ETL tools as a support for
harmonization of spatial data
according to INSPIRE

90% of the data in the world today has been created in the last two years alone

Every day, we create 2.5 quintillion bytes of data

$$10^{18}$$

Exabyte

1 000 000 000 gigabytes
1 000 000 terabytes

Spatial ETL tools

FME, GeoKettle, Sextante, SDI Talend Open Studio
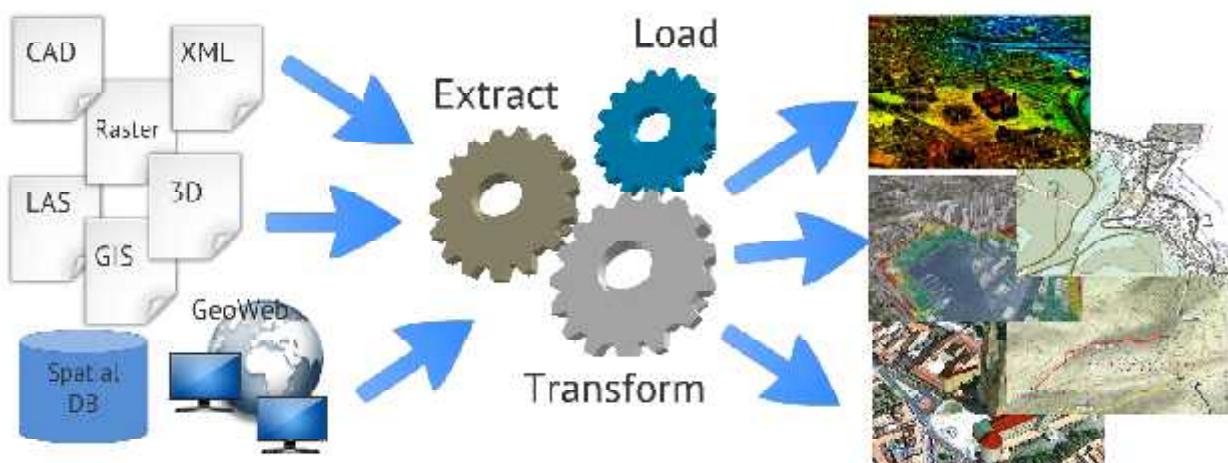
Live Demonstration

Data inspection
Transformation
Format conversion
Schema mapping
Cataloging
Quality control
Spatial analysis
XML creation
HTTP request
Command line execution

