RDA & In-situ data access: sharing experience with Copernicus Land

Hans.dufourmont@eea.europa.eu
19.06.2014
Copernicus land service
pan-European & local components

Corine Land Cover 2012
Sealed soil
Forest type
Tree cover density
Grassland
Wetlands
Urban Atlas
Image mosaics
LC Biodiversity
Water bodies
Experience with in-situ data sources for Land Monitoring services

- On **24 project plans**, access offered to in-situ data:
  - 12 topographic DBs (average)
  - 14 ortho-photo datasets
  - 11 forest inventories
  - 6 grassland inventories
  - 12 DEMs
  - 13 hydrographic data
  - 10 city maps
  - 10 LPIS
  - 10 national LC DB
  - 13 protected area data
  - 11 soil data
  - 12 road networks

- **Type of use:**
  - Training classifiers
  - Verification & enhancement
  - Validation
Issues with in-situ data for Copernicus Land services

**Solution implemented for GIO land:**
- For service providers: clause in the contracts: access to in-situ under their responsibility;
- For NRCs: annex to the grant files indicate data accessibility;
- Access to pan-European datasets.

**Problems encountered and justifying a need for some level of in-situ coordination:**
- Only partial coverage of EEA39;
- Different levels of fitness for purpose;
- No full clarity on access for service providers producing HRLs !!!

**Despite...**
- Access conditions shall be harmonised to EU institutions: INSPIRE directive, art. 17 § 8;
- Application of GMES regulation art. 5 § 2;

So, all in all: access to in-situ: not **SO** straightforward ????
Availability of information on access to in situ data for Land

GIO Land In-Situ analyses:

- Access information available from Reference Data Assessments (RDA)
- Access information obtained from Country visits
- Access information obtained from the ANNEX III - GIO Land
- EEA 39 coverage
- No information available
Country responses on data access: ortho-photos
Coverage: the (then) EU 27 + Norway and Switzerland*

Type of data access: Oriented to Network service URL as provision of data access

Tools: Web services validated with Open Source and commercial GIS software

Web surveys

Target: technical specifications / access points (URL) / data access conditions

Datasets: Geospatial Reference data (shared between Land and EMS services)

Reporting period: Nov 2012 – Jun 2013

Feedback received: 13 countries
INSPIRE created the Framework for documenting and reporting geospatial data and data access.

Use of INSPIRE, discovery services, web interaction as base to complement Previous surveys.

Inventory of Network services (secured and open)
http://wms.jpl.nasa.gov/wms.cgi
http://columbia.edu/wfs/gpw
...
Analysis criteria

- **Catalogue of Network services for two layers:**
  - Roads (Transport Network – ANNEX I.7)
  - Rivers (Hydrology - ANNEX I.8)

- **Criteria for selection of network services**
  - Preference to scales/resolutions matching In-Situ Copernicus requirements (target and threshold scales)
  - Data formats compliant with In-Situ requirements
  - National coverage

- **Criteria for selection of URL:**
  - Preference to service’s URL
  - When not possible: GetCapabilities XML document, WSDL document, Client application URL (giving access to the service), http page.
## Gap analysis against required fields

<table>
<thead>
<tr>
<th>Field name</th>
<th>Match with reporting obligation</th>
<th>Match with Metadata implementation rules</th>
<th>XPATH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product</td>
<td>Spatial data sets (Data input sheet)</td>
<td>2.2.3 Resource title</td>
<td>identificationInfo[1]/citation/*/title</td>
</tr>
<tr>
<td>Country</td>
<td>Member State (Contributor)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RespAut</td>
<td>RA - Responsible authority within the MS (Data input sheet)</td>
<td>2.10.1 Responsible party</td>
<td>identificationInfo[1]/pointOfContact</td>
</tr>
<tr>
<td>SpType</td>
<td></td>
<td>2.7.1 Lineage</td>
<td>dataQualityInfo/<em>/lineage/</em>/statement</td>
</tr>
<tr>
<td>SpCoverage</td>
<td>Spatial data sets/Extend (Data input sheet)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scale</td>
<td></td>
<td>2.7.2 Spatial resolution</td>
<td>identificationInfo[1]/spatialResolution/<em>/equivalentScale/</em>/denominator</td>
</tr>
<tr>
<td>FreqUpdate</td>
<td></td>
<td>2.6.3 Date of last revision</td>
<td>identificationInfo[1]/citation/<em>/date[</em>/<em>/dateType/</em>/text()='revision']/*/date</td>
</tr>
<tr>
<td>TopConsis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RefSystem</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MetaType</td>
<td></td>
<td>2.2.3 Resource Type</td>
<td>hierarchyLevel</td>
</tr>
<tr>
<td>MetaStand</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MetaINSPIRE</td>
<td>Spatial data sets/compliance (Data input sheet)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MetaURL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoupledServices</td>
<td>Spatial data sets/view services - download services (Data input sheet)</td>
<td>2.2.4 Resource locator</td>
<td>distributionInfo/<em>/transferOptions/</em>/onLine/*/linkage</td>
</tr>
<tr>
<td>ServName</td>
<td>Services/name (Data Input sheet)</td>
<td>2.2.3 Resource title</td>
<td>identificationInfo[1]/citation/*/title</td>
</tr>
<tr>
<td>ServAcc</td>
<td>Services/name (Data Input sheet)</td>
<td>2.9.2 Conditions applying to access and use</td>
<td>identificationInfo[1]/resourceConstraints/*/useLimitation</td>
</tr>
<tr>
<td>ServSpec</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TypeURL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ServURL</td>
<td>Services/URL (Data Input sheet)</td>
<td>2.2.4 Resource locator</td>
<td>distributionInfo/<em>/transferOptions/</em>/onLine/*/linkage</td>
</tr>
</tbody>
</table>
Use of INSPIRE Monitoring & Reporting templates

