



Metadata for OGC-based Geospatial Services

Part 1: Metadata

Geoff Mortson



Training Outcomes

- Explain the concept of metadata
- Understand the value of metadata
- Understand standards and profiles
- Learn about metadata creation and validation tools



Goals

- Learn
- Leverage
- Leave and Lead



WHAT IS METADATA?

What is metadata?

title
supplemental information
abstract



time period
author
sources
(file) size

What is metadata?



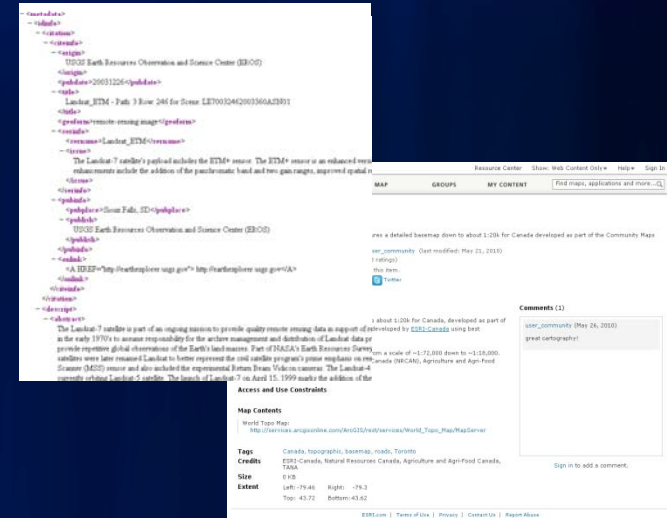
entity

attributes

| Nutrition Facts | |
|---|---------------------------|
| Serving Size ½ cup (114g) | |
| Servings Per Container 4 | |
| Amount Per Serving | |
| Calories 90 | Calories from Fat 30 |
| % Daily Value* | |
| Total Fat 3g | 5% |
| Saturated Fat 0g | 0% |
| Cholesterol 0mg | 0% |
| Sodium 300mg | 13% |
| Total Carbohydrate 13g | 4% |
| Dietary Fiber 3g | 12% |
| Sugars 3g | |
| Protein 3g | |
| Vitamin A 80% | Vitamin C 60% |
| Calcium 4% | Iron 4% |
| * Percent Daily Values are based on a 2,000 calorie diet. Your daily values may be higher or lower depending on your calorie needs: | |
| | Calories: 2,000 2,500 |
| Total Fat | Less than 65g 80g |
| Sat Fat | Less than 20g 25g |
| Cholesterol | Less than 300mg 300mg |
| Sodium | Less than 2,400mg 2,400mg |
| Total Carbohydrate | 300g 375g |
| Dietary Fiber | 25g 30g |
| Calories per gram: | |
| Fat 9 • Carbohydrate 4 • Protein 4 | |

Importance of Metadata

- Protects investment in data
- Helps users understand data
- Enables discovery
- Limits liability
- Provides evidence of prudent data stewardship
- Reduces workload associated with questions about data
- Cuts overall costs



Metadata and Sharing



Data Sharing Results in a Continued Process



STANDARDS AND PROFILES

Metadata Standards

- **Dublin Core Metadata Initiative** – support a broad range of purposes, not just geospatial resources.
- **Content Standard for Digital Geospatial Metadata (CSDGM)** – commonly referred to as “FGDC Metadata standard”
- **International Organization of Standards (ISO)** – several geospatial metadata standards exist under this organization:
 - ISO 19115 Geographic Information – Metadata
 - ISO 19119 Geographic Information – Services
 - ISO 19139 Geographic Information – Metadata - XML schema implementation

Metadata Profiles

- **Modifications to a metadata standard that has been adopted by a standards body, agency, or organization.**
- **A profile may:**
 - **reduce the overall number of metadata elements that were originally included in a standard.**
 - **further restrict optional metadata elements, making it mandatory where before it was optional; or not make mandatory elements optional.**
 - **further restrict the values allowed in a metadata element.**

Common Metadata Related Terms

- **Metadata element** - An individual piece of information in an item's metadata. A metadata element may contain a value such as a title or date, or it may contain other metadata elements.
- **Stand-alone metadata** - An XML file containing geospatial metadata.
- **Metadata catalog** - A searchable online collection of metadata describing geospatial resources. One example of a metadata catalog is the U.S. collection of maps and data geodata.gov. These catalogs can also be referred to as repositories.

DUBLIN CORE

Dublin Core

- **Simple and generic resource description**
- **Simple: 15 elements; Qualified: 18 elements**
- **It is a very minimal standard that allows user to quickly create metadata to describe a resource.**
- **Dublin Core has become one of most popular vocabularies for use with Resource Description Framework (RDF), more recently in the context of the Linked Data movement.**

Dublin Core Schema

```
- <rdf:RDF>
  - <rdf:Description rdf:about="http://purl.org/dc/elements/1.1/">
    + <dcterms:title xml:lang="en-US"></dcterms:title>
    + <rdfs:comment></rdfs:comment>
      <dcterms:publisher xml:lang="en-US">The Dublin Core Metadata Initiative</dcterms:publisher>
      <dcterms:modified>2008-01-14</dcterms:modified>
    </rdf:Description>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/title"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/creator"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/subject"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/description"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/publisher"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/contributor"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/date"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/type"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/format"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/identifier"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/source"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/language"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/relation"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/coverage"></rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/rights"></rdf:Property>
</rdf:RDF>
```

