- Metadata is a summary document providing
  - Content
  - Quality
  - Type
  - Creation
  - Spatial information about a data set

In plain language, this means that it is supposed to provide 4Ws (**what**, **where**, **when**, and **why)** information of spatial data…

- It can be stored in any formats:
  - Text file
  - Extensible Markup Language (XML)
  - Database record

19

Quote from

GI Framework Data Standard, Base Document

- All datasets shall have metadata that conforms to at least the minimal set of mandatory elements of either ISO 19115-1, Geographic Information Metadata, or FGDC-STD-001-1998, Content Standard for Digital Geospatial Metadata
  - http://www.fgdc.gov/standards/projects/FGDC-standards-projects/framework-data-standard/GI_FrameworkDataStandard_Part0_Base.pdf (p12)
  - "Must" and "Optional" (see the next slide)
- However, more extensive metadata should be provided (http://www.fgdc.gov/standards/standards_publications/index_html)

20

“M” and “O” in ISO19115-1

An Example

B.2.4.2.2 Process step information

	Name / Role Name	Short Name	Definition	Obligation / Condition	Maximum occurrence	Data type
86.	LI_ProcessStep	PrcessStep	information about an event or transformation in the life of a dataset including the process used to maintain the dataset	Use obligation from referencing object	Use maximum occurrence from referencing object	Aggregated Class (LI_Lineage and LI_Source)
87.	description	stepDesc	description of the event, including related parameters or tolerances	M	1	CharacterString
88.	rationale	stepRat	requirement or purpose for the process step	O	1	CharacterString
89.	dateTime	stepDateTm	date and time or range of date and time on or over which the process step occurred	O	1	Class
90.	processor	stepProc	identification of, and means of communication with, person(s) and organization(s) associated with the process step	O	N	Class