- Reports on yearly basis (from May 2013 EEA is coordinating national submissions to the CDR)

Use of SPARQL endpoint

- query specific INSPIRE themes, datasets and country
  http://cr.eionet.europa.eu/sparql
Results

- Number of records for all countries:
  - **98 records** collected for **Road Networks**
  - **60 records** collected for **Rivers**

- Ex.: Spain

<table>
<thead>
<tr>
<th>Products</th>
<th>Spatial Type</th>
<th>Spatial Coverage</th>
<th>Scale</th>
<th>Frequency of update</th>
<th>Top consistency</th>
<th>Reference system</th>
<th>Network services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Road Network dataset</td>
<td>vector</td>
<td>national</td>
<td>1:25 000 or larger</td>
<td>2 years</td>
<td>provided</td>
<td>EPSG:4258</td>
<td>Available</td>
</tr>
<tr>
<td>Responsible party</td>
<td>CENTRO NACIONAL DE INFORMACION GEOGRÁFICA (CNIG) - CartoCiudad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metadata</th>
<th>INSPRE compliance</th>
<th>Metadata type</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO 19139</td>
<td>yes</td>
<td>Dataset / data series / service</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Associated services</th>
<th>Specifications</th>
<th>URL type</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>View service</td>
<td>Unrestricted</td>
<td>OGC WMS v 1.3.0</td>
<td>Service's URL</td>
</tr>
<tr>
<td>Download service</td>
<td>Atom feed</td>
<td></td>
<td><a href="http://centrodescargas.cniq.es/CentroDescargas/buscadorCatalogo.do?codFamilia=02122">http://centrodescargas.cniq.es/CentroDescargas/buscadorCatalogo.do?codFamilia=02122</a></td>
</tr>
</tbody>
</table>
Layer: Roads; **INSPIRE theme:** Transport Network; **geographic extent:** National; **scale:** < 1:50 000; **Data format:** vector geometries
Lessons learned

- Non-direct association between geospatial network services and scale or other data specifications; this creates sometimes difficulties to validate web services against matching In-Situ requirements, the scale being one of the primary factors.

- Provision of multi products: countries report more than one product per dataset whenever more products (or none) are matching the service requirements.

- Mismatch between data, services and specifications (metadata): not always there is correspondence of specifications required and those provided by the actual data.

- Web service validation required: there is need of validation between datasets and derived network services for both download and view services.
Solution strategy: A 2-track approach

1. **Centralised approach:**
   - Procure existing European datasets if service requirements are met;
   - Upgrade existing or produce new European datasets if service requirements of existing datasets are not met.

2. **Decentralised approach:**
   - MS implement INSPIRE in order to:
     - a) harmonise national datasets to European specifications (on-going);
     - b) set up national/regional nodes of services (Catalog, View, Download, Transformation);
     - c) harmonise access and use conditions (art. 17 § 8).

**EEA coordinates:**
- Minimal functional linkage between European and national datasets, e.g. through commonly shared identifiers
- Maintenance arrangements to ensure that European data do follow the pace of national updates (cfr. ESS: European Statistical System)
- Stepwise implementation: priority setting: a) reference data access (cfr. INSPIRE annex I), b) supporting reference data (cfr. INSPIRE annex II, including VHR ortho-satellite imagery, c) thematic in-situ data
Challenges and proposed schema for in-situ data access

Centralised data procurement / production

Decentralised data access

e.g. EBM, EU-DEM, EU-hydro...

Link through IDs Maintenance

Time

2013

fading in

fading out

2019*

e.g. national LC maps, national forest DBs...

* Final deadline INSPIRE implementation
RDA activities

- Analysis of feasibility of improved access mechanisms to the national and regional reference data according to INSPIRE data themes (Annex I&II)
- The set-up of an improved and operational access to the selected reference data
- EU-Hydro and EU-DEM upgrade
- EU-Hydro and EU-DEM statistical validation
- Dissemination of EU-Hydro and EU-DEM
Service Providers of Copernicus services

Single Sign-on

Middleware software

Modular plugin
Modular plugin
Modular plugin

National and European Data Providers

Administrator and operator

Administration

Monitoring

(Secured and/or Public Data)
CORDA will be:

- Simple to use – will only require an internet connection to access CORDA with a single login from the GIS software of the choice

- Available 24/7 – ensuring access to the national reference data in cases where the respective national services are temporarily not available

- Reliable – providing access to the timely and up-to-date national reference data

- Efficient – everything discoverable in one place

- Sustainable – access to the data ensured through the whole production and update cycle
Technically a reference dataset available through CORDA may be a single mosaic dataset compiled from multiple national datasets.

It is important to note that it is not under the CORDA mandate to harmonise and ‘fix’ data.
European in-situ datasets still important:
- EU-hydro
- EU-DEM
- LUCAS ground survey
Conclusions and Recommendations

- On-going INSPIRE implementation provides first components for a distributed access to in situ data.

- However:
  - still pretty much suffering from partial coverages
  - Copernicus services are in need of data access now.

- Therefore: Continuation of 2-track approach for adequate in-situ data access is essential to address Copernicus services data access needs.

- Further development of the CORDA access node, including the establishment of (harmonised) access agreements

- Maintenance of some key pan-European datasets

- Linkage between national/regional datasets and pan-European datasets essential as mid-term target
Thank you for your attention