Source: <http://purl.org/dc/elements/1.1>

Dublin Core Samples

```
- <rdf:RDF>
- <rdf:Description about="http://dublincore.org">
  <dc:title xml:lang="en">California_Geology<
- <dc:description>
  USGS Web Mapping Service: Lithologic ma
  </dc:description>
  <dc:contributor>Peter N. Schweitzer</dc:co
- <dc:identifier>
  http://mrdata.usgs.gov/cgi-bin/state/ca?SERV
  </dc:identifier>
  <dc:type>WMS</dc:type>
  <dc:subject>geoscientificInformation</dc:subject>
  <dc:subject>WMS</dc:subject>
- <dc:references>
  http://mrdata.usgs.gov/cgi-bin/state/ca?SERVICE=WMS&VERSION=1.1.1&REQUEST=GetCapabilities&
  </dc:references>
  <dc:references>http://mrdata.usgs.gov/cgi-bin/state/ca?</dc:references>
- <ows:WGSS4BoundingBox>
  <ows:LowerCorner>-126.4 32</ows:LowerCorner>
  <ows:UpperCorner>-114.1 42.1</ows:UpperCorner>
  </ows:WGSS4BoundingBox>
  </rdf:Description>
</rdf:RDF>
```

```
- <rdf:RDF>
- <rdf:Description about="http://discomap.eea.europa.eu/ArcGIS/rest/services/Water/EprtrDiffuseWaterPaggr_Dyna_WGS84/MapServer">
- <dc:identifier>
  http://discomap.eea.europa.eu/ArcGIS/rest/services/Water/EprtrDiffuseWaterPaggr_Dyna_WGS84/MapServer
  </dc:identifier>
  <dc:title>facilities</dc:title>
  <dc:creator>hg</dc:creator>
- <dc:references scheme="urn:x-esri:specification:ServiceType:ArcGIS:MapServer">
  http://discomap.eea.europa.eu/ArcGIS/rest/services/Water/EprtrDiffuseWaterPaggr_Dyna_WGS84/MapServer
  </dc:references>
- <dc:references scheme="urn:x-esri:specification:ServiceType:ArcIMS:Metadata:Thumbnail">
  http://discomap.eea.europa.eu/ArcGIS/rest/services/Water/EprtrDiffuseWaterPaggr_Dyna_WGS84/MapServer/export?size=256,256&f=irnage
  </dc:references>
  <dc:subject>MapServer</dc:subject>
  <dc:subject>liveData</dc:subject>
  <dc:subject>service</dc:subject>
- <ows:WGSS4BoundingBox>
  <ows:LowerCorner>-69.10316148884661 -26.019455147057393</ows:LowerCorner>
  <ows:UpperCorner>61.786289498096856 75.8317921274247</ows:UpperCorner>
  </ows:WGSS4BoundingBox>
  </rdf:Description>
</rdf:RDF>
```

**CONTENT STANDARD FOR
DIGITAL GEOSPATIAL
METADATA (FGDC)**

Federal Geographic Data Committee

- **“FGDC tasked by executive order to develop procedures and assist in the implementation of a distributed discovery mechanism for national digital geospatial data”**
- **Content Standard for Digital Geospatial Metadata (CSDGM) is the current U.S. Federal Metadata Standard. It is commonly referred to as “FGDC Metadata”**
- **All geospatial data created as of January 1995 has to be documented using this standard.**
- **FGDC comprised of 334 elements**
 - **Of the 334, 119 elements contain others elements**

FGDC Metadata Standard in detail

- **The standard is organized into a series of elements that define the content for the metadata**
 - **Sections**
 - Main chapter of the standards
 - **Compound Elements**
 - A group of related data elements
 - **Data Elements**
 - A logically primitive item of data – ‘text you fill in’

Sections

- **The main ‘chapters’ of the standard**
- **Composed of;**
 - **Section definition**
 - **List of elements, definitions, types and values**
 - **Information about what is mandatory and repeatable**
- **There are 11 sections of the FGDC standard:**
 - **Metadata**
 - **Metadata (0)**
 - **Main Sections**
 - **Identification (1)**
 - **Data Quality (2)**
 - **Spatial Data Organization (3)**
 - **Spatial Reference (4)**
 - **Entity and Attribute (5)**
 - **Distribution (6)**
 - **Metadata Reference (7)**
 - **Supporting Sections**
 - **Citation (8)**
 - **Time Period (9)**
 - **Contact (10)**

