GeoPackage as an encoding format for INSPIRE data

Thorsten Reitz, wetransform
Heidi Vanparys, SDFI Denmark
Lars Erik Storgaard, SDFI Denmark
Stefania Morrone, Epsilon Italia
Jari Reini, NLS Finland

Darja Lihteneger / INSPIRE MIG and MIG-T Meeting / 24 - 25.11.2022

© Dominika Hrašková, REDISCOVER Nature/EEA

European Environment Agency

Outline

- INSPIRE Good practice on GeoPackage process status
- Implementation in domains (GeoPackage Webinar)
- Lessons learned and conclusions

Good Practice on GeoPackage: aim

Aim:

- To **describe a mechanism** to create **INSPIRE datasets encoded using the OGC GeoPackage encoding standard**.
- Ensure compliance: These datasets will be **compliant with the INSPIRE Implementing Rules (IR), and technical compliance can be shown through transformation to the default encoding (GML)**.
- **Alternative encoding**: In this perspective, the GeoPackage can be used both as an additional and an alternative encoding for INSPIRE datasets.

Main benefits of the Good Practice

to Data Providers:

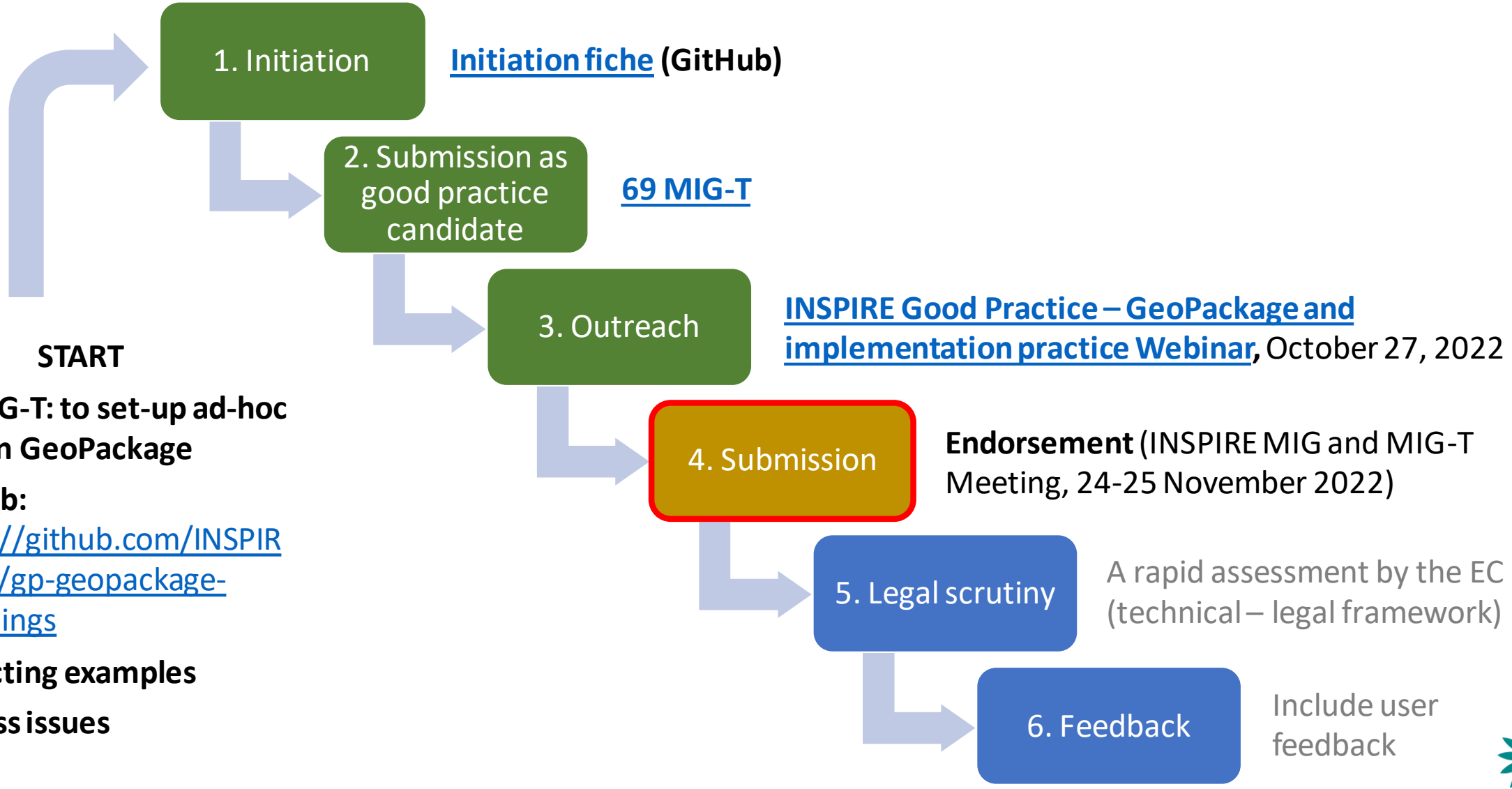
- the relatively simple data models focus on key values, which simplify data harmonisation as compared to creating full INSPIRE GML
- more robust encoding based on the use of templates reduces number of schema and encoding errors
- usability and re-usability of provided data is considerably improved
- GeoPackage files can handle large number of features in a smaller file size, hence the sharing of entire data sets, sized medium to large is facilitated.

to Data Users:

- the simplified data model aligns directly with use cases
- GeoPackage encoded data can be directly used in GIS tools without any ETL process, with good performance. The data consumption experience is considerably improved with possibility to directly edit/update data, create views and store styles
- loading data on GIS or making queries on the file database is faster, since the vector layers in geopackage are inherently rtree indexed (spatial indexing).

Source: JRC INSPIRE Team (69th MIG-T Meeting)

INSPIRE Good Practice for GeoPackage – process status



START

- **63 MIG-T: to set-up ad-hoc WG on GeoPackage**
- **GitHub:** https://github.com/INSPIRE-MIF/gp-geopackage-encodings
- **Collecting examples**
- **Discuss issues**

About GeoPackage

- ***"GeoPackage is an open, standards-based, platform-independent, portable, self-describing, compact format for transferring geospatial information"***

- Based on database system SQLite, can store:
  - Vector features with attributes (including spatial data and other tabular data)
  - Tile matrix sets of imagery, raster maps
  - Tile gridded coverage data (Extension)
  - Metadata and schema constraints
- Data sharing without information loss between GIS and direct use in different GIS
- Supported by OGC GeoPackage validation tools, https://github.com/opengeospatial/ets-gpkg12
- Maximum size ~140TB, good read performance
- Suitable for desktops, network servers, and mobile applications

Source: Thorsten Reitz, Heidi Vanparys (GeoPackage Webinar)

Evidence of GeoPackage implementation (outreach webinar)

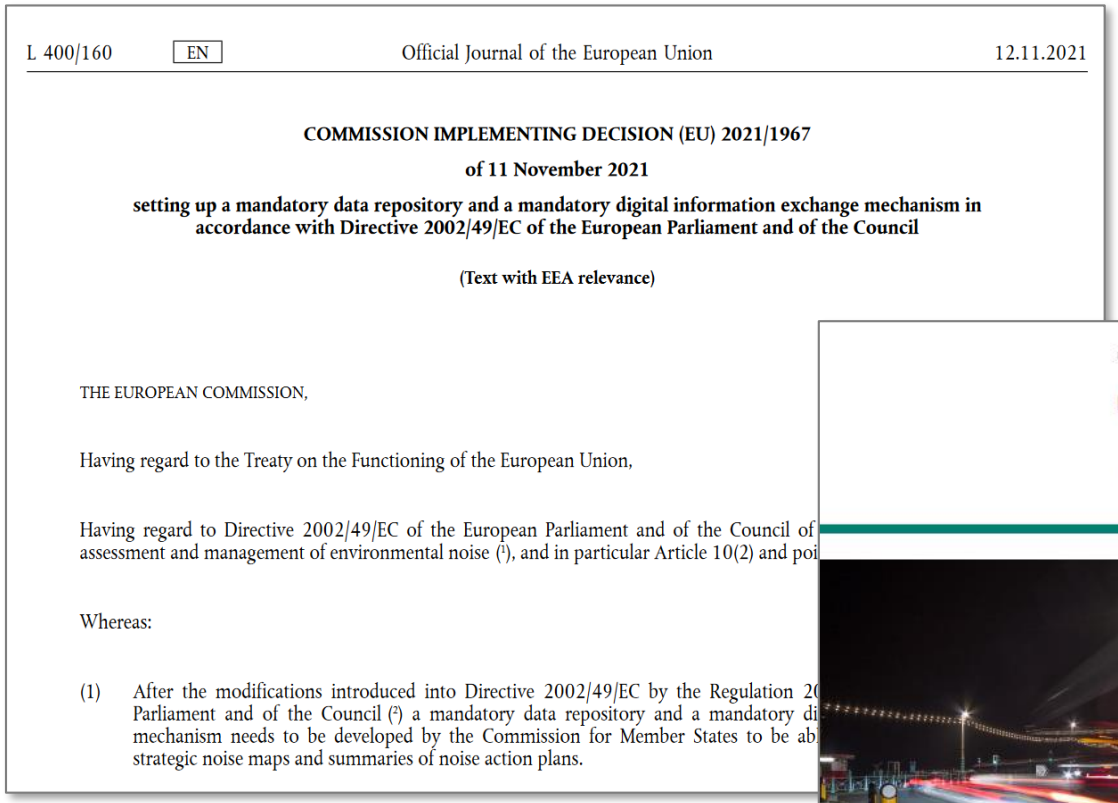
Thematic domain	INSPIRE data model covered by GeoPackage implementation
Environmental Noise Directive reporting of noise sources, strategic noise maps and noise action plans	• **Area Management** • **Transport Networks: roads, railways and airports** • **Human Health: environmental health determinant measure**
Addresses, Denmark **Addresses, Finland**	• **Addresses**
Geological Map of Italy 1:100k **Geology – extended model** for easy-to-use, interoperable subsurface data, Italy	• **Geology**

Examples: https://github.com/INSPIRE-MIF/gp-geopackage-encodings/blob/main/examples/overview.md
GeoPackage Webinar: https://inspire.ec.europa.eu/events/inspire-good-practice-geopackage-and-implementation-practice-webinar

Example: Environmental Noise Directive reporting

Wide stakeholders' cooperation

- European Environment Agency (EEA)
- ETC and Eionet (Noise experts and reporters)
- European Commission
- EEA consultancy (wetransform, Epsilon Italia, KU Leuven, other EEA consultancy services)
- National implementations of Noise portals and data repositories



L 400/160 EN Official Journal of the European Union 12.11.2021

COMMISSION IMPLEMENTING DECISION (EU) 2021/1967

of 11 November 2021

setting up a mandatory data repository and a mandatory digital information exchange mechanism in accordance with Directive 2002/49/EC of the European Parliament and of the Council

(Text with EEA relevance)

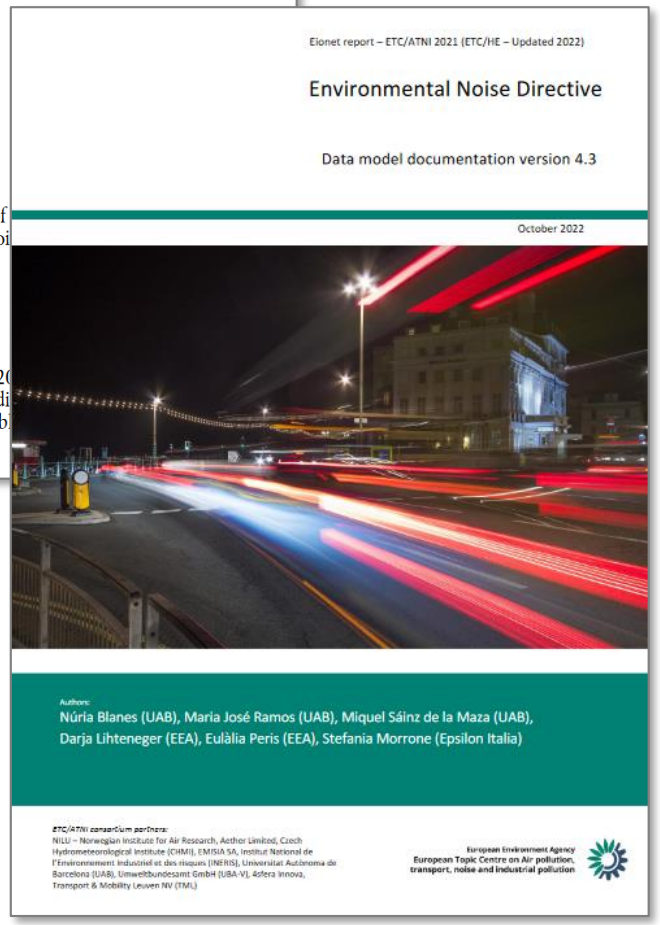
THE EUROPEAN COMMISSION,

Having regard to the Treaty on the Functioning of the European Union,

Having regard to Directive 2002/49/EC of the European Parliament and of the Council of
assessment and management of environmental noise (1), and in particular Article 10(2) and poi

Whereas:

(1) After the modifications introduced into Directive 2002/49/EC by the Regulation 2
Parliament and of the Council (2) a mandatory data repository and a mandatory di
mechanism needs to be developed by the Commission for Member States to be ab
strategic noise maps and summaries of noise action plans.



Eionet report – ETC/ATNI 2021 (ETC/HE – Updated 2022)

Environmental Noise Directive

Data model documentation version 4.3

October 2022

Authors:
Núria Blanes (UAB), María José Ramos (UAB), Miquel Sáinz de la Maza (UAB),
Darja Lihteneger (EEA), Eulàlia Peris (EEA), Stefania Morrone (Epsilon Italia)

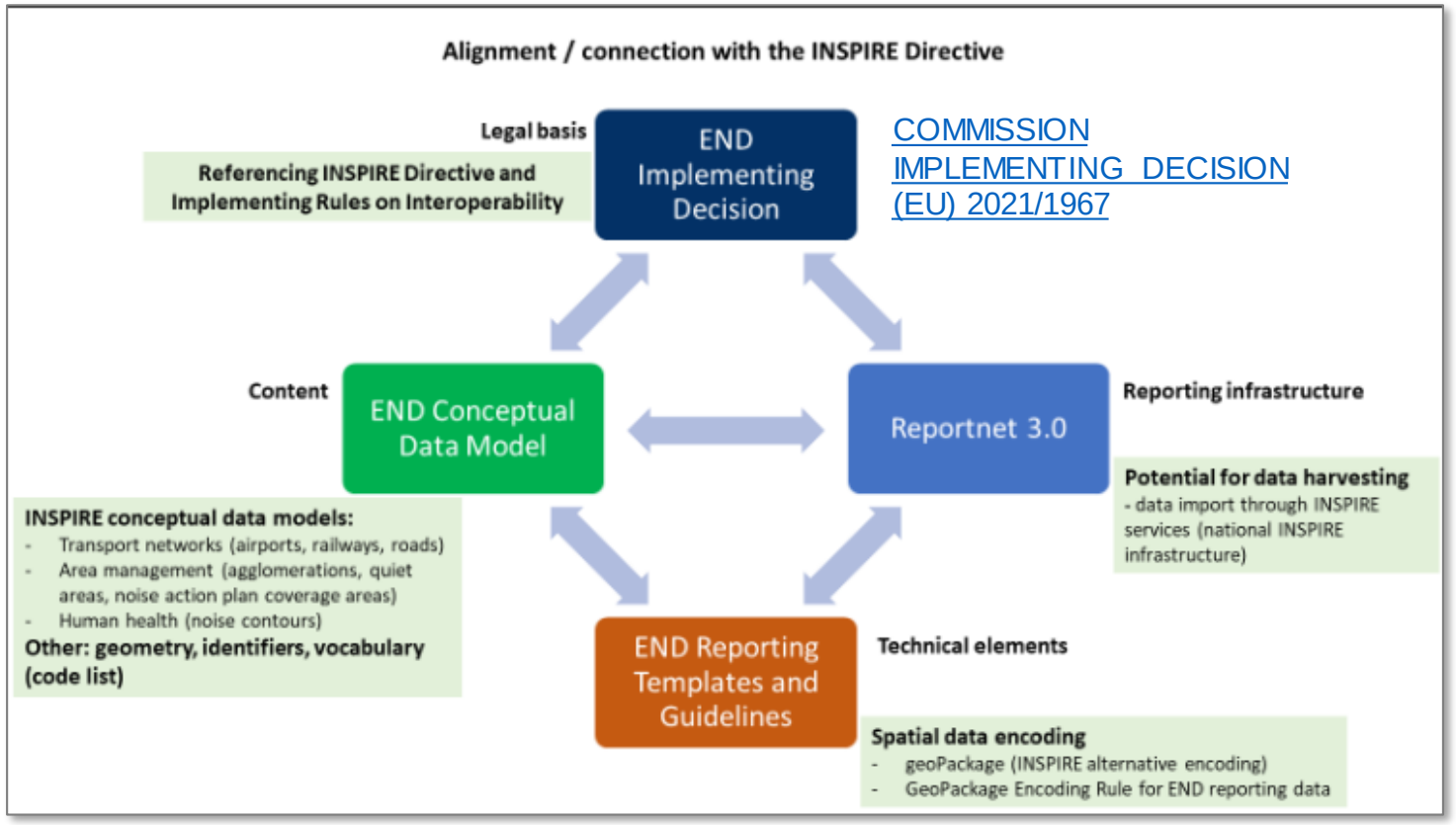
ETC/ATNI consortium partners:
NILU – Norwegian Institute for Air Research, Aether Limited, Czech Hydrometeorological Institute (CHMI), EMISIA SA, Institut National de l'Environnement Industriel et des risques (INERIS), Universitat Autònoma de Barcelona (UAB), Umweltbundesamt GmbH (UBA-V), 4sfera Innova, Transport & Mobility Leuven NV (TML)

European Environment Agency
European Topic Centre on Air pollution, transport, noise and industrial pollution

Darja Lihteneger
Thorsten Reitz

Example: Environmental Noise Directive reporting

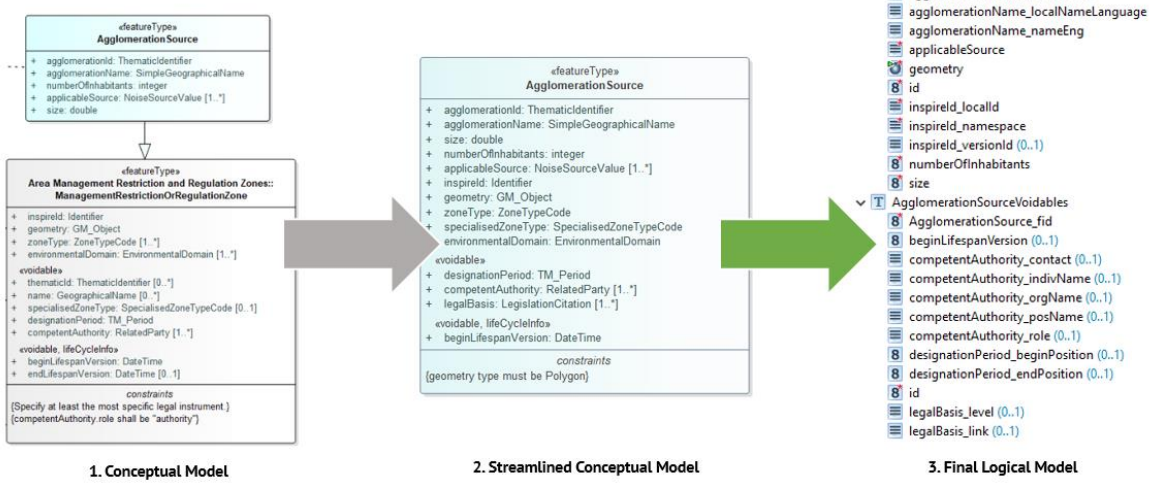
- **Aligning directives: END and Noise**
- **END conceptual data model**: re-using or extending INSPIRE data models
- **GeoPackage encoding**
- **Reporting into Reportnet3**
- **National implementations**: Noise portals and repositories supporting national data collection, validation, reporting and INSPIRE implementation



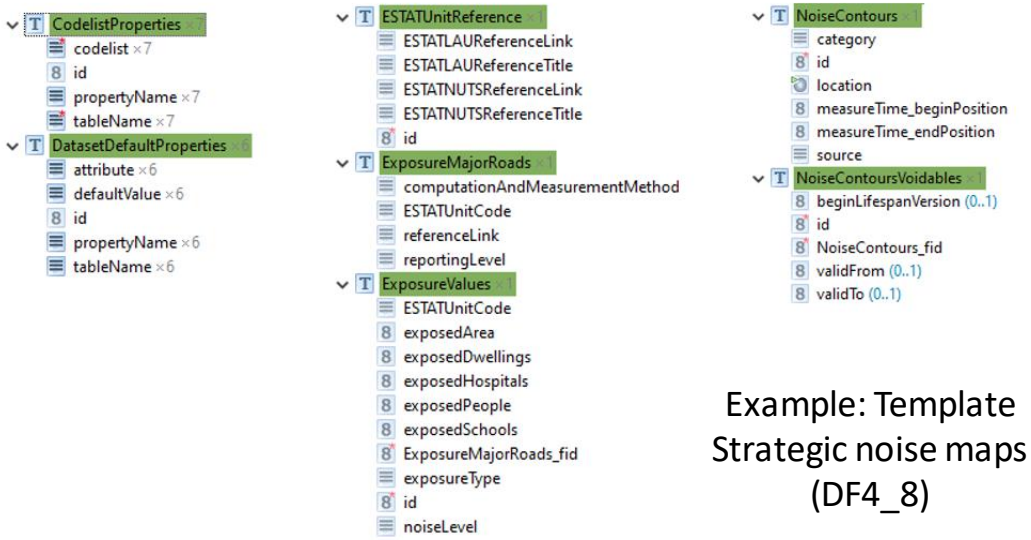
Noise Eionet Portal: https://www.eionet.europa.eu/reportnet/docs/noise

Example: Environmental Noise Directive reporting

Streamlined conceptual data model to logical model

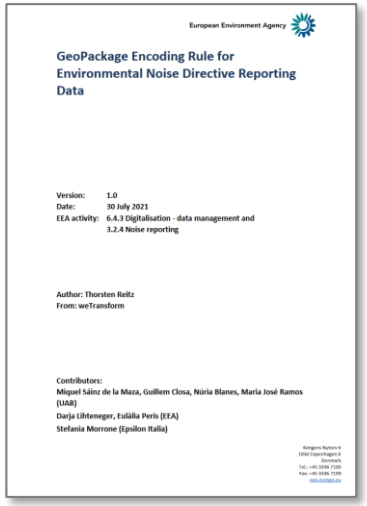


GeoPackage template (executable transformation project in HALE)
Data transformation with HALE Studio 4.0+



Example: Template Strategic noise maps (DF4_8)

Model transformation rules (UML to GeoPackage)



- **Primary (core) table(s)**: key properties that have per-feature values
- Move properties that are nillable to "**Voidables**" table
- "**DatasetDefaultProperties**": move values that are the same for all features to DatasetDefaultProperties
- "**CodelistProperties**": properties with code lists and URL of code lists (INSPIRE + extended, Noise specific code lists)

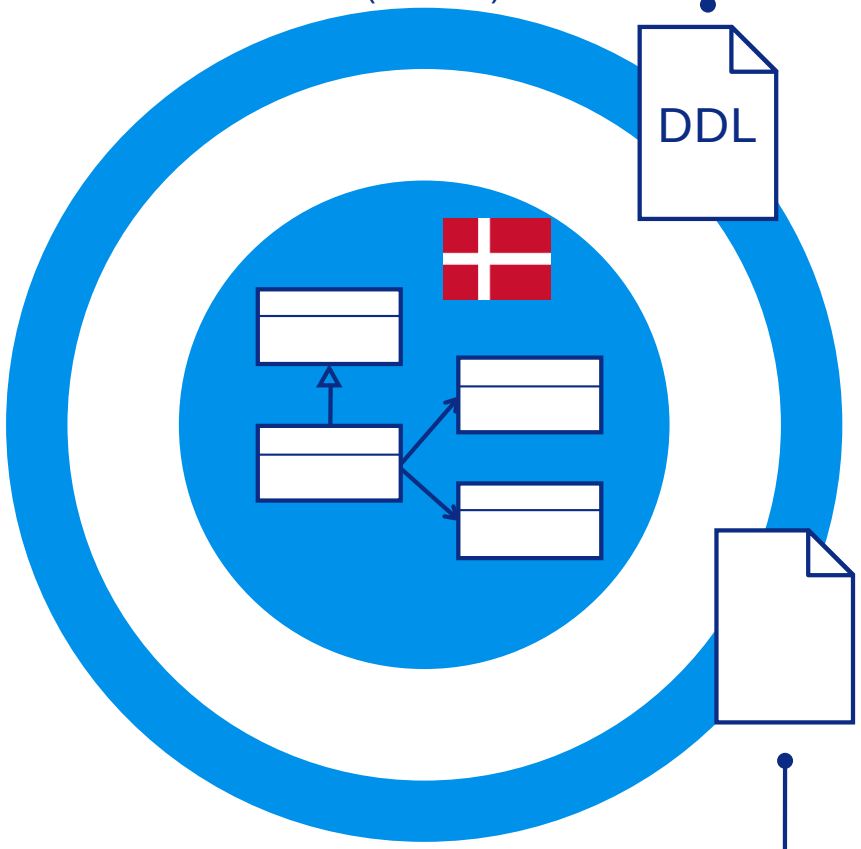
Source: Thorsten Reitz (GeoPackage Webinar)

Example: Addresses, Denmark

Heidi Vanparys
Lars Erik Storgaard

Example: Addresses, Denmark – model-driven approach ×2

DDL (tables)
defines the structure of the data in our ”data warehouse” (Oracle).

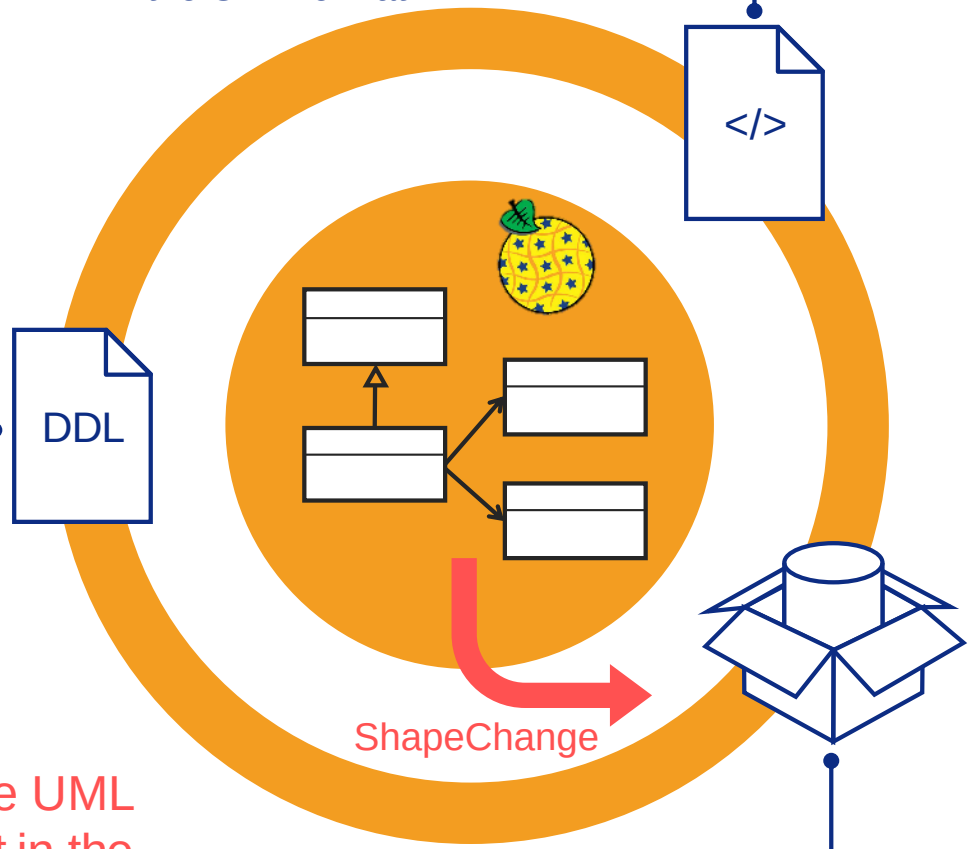


(other artefacts derived from the UML model)

manual creation; structure can be based on the GeoPackage template

DDL (views)
defines the structure of the INSPIRE data in our ”data warehouse” (Oracle).

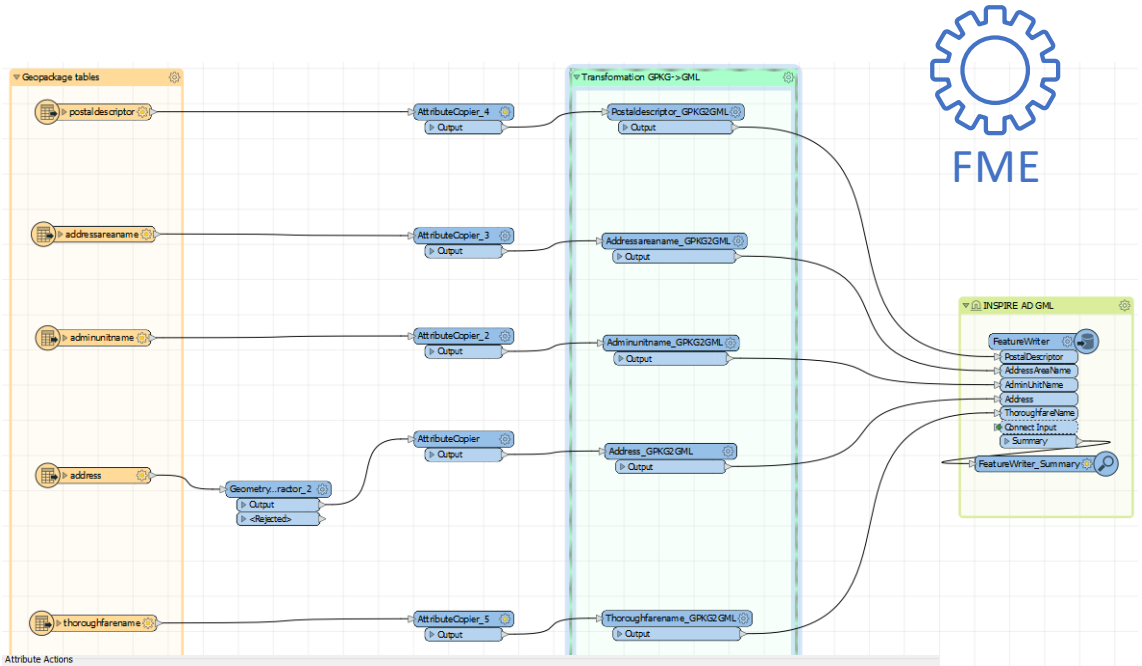
INSPIRE GML application schema
defines the structure of INSPIRE data in the GML format.



definitions from the UML model are present in the GeoPackage template
→ self-documenting template

INSPIRE GeoPackage template
defines the structure of INSPIRE data in the GeoPackage format.

Example: Addresses, Denmark



FME

Attribute Actions

Input Attribute	Output Attribute	Attribute Value	Action
inspireid	inspireId.Identifier.localId	@TrimLeft(@Value(inspireId.Identifier.localId),https://geo.data.gov.dk/dataset/)	Rename
	inspireId.Identifier.namespace	https://geo.data.gov.dk/dataset/50b921ea-935e-d605-2287-4ee364046795	Set Value
alternativeidentifier	alternativeIdentifier.xsi_nil	true	Rename
	alternativeIdentifier.nilReason	other:unpopulated	Set Value
position_default	position{0}.GeographicPosition.default	<Enter new value (optional)>	Rename
position_method	position{0}.GeographicPosition.method.xlink_href	<Enter new value (optional)>	Rename
position_specification	position{0}.GeographicPosition.specification.xlink_href	<Enter new value (optional)>	Rename
status	status.xlink_href	<Enter new value (optional)>	Rename
validfrom	validFrom	@DateTimeFormat(@Value(validFrom), ISO)	Rename
validto	validTo	@DateTimeFormat(@Value(validTo), ISO)	Rename
beginlifespanversion	beginLifespanVersion	@DateTimeFormat(@Value(beginLifespanVersion), ISO)	Rename
endlifespanversion	endLifespanVersion	@DateTimeFormat(@Value(endLifespanVersion), ISO)	Rename
component_adminunitname_1	component_adminunitname_1a	@Trim(@Value(component_adminunitname_1a),https://geo.data.gov.dk/dataset)	Rename
component_adminunitname_1a	component_adminunitname_1b	@ReplaceString(@Value(component_adminunitname_1b),/,_)	Rename
component_adminunitname_1b	component_adminunitname_1b_#	#_@Value(component_adminunitname_1b_#)	Rename
component_adminunitname_1b_#	component{0}.xlink_href	<Enter new value (optional)>	Rename
component_postaldescriptor	component_postaldescriptor_a	@Trim(@Value(component_postaldescriptor_a),https://geo.data.gov.dk/dataset)	Rename
component_postaldescriptor_a	component_postaldescriptor_b	@ReplaceString(@Value(component_postaldescriptor_b),/,_)	Rename
component_postaldescriptor_b	component_postaldescriptor_b_#	#_@Value(component_postaldescriptor_b_#)	Rename
component_postaldescriptor_b_#	component{1}.xlink_href	<Enter new value (optional)>	Rename
component_thoroughfarename	component_thoroughfarename_a	@Trim(@Value(component_thoroughfarename_a),https://geo.data.gov.dk/dataset)	Rename
component_thoroughfarename_a	component_thoroughfarename_b	@ReplaceString(@Value(component_thoroughfarename_b),/,_)	Rename
component_thoroughfarename_b	component_thoroughfarename_b_#	#_@Value(component_thoroughfarename_b_#)	Rename
component_thoroughfarename_b_#	component{2}.xlink_href	<Enter new value (optional)>	Rename
component_addressareaname_1	component_addressareaname_1a	@Trim(@Value(component_addressareaname_1a),https://geo.data.gov.dk/dataset)	Rename

See https://github.com/INSPIRE-MIF/gp-geopackage-encodings/pull/24/files for the GPKG-template and the FME workspace.

Test run on 23:33 - 26.10.2022 with test suite Conformance class: Application schema, Addresses and 2 more test suites

Status Failed
Started 26/10/2022 21:33:44 GMT
Duration 3 s

	Total	Count	Skipped	Failed	Warnings	Manual
Test suites	8	0	1	0	3	
Test cases	22	0	1	0	6	
Assertions	54	0	2	0	8	

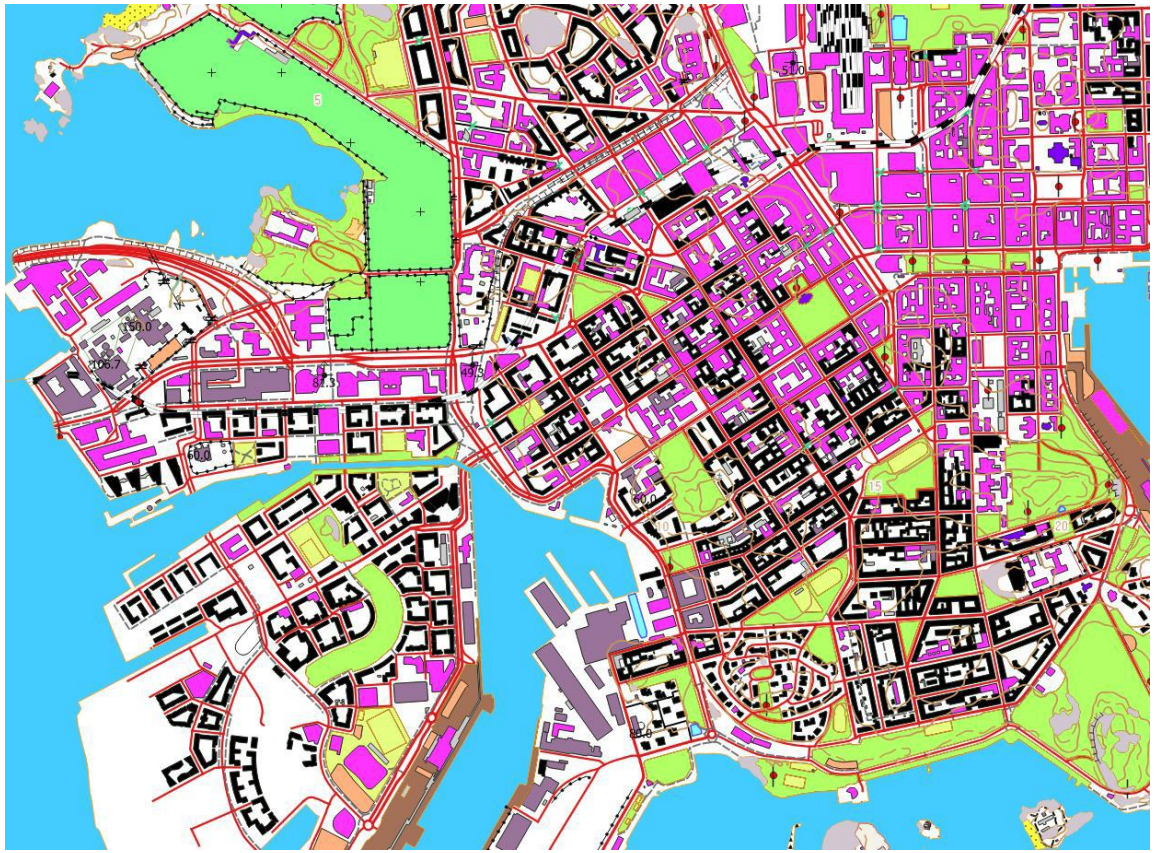
- Conformance class: INSPIRE GML encoding
- Conformance class: INSPIRE GML application schemas, General requirements
- Conformance class: GML application schemas, Addresses
- Conformance class: Application schema, Addresses
- Conformance class: Data consistency, General requirements
- Conformance class: Data consistency, Addresses
- Conformance class: Reference systems, General requirements
- Conformance class: Reference systems, Addresses

XML document 'DK_INSPIRE_AD.gml', Address '_50b921ea-935e-d605-2287-4ee364046795_0a3f508d-d95b-32b8-e044-0003ba298018': A spatial geometry uses an unexpected coordinate reference system 'http://www.opengis.net/def/crs/EPSG/0/25832'. Please refer to table 2 of the data specification for the list of expected coordinate reference systems.

Example: Addresses and Topographic Map Database, Finland

OGC API Features – Encoding in INSPIRE context

Delivering GeoPackage through API with embedded styles



Jari Reini

Example: Addresses, Finland

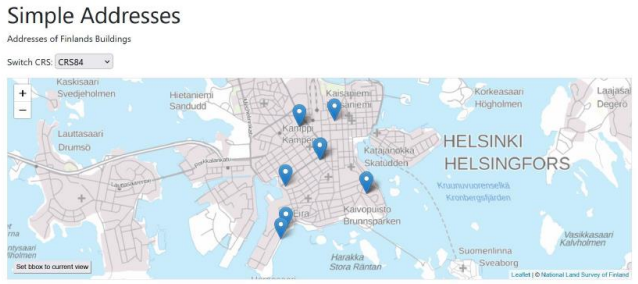
OGC API Features – Encoding in INSPIRE context:

- OGC API Features *usually* supports GeoJSON and HTML.
- Mandatory encodings are not defined.
- MIG has endorsed interface (OGC API Features) as Good Practice
- GeoJSON/Geopackage as an encoding for Download Services
  - Relies on "INSPIRE UML-to-GeoJSON encoding rule" and "Template for Alternate Encodings for INSPIRE Data"
  - Follows principles as in "Best practices for GeoPackage encodings"

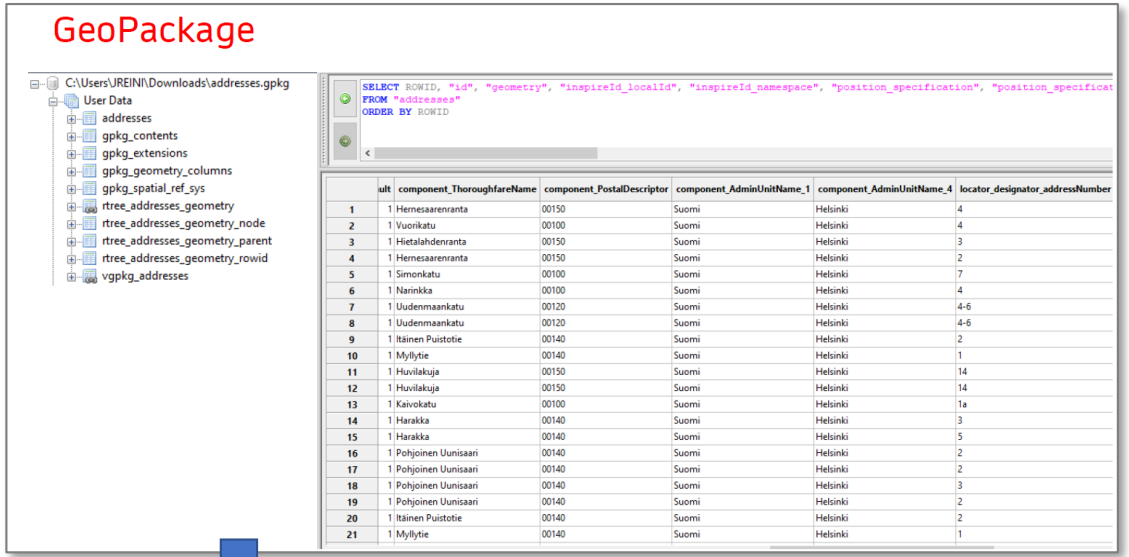
Geopackage via OGC API Features

OGC API Features

.../inspire-addresses/features/v1/collections/addresses/items?bbox=24.928,60.151,24.957,60.172&**f=gpkg**&limit=-1



GeoPackage

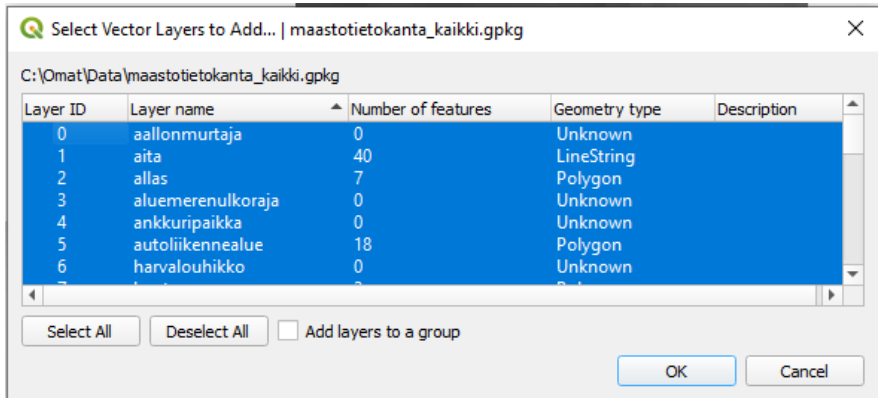


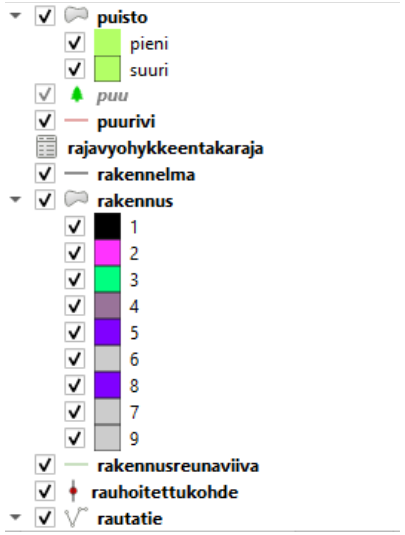
Feature	Value
addresses [4]	
inspireId_namespace	
(Derived)	
(Actions)	
id	281128
inspireId_localId	281128
inspireId_namespace	https://github.com/INSPIRE-MIF/2017.2/GeoJSON/ads/examples/
position_specification	building
position_specification_href	https://inspire.ec.europa.eu/codelist/GeometrySpecificationValue/buildin
position_method	byOtherParty
position_method_href	https://inspire.ec.europa.eu/codelist/GeometryMethodValue/byOtherPart
position_default	true
component_ThoroughfareName	Tehtaankatu
component_PostalDescriptor	00150
component_AdminUnitName_1	Suomi
component_AdminUnitName_4	Helsinki
locator_designator_addressNumber	21
locator_level	siteLevel
locator_level_href	https://inspire.ec.europa.eu/codelist/LocatorLevelValue/siteLevel
building	103103877H
parcel	09100701260021
inspireId_namespace	
inspireId_namespace	
inspireId_namespace	

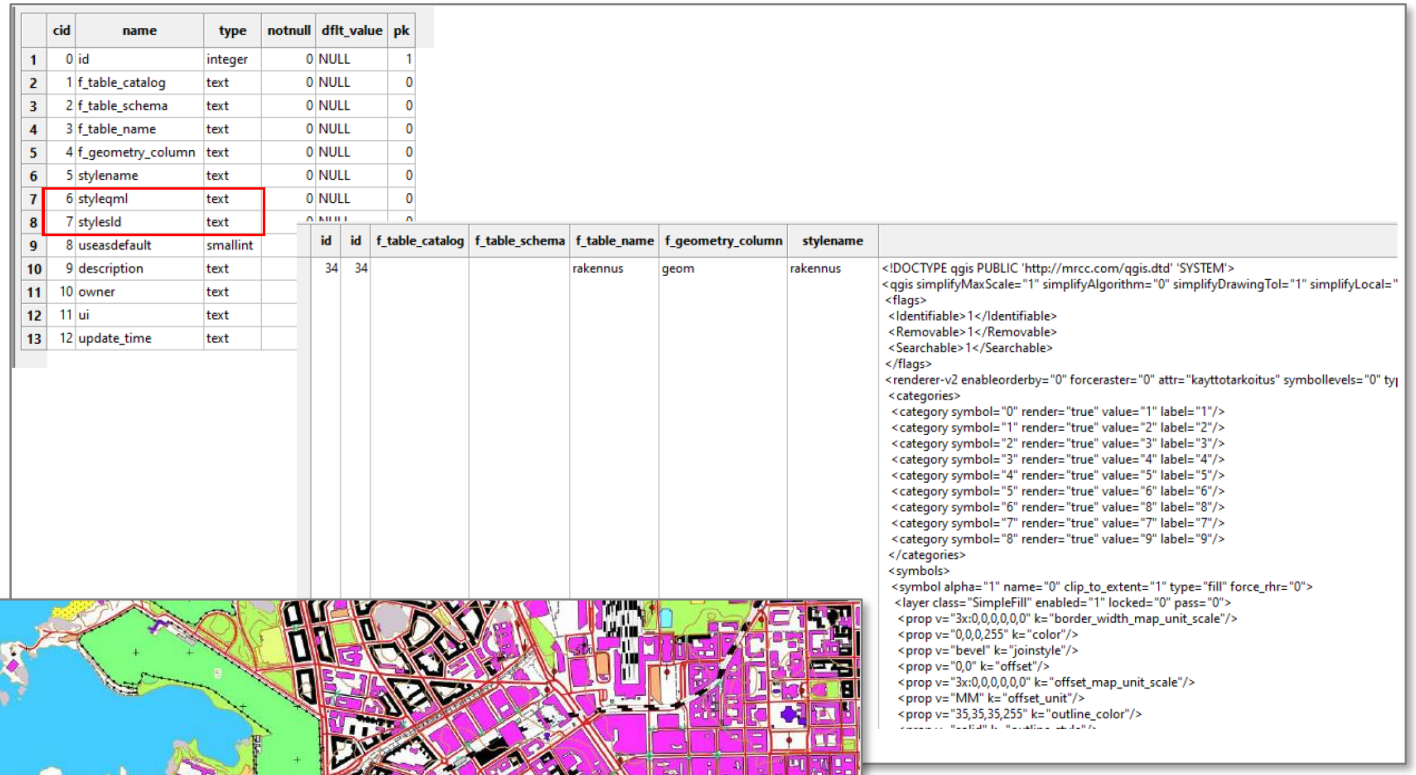
Example: Topographic Map Database, Finland

GeoPackage with styles

- Topographic Map Database of Finland (112 layers)



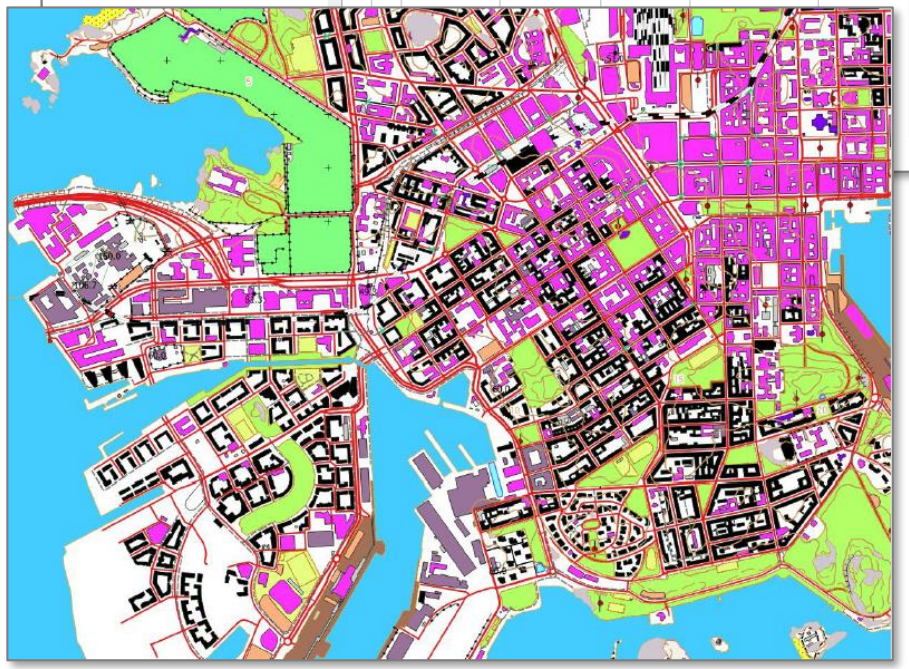




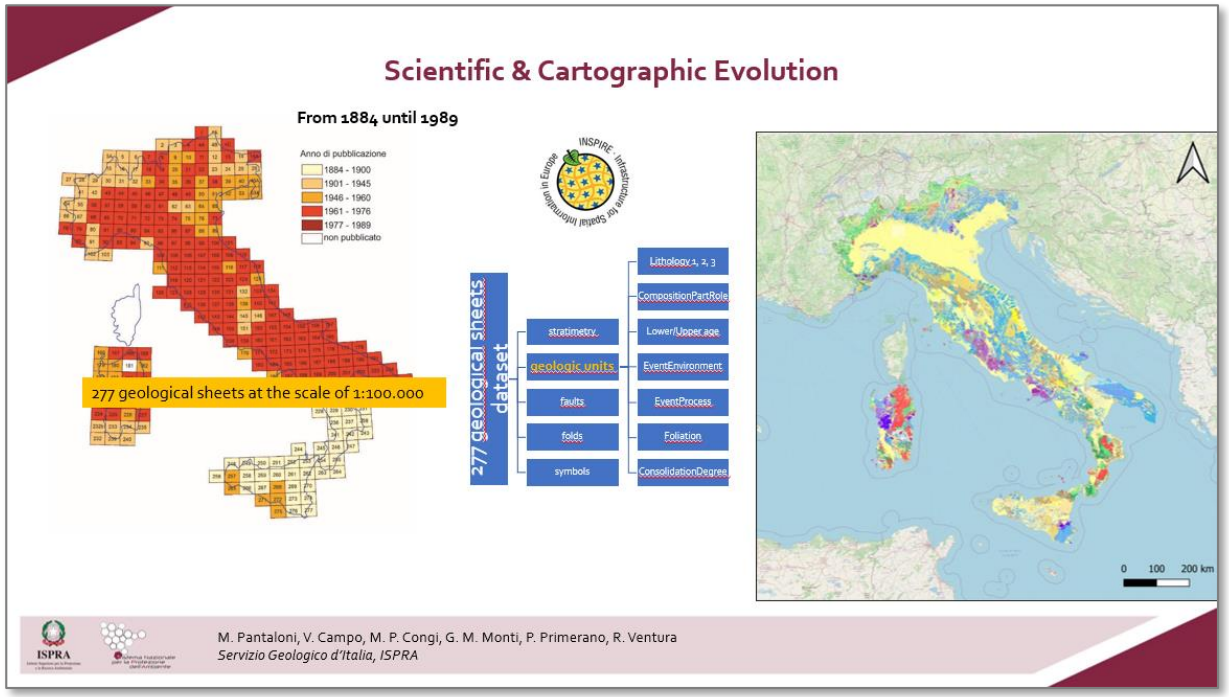
	cid	name	type	notnull	dflt_value	pk
1	0	id	integer	0	NULL	1
2	1	f_table_catalog	text	0	NULL	0
3	2	f_table_schema	text	0	NULL	0
4	3	f_table_name	text	0	NULL	0
5	4	f_geometry_column	text	0	NULL	0
6	5	stylename	text	0	NULL	0
7	6	styleqml	text	0	NULL	0
8	7	stylesld	text	0	NULL	0
9	8	useasdefault	smallint			
10	9	description	text			
11	10	owner	text			
12	11	ui	text			
13	12	update_time	text			

id	id	f_table_catalog	f_table_schema	f_table_name	f_geometry_column	stylename
34	34			rakennus	geom	rakennus

<!DOCTYPE qgis PUBLIC 'http://mrcc.com/qgis.dtd' 'SYSTEM'>
<qgis simplifyMaxScale="1" simplifyAlgorithm="0" simplifyDrawingTol="1" simplifyLocal="
<flags>
<Identifiable>1</Identifiable>
<Removable>1</Removable>
<Searchable>1</Searchable>
</flags>
<renderer-v2 enableorderby="0" forceraster="0" attr="kayttotarkoitus" symbollevels="0" ty
<categories>
<category symbol="0" render="true" value="1" label="1"/>
<category symbol="1" render="true" value="2" label="2"/>
<category symbol="2" render="true" value="3" label="3"/>
<category symbol="3" render="true" value="4" label="4"/>
<category symbol="4" render="true" value="5" label="5"/>
<category symbol="5" render="true" value="6" label="6"/>
<category symbol="6" render="true" value="8" label="8"/>
<category symbol="7" render="true" value="7" label="7"/>
<category symbol="8" render="true" value="9" label="9"/>
</categories>
<symbols>
<symbol alpha="1" name="0" clip_to_extent="1" type="fill" force_rhr="0">
<layer class="SimpleFill" enabled="1" locked="0" pass="0">
<prop v="3x:0,0,0,0,0,0" k="border_width_map_unit_scale"/>
<prop v="0,0,0,255" k="color"/>
<prop v="bevel" k="joinstyle"/>
<prop v="0,0" k="offset"/>
<prop v="3x:0,0,0,0,0,0" k="offset_map_unit_scale"/>
<prop v="MM" k="offset_unit"/>
<prop v="35,35,35,255" k="outline_color"/>

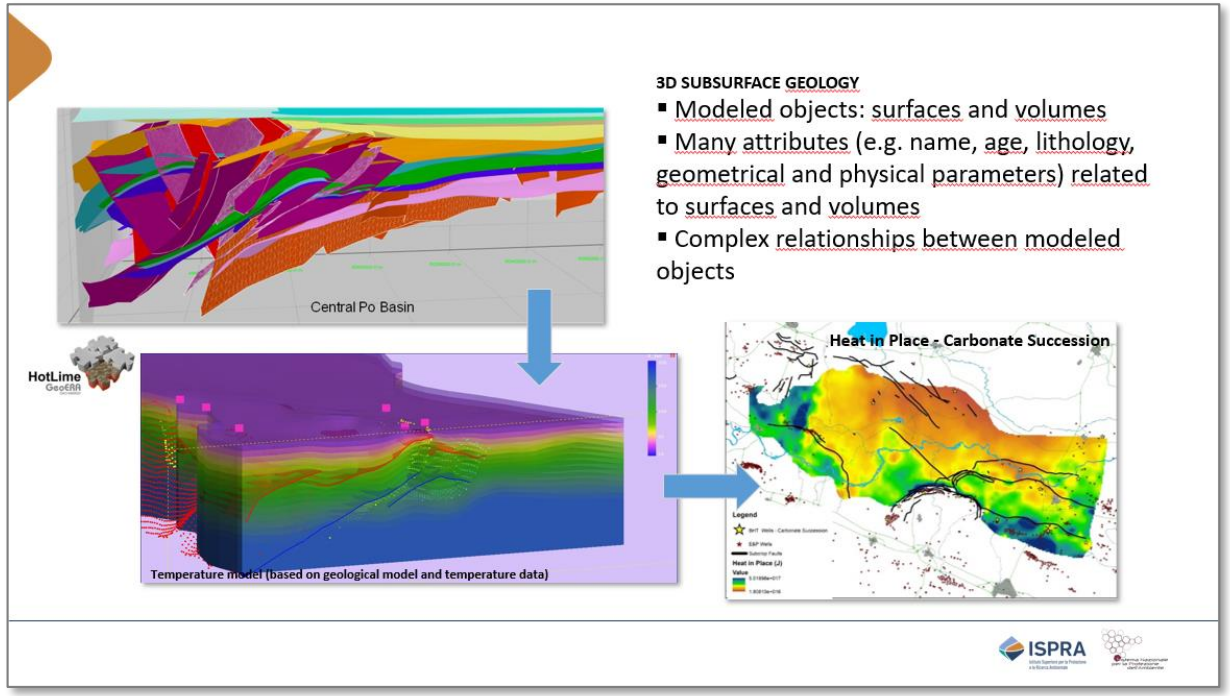


Example: Geology, Italy



GPKG vs GML encoding for the Geological Map of Italy 1:100K

INSPIRE-extended data model and GPKG encoding for interoperable geological subsurface data



Stefania Morrone

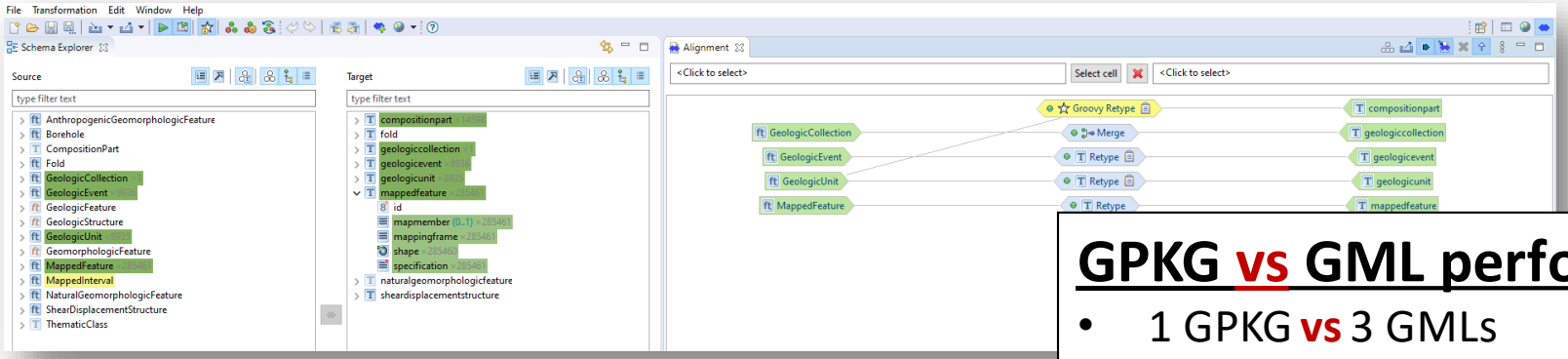
Example: Geology, Italy

GPKG logical model for INSPIRE GE

- process based on hale studio: INSPIRE Geology core XSD → hale schema definition file* → GPKG template

From harmonized GMLs to GPKG dataset

- hale studio project:
  - source schema: Geology XSD → target schema: Geology GPKG template
  - source data: 3 INSPIRE GMLs already published in the Geoportal → output data: 1 GPKG file



GPKG vs GML performances in (Q)GIS enviroment:

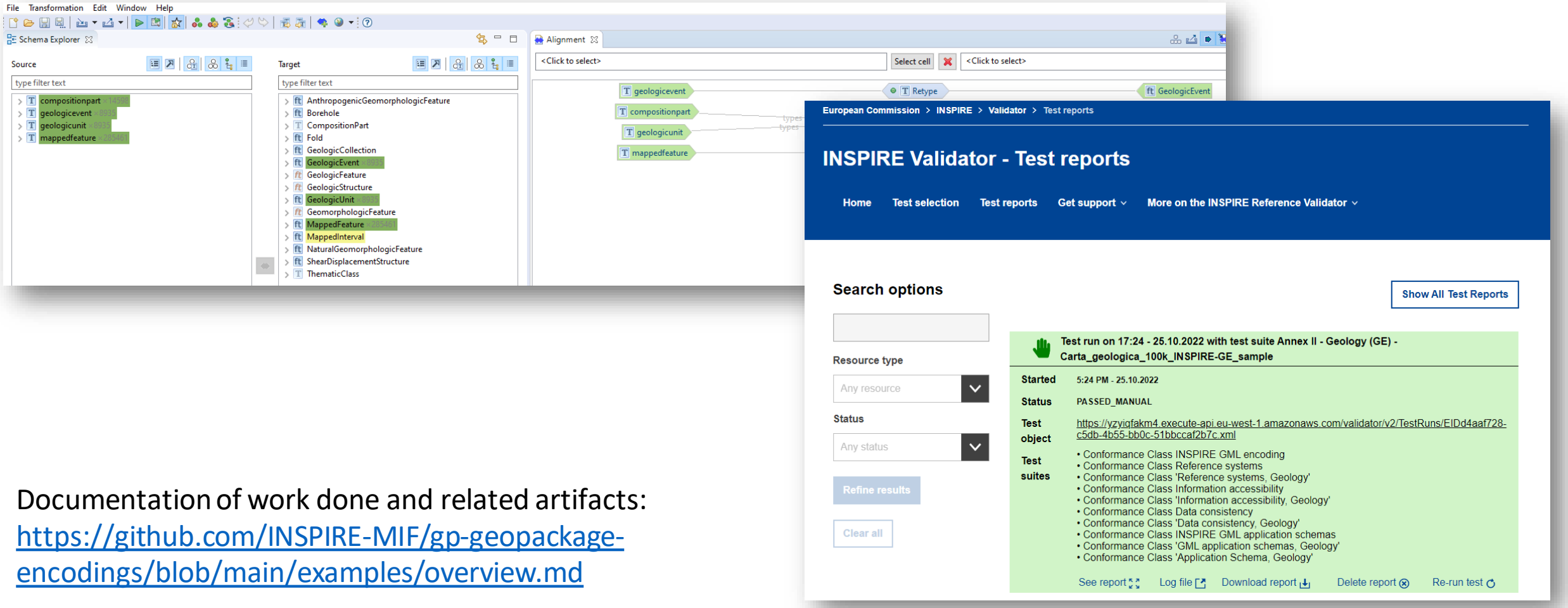
- 1 GPKG **vs** 3 GMLs
- loading times: 3 - 4 seconds **vs** 3 - 4 minutes per single GML file
- possibility to edit and save also portions of the data preserving GPKG structure **vs** no possibility
- possibility to categorize and include style directly in the GPKG + provide views **vs** no possibility

* some manual work is required to comply with GPKG GP specifications

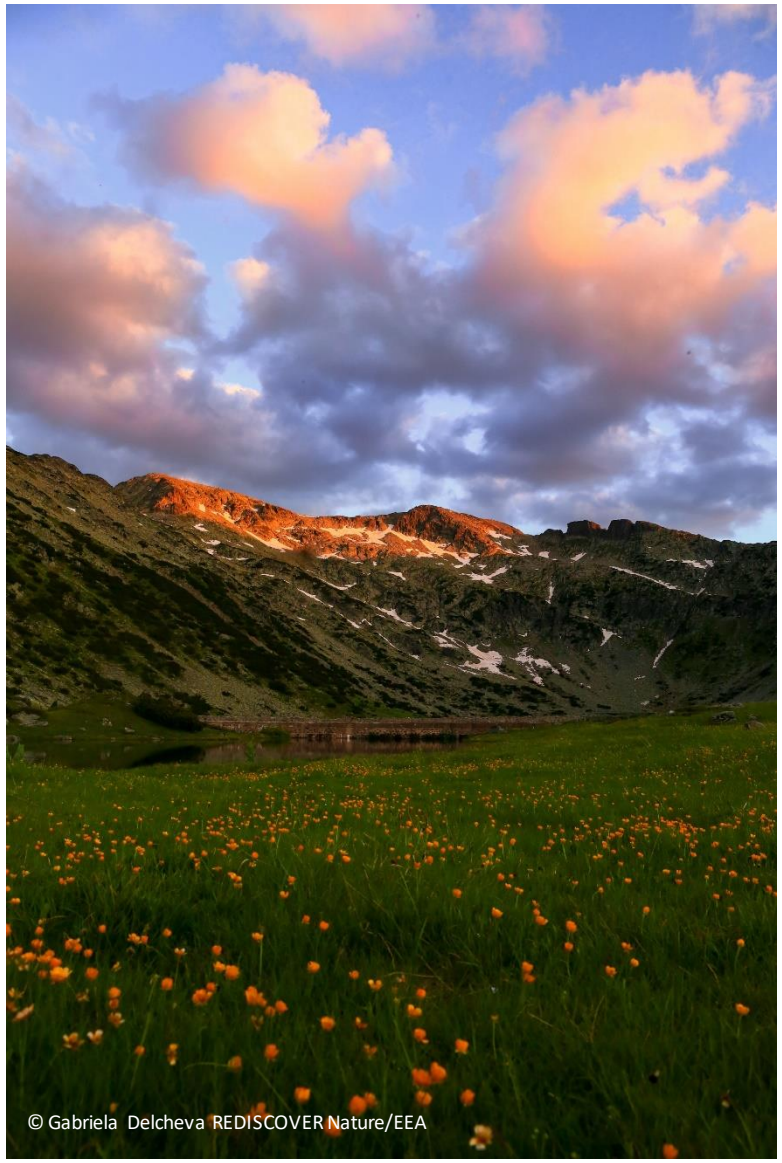
Example: Geology, Italy

Validation: deriving back the INSPIRE GML dataset from the GPKG dataset **without information loss**

- hale studio project:
  - source schema: Geology GPKG template → target schema: Geology XSD
  - source data: GPKG dataset → output data: INSPIRE GML (successfully validated with INSPIRE Validator)



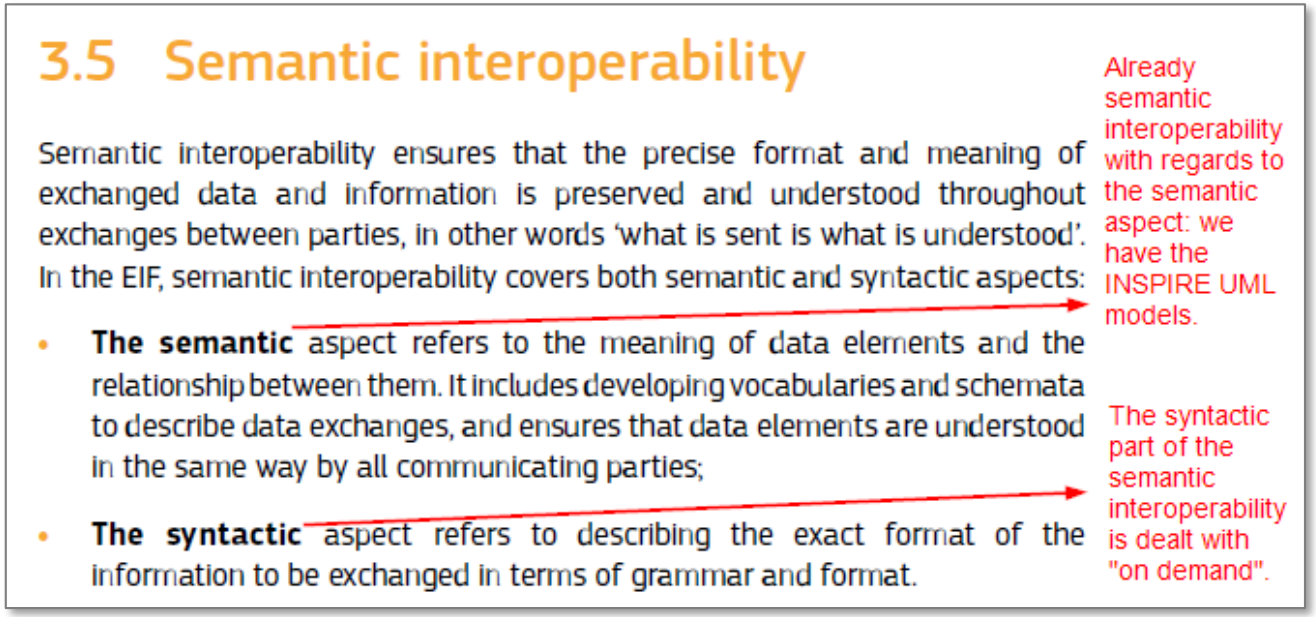
Documentation of work done and related artifacts:
https://github.com/INSPIRE-MIF/gp-geopackage-encodings/blob/main/examples/overview.md



Lessons learned

Lessons learned

- **Theme by theme** implementation based on real use cases
- **Interoperability on demand**
  - Semantic interoperability: based on INSPIRE UML
  - Syntactic: based on GeoPackage standard, transformation rules, use case requirements
- **Re-use** from existing alternative good practices schemas
- **Re-use & extend standard code lists** wherever possible
- **Executable transformation rules** can be re-used (lower transformation or validation burden)



3.5 Semantic interoperability

Semantic interoperability ensures that the precise format and meaning of exchanged data and information is preserved and understood throughout exchanges between parties, in other words 'what is sent is what is understood'. In the EIF, semantic interoperability covers both semantic and syntactic aspects:

- **The semantic** aspect refers to the meaning of data elements and the relationship between them. It includes developing vocabularies and schemata to describe data exchanges, and ensures that data elements are understood in the same way by all communicating parties;
- **The syntactic** aspect refers to describing the exact format of the information to be exchanged in terms of grammar and format.

Already semantic interoperability with regards to the semantic aspect: we have the INSPIRE UML models.

The syntactic part of the semantic interoperability is dealt with "on demand".

- **Use multiple encodings** to best support different applications for your data
- **Validation:** through INSPIRE GML with INSPIRE Reference Validator (executable transformation rules to default encoding, or GeoPackage and INSPIRE GML datasets)

Good Practice requirements and GPKG implementation practice

GOOD PRACTICE REQUIREMENTS

- Logical model (for GeoPackage)
- GeoPackage template
- Executable specification UML-to-GeoPackage or GML-to-GeoPackage
- Executable model for data transformation
- Validation

GPKG IMPLEMENTATION PRACTICE: fulfils requirements

Theme-by-theme approach based on real use cases
Data simplification – focus on key values (e.g. flattening, simplified complex data types)

GeoPackage **template or files**, description part of GPKG or as separate documentation

Applicable **model transformation rules**, re-using model-transformation repository

Executable transformation with ETL tools (e.g. HALE, FME) (automatization)

Implicit compliance: involved through the full process
Validation: generating INSPIRE GML and validation with INSPIRE Reference Validator (central tooling for GeoPackage is not available)

Conclusions

GeoPackage is a mature encoding and suitable for different thematic domains and applications

Proven implementation practices in providing INSPIRE datasets in GeoPackage format

Endorsement of INSPIRE Good Practice for GeoPackage as alternative encoding provides a ground and knowledged for further implementation

Thank you

© Dominika Hrašková, REDISCOVER Nature/EEA