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

Irena Mitton, **Dragan Divjak**
LIST geoinformatika d.o.o., Croatia

Contact e-mail: irena.mitton@gmail.com

Abstract:
The fundamental driver for INSPIRE is the need for harmonized spatial data to better support decision making in areas such as environment, sustainable development and disaster response. European National Mapping and Cadastral Agencies - NMCA modify their data according to the INSPIRE requirements in order to be able to use spatial data in identical manner in the entire European Union, thus contributing to the establishment of the European Spatial Data infrastructure – ESDI. Customizing existing data by mapping instead of requiring re-collection by specific rules will achieve interoperability with minimal costs for the Member States of the European Union. Specialized solution that allows simple and fast harmonization of spatial data is referred as spatial ETL (Extract, Transform and Load) tools. FME (Feature Manipulation Engine) is today's most advanced spatial ETL solution. Using FME facilitates harmonizing of spatial data, as well as their integration in the INSPIRE data structure. Thereby there is a complete control to preserve semantics, geometric structure and metadata that ensure overall quality and compliance with INSPIRE Implementing Rules. The harmonization process of spatial data involves five stages: evaluation, assembly, transformation, validation and publication. In the evaluation stage it is essential to have a good knowledge of the existing data models in conceptual and semantic sense. Often data required for specific INSPIRE themes comes from multiple sources in different formats and therefore there is a need for assembly. FME enables automatic conversion of more than 275 different formats: CAD, GIS, vectors, raster, database, text files, XML, network services, LiDAR, 3D and non-spatial source data. The core of the harmonization workflow is the transformation process which reshapes source schema and geometry to match the required destination INSPIRE data structure. Disparate data sources imply different data models and different mapping schemas to be customized to a common destination model. The configuration of mapping scheme includes processes such as creation of new feature classes, new attributes, conditional value mapping, code list mapping as well as geometric transformation. Geometric transformation often includes coordinate system reprojection, type conversion (CAD lines to GIS polygons, non-spatial text coordinates to geometric objects), generalization and interpolation of spatial data. Validation process is essential prior to the publication of spatial data. Validation of schemas, unique ID-s, geometry types and spatial bounds ensures quality requirements of INSPIRE. The presentation will demonstrate capabilities of spatial ETL tool FME in customizing multiple data sources of different complexity according to INSPIRE.

Keywords: INSPIRE, spatial ETL tools, FME, interoperability, harmonization of spatial data