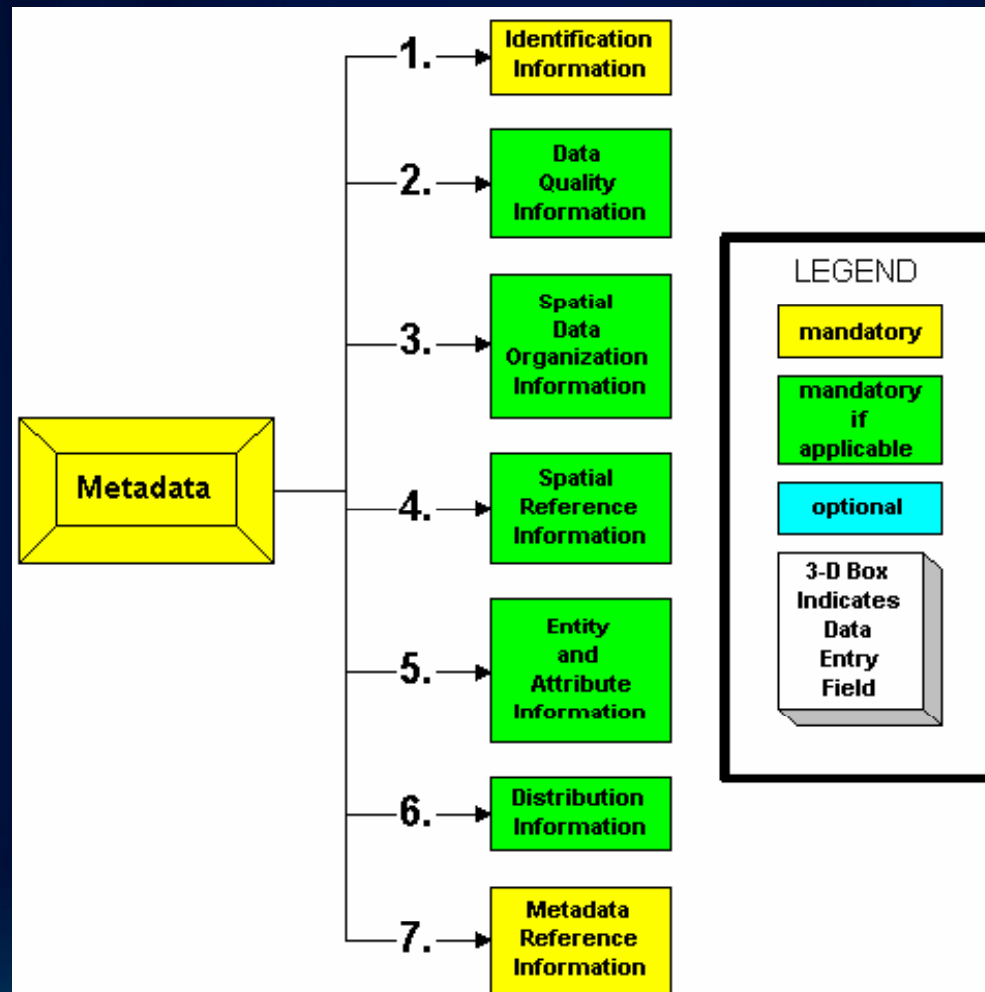
Compound element

- **A group of related data elements or other compound elements.**
 - All compound elements ultimately are comprised of data elements.
- **Form:**
 - **Compound element name -- definition.**
 - Type: compound
- **Example:**
 - **Description -- a characterization of the data set, including its intended use and limitations.**
 - Type: compound

Data element

- **A logically primitive item of data.**
 - Data elements are the things that you "fill in."
- **Form:**
 - Data element name -- definition.
 - Type: (choice of "integer", "real", "text", "date", or "time")
 - Domain: (describes valid values that can be assigned)
- **Example:**
 - Abstract -- a brief narrative summary of the data set.
 - Type: text
 - Domain: free text

FGDC Metadata Standard Resource



Profiles of FGDC

- **Biological Data Profile**
 - Changes mandatory and conditional elements
 - Adds additional elements
 - Uses specified taxonomical vocabulary
- **Metadata profile for Shoreline Data**
 - Provides a standardized set of terms and data elements to support shoreline and coastal datasets.

**INTERNATIONAL
ORGANIZATION OF
STANDARDS (ISO)**

International Organization for Standards (ISO)

- **Publisher of international standards**
- **ISO/TC 211 Geographic information/Geomatics is responsible for the ISO geographic information series of standards.**
 - **ISO 19115:2003 Geographic Information – Metadata**
 - **ISO 19119:2005 Geographic Information – Services**
 - **ISO 19139:2007 Geographic Information – Metadata – XML Schema Implementation**

ISO 19115:2003 Geographic Information – Metadata

- **Objective: “provide a structure for describing digital geographic data.”**
- **This international standard defines:**
 - **Mandatory and conditional metadata sections, metadata entities and metadata elements;**
 - **Minimum set of metadata required**
 - **Optional metadata elements**
 - **The ability to extend the metadata standards to meet specific organizational needs.**

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) |

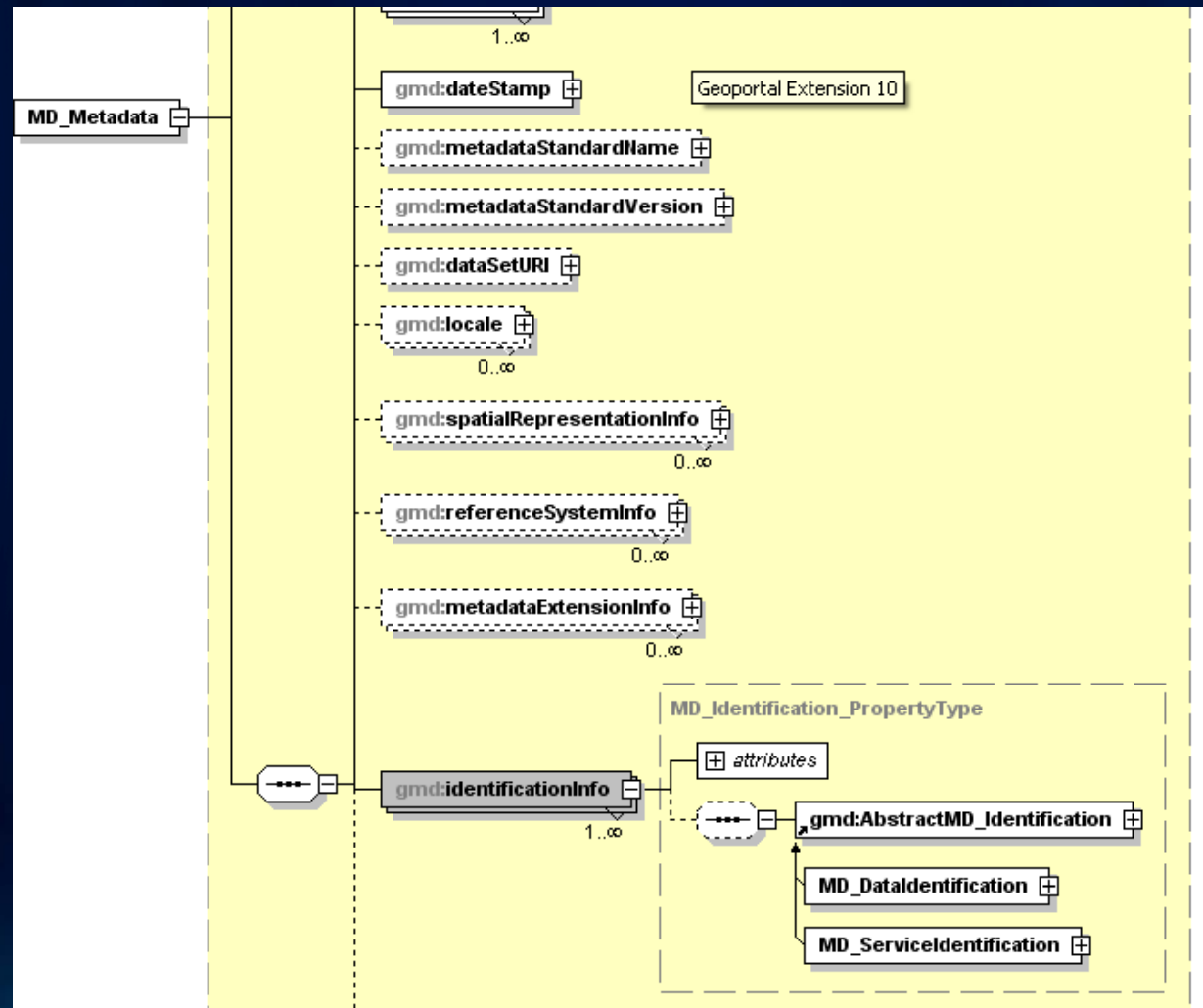
ISO 19119:2005 Geographic Information – Services

