



Lineage and the importance of the ancestors and descendants, specially if you are noble

Why lineage is important in metadata

Agenda

What is lineage

Lineage at INSPIRE

State of art in Europe

Conclusions



Lineage at ISO

ISO 19115-1:2014 "Geographic Information - Metadata - Part 1: Fundamentals" Standard, lineage is defined as provenance, source(s) and production process(es) used in producing a resource. It is one of the 13 packages used to define and provide metadata.



Well, boring

Lineage at EC

According to the Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata, lineage means the history of a data set, and the life cycle from collection and acquisition through compilation and derivation to its current form, in accordance with EN ISO 19101 rule.

Ah, datasets...

Lineage and Quality

Lineage is part of the quality documentation of the data

- ISO 19115 (2014-04-01)

The Lineage information metadata class is a not compulsory element that reflects the qualitative(descriptive) aspect of the quality of a dataset. Lineage Information Package supports the provision of metadata concerning the sources and production processes used in producing a resource. It is covered by the LI_Lineage class who aggregates two classes: "LI_Source" and "LI_ProcessStep". None of them are mandatory.

Not compulsory, good!

Tags

- “LI_Source” has the purpose of offering the source of data used in creating the data set received by the user, while the
- “LI_ProcessStep” has the objective of informing about events or transformations in the life of the dataset.



Not compulsory, good!

Lineage at INSPIRE

In metadata, lineage includes free text about information of data sources and transformation steps followed in the creation process. Following INSPIRE Implementing Rules for Metadata, lineage is mandatory for datasets and dataset series, and it should include quality information for interoperability. It is not applicable to services.



It was too good to be true

Lineage at IR

- Apart from describing the process history, the overall quality of the dataset (series) should be included in the Lineage metadata element. This statement should contain any quality information required for interoperability and/or valuable for use and evaluation of the data set (series).
- A single ISO 19115 metadata set may comprise more than one set of quality information, each of them having one or zero lineage statement. There shall be one and only one set of quality information scoped to the full resource and having a lineage statement (INSPIRE only considers the instance of MD_Metadata.dataQualityInfo applicable to the whole Resource)

OK, only one and only

Lineage at TG

Going in the INSPIRE Technical Guidelines, it is stated that lineage's value domain is free text and it is added a recommendation to use the sub-elements of LI_Lineage as follows:

- LI_ProcessStep sub-element should be used for the description of the transformation process of the local to the common INSPIRE data structures (reflect the transformations carried out to make this local/national dataset compliant with INSPIRE).
- LI_Source sub-element should be used for the description of the source data (local/national data base used to derive the INSPIRE data set).

Is this easy?

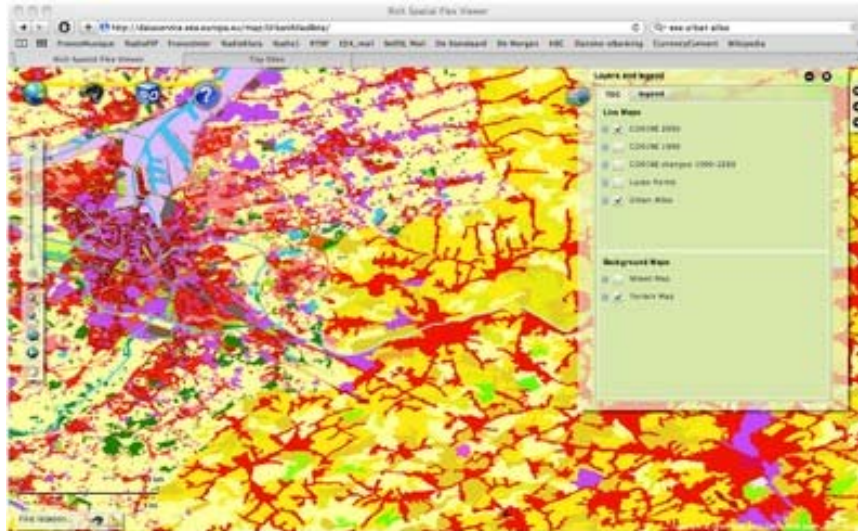
Elevation

It is suggested to provide information about the Coverage of the Elevation Grid within the lineage



Land Cover

It is established that the observation date shall be provided at the coverage level through metadata within processStep information provided in the lineage



Hidrography

It is recommended to document in the lineage the selection of rules (based on the guides provided in this TG) decided by each Member State for the definition of the objects to be included under this INSPIRE thematic frame.



Cadastral Parcels

It is recommended to document in the lineage the selection rules in the Data Capture of this theme. It has been created by the Thematic Working Group a template for lineage, in order to help cadastral producers to document the remaining national specificities that are inevitable even when data is published in an interoperable way for INSPIRE. Other uses recommended for Lineage are to reflect:

- "Cadastral gaps" and "Cadastral overlaps" at borders.
- the structure of national cadastral reference.
- the Life-cycle rules of cadastral parcel spatial objects that applies in a Country.
- the positional accuracy when it is estimated from the knowledge of source data and of production processes.

Addresses

It is recommended to document the life-cycle rules for addresses in the data set.



Lineage nowadays

Annex I. Administrative Units - INSPIRE dataset (GIS-Centras, Some country):

```
<gmd:lineage>
  <gmd:LI_Lineage>
    <gmd:statement xsi:type="gmd:PT_FreeText_PropertyType">
      <gco:CharacterString>LT</gco:CharacterString>
      <gmd:PT_FreeText>
        <gmd:textGroup>
          <gmd:LocalisedCharacterString locale="#en">
            LT
          </gmd:LocalisedCharacterString>
        </gmd:textGroup>
      </gmd:PT_FreeText>
    </gmd:statement>
  </gmd:LI_Lineage>
</gmd:lineage>
```

Lineage nowadays

DTM 50 (Kartverket, Some Country):

```
<gmd:lineage>
  <gmd:LI_Lineage>
    <gmd:statement>
      <gco:CharacterString xmlns:gco="http://www.isotc211.org/2005/gco">
        Ingen prosseshistorie tilgjenglig.
      </gco:CharacterString>
    </gmd:statement>
  </gmd:LI_Lineage>
</gmd:lineage>
```


Lineage nowadays

Ortophoto in some country

```
<gmd:lineage>
  <gmd:LI_Lineage>
    <gmd:statement>
      <gco:CharacterString>
        National Cadastre and Mapping Agency S.A.
      </gco:CharacterString>
    </gmd:statement>
  </gmd:LI_Lineage>
</gmd:lineage>
```

Lineage nowadays

```
<gmd:lineage>
<gmd:LI_Lineage>
  <gmd:statement>
    <gco:CharacterString>
      1.Preparación de los datos: Las capas "Aeropuertos (SIGNA), Autopista (SIGNA), Autovía (SIGNA),
      Carretera autonómica (SIGNA), Carretera nacional (SIGNA), Ferrocarril convencional (SIGNA),
      Ferrocarril AVE (SIGNA), Puerto (SIGNA)" proceden del SIGNA (Sistema de Información Geográfico
      Nacional de España). La capa Viales Urbanos (Cartociudad) procede de Cartociudad. Las capas Área
      de pista, Autopista y autovía (BTN25), Carretera convencional (BTN25), Enlace de carreteras
      (BTN25), Ferrocarril convencional (BTN25), Puerto (BTN25)" proceden de la BCN25 y BTN25. A partir
      de procesos de transformación realizados con la herramienta FME, sobre los datos originales
      almacenados en Oracle Spatial, se ha realizado la carga final en PostGIS ofreciendo como Sistema
      de Referencia: ETRS89, latitud y longitud. 2.Configuración y creación de las capas con Geoserver
      2.1.0. 3.Edición y creación de los archivos de estilos (SLD) a partir de la simbología establecida por
      Inspire y por el IGN, obteniendo como resultado dos estilos de visualización. 4.Publicación de la
      capa.
    </gco:CharacterString>
  </gmd:statement>
  <gmd:processStep gco:nilReason="missing"/>
  <gmd:source gco:nilReason="missing"/>
</gmd:LI_Lineage>
</gmd:lineage>
```

Lineage nowadays

LiDAR based Digital Surface Model (DSM)

```
<gmd:lineage>  
  <gmd:LI_Lineage>  
    <gmd:statement>  
      <gco:CharacterString>
```

The British Antarctic Survey (BAS) acquired a new high resolution (circa 1 point/meter) and high accuracy (25 cm vertical) LiDAR dataset over Cornwall and Devon during the month of July and August 2013 for the Tellus South West project. The conversion of the raw LiDAR data into the digital terrain and surface model was carried out by the Environment Agency (contract). The spatial dataset consists of 2 cm height layers representing the terrain and surface height respectively at a spatial resolution of 1 m (grid cell) and an average height accuracy of 25 cm. The data is available as individual files representing sections of Ordnance Survey tiles in ESRI grid format (e.g. SX37ne; SX37nw; SX37se; SX37sw).

```
      </gco:CharacterString>  
    </gmd:statement>  
  </gmd:LI_Lineage>  
</gmd:lineage>
```

Desirable information for lineage

```
<gmd:lineage>
<gmd:LI_Lineage>
  <gmd:statement>
    <gco:CharacterString>
      Description of source data used in data creation (included its date and accuracy) and the event or
      transformation in the life of a dataset, included the process used to maintain the dataset (enhancing
      the detail of the process and avoiding the commercial trademarks as far as possible).
    </gco:CharacterString>
  </gmd:statement>
  <gmd:processStep>
    <gmd:LI_ProcessStep>
      <gmd:description>
        <gco:CharacterString>
          Information about a process transformation of a source dataset into a dataset compliant to
          the technical requirements of the applicable Implementing Rules and Technical Guidelines of
          INSPIRE.
        </gco:CharacterString>
      </gmd:description>
    </gmd:LI_ProcessStep >
```

Desirable information for lineage

```
<gmd:LI_Source>  
  <gmd:description>  
    <gco:CharacterString>  
      Description of source database used to derive the INSPIRE data set, denomination of the data.  
    </gco:CharacterString>  
  </gmd:description>  
</gmd:LI_Source>  
</gmd:source>  
</gmd:LI_Lineage>  
</gmd:lineage>
```




Muchas Gracias

ALMANZA, Josu
BUSTO, Natalia
GONZÁLEZ, Nagore
GUINEA, Alejandro
LÓPEZ DE TURISO, Olga
PASCUAL, Estibaliz

INSPIRE Conference 2016