Information Around Us: Questions Connected to Information and Data Heterogeneities in Planning Activities

Otakar Čerba
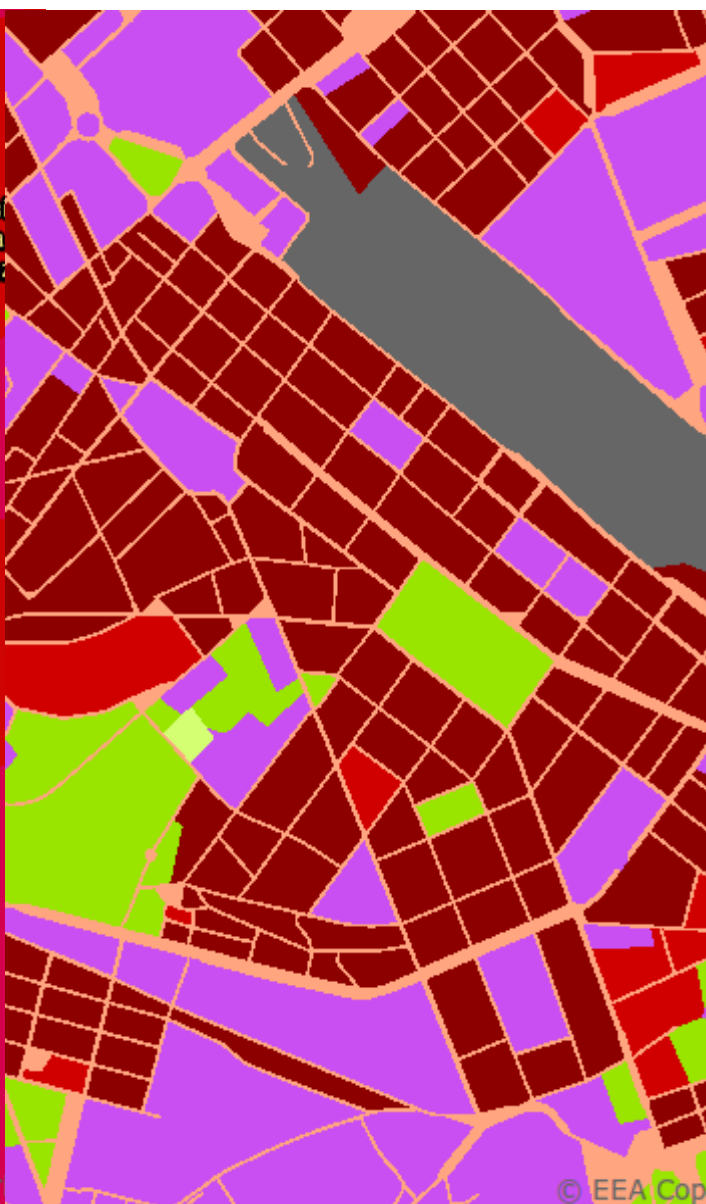
University of West Bohemia
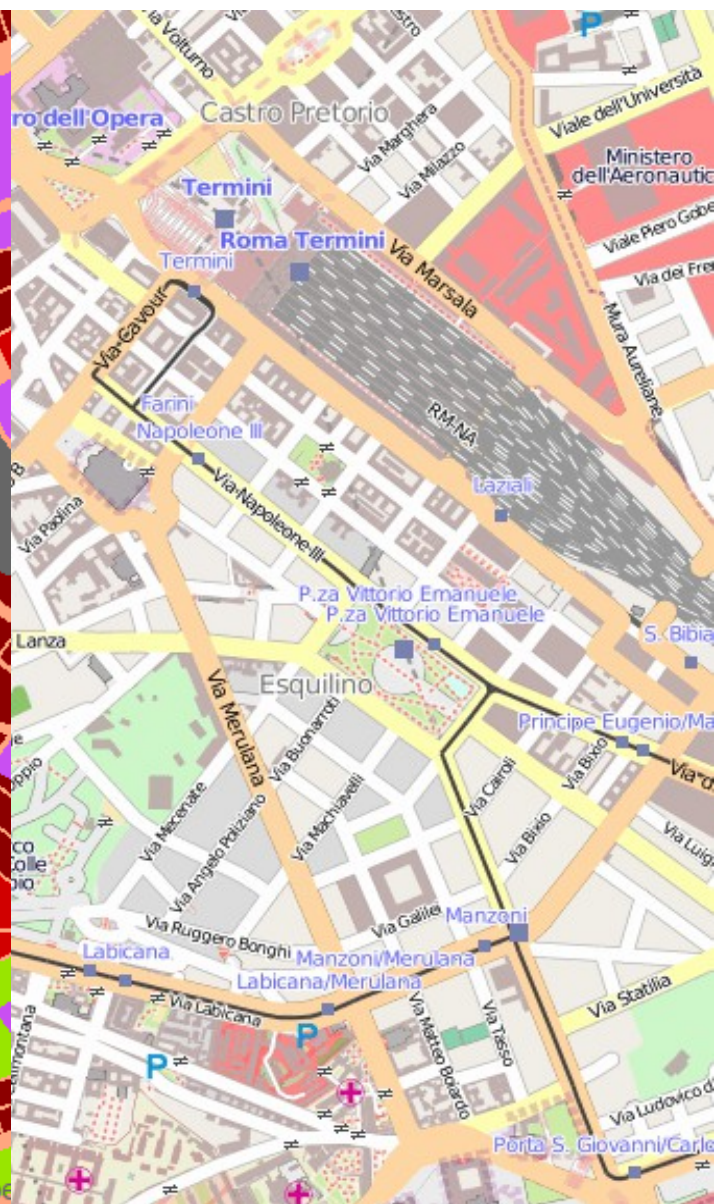Plzeň, Czech Republic

