

# “XML”

Master di II livello "Sistemi informativi geografici per il monitoraggio  
e la gestione del territorio"

Claudio Rocchini

claudio@rockini.name

Istituto Geografico Militare

2013

# Introduzione a XML

- eXtensible Markup Language = Linguaggio a marche estensibile.
- Sistema di codifica facile da leggere sia dal computer (correttezza formale) sia dagli esseri umani (file di testo semplice).
- Ci interessa perché molto utilizzato in ambiente GIS.

# Caratteristiche di XML

- Specifica multiuso per creare nuovi *linguaggi a marche*.
- Non esiste un formato XML, ma una serie di sue sotto-specifiche.
- E' la struttura base di molti linguaggi di definizione, es: *SVG* (disegni vettoriali), *HTML* (siti web), *GML* (dati geografici)
- Es. i Metadati Italiani (CNIPA) hanno una struttura XML

# Struttura di una marca

- Un file XML è un file di testo (si può vedere col blocco notes), con una particolare struttura

- Esempio di marca:

```
<libro>Questo è un libro...</libro>
```

- In generale, un file XML è un insieme di marche

```
<libro>Questo è il primo</libro>
```

```
<libro>Questo è il secondo</libro>
```

# Dettagli sulla marca

- La marca inizia con il suo *nome* fra parentesi angolate (es. `<marca>` )
- Continua con il suo contenuto.
- Finisce con il tag di fine marca (es. `</marca>` ).
- La marca è identificata attraverso il suo nome.

# Intestazione del File XML

- Oltre a contenere le marche, il file XML deve iniziare con una intestazione che dichiara la versione del formato utilizzata e la codifica dei caratteri (encoding), es:

```
<?xml version="1.0" encoding="UTF-8" ?>
```

- Come in molti linguaggi informatici, è possibile inserire commenti, cioè parti descrittive che non fanno parte del dato:

```
<!-- Questo è un commento. -->
```

# Un inciso: codifica carattere

- In passato i file di testo (.txt) erano *tutti uguali*.
- La codifica utilizzata era *ASCII*:  
corrispondente ai caratteri presenti sulla tastiera americana più le lettere accentate.
- Al giorno d'oggi il mondo è *globalizzato*:  
necessità di rappresentare testi cinesi,  
giapponesi, arabi, russi, ma anche caratteri  
norvegesi: nascita di varie codifiche carattere.

# Attributi

- Oltre al corpo dell'oggetto, la marca iniziale può contenere una serie di attributi ed il loro relativo valore:

```
<elem nome_attributo="valore_attributo">  
Contenuto Elemento </elem>
```

- Esempio reale in HTML titolo principale con colore associato:

```
<h1 color="red">Capitolo Primo</h1>
```



# Nidificazione delle Marche

- Il corpo di una marca può contenere al suo interno sotto-marche, e così via, es:

```
<carta>
```

```
  <tematismo nome="trasporti">
```

```
    <classe>Strade</classe>
```

```
    <classe>Ponti</classe>
```

```
  </tematismo>
```

```
</carta>
```

- L'*indentazione* della scrittura, anche se non obbligatoria, rende più chiara la struttura gerarchica

# Un esempio di fantasia

```
<?xml version="1.0" encoding="UTF-8" ?>
<!-- Esempio culinario -->
<ricetta nome="pane" tempo_cott="3 ore">
  <title>Pane</title>
  <ingrediente quant="8" unit="dL">Farina</ingrediente>
  <ingrediente quant="4" unit="dL"
  stato="fredda">Acqua</ingrediente>
  ...
  <istruzioni>
    <passo>Mescola gli ingredienti</passo>
    ...
    <passo>Cuoci a 180 gradi per 3 ore.</passo>
  </istruzioni>
</ricetta>
```

# Marche vuote

- Spesso le marche finali sono vuote e non hanno corpo, es:

```
<marca_vuota></marca_vuota>
```

- In questo caso la scrittura può essere contratta nel modo:

```
<marca_vuota/>
```

