Model Generalisation in the Context of National Infrastructure for Spatial Information

Tomas MILDORF and Vaclav CADA, Czech Republic

Key words: NSDI, INSPIRE, model generalization, cadastre, spatial planning

SUMMARY

Human activity is one of the most important factors that drive the changes of our living environment. Sustainable development aims to eliminate the negative activities and to support the positive ones in a long term approach. The role of infrastructure for spatial information and the advantage of its use to support sustainable development have been demonstrated in a range of projects and studies. Many specialisms including policy, economics, industry and environment profit from the extensive use of infrastructure for spatial information and its components. These are metadata, spatial data sets and spatial data services including the network services and technologies; agreements on sharing, access and use; and coordination and monitoring mechanisms, processes and procedures.

This contribution introduces use case from the Czech Republic where model generalisation should play a crucial role – cadastral data model. In the context of the INSPIRE (Infrastructure for Spatial Information in the European Community) and related initiatives it is shown how model generalisation can contribute to better exploitation of cadastral data in connection with spatial planning and other activities. The stress is on model generalisation of cadastral data model into middle scale data model. The aim is to present the ideas of the authors to address an important aspect of the infrastructure for spatial information – model generalisation.

Model Generalisation in the Context of National Infrastructure for Spatial Information

Tomas MILDORF and Vaclav CADA, Czech Republic

1. INTRODUCTION

"An idea is always a generalization, and generalization is a property of thinking. To generalize means to think." Hegel, Georg

Human activity is one of the most important factors that drive the changes of our living environment. Sustainable development aims to eliminate the negative activities and to support the positive ones in a long term approach. The role of infrastructure for spatial information and the advantage of its use to support sustainable development have been demonstrated in a range of projects and studies; e.g. (Corbin 2008) or (ACIL Tasman 2008). Many specialisms including policy, economics, industry and environment profit from the extensive use of infrastructure for spatial information and its components.

This contribution introduces use case from the Czech Republic where model generalisation should play a crucial role – cadastral data model. In the context of the INSPIRE (Infrastructure for Spatial Information in the European Community) and related initiatives it is shown how model generalisation can contribute to better exploitation of cadastral data in connection with spatial planning and other activities. The stress is on model generalisation of cadastral data model into middle scale data model. Model generalisation is an important aspect of the infrastructure for spatial information. The study should contribute to long-term research goal of generalisation which is formalising the generalisation problem. As described by (van Oosterom 2009)) the generalisation problem is specifying user needs from the geo-information producers and geo-information users points of view.

2. INSPIRE & CADASTRAL PARCELS

INSPIRE (European Parliament 2007) is one of the most important current initiatives on the EU level. It sets the legal framework for establishment of the Infrastructure for Spatial Information in the European Community (INSPIRE). The main purpose is to help Community environmental policies and policies or activities which may have an impact on the environment.

But there are also other initiatives; INSPIRE is not the only one. GEOSS (Global Earth Observation System of Systems) is orchestrating systems for monitoring and forecasting changes in the global environment. Service oriented GMES (Global Monitoring for Environment and Security) provides reliable and up-to-date information on how our planet and its climate are changing. SEIS (Shared Environmental Information System) concerts reporting obligations of the EU Member States to avoid duplication of information collection. All of them are interconnected and complementing each other up to a certain level. The aim of all these activities is to serve users spatial and also non-spatial information with required quality, thematic content and level of detail. The information must be up-to-date and provided in a short time through modern technologies including web services.

The INSPIRE framework is based on the following building blocks: metadata; spatial data sets and spatial data services; network services and technologies; agreements on sharing, access and use; and coordination and monitoring mechanisms, processes and procedures. By defining these components, the mechanisms for sharing and exchange of information will be secured.

Cadastral data play crucial role in the implementation of INSPIRE. Cadastral parcels, as one of the spatial data themes in scope of INSPIRE, is considered as a reference layer for other environmental data. Reference data were defined by (DPLI Working Group 2002) as series of datasets that everyone involved with geographic information uses to reference his/her own data as part of their work. It provides also a common link between applications and thereby provides a mechanism for the sharing of knowledge and information amongst people. Reference data must fulfil three functional requirements:

- provide an unambiguous location for a user's information;
- enable the merging of data from various sources;
- provide a context to allow others to better understand the information that is being presented.

The (INSPIRE Thematic Working Group Cadastral Parcels 2010) defines reference data as data that constitute the spatial frame for linking and pointing at other information that belong to specific thematic field.