1. Data type?

format
size
level names
bands
histogram
CRS

Raster

TK25

Map1.tif   Map2   Map3   Map4

DOP5

DOP1.tif   DOP2.tif   DOP3.tif   DOP4.tif

Vector

Topographic Data

Archiving

1. Data type?

format
size
level names
bands
histogram
CRS

Raster

TK25

DOP5

Vector

Topographic Data

CS/W preparation
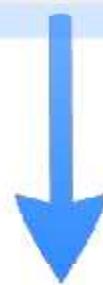
Schema mapping

## Schema mapping

OSM topographic data → Rules → INSPIRE application schema

Error
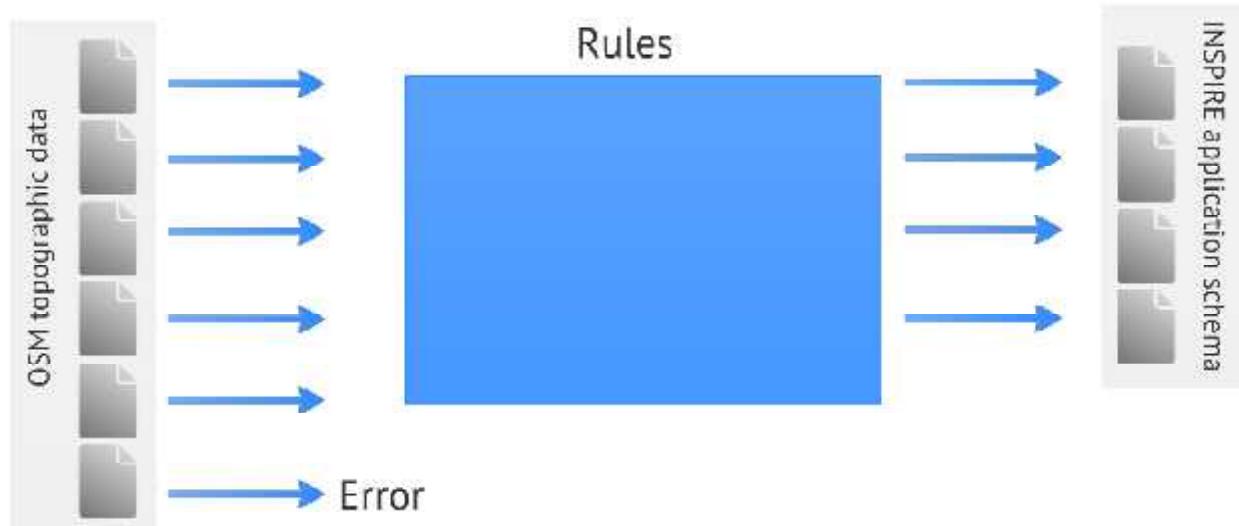
Rules
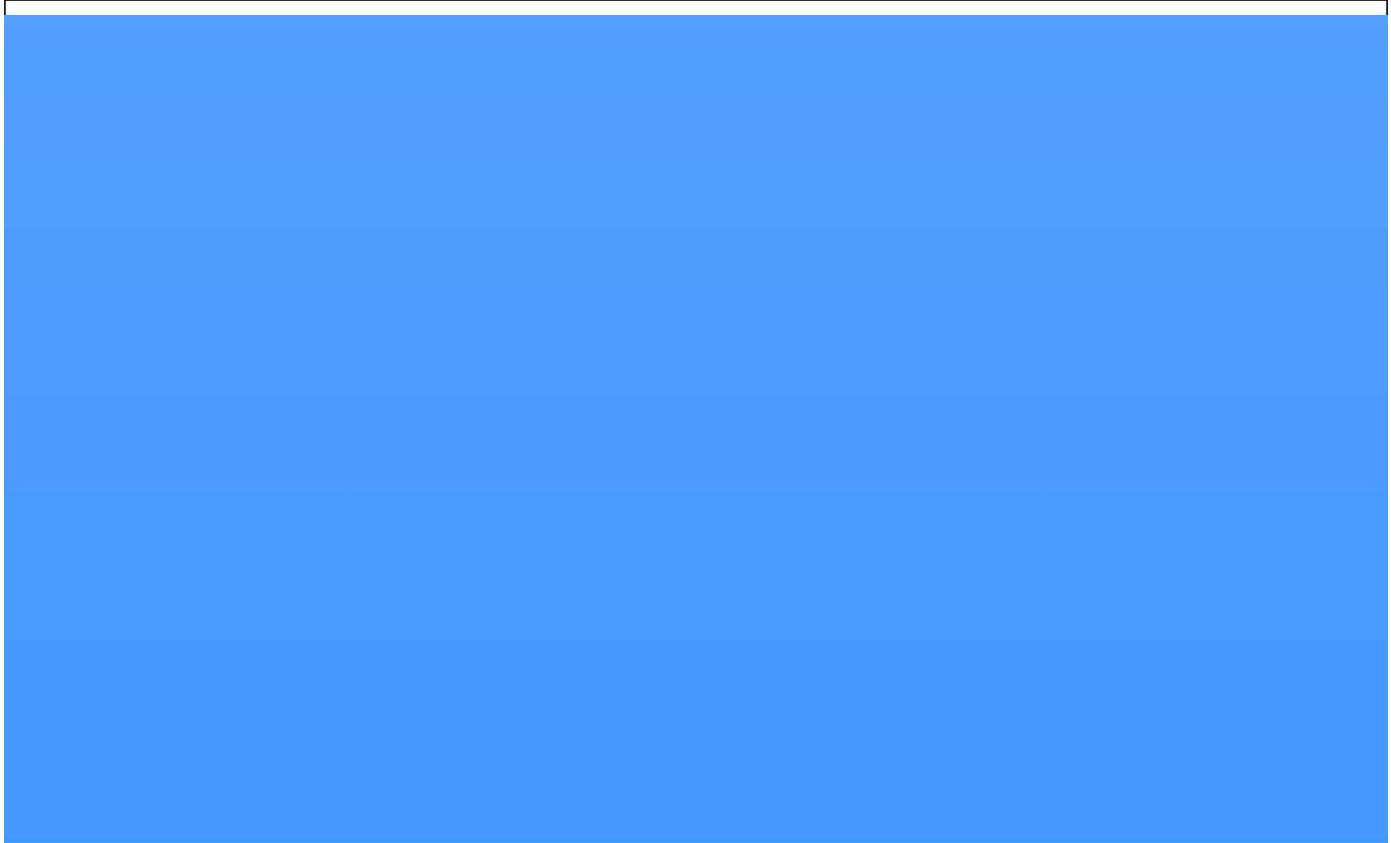
Compression
WMS preparation
Reprojection
Raster pyramid