- Esempio HTML:

```
<input type="button" value="Premi"/>
```

# Namespace 1

- In file molto complessi, ci potrebbero essere collisioni nei nomi delle marche (cioè nomi uguali).
- Si introduce il concetto di *namespace*, ovvero il *cognome* di una marca; il cognome precede il nome separato da due-punti.
- Utilizzo: `<cognome : nome>...</cognome : nome>`

# Namespace 2

- I namespace vanno dichiarati prima dell'uso come attributo e devono essere associati ad un indirizzo *web unico*, per evitare ambiguità.
- Es per definire il conome gmd:  
`xmlns:gmd=http://www.isotc211.org/2005/gmd`
- xmlns sta per XML NameSpace

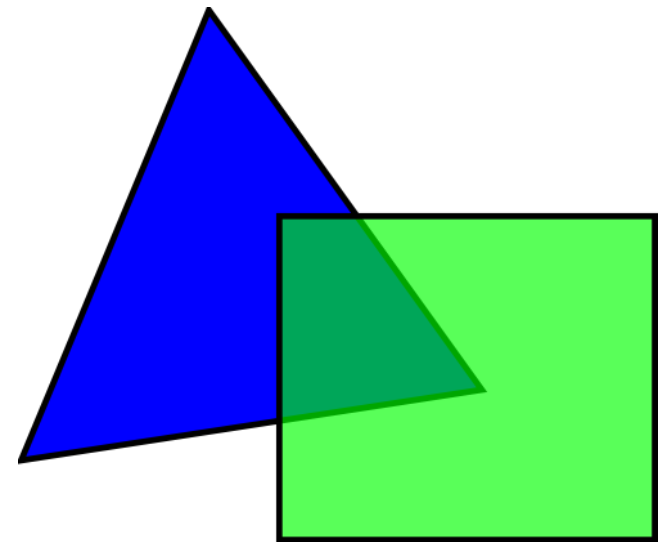
# Namespace: Esempio Metadati

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<rndt:Request
  xmlns:rndt="http://www.cnipa.gov.it/RNDT/ITgmd"
  xmlns:gmd="http://www.isotc211.org/2005/gmd"
  xmlns:gml="http://www.opengis.net/gml/3.2"
  ...
  ...
  <gmd:cornerPoints>
    <gml:Point gml:id="V_DS_1">
      <!--<gml:name>Vertice alto a sinistra</gml:name>-->
      <gml:coordinates cs=","
        decimal=".">47.09114,9.21459</gml:coordinates>
    </gml:Point>
  </gmd:cornerPoints>
```



# Esempio SVG

```
<?xml version="1.0" encoding="UTF-8"
  standalone="no"?>
<!-- Created with Inkscape -->
<svg width="800" height="800" id="svg2"
  version="1.0" >
<g>
  <path
  style="fill:#0000ff;stroke:#000000;
  stroke-width:6;"
  d="M 248,145 L 65,582 L 514,514
    L 248,145 z "/>
  <rect style="fill:#00ff00; fill-
  opacity:0.6;stroke:#000000;
  stroke-width:6;"
  width="365"
  height="314"
  x="317" y="345" />
</g>
</svg>
```



# Esempio HTML

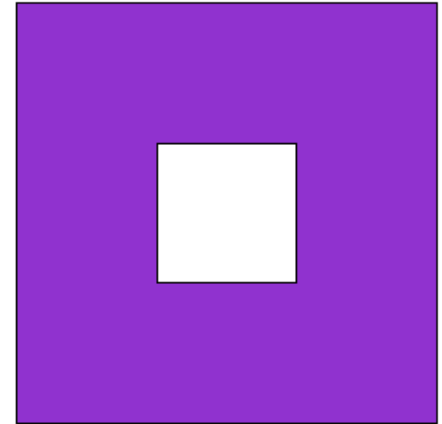
```
<!DOCTYPE HTML PUBLIC "-//
  //W3C//DTD XHTML 1.0
  Transitional//EN">
<html>
  <head>
    <title>Master 2009</title>
  </head>
  <body bgcolor="#E0FFE0">
    <h1>Capitolo 1</h1>
    <p>
      Questo è il primo paragrafo
      <a href="indirizzo">Questo è
      un link</a>
    </p>
  </body>
</html>
```





# Esempio GML (parte)