3. DATA QUALITY

The generation of complex infrastructure for spatial information is subjected to well-defined data from various sources, institutions or even by various geosciences. The data are essential for decision-making processes in various specialisms including policy, economics, industry and environmental planning. The quality of decision-making processes is highly dependent on the quality of geodata and geospatial information. The International Organization for Stardardization sets out the following data quality elements (International Organization for Standardization 2002) that shall be used to describe how well a dataset meets the criteria set forth in its product specification:

- completeness: presence and absence of features, their attributes and relationships;
- logical consistency: degree of adherence to logical rules of data structure, attribution and relationships (data structure can be conceptual, logical or physical);
- positional accuracy: accuracy of the position of features;
- temporal accuracy: accuracy of the temporal attributes and temporal relationships of features;
- thematic accuracy: accuracy of quantitative attributes and the correctness of non-quantitative attributes and of the classifications of features and their relationships.

4. GENERALISATION

The International Cartographic Association defined in 1973 generalisation as “the selection and simplified representation of detail appropriate to scale and/or the purpose of a map”. Another definition by (McMaster & Shea 1992)) says that “digital generalization can be defined as the process of deriving, from a data source, a symbolically or digitally-encoded

cartographic data set through the application of spatial data and attribute transformation".
Generalisation in terms of spatial information can be described as a process responsible for generating visualisations or geographic databases at a coarser level of detail than the original source database, while retaining essential characteristics of the underlying geographic information (Sester et al. 2009). There were many attempts to improve generalisation processes and to integrate them into automated systems in the last decades. Generalisation has created many innovative solutions due to research activities performed in this field.

5. MODEL GENERALISATION

Several theoretical generalisation frameworks have been described in literature within last 40 years. They include division of generalisation, terminology, description of generalisation processes and other aspects. The generalisations concepts differ in these aspects.
There are two main generalisation approaches - cartographic generalisation and model (database) generalisation. While cartographic generalisation is focused on the depiction of objects in map (graphical and visualisation aspects), model generalisation aims at deriving model of lower level of detail that can be used for specific purposes. However, there is not clear distinction between cartographic and model generalisation.
In some literature can be found also object generalisation. It comprises generalisation process while defining and building original database (primary model) of the real world.
"Model generalization as a process of geospatial abstraction aims at modelling real world at different levels of semantic and geometric complexity" ((Basaraner 2002).
The role of model generalisation is depicted in Figure 2.

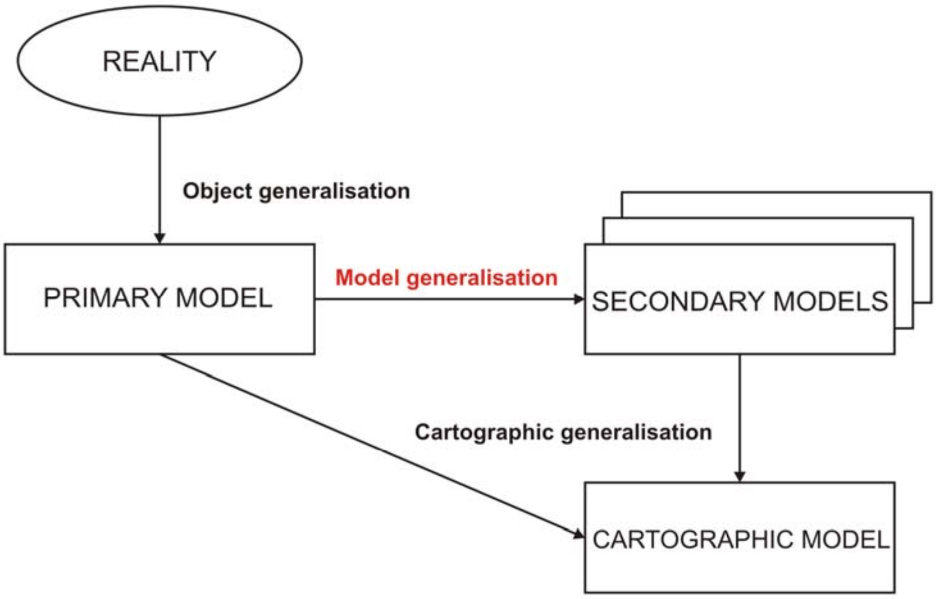

Fig. 2. Role of model generalisation (Gruenreich 1992).