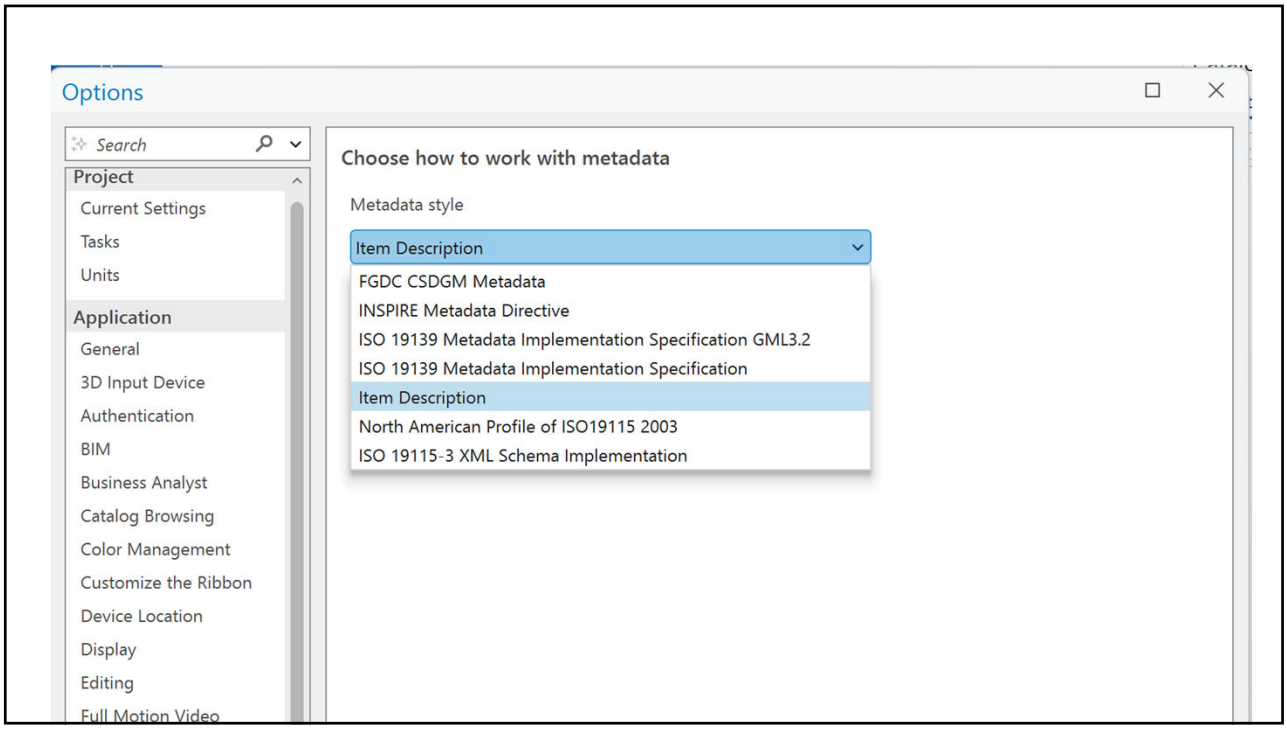
21

Metadata Standards

- In the US, Federal Geographic Data Committee (FGDC) content standard for digital geospatial metadata (CSDGM)
  - FGDC-STD-001-1998
    https://www.fgdc.gov/standards/projects/metadata/base-metadata/v2_0698.pdf
- The International Organization of Standardization (ISO) has also created a spatial metadata standard
  - ISO 19115:2003 Geographic Information (GI)
  - ISO 19115-1:2014 GI Metadata Part 1 Fundamentals
    - Data quality components moved to ISO 19157:2013 Data Quality
  - ISO 19115-3:2023 Geographic information - Metadata – XML schema implementation

Other related metadata include Dublin Core, Marc Record, and PROV (i.e., provenance)