- **ISO 19119 extends the architecture reference model of ISO 19101 to provide a framework for specifying individual geographic information services.**
- **This standards was designed to support the following items:**
 - **provide an abstract framework to allow coordinated development of specific services;**
 - **enable interoperable data services through interface standardization;**
 - **support development of a service catalogue through the definition of service metadata;**
 - **allow separation of data instances and service instances;**
 - **enable use of one provider's service on another provider's data;**
 - **define an abstract framework which can be implemented in multiple ways.**

ISO 19139:2007 Geographic Information – Metadata – XML Schema Implementation

- **Encoding standards are needed to support the interchange of geographic information between systems**
 - **Provided the framework to organize and share metadata**
- **ISO/TS 19139 defines Geographic Metadata XML (gmd) encoding, an XML schema implementation derived from ISO 19115 and compliant with ISO 19118**

Graphical Representation of ISO metadata standard



Profiles of ISO

- **INSPIRE**

- **Changes mandatory and conditional elements**
- **Adds additional elements**
- **Uses specified taxonomical vocabulary**

- **North American Profile (NAP)**

- **Extends standardization across borders**
- **Fewer mandatory elements and more optional elements**
- **Extended and new elements to capture more specific information**

FGDC and ISO Example Mandatory Elements

| FGDC | ISO 19139/19115 |
|-----------------------------|-----------------------------|
| Title | Datestamp |
| Originator | Organization |
| Publication Date | Organization Role |
| Abstract | Title |
| Purpose | Publication Date |
| Progress | Abstract |
| Frequency | Data Theme |
| Time Period of Content | Extent |
| Currentness | Metadata Standard Reference |
| Extent | |
| Theme Keywords | |
| Keywords | |
| Metadata Creation Date | |
| Contact Information ... | |
| Metadata Standard Reference | |

CAPTURING METADATA

When to write Metadata

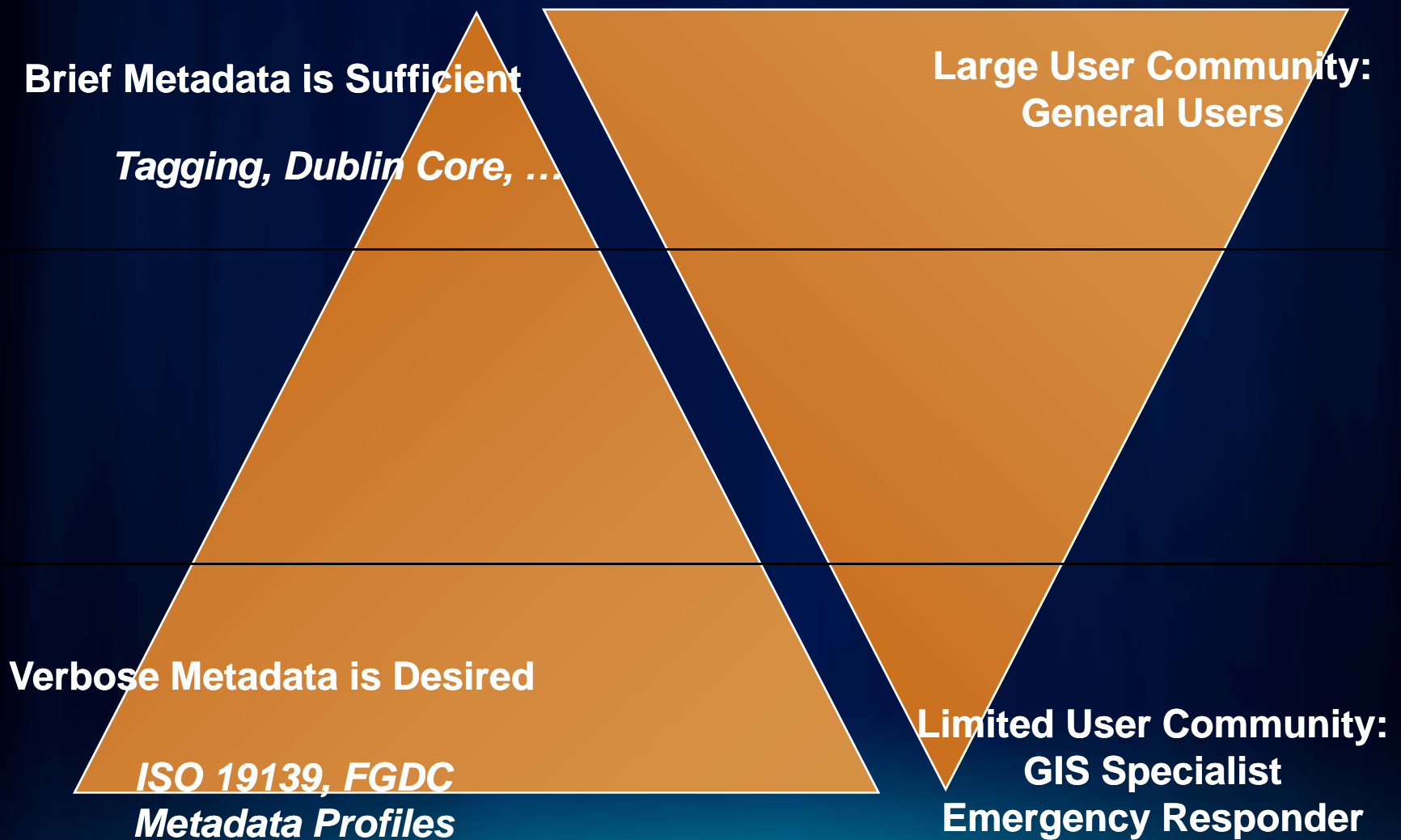
- **In most cases metadata is captured at the end of a project.**
 - **Cumbersome**
 - **Rushed**
 - **Elements can be forgotten or documented incorrectly**
- **Better to capture metadata during the process**
 - **More accurate**
 - **Integrated into the process**
 - **Greater functionality**