Case study – Spatial Data

CORINE land cover

Urban Atlas

OpenStreetMap

Why CLC, UA and OSM?

Combination of land use and land cover data
Well standardized or documented
Accessible data

| CORINE land cover	Urban Atlas	OpenStreetMap
Pan-European data Respected nomenclature Seamless	Pan-European data Detailed Extension of CLC nomenclature	Global data Updated
Quite old data Not very detailed	Only urbanized areas Discrete regions Quite old data	Various quality and another data properties Complicated tag system

Why spatial data harmonization?

Theme	Land use / land cover	Land use / land cover	All spatial features
Provider	European Environment Agency	European Environment Agency	OpenStreetMap Foundation
Geometry	Areas (in vector version)	Areas	Points (nodes), lines (ways), areas(or polygons as closed ways)
Nomenclature	CLC nomenclature; three-levels hierarchy containing code numbers and explanations	CLC nomenclature; four-levels hierarchy containing code numbers and explanations	Tag containing key (type of element) and value (detail specification of key)
Original format	GeoTIFF	ESRI shapefile	.osm (XML based format)
Granularity	100 m resolution	100 times higher than CORINE land cover	Depend on each contributor
Updating / Last version	Temporal coverge: 2006; Last upload 7 May 2012	Temporal coverge: 2005-2007; Last upload 28 May 2010	Every day (but not systematic, it depends on an activity of particular contributors)
Download	http://www.eea.europa.eu/data-and-maps; including seamless vector versions in ESRI shapefile format	http://www.eea.europa.eu/data-and-maps; ZIP archives containing ESRI shapefile data	http://wiki.openstreetmap.org/wiki/Download; including links to external sources containing ESRI shapefile data
Licence	EEA standard re-use policy: unless otherwise indicated, re-use of content on the EEA website for commercial or non-commercial purposes is permitted free of charge	EEA standard re-use policy: unless otherwise indicated, re-use of content on the EEA website for commercial or non-commercial purposes is permitted free of charge	Open Data Commons Open Database License (data); Creative Commons CC-BY-SA (maps)