Archiving
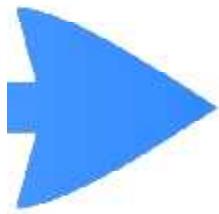
TK25 metadata

Default elements:
- Resource title
- Resource abstract
- Resource Type
- Resource language
- Topic category
- Keyword value
- Originating controlled vocabulary
- Lineage
- Spatial resolution
- Degree
- Specification
- Limitations on public access
- Conditions applying to access and use

Dynamic elements:
- Unique resource identifier
- westBoundLongitude
- eastBoundLongitude
- southBoundLatitude
- northBoundLatitude
- Date of publication
- Metadata date

Variable elements:
- Responsible party
- Responsible party role
- Metadata point of contact

Metadata hierarchy
Resource type

Dataset series
TK25
1995 - 2012

Dataset
TK25 Split
2002

TK25 Pula
1998

TK25 Osijek
2005
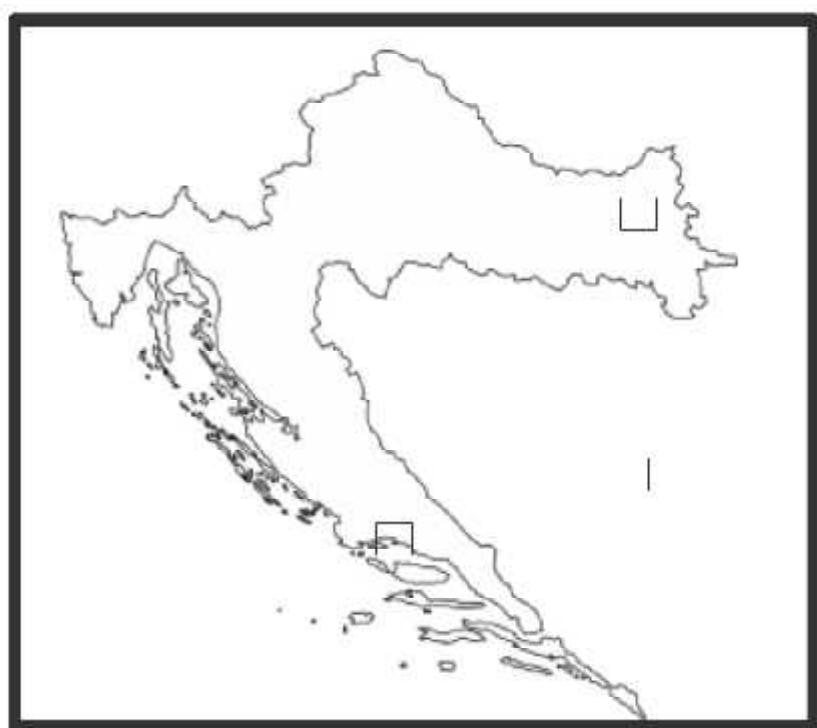
Raster info

TK25
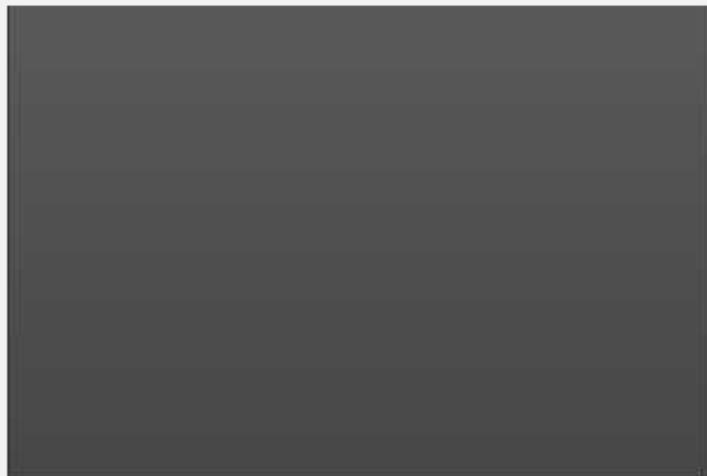
CRS
HTRS96

Resource title
TK25

Unique ID
527c4cac-070c-4bca

TimeStamp
2010-05-27

Raster extents

CRS	Unique ID	TimeStamp	Resource title
HTRS96	527c4cac-070c-4bca	2010-05-27	TK25 Jarmina

6   25-4-107-16   25-3-107-17
   Ernestinovo   Borovo

northBoundLatitude

northBoundLatitude
47.459756

25-2-108-16
Jarmina

southBoundLatitude
42.384073

eastBoundLongitude
19.756992

westBoundLongitude
12.813632

6   westBoundLongitude   eastBoundLongitude

southBoundLatitude

6   25-4-108-16   25-3-108-17

```
      </gmd:topicCategory>
      <gmd:extent>
        <gmd:EX_Extent>
          <gmd:geographicElement>
            <gmd:EX_GeographicBoundingBox>
              <gmd:westBoundLongitude>
                <gco:Decimal>12.813632</gco:Decimal>
              </gmd:westBoundLongitude>
              <gmd:eastBoundLongitude>
                <gco:Decimal>19.756992</gco:Decimal>
              </gmd:eastBoundLongitude>
              <gmd:southBoundLatitude>
                <gco:Decimal>47.459756</gco:Decimal>
              </gmd:southBoundLatitude>
              <gmd:northBoundLatitude>
                <gco:Decimal>42.384073</gco:Decimal>