Common Metadata Obstacles

- **Metadata standards are too detailed**
- **Metadata production requires time and resources**
- **There are few tangible benefits and incentives to produce metadata.**



Take Another Look at Metadata



Improve Metadata Capture

- **Develop organizational processes and procedures**
- **Understand metadata standards fit**
- **Map metadata fields to workflows**
- **Utilize metadata creation and validation tools**

Process and Procedures

- **Assign responsibilities for metadata creation**
- **Assign priorities to metadata production**
- **Establish guidelines**
- **Implement**
- **Educate**

Metadata Standard and the Organization

- **Understand the standard and what is required for your organization**
 - **Which fields are mandatory?**
 - **Are optional fields really mandatory for your organization?**
- **What are other organizations using as their standard/profile?**
- **What fields in the standard or profile would have the same values?**

Map Metadata to Workflow

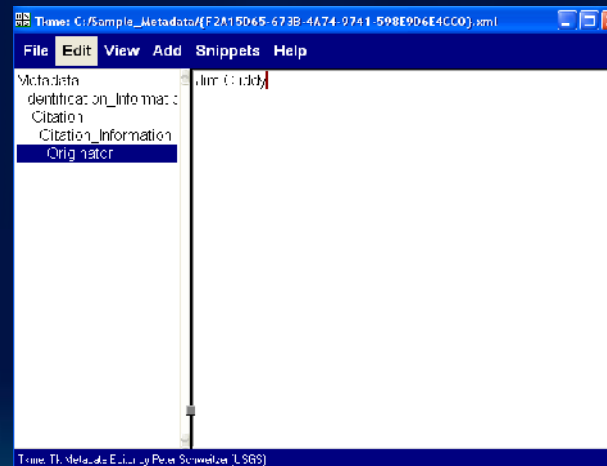
- **As the geospatial resource is being created capture metadata attributes**
 - **This saves time and energy at the end**
- **So in the initial phase you could capture title, abstract, point of contact, purpose, etc.**
- **The data processing could capture Data Quality, geoprocessing steps, etc.**
- **Data analysis phase could capture attribute accuracy, metadata reference, etc.**

Metadata Tools

- **Creation**
 - Xtme/Tkme
 - ArcCatalog
 - EPA Metadata Editor
 - ArcGIS Server Geoportal Extension
- **Validation**
 - Metadata parser (mp)
 - Geospatial Metadata Validation Service
 - ArcGIS Server Geoportal Extension
 - PLTS Metadata Check
- **Transform tools**
 - Keep up to date on www.fgdc.gov

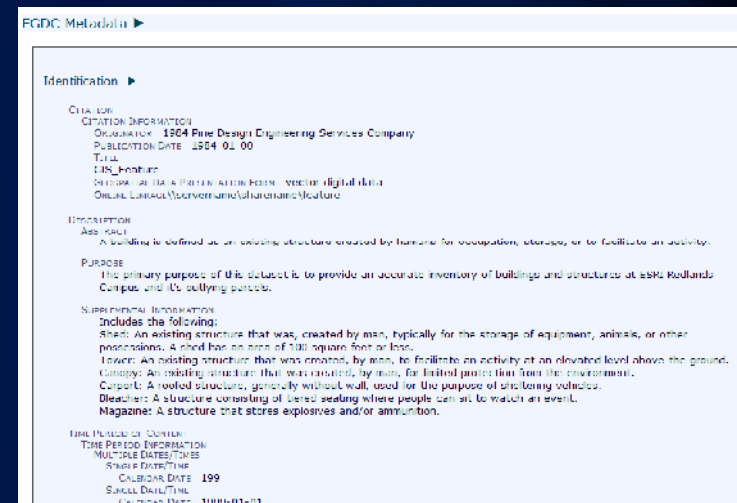
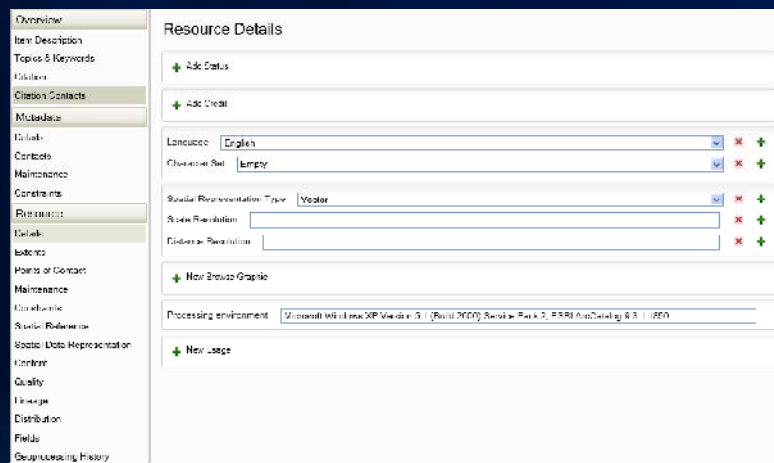
Tkme

- Windows based tool
- Also available as a unix based tool – ‘xtme’
- Developed and actively maintained by Peter Schweitzer of the USGS Geology Discipline



ArcCatalog

- ArcCatalog helps GIS users by providing an integrated and unified view of all the data files, databases, and ArcGIS documents available to ArcGIS users.



