INSPIRE Good Practice: GeoPackage encoding of INSPIRE datasets

Name of the GP

GeoPackage encoding of INSPIRE datasets

Description of the GP

This Good Practice (GP) describes a mechanism to create INSPIRE data sets encoded using the OGC GeoPackage encoding standard. These data sets will be compliant with the INSPIRE Implementing Rules (IR), and technical compliance can be shown through transformation to the default encoding (GML). In this perspective, the GeoPackage can be used both as an additional and an alternative encoding for INSPIRE data sets.

The process underpinning this GP focuses on creating GeoPackage models optimised for usability in GIS environments. It is recognised that GeoPackage is much faster (compared to other encodings such as GML and GeoJSON) in traditional reading, writing and processing operations performed in GIS environments. It can be used on core INSPIRE data specifications, or on extensions, and can also be use-case specific. This means that even for one INSPIRE theme, multiple logical schemas could, over time, become available.

An example is the European Noise Directive (END) use case. For delivery of reporting data for the END, a conceptual model was developed that builds on INSPIRE data specifications, such as Human Health and Safety and Area Management/Restriction/Regulation Zones and Reporting Units. The END conceptual model extends and re-uses feature types from these data specifications and then applies specific model transformation rules to derive a logical model optimised for the END purposes. This optimised model is not a generic model for that INSPIRE data specifications though.

GeoPackage models related to other data themes/use cases can be described in the INSPIRE-MIF [repository for the GeoPackage encoding](#).

For any GeoPackage logical schema to be accepted as specific implementation of this GP, the following resources must be created:

- A description of the logical model, its relevance and expected benefits, its implementations and its limitations compared to the default encoding

- An empty GeoPackage template file

- A formal, executable specification of the *UML-to-Geopackage* model transformation or *the GML-to-GeoPackage* model transformation referencing and conforming to the generic INSPIRE model transformation rules in the INSPIRE-MIF [model-transformation-rules](#) repository. These generic INSPIRE model transformation rules are encoding-agnostic and will/may be underpinning the development of different encodings for INSPIRE data. These will in turn be described in their own, encoding-specific INSPIRE-MIF repositories. The [gp-geopackage-encodings](#) repository is the encoding-specific repository for this GP.

  The executable specification can be delivered in any form that takes a computer-readable model as input and creates a computer-readable model as output. These specifications can be programs. They need to be documented in such a way that the mapping is human-readable. Examples include an annotated ShapeChange configuration, an XSLT, or a hale studio model transformation.

- A formal, executable model for data transformation that allows to transform a data set encoded as GeoPackage to the default encoding. This can be delivered as an ETL workbench (e.g. hale studio transformation project, FME workbench), as a standalone program or as a service. In all cases, the executable model must be documented in a human-readable form.

This Good Practice was developed with a focus on maximal interoperability. It does not rely on existing extensions to the GeoPackage standard, such as the Related Tables extension, or the GeoPackage Metadata Profiles extension. Should these become widely supported later on, this GP can be updated.

INSPIRE component(s)

Data: this GP describes how data falling under any INSPIRE theme can be encoded in GeoPackage format in such a way that the abstract requirements from related specifications are met. The GP relies on INSPIRE code lists, and transformation services.

Note: This document does not describe a formal mechanism to publish data sets as a network service. The default option would be to use a Predefined Download service implemented via OpenSearch/ATOM Feed. Upcoming alternatives, such as STAC or direct file linking from metadata, would also work.

References

Normative reference

- GeoPackage Encoding Standard, v. 1.3.0
- Common Model Transformation Rules

Other references

The GP approach builds on:

- the outcome of MIWP Action 2017.2 on alternative encodings for INSPIRE data. Specifically, it is based on the initial template for proposing new encodings and the work in the INSPIRE UML-to-GeoJSON encoding rule described in the INSPIRE-MIF 2017.2 repository

- the INSPIRE UML-to-GeoPackage encoding rule available in the IAAA-Lab/U2G repository.

The GP has been applied to create the GeoPackage Encoding Rule for Environmental Noise Directive Reporting Data, v. 1.0 (2021-07-30).

Relevance & expected benefits

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

The GeoPackage Encoding Standard has been developed by the Open Geospatial Consortium and is built on SQLite. More specifically, a GeoPackage is the SQLite container and the GeoPackage Encoding Standard governs the rules and requirements of content stored in a GeoPackage container:

- vector features
- tile matrix sets of imagery and raster maps at various scales
- attributes (non-spatial data)
- extensions

Since a GeoPackage is a database container, it supports direct use. This means that the data in a GeoPackage can be accessed and updated in a "native" storage format without intermediate format translations.

GeoPackage is well supported by common desktop, mobile and Web GIS platforms, as well as by ETL tools.

In this GP the transformation rules for GeoPackage templates have been developed for optimal usability and interoperability. They can be visualised and processed directly in any standard GIS environment. It is also possible to derive the data sets in the default encoding (GML) automatically to perform compliance tests.

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).

Intended Outcome

The expected outcome is that the various communities of INSPIRE implementers create logical data models for GeoPackage for a wide range of INSPIRE themes and use cases. To make sure interoperability and compliance are guaranteed, all these implementers rely on the common process and set of rules described in this good practice.

Evidence of implementation & support

The good practice does not require any specific tool support beyond basic support for the GeoPackage specification. GeoPackage is supported by all major GIS tools such as ArcGIS and QGIS, ETL tools such as hale studio and FME, and libraries such as GDAL/OGR. Since the good practice is built without dependencies on any extensions, and takes limitations of individual tools (such as maximum attribute name lengths or dependencies on the presence of spatial indexes) into account, it is compatible with these environments.

The Good Practice also requires the submission of formal model transformation and data transformation rules for each data model. A simple mapping table is not considered sufficient. To create these formal rules, model transformation tools such as ShapeChange as well as data transformation tools such as hale studio and FME can be used.

Concrete implementations

The following concrete implementations of the good practice are known:

- **EEA END reporting** (extension of *Transport Networks*, *Human Health and Safety*, and *Area Management/Restriction/Regulation Zones and Reporting Units* data themes). Guidelines, Video tutorials, Reporting examples and Validation Rules available in dedicated section of the Eionet Portal ; available also pre-defined GeoPackage templates and the END Data model documentation aligning END reporting requirements with the INSPIRE data models and the source for creating GeoPackage encoding rules and templates

- **HaDEA GO-PEG project**: GO-DEPTH use case (extension of *Geology* data theme), building on the EEA END reporting approach and developed in cooperation with ISPRA (Italian Institute for Environmental Protection and Research). Video and presentation used during the GO-PEG workshop "*Keeping the pace with INSPIRE developments: alternative encodings for INSPIRE and new features in hale studio*" available here. A geological 3D model of the Po Basin data set, which implements this GP, is downloadable via OGC API service here

Limitations

This Good Practice only describes the framework in which specific GeoPackage logical models for INSPIRE data sets are created. The theme-specific or use-case specific encoding rules will be made available in the INSPIRE-MIF GeoPackage repository.

Optimal use case: delivery of entire datasets sized medium to large (overhead could affect delivery of small data sets)