              </gmd:northBoundLatitude>
            </gmd:EX_GeographicBoundingBox>
          </gmd:geographicElement>
        </gmd:EX_Extent>
      </gmd:extent>
      <gmd:extent>
```

```xml
<gmd:citation>
  <gmd:CI_Citation>
    <gmd:title>
      <gco:CharacterString>ResourceTitle</gco:CharacterString>
    </gmd:title>
    <gmd:date>
      <gmd:CI_Date>
        <gmd:date>
          <gco:Date>2010-05-17</gco:Date>
        </gmd:date>
        <gmd:dateType>
          <gmd:CI_DateTypeCode codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelist.xml/CI_Date
        </gmd:dateType>
      </gmd:CI_Date>
    </gmd:date>
    <gmd:date>
      <gmd:CI_Date>
        <gmd:date>
          <gco:Date>TimeStamp</gco:Date>
        </gmd:date>
        <gmd:dateType>
          <gmd:CI_DateTypeCode codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelist.xml/CI_Date
        </gmd:dateType>
      </gmd:CI_Date>
    </gmd:date>
    <gmd:date>
      <gmd:CI_Date>
        <gmd:date>
          <gco:Date>TimeStamp</gco:Date>
        </gmd:date>
        <gmd:dateType>
          <gmd:CI_DateTypeCode codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelist.xml/CI_Date
        </gmd:dateType>
      </gmd:CI_Date>
    </gmd:date>
    <gmd:identifier>
      <gmd:RS_Identifier>
        <gmd:code>
          <gco:CharacterString>527...Unique ID...of-92bebe7be902</gco:CharacterString>
        </gmd:code>
      </gmd:RS_Identifier>
    </gmd:identifier>
  </gmd:CI_Citation>
</gmd:citation>
<gmd:abstract>
  <gco:CharacterString>Topografiska kartan mjerila 1:25000</gco:CharacterString>
</gmd:abstract>
<gmd:pointOfContact>
  <gmd:CI_ResponsibleParty>
    <gmd:organisationName>
      <gco:CharacterString>DGU</gco:CharacterString>
    </gmd:organisationName>
    <gmd:contactInfo>
      <gmd:CI_Contact>
```


HTTP Post

http://localhost:8080/geonetwork/srv/en/metadata.insert

http://localhost:8080/geonetwork/srv/en/xml.user.login

CS/W