```
... <gml:Polygon srsName="EPSG:4623">
  <gml:outerBoundaryIs>
    <gml:LinearRing>
      <gml:coordinates>10,43 13,43 13,46 10,46
      10,43</gml:coordinates>
    </gml:LinearRing>
  </gml:outerBoundaryIs>
  <gml:innerBoundaryIs>
    <gml:LinearRing>
      <gml:coordinates>11,44 12,44 12,45 11,45
      11,44</gml:coordinates>
    </gml:LinearRing>
  </gml:innerBoundaryIs>
</gml:Polygon> ...
```

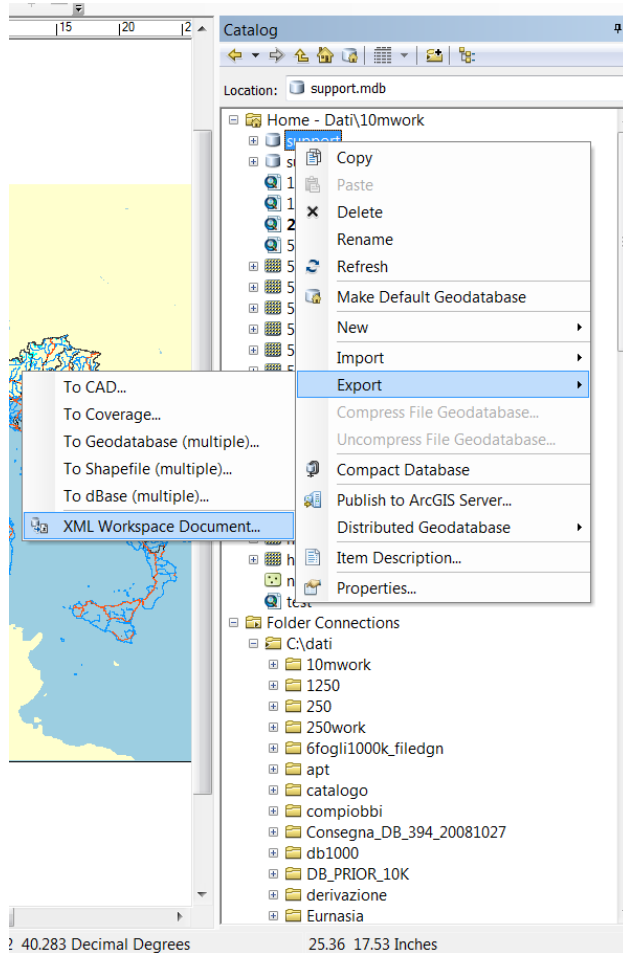


# Esempio: KML (Google Earth)



```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2" xmlns:gx="http://www.google.com/kml/ext/2.2" xmlns:xi="http://www.w3.org/2001/XMLSchema-instance">
  <Document>
    <name>Campo Gallina.kml</name>
    <Style id="sh_ylw-pushpin0">
      <IconStyle>
        <scale>1.3</scale>
        <Icon>
          <href>http://maps.google.com/mapfiles/kml/pushpin/ylw-pushpin.png</href>
        </Icon>
        <hotSpot x="20" y="2" xunits="pixels" yunits="pixels"/>
      </IconStyle>
    </Style>
    <Placemark>
      <name>Campo Gallina</name>
      <LookAt>
        <longitude>12.83333077593442</longitude>
        <latitude>42.83333496258149</latitude>
        <altitude>0</altitude>
        <heading>-1.738693981054448e-006</heading>
      </LookAt>
    </Placemark>
  </Document>
</kml>
```

# Esempio: ArcGIS (Schema DB)