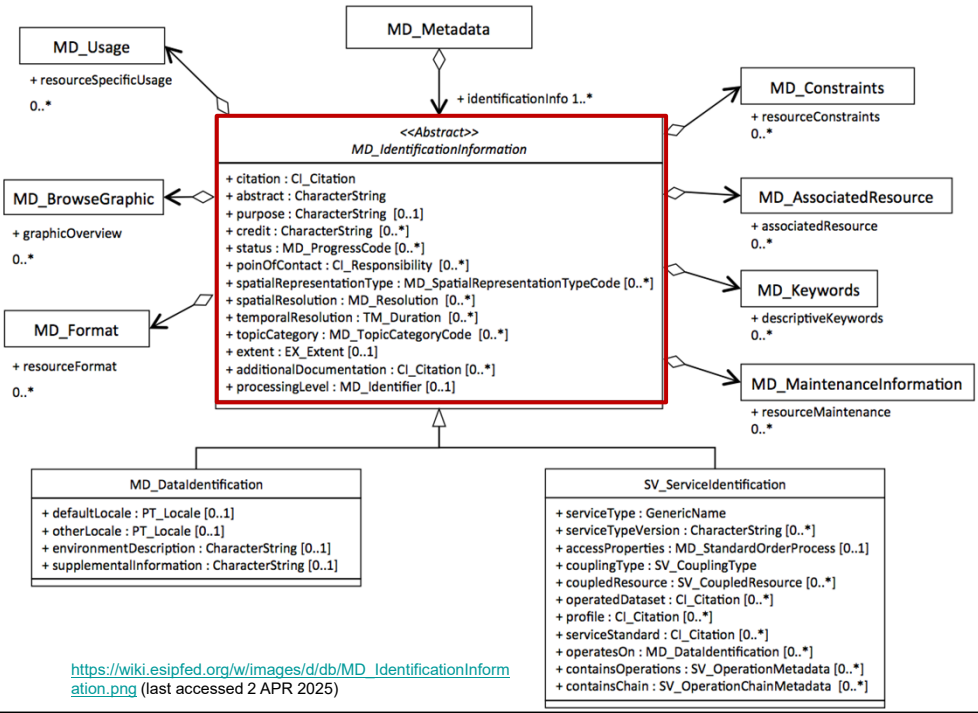
22



23

ISO 191

- Citati
- Abstr
- Purpc
- Eleme data identi
  - Lan
  - cha
  - Top
  - Exte

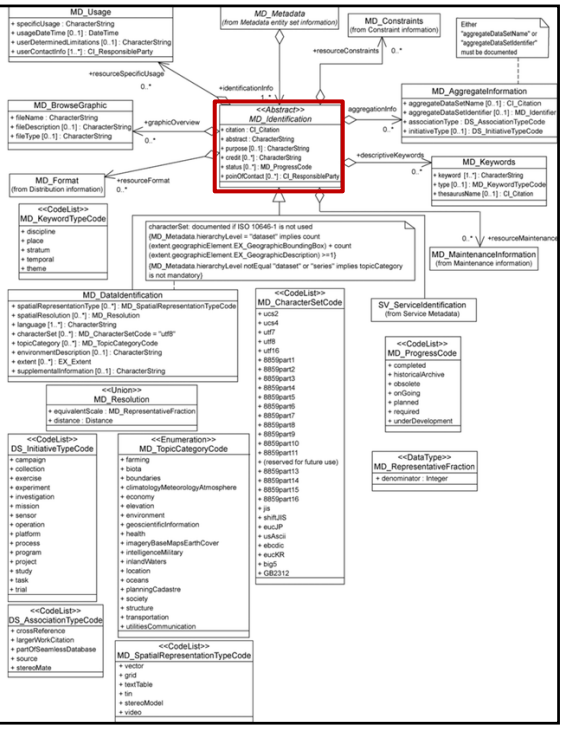


https://wiki.esipfed.org/w/images/d/db/MD_IdentificationInformation.png (last accessed 2 APR 2025)

24

ISO 19115 Identification

- Citation
- Abstract
- Purpose, credit, status, etc
- Elements for data identification
  - Language
  - character set
  - Topic category
  - Extent
- Elements for service identification
  - Service type



25

ISO 19115 Core Elements

For Geographic Datasets

Metadata Language
Metadata Date Stamp
Geographic Location of the Dataset
Dataset Topic Category

Dataset Language
Abstract Describing the Dataset
Dataset Character Set
Dataset Reference Date
Dataset Title

Metadata Character Set

Metadata Point of Contact

Must | Conditional Must

ISO 19115:2003(E)

Table 3 — Core metadata for geographic datasets

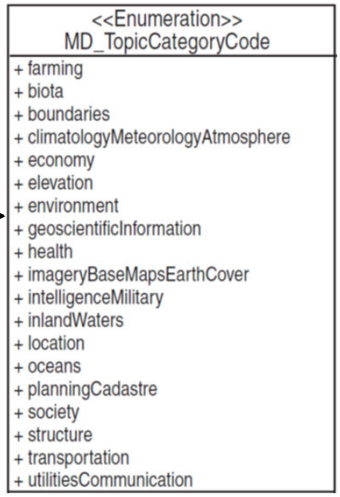
Dataset title (M) (MD_Metadata > MD_DataIdentification citation > CI_Citation title)	**Spatial representation type** (O) (MD_Metadata > MD_DataIdentification spatialRepresentationType)
Dataset reference date (M) (MD_Metadata > MD_DataIdentification citation > CI_Citation date)	**Reference system** (O) (MD_Metadata > MD_ReferenceSystem)
Dataset responsible party (O) (MD_Metadata > MD_DataIdentification pointOfContact > CI_ResponsibleParty)	**Lineage** (O) (MD_Metadata > DQ_DataQuality lineage > LI_Lineage)
Geographic location of the dataset (by four coordinates or by geographic identifier) (C) (MD_Metadata > MD_DataIdentification extent > EX_Extent > EX_GeographicExtent > EX_GeographicBoundingBox or EX_GeographicDescription)	**On-line resource** (O) (MD_Metadata > MD_Distribution > MD_DigitalTransferOption onLine > CI_OnlineResource)
Dataset language (M) (MD_Metadata > MD_DataIdentification language)	**Metadata file identifier** (O) (MD_Metadata fileIdentifier)
Dataset character set (C) (MD_Metadata > MD_DataIdentification characterSet)	**Metadata standard name** (O) (MD_Metadata metadataStandardName)
Dataset topic category (M) (MD_Metadata > MD_DataIdentification topicCategory)	**Metadata standard version** (O) (MD_Metadata metadataStandardVersion)
Spatial resolution of the dataset (O) (MD_Metadata > MD_DataIdentification spatialResolution > MD_Resolution equivalentScale or MD_Resolution distance)	**Metadata language** (C) (MD_Metadata language)
Abstract describing the dataset (M) (MD_Metadata > MD_DataIdentification abstract)	**Metadata character set** (C) (MD_Metadata characterSet)
Distribution format (O) (MD_Metadata > MD_Distribution > MD_Format name and MD_Format version)	**Metadata point of contact** (M) (MD_Metadata contact > CI_ResponsibleParty)
Additional extent information for the dataset (vertical and temporal) (O) (MD_Metadata > MD_DataIdentification extent > EX_Extent > EX_TemporalExtent or EX_VerticalExtent)	**Metadata date stamp** (M) (MD_Metadata dateStamp)