EPA Metadata Editor

- Extension to ArcCatalog 9.2
- Allows user to create and edit metadata that meets the EPA Geospatial Metadata Technical Specification

The screenshot shows the EPA Metadata Editor application window. The title bar reads "EPA Metadata Editor". The interface is divided into several sections:

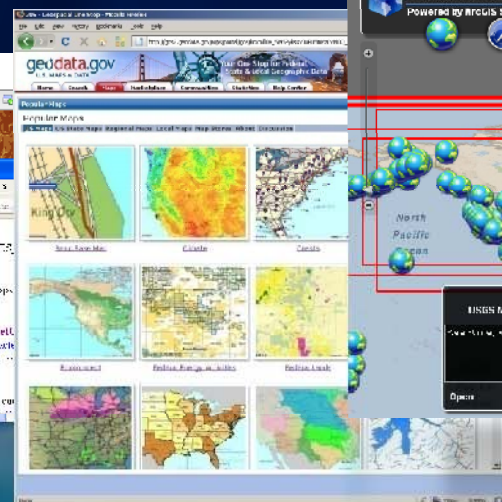
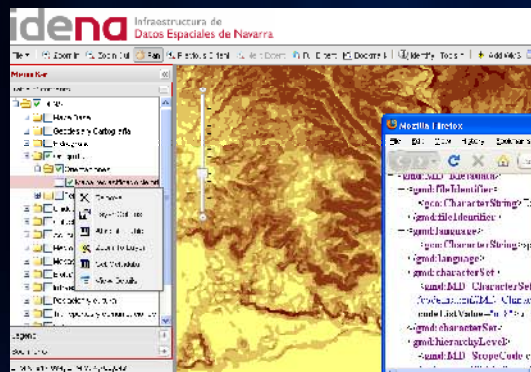
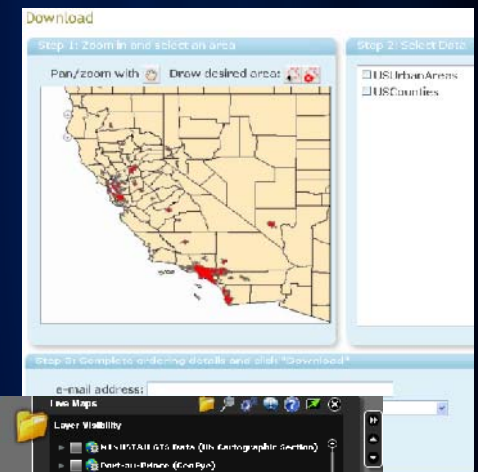
- Citation:** Fields for Origin (United States Environmental Protection Agency - Office of Environment), Title (EPA Metadata Editor Version 3.0), Publisher (U.S. Environmental Protection Agency, Headquarters), and Published at (Washington, DC).
- Online Linkage:** Fields for Primary Linkage (http://www.epa.gov/geospatial/eme.html) and Secondary Linkage.
- Description:** Fields for Abstract and Purpose.
- Time Period:** Fields for Date of Data Set, Range of dates, Progress of data, Data currency, and Update frequency.
- Bounding Box:** Fields for North, South, East, and West coordinates.
- Keywords:** A list of keywords including "boundaries", "climatology/Meteorology/Atmosphere", "economy", "planning", "environment", "farming", "genetics/Biotechnology", "health", "imagery/Basemap/EarthCover", and "land/Water".
- Default Constraints:** Fields for Access, Use, and Security.
- Contact:** Fields for Primary Person and Primary Organization.

At the bottom, there is a legend for field requirements: *YELLOW - mandatory, GREEN - mandatory if applicable, BLUE - optional. A note says "Click on text to link to element description". Buttons for "Save", "Save & Close", "Cancel", and "Help" are visible.

<http://www.epa.gov/geospatial/eme.html>

ArcGIS Server Geoportal Extension

- Customizable web access point
- Search and view geospatial resources
- Manage, publish, and store metadata
- View live map services
- Download data



Metadata Parser (MP)

- **Metadata for the MP**

<http://geology.usgs.gov/tools/metadata/tools/doc/metadata/mp.html>

- **“MP is a program for validating the syntactical structure of formal metadata, testing the structure against the Content Standards for Digital Geospatial Metadata devised by the Federal Geographic Data Committee (FGDC).”**

METADATA'S ROLE IN SHARING

Using Resources from Multiple Organizations

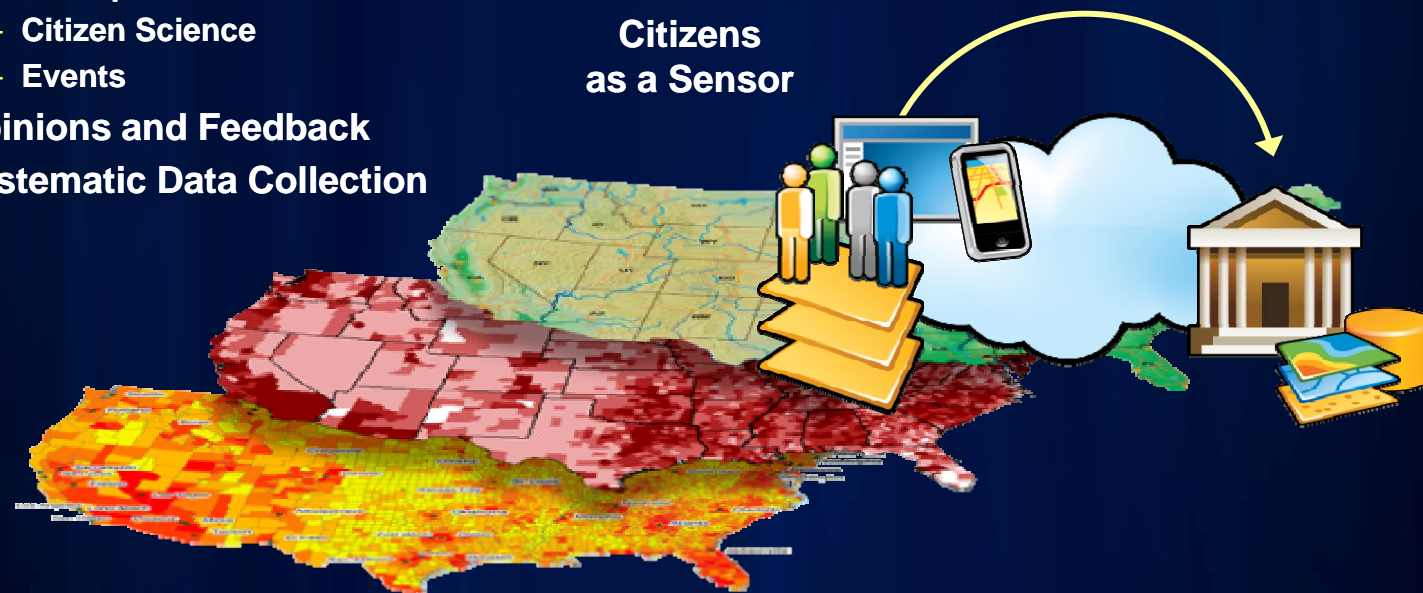
- Metadata helps describe the data to help determine fitness of use.
- Confidence in authoritative data