Spatial data harmonization
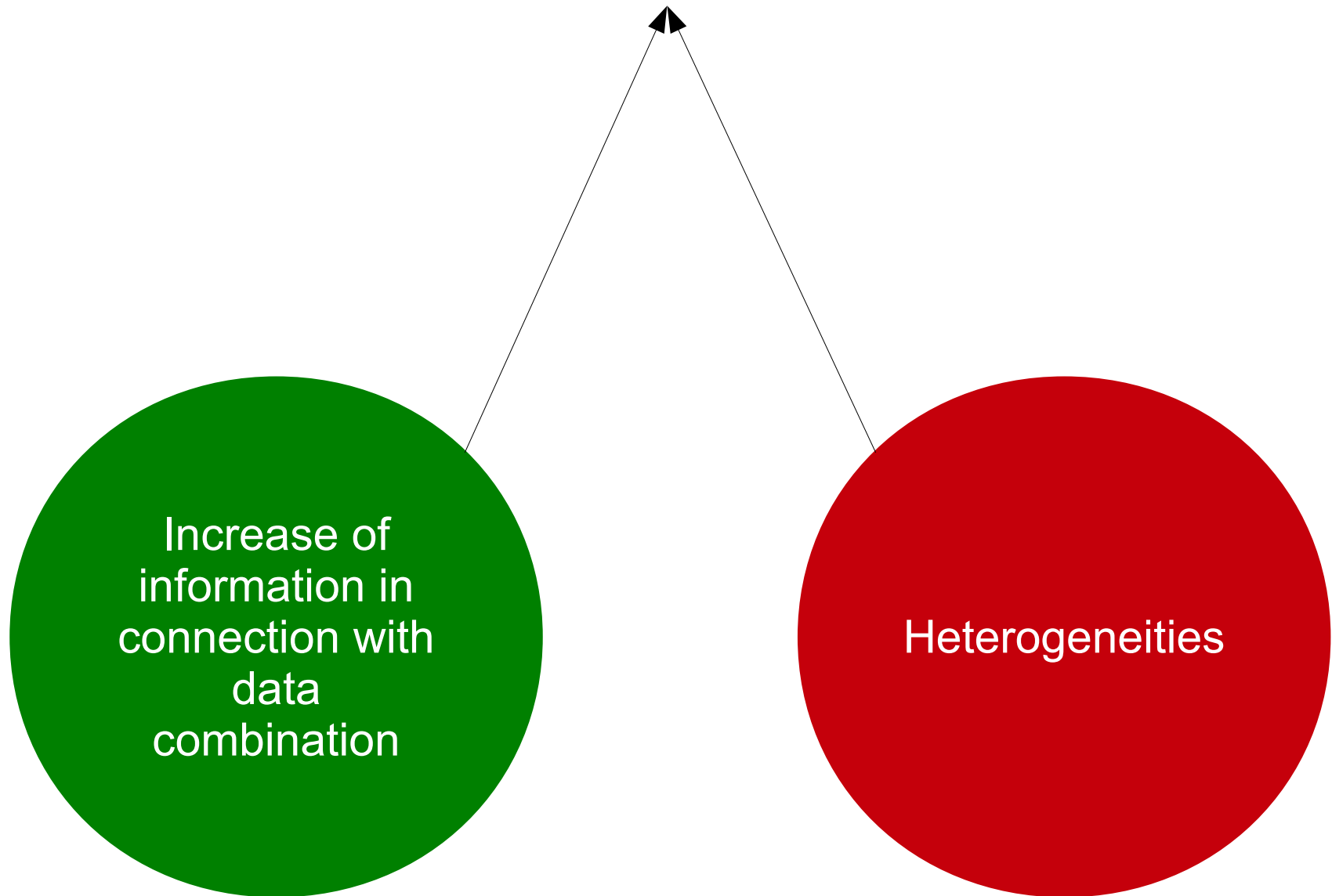
Spatial data harmonization

Activity for elimination or reduction of heterogeneities of various properties of spatial data to support interoperability.

Spatial data harmonization process

(1) CLC as fundamental reference data set

(2) To add on UA elements to CLC data

(3) To replace CLC elements that are covered or overlapped by UA elements

(4) To decide if OSM in particular locality is appropriate to improve data

(5) If yes, to realize steps 2 and 3 with OSM data

(6) To transform nomenclature to the CLC taxonomy with UA classes for urbanized areas

Harmonization requirements

- Development of common data model (id, geometry, nomenclature)

- Transformation OSM elements (nodes, ways, closed ways) to polygons

- Matching table or ontology to harmonize nomenclatures

- Harmonization tools (RDBMS, GIS, ETL...)