6. GENERALISATION IN THE SCOPE OF INSPIRE

Generalisation in terms of spatial information can be described as a process responsible for generating visualisations or geographic databases at a coarser level of detail than the original source database, while retaining essential characteristics of the underlying geographic information (Sester et al. 2009). There were many attempts to improve generalisation processes and to integrate them into automated systems in the last decades. Generalisation has created many innovative solutions due to research activities performed in this field.

INSPIRE addresses the issue of generalisation but it is not in its main scope. Generalisation and its involvement in the Infrastructure for Spatial Information in the European Community were discussed during the Vienna meeting of the Network Services Drafting Team on 19-20 April 2006. The initial idea was to make generalisation as a part of the INSPIRE data transformation services. Because of the immaturity of generalisation the Network Services Drafting Team did not recommend the use of generalisation services, but supported multiple-representation (INSPIRE Drafting Team "Data Specifications" 2008b).

A workshop on multiple-representation and data consistency was held in JRC Ispra on 7-8 November 2006. The main objective was to provide input to the INSPIRE Implementing Rules in the field of data harmonisation. The following recommendations from the workshop should be mentioned:

- “Automatic generalisation methods are not mature enough to be considered as a service in the ESDI. For practical reasons multiple scale representation is generally needed, which can be completed by generalisation.
- Establishing and/or preserving links between different representations contributes to update propagation, thus to data consistency.
- There are different modelling approaches for MRDB (Multiple Representation Databases) that are usually used for linking different Levels of Detail.
- If required, correspondences between the databases can be established by various tools of data matching (data mining, ontologies and formal specifications) and transformations (conflation, generalisation, matching geometries)” (INSPIRE Drafting Team "Data Specifications" 2008b).

As a result Multiple-representations became one of the Data interoperability components (INSPIRE Drafting Team "Data Specifications" 2008a) listed in Figure 1 (component ‘R’) which describes best practices for how data can be aggregated:

- Across time and space;
- Across different resolutions (“generalisation” of data).

Such aggregation processes are used in particular to create the following results:

- Multiple representations
- Derived reporting (example: typically water samples at 1 km intervals are reported to the European level)

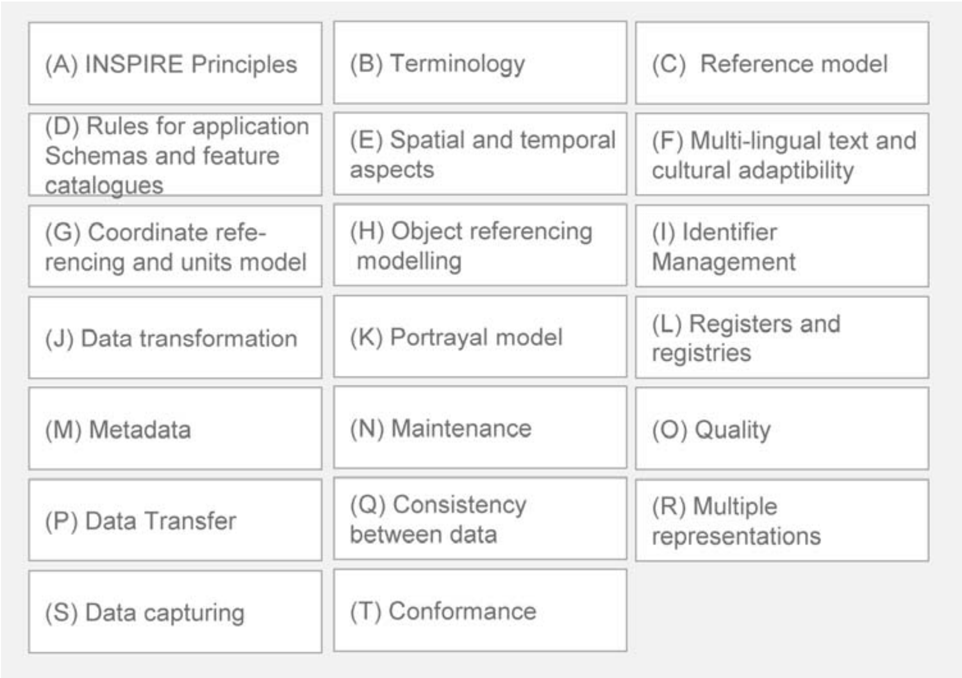

Fig. 1 INSPIRE Data interoperability components (INSPIRE Drafting Team "Data Specifications" 2008a)

7. CASE STUDY

The section of Geomatics[1] at the University of West Bohemia in Pilsen contributes to the development of infrastructure for spatial information and its components. Long-term research and development activity of the section of Geomatics is the inclusion of cadastral data