26

ISO 19115 Core Elements

For Geographic Datasets

Metadata Language
Metadata Date Stamp
Geographic Location of the Dataset
Dataset Topic Category → *Main theme of the dataset*
Dataset Language
Abstract Describing the Dataset
Dataset Character Set
Dataset Reference Date
Dataset Title
Metadata Character Set
Metadata Point of Contact

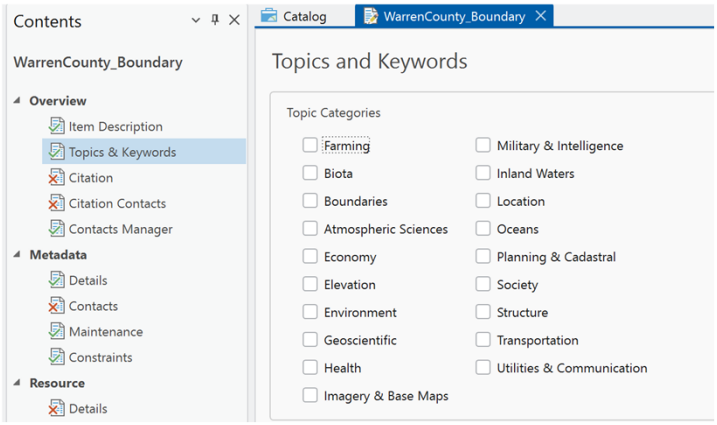


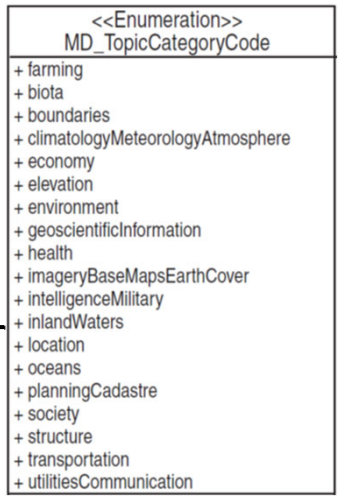
Must | Conditional Must

27

ISO 19115 Core Elements

For Geographic Datasets





For definitions, see https://wiki.esipfed.org/ISO_19115_and_19115-2_CodeList_Dictionaries#MD_TopicCategoryCode (last accessed 2 APR 2025)

28

ISO 19115 Core Elements (22 in total)

For Geographic Datasets

Spatial Resolution of the Dataset
Reference System
Metadata Standard Name
Metadata Language
Metadata Date Stamp
Geographic Location of the Dataset
Dataset Topic Category
Dataset Responsible Party
Dataset Language
Abstract Describing the Dataset
Lineage
Dataset Character Set
Dataset Reference Date
Dataset Title
Distribution Format
Metadata Character Set
Metadata File Identifier
Online Resource
Metadata Point of Contact
Metadata Standard Version
Spatial Representation Type
Vertical and Temporal Extent Information for the Dataset

Must | Conditional Must | Optional

29

Data Quality (DQ_DataQuality)

- Data quality in ISO19115 metadata
  - Position Accuracy
  - Thematic Accuracy
  - Temporal Accuracy
  - Logical Consistency
  - Completeness

The polygon data around the coastal strip are spatially correct and consistent at 1:1 million scale.

Attributes are assumed to be correct, and attributes in each of the coverages are: Record, Area (for polygon coverage), Perimeter, ID.

The example is drawn from https://www.ntlis.nt.gov.au/metadata/export_data?type=html&metadata_id=2DBCB771205206B6E040CD9B0F274EFE (last accessed 2 APR 2025)

30

Data Quality (DQ_DataQuality)

- Data quality in ISO19115 metadata
  - Position Accuracy
  - Thematic Accuracy
  - Temporal Accuracy
  - Logical Consistency
  - Completeness

One attribute is mapped per coverage. Coverages are topologically consistent. No particular tests conducted by ERIN.

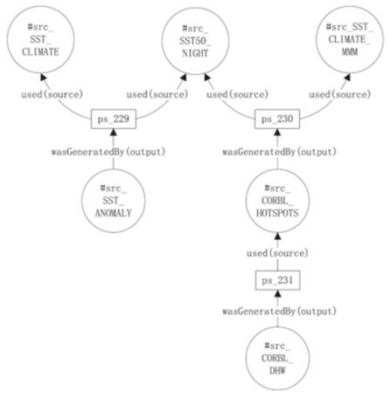
Complete for the Australian continent.

The example is drawn from https://www.ntlis.nt.gov.au/metadata/export_data?type=html&metadata_id=2DBCB771205206B6E040CD9B0F274EFE (last accessed 2 APR 2025)

31

Data Quality (DQ_DataQuality)

- Data quality in ISO19115 metadata
  - Position Accuracy, Thematic Accuracy, Temporal Accuracy, Logical Consistency, and Completeness
  - Lineage
    - Statement
    - Process step
    - Source



From: US NOAA CORBL dataset

32

Authoring Metadata

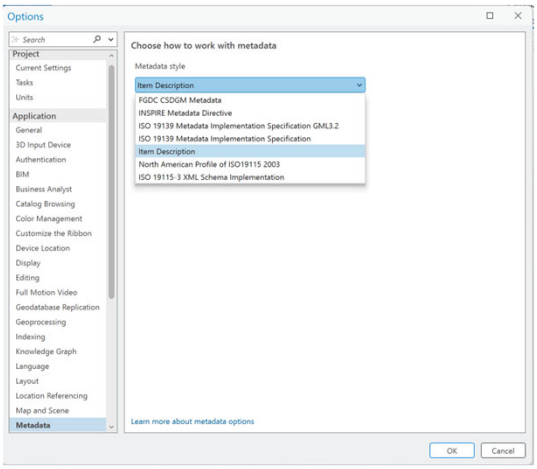
- Can be done by a text editor, but usually done in specialized software tools because
  - Complex linkages between metadata elements
  - Redundant information
  - The standard itself cannot enforce internal consistency

	Name / Role Name	Short Name	Definition	Obligation / Condition
78.	DQ_DataQuality	DataQual	quality information for the data specified by a data quality scope	Use obligation from referencing object
79.	scope	dqScope	the specific data to which the data quality information applies	M
80.	*Role name:* report	dqReport	quantitative quality information for the data specified by the scope	C / lineage not provided?
81.	*Role name:* lineage	dataLineage	non-quantitative quality information about the lineage of the data specified by the scope	C / report not provided?

33

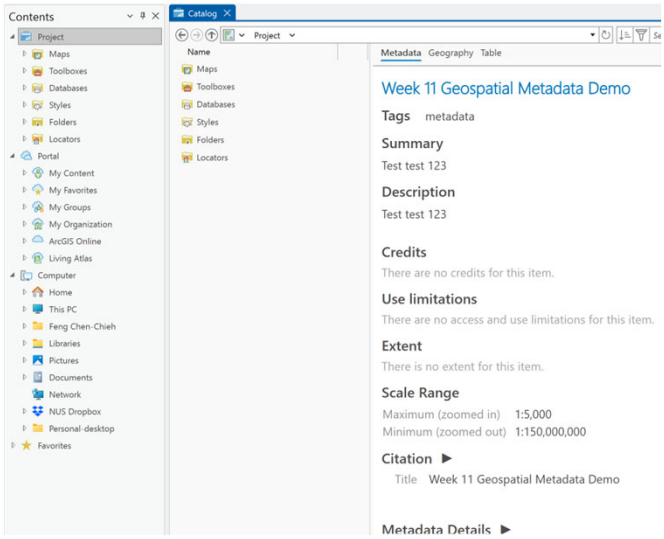
Metadata in ArcGIS Pro

- Styles
  - Item (Esri specific)
  - FGDC CSDGM
  - ISO 19115
    - North American Profile
  - INSPIRE
  - Simplify the overall metadata workflow
  - Support validating against a metadata standard



34

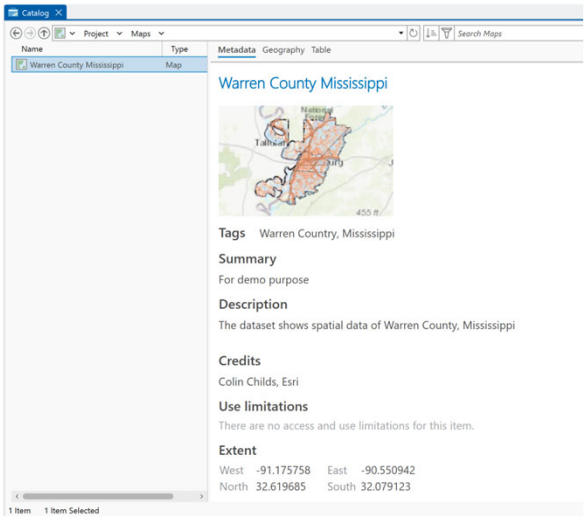
Geospatial Metadata in Catalog View



Where Are Metadata?

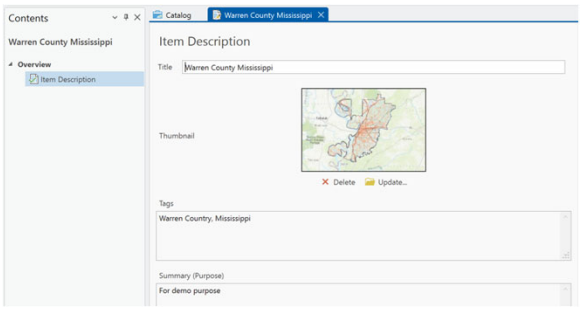
- Most ArcGIS items can have metadata
- Termed item description in ArcCatalog view
  - Project
  - Maps / Layers
  - Datasets
  - Tools
  - Files
  - Etc

See previous slide

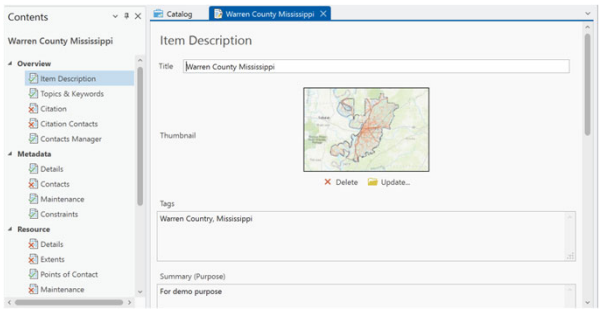


Metadata Styles in ArcGIS Pro

- Metadata presented in HTML; The chosen style determines how metadata items are rendered



Esri Item



FGDC CSDGM

37

Summary

- Metadata
  - Data about data, but more broadly, data about resources
  - Resource documentation, identification/discovery, and evaluation (esp. fitness-for-use)
  - Attribution and legal issues
- Reference for encoding metadata for geospatial data for ensuring interoperability?
  - ISO 19115-1:2014 Geographic Information – Metadata
  - Other extensions and profiles

38

Supplement Slides

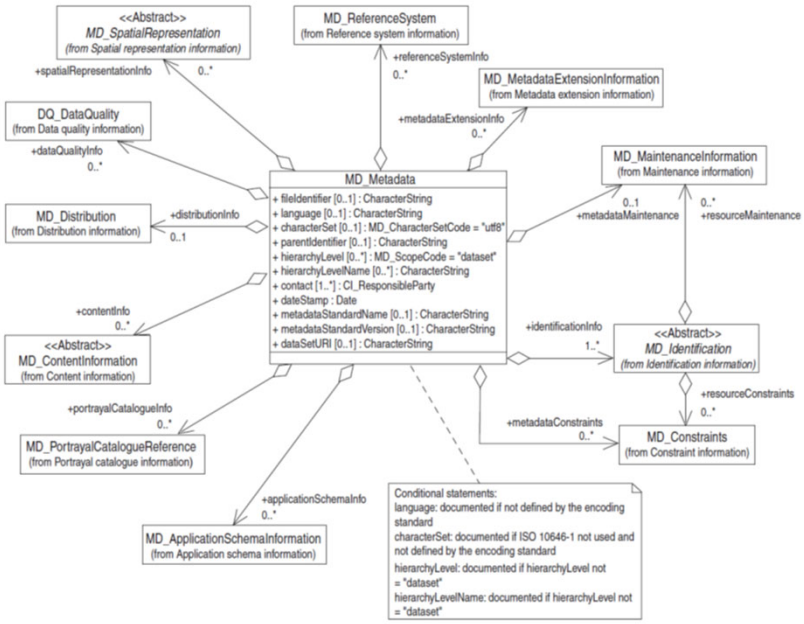
39

https://pro.arcgis.com/en/pro-app/latest/help/metadata/best-practices-for-editing-metadata.htm

- Will you follow a metadata standard or profile? If so, do you require metadata to be valid for that standard or profile?
- What legal matter must be included in the metadata?
- Will you publish items to ArcGIS Online or an ArcGIS Enterprise portal? Information required on publication may be considered optional for a metadata standard.
- Will you publish metadata to a separate metadata catalog or website? If so, does that catalog or site have requirements? Is specific content required to support searches on that catalog or site?
- What information is already available on a website that can be referenced, such as contact and download, or ordering instructions? Reference the online content instead of adding a static copy of it to the metadata.
- What content can be standardized in a metadata template?

40

Supplement slide: (UML) diagram of the top metadata elements in ISO 19115:2003



41

Supplement Slide: Structure of CSDGM Schema

- Standard is organized into 10 sections
  - Section 1: Identification Information
    - Basic information about the data set
  - Section 2: Data Quality Information
    - A general assessment of the quality of the data set
  - Section 3: Spatial Data Organization Information
    - The mechanism used to represent spatial information in the data set
  - Section 4: Spatial Reference Information
    - The description of the reference frame for, and the means to encode, coordinates in the data set
  - Section 5: Entity and Attribute Information
    - Details about the information content of the data set, including their entity types, their attributes, and the domains from which attribute values may be assigned

42

Supplement Slide: Structure of CSDGM Schema (cont.)

- Standard is organized into 10 sections
  - Section 6: Distribution Information
  - Section 7: Metadata Reference Information
  - Section 8: Citation Information
  - Section 9: Time Period Information
  - Section 10: Contact Information

43