Harmonization & Ontology
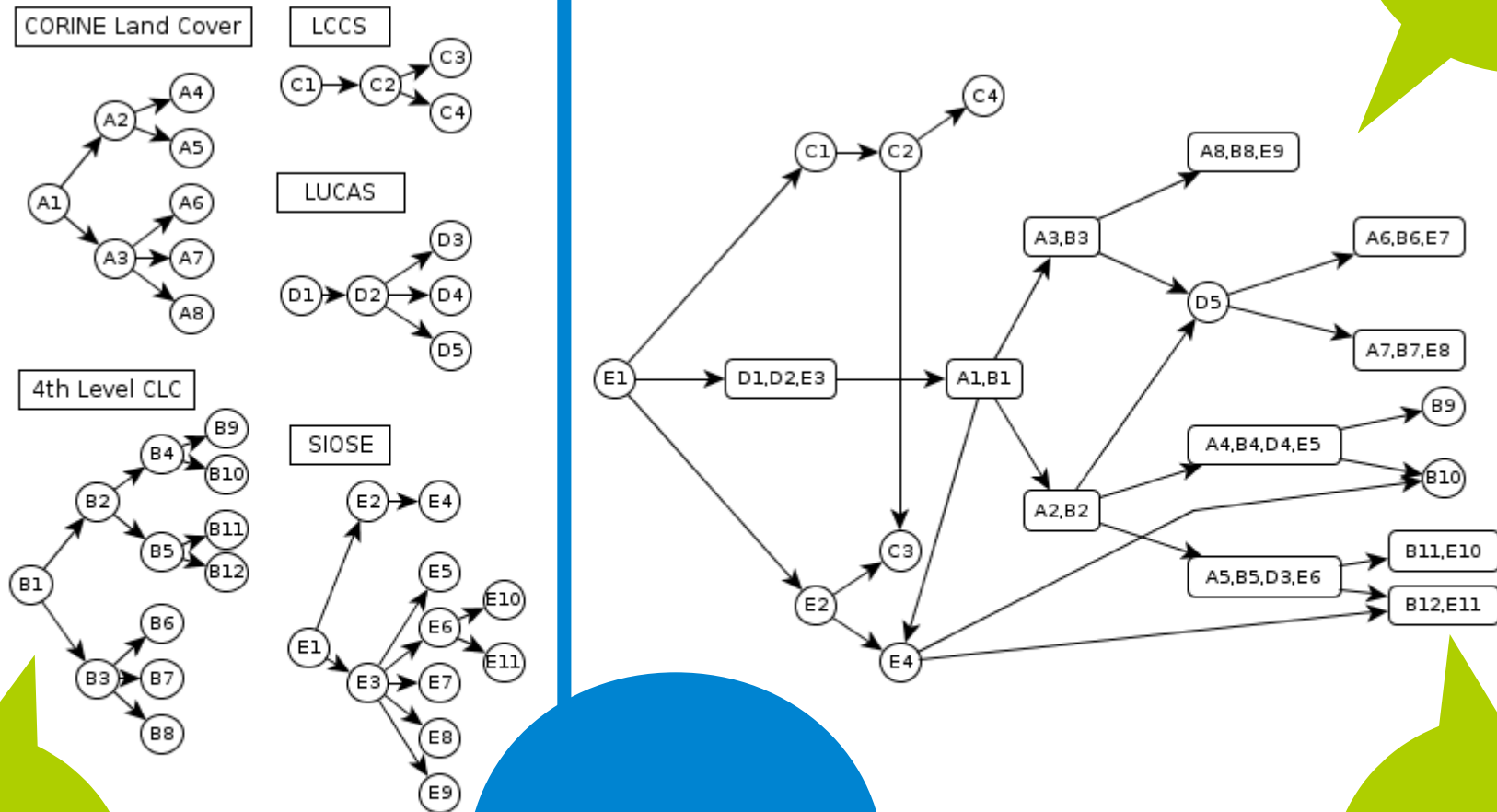
Harmonization results

- Very heterogeneous data set or map

- Maximal utilization of advantages of source data sets

- The most detailed and updated data in particular localities

- Possibility of addition new data sources

Opening question

...much information around us, but users feel a lack of information and call for new data collection...

Why?

Reasons & Solutions

- Self-standing, isolated products → Open Linked Data

- Little information about data → metadata, more comfort metadata catalogues

- Conditions of use → clear licences, open-source, CC...

- Exchange format → simplicity and standardization (e.g. KML)

- Terminology → controlled vocabularies, gazetteers, thesauruses

- Low attractiveness → new forms of data presentation and commercialization

Thank you for your attention

http://cz.linkedin.com/in/otakarcerba/