```
<?xml version="1.0"?>
<esri:Workspace xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:esri="http://www.esri.com/schemas/ArcGIS/10.0">
  <WorkspaceDefinition xsi:type="esri:WorkspaceDefinition">
    <WorkspaceType>esriLocalDatabaseWorkspace</WorkspaceType>
    <Version/>
  </WorkspaceDefinition>
  <Domains xsi:type="esri:ArrayOfDomain">
    <Domain xsi:type="esri:CodedValueDomain">
      <DomainName>HorizontalAlignment</DomainName>
      <FieldType>esriFieldTypeSmallInteger</FieldType>
      <MergePolicy>esriMPTDefaultValue</MergePolicy>
      <SplitPolicy>esriSPTDuplicate</SplitPolicy>
      <Description>Valid horizontal symbol alignment values.</Description>
      <Owner/>
      <CodedValues xsi:type="esri:ArrayOfCodedValue">
        <CodedValue xsi:type="esri:CodedValue">
          <Name>Left</Name>
          <Code xsi:type="xs:short">0</Code>
        </CodedValue>
        <CodedValue xsi:type="esri:CodedValue">
          <Name>Center</Name>
          <Code xsi:type="xs:short">1</Code>
        </CodedValue>
        <CodedValue xsi:type="esri:CodedValue">
          <Name>Right</Name>
          <Code xsi:type="xs:short">2</Code>
        </CodedValue>
        <CodedValue xsi:type="esri:CodedValue">
          <Name>Full</Name>
          <Code xsi:type="xs:short">3</Code>
        </CodedValue>
      </CodedValues>
    </Domain>
    <Domain xsi:type="esri:CodedValueDomain">
      <DomainName>VerticalAlignment</DomainName>
      <FieldType>esriFieldTypeSmallInteger</FieldType>
      <MergePolicy>esriMPTDefaultValue</MergePolicy>
      <SplitPolicy>esriSPTDuplicate</SplitPolicy>
      <Description>Valid symbol vertical alignment values.</Description>
      <Owner/>
      <CodedValues xsi:type="esri:ArrayOfCodedValue">
        <CodedValue xsi:type="esri:CodedValue">
          <Name>Top</Name>
          <Code xsi:type="xs:short">0</Code>
        </CodedValue>
        <CodedValue xsi:type="esri:CodedValue">
          <Name>Center</Name>
          <Code xsi:type="xs:short">1</Code>
        </CodedValue>
      </CodedValues>
    </Domain>
  </Domains>
</esri:Workspace>
```

# Note su XML-Schema (XSD)

- XML-Schema (file con estensione XSD) è una meta-specifica XML, vale a dire una specifica di file XML che descrivono in modo dettagliato un'altra specifica XML.
- Con XML-Schema è possibile definire e distribuire sotto-specifiche XML in modo esatto: esistono software automatici di validazione che leggono la specifica ed un dato da controllare.

# Esempio XSD: ISOTC211

- ISOTC211 è una specifica su come si scrivono metadati geografici (es. utilizzata in ambito europeo)
- I metadati veri e propri sono in XML
- Esiste un XML-Schema (una serie di file XSD), distribuito pubblicamente, che definisce la specifica del metadato.
- Chiunque può controllarne la correttezza.

# ISOTC211: XML-Schema

The screenshot displays the Liquid XML Studio interface. The main window shows a graphical representation of an XML Schema (XSD) for a bookstore. The schema is structured as follows:

- bookstoreType** (Complex Type):
  - Contains one or more (**0..\***) **book** elements.
  - book** (Element):
    - Complex Type **bookType**:
      - Contains **title** (Element, type **string**), **author** (Element, type **authorName**), and **genre** (Element, type **string**).
      - Attributes: **price** (type **double**), **publicationdate** (type **date**), and **ISBN** (type **string**).

- authorName** (Complex Type):
- Contains **first-name** (Element, type **string**) and **last-name** (Element, type **string**).

The Properties window on the right shows the details for the **title** element:

Properties	
<b>Cardinality</b>	
Min Occurs	
Max Occurs	
<b>Content</b>	
Name	<b>title</b>
Simple Type Content	<b>Restriction</b>
Type	<b>string</b>
<b>Facets</b>	
Enumerations	<empty>
Length	
Min Length	
Max Length	50
Pattern	
Whitespace	None
<b>Properties</b>	
Abstract	False
Block	
Default	
Final	
Fixed	
Form	None
Id	
NotNullable	False

The Id property is defined as: "A unique ID identifying the item".

The Documentation window shows the following text: "The title of the book. Max 50 characters."

The XSD code in the main editor is as follows:

```
11 <xsd:element name="title">
12 <xsd:annotation>
13 <xsd:documentation>The title of the book.
14 Max 50 characters.</xsd:documentation>
15 </xsd:annotation>
16 <xsd:simpleType>
17 <xsd:restriction base="xsd:string">
18 <xsd:maxLength value="50" />
19 </xsd:restriction>
20 </xsd:simpleType>
21 </xsd:element>
22 <xsd:element name="author" type="authorName" />
23 <xsd:element minOccurs="0" name="genre" type="xsd:string" />
24 </xsd:sequence>
```

# ISOTC211: XML-Schema (estratto)

```
<!-- ===== -->
<xs:complexType name="EX_VerticalExtent_Type">
  <xs:annotation>
    <xs:documentation>Vertical domain of dataset</xs:documentation>
  </xs:annotation>
  <xs:complexContent>
    <xs:extension base="gco:AbstractObject_Type">
      <xs:sequence>
        <xs:element name="minimumValue" type="gco:Real_PropertyType"/>
        <xs:element name="maximumValue" type="gco:Real_PropertyType"/>
        <xs:element name="verticalCRS" type="gsr:SC_CRS_PropertyType"/>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
<!-- ..... -->
```

# ISOTC211: Scheda Esempio

```
<?xml version="1.0" encoding="UTF-8"?>
<gmd:MD_Metadata xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:gml="http://www.opengis.net/gml"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:gco="http://www.isotc211.org/2005/gco"
xmlns:gmd="http://www.isotc211.org/2005/gmd" xsi:schemaLocation="http://www.isotc211.org/2005/gmd
http://schemas.opengis.net/iso/19139/20060504/gmd/gmd.xsd">
- <gmd:fileIdentifier>
  <gco:CharacterString>5bafd8e</gco:CharacterString>
</gmd:fileIdentifier>
- <gmd:language>
  <gmd:LanguageCode codeListValue="ita"
codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodellists.xml#Li
</gmd:language>
- <gmd:characterSet>
  <gmd:MD_CharacterSetCode codeListValue="MD_CharacterSetCode_utf8"
codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodellists.xml#MD_CharacterSetCode"
codeSpace="ISOTC211/19115">MD_CharacterSetCode_utf8</gmd:MD_CharacterSetCode>
</gmd:characterSet>
- <gmd:hierarchyLevel>
  <gmd:MD_ScopeCode codeListValue="dataset"
codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodellists.xml#M
</gmd:hierarchyLevel>
- <gmd:contact>
  - <gmd:CI_ResponsibleParty>
    - <gmd:organisationName>
      <gco:CharacterString>Istituto Geografico Militare</gco:CharacterString>
</gmd:organisationName>
    - <gmd:contactInfo>
      - <gmd:CI_Contact>
        - <gmd:address>
          - <gmd:CI_Address>
            - <gmd:electronicMailAddress>
              <gco:CharacterString>info@geomil.difesa.it</gco:CharacterString>
</gmd:electronicMailAddress>
</gmd:CI_Address>
</gmd:address>
</gmd:CI_Contact>
</gmd:contactInfo>
        - <gmd:role>
          <gmd:CI_RoleCode codeListValue="pointOfContact"
codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodellists
</gmd:role>
</gmd:CI_ResponsibleParty>
      </gmd:contact>
    - <gmd:dateStamp>
      <gco:Date>2011-03-31</gco:Date>
</gmd:dateStamp>
  - <gmd:metadataStandardName>
```





# Bibliografie e Riferimenti

- [www.w3.org/MarkUp](http://www.w3.org/MarkUp) : gruppo sul Markup
- [www.w3.org/XML](http://www.w3.org/XML) : XML
- [www.w3.org/TR/html4](http://www.w3.org/TR/html4) : HTML 4
- [www.w3.org/Graphics/SVG](http://www.w3.org/Graphics/SVG) : SVG
- [www.opengeospatial.org/standards/gml](http://www.opengeospatial.org/standards/gml) : GML