Metadata and Crowd Sourcing

Engaging Citizens in Spatial Data Collection and Civic Action

- **Observations**
 - Conditions (i.e. 311 . . .)
 - Delinquent Behavior
 - Citizen Science
 - Events
- **Opinions and Feedback**
- **Systematic Data Collection**

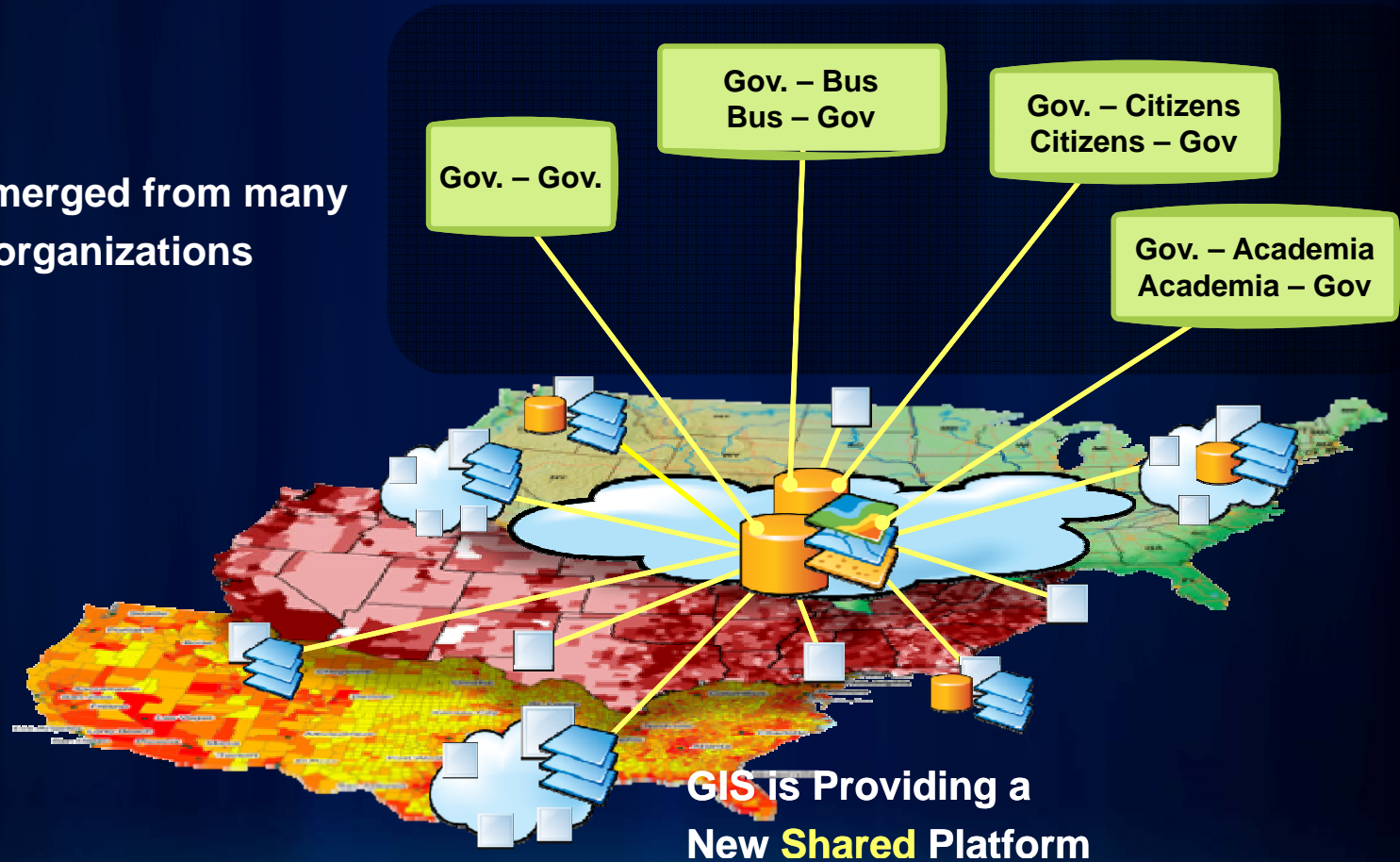


Empowering Everyone to Participate and Contribute

Participatory GIS Is Emerging

Location Enabled Applications

Data is being merged from many different organizations

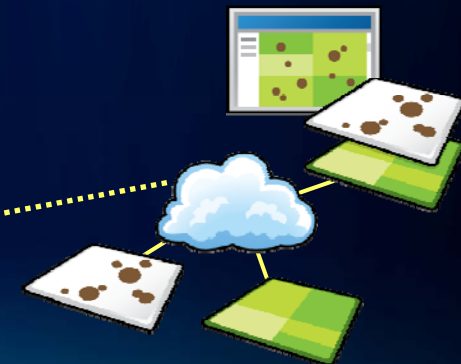


Highlighting the Importance of Metadata

- **Emergency response (hours/days):**
 - Effective delivery of emergency services and supplies
 - All sources of information may help
- **Reconstruction (weeks/months):**
 - Rebuilding of essential services and infrastructures
 - Continuously evolving decentralized geographic data sources
- **Development (months/years):**
 - Planning, development of services to support the society
 - From an unstructured societal contribution of observations to a managed operational framework to support



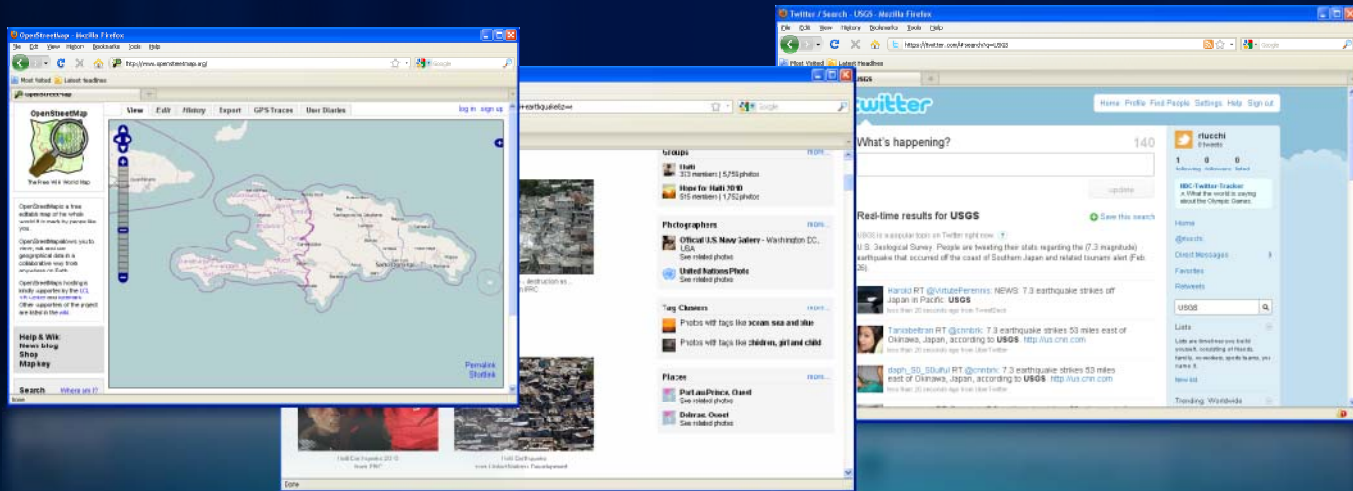
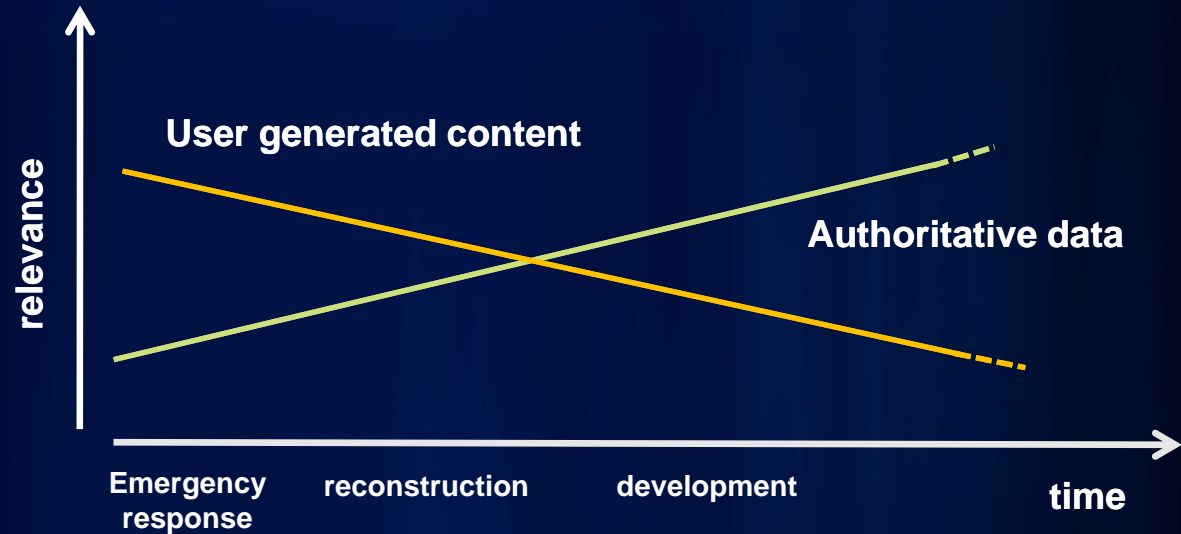
**Find
Use
Combine
Collaborate**



Where are the data and services?

Data sources:

- Authoritative
 - Agencies
- Social networks:
 - Users as sensors



Sources and Further Reading

- FGDC publication list website - <http://www.fgdc.gov/metadata/metadata-publications-list>
 - “Content Standard for Digital Geospatial Metadata Workbook”
 - “Institutionalize Metadata Before it Institutionalizes You”
- ISOTC/211 website - <http://www.isotc211.org/>
 - “ISO_TC_211_Standards_Guide”

EXERCISE: PART 1

Metadata Exercise

- Create FGDC metadata using the ArcGIS Server Geoportal Extension web-based Metadata Editor. <http://gptogc.esri.com>
- If possible the metadata should be describing a real geospatial resource.
- If you can't use real data use this map service as a resource:
 - <http://www6.city.newport-beach.ca.us/ArcGIS/rest/services/Base/MapServer>



Exercise Questions

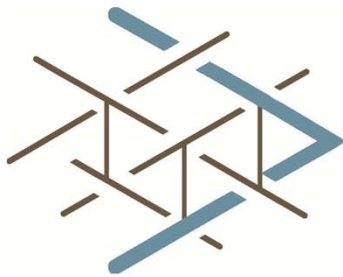
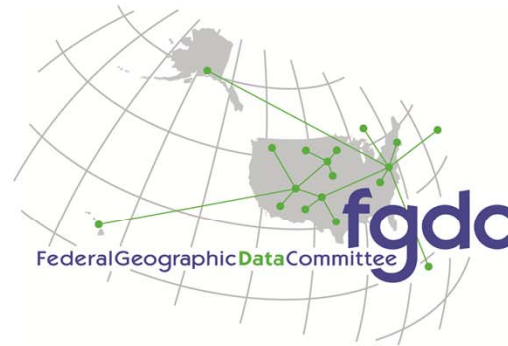
- What is the data element tag that holds the url for the resource?
- What is the last compound element tag prior to the following data element tags?
 - `<westbc>`, `<eastbc>`, `<northbc>`,
`<southbc>`
- What is the tag of the Theme Topics?
 - Where do those theme topics come from?
- What is the last section tag in the xml document?



QUESTIONS?



ESRI



NARC

Building Regional Communities

Acknowledgement

The project described in this publication was supported by Cooperative Agreement Number G10AC00236 from the United States Geological Survey. Its contents are solely the responsibility of the authors and does not necessarily represent the official views of the USGS.