

## TECHNOLOGICAL ASPECT OF SPATIAL DATA HARMONIZATION

Jan Ježek

Západočeská univerzita v Plzni

**Co-authors / Spoluautoři:** Ing. Mgr. Otakar Čerba

### **Abstract:**

Data harmonization is necessary for creating the possibility to combine data from heterogeneous sources (e.g. regional datasets) into integrated, consistent and unambiguous information products (e.g. European datasets). Such datasets can be then easily used with combination with other harmonized data for viewing as well as querying and analyzing. Data harmonization is complex task that has not a universal solution that will cover all possible scenarios. Ideal technical solution (system architecture, software) is always determined by many specific facts like the way how are the original data stored, size of the data and type of harmonization. Ideal solution depends also on current technical solutions used by particular organization and technical competence of IT staff. The aim of this contribution is to describe two options how to handle data harmonization. One of the approach are ETL Tools (Spatial Data Integrator, Hale). Other solution is based on usage of current capabilities provided by relation database management systems -specifically PostgreSQL with PostGIS.

Main focus will be put on performance and requirements of each solution. Technical aspects will be also described from INSPIRE perspective. Both solution will be shown on particular scenarios on real datasets. One of these harmonization scenario is conversion of spatial planning data from Latvian region Zemgale that has been converted into the Corine land cover data model.

## TECHNOLOGICKÉ ASPEKTY ŘEŠENÍ HARMONIZACE PROSTOROVÝCH DAT

### **Abstrakt:**

Harmonizace dat je důležitým krokem pro docílení možnosti kombinovat heterogenní datové sady (např. data na regionální úrovni) do integrované, konzistentní a jednoznačné datové sady (např. data na evropské úrovni). Výsledná data mohou tak být pak jednoduše použita v kombinaci s jinými harmonizovanými daty a to jak pro prohlížení, tak pro dotazování a vytváření analýz. Samotné harmonizování dat je z technického hlediska velice komplexním problémem, jehož řešení není nikdy zcela univerzálně aplikovatelné. Důležitým faktorem vhodného technologického postupu (volby softwaru, architektura řešení) je způsob uložení výchozích dat, velikost dat a typ harmonizace apod. Ideální řešení taktéž podmiňují stávající zavedené postupy a použité technologie organizace, stejně jako kompetence jejich technických pracovníků. Cílem příspěvku je popsat základní dva způsoby možné harmonizace dat. Jedním z popsanych přístupů je postup harmonizace pomocí ETL nástrojů (Hale, Spatial Data Intergrator). Dalším příkladem je řešení založené na možnostech současných relačních databází, konkrétně databáze PostgreSQL s rozšířením PostGIS. Důraz bude kladen především na výkonnost jednotlivých řešení, ale také na náročnost obsluhy těchto systémů. Řešení budou taktéž zasazeny do kontextu s direktivou INSPIRE.

Představené řešení pro harmonizaci bude demonstrováno na konkrétním modelovém případě. Tímto případem je harmonizace dat územního plánování v lotyšském regionu Zemgale, přičemž tato data byla harmonizována s datovým modelem Corine Land Cover.

Autor nedodal plný text příspěvku.

Author did not supply full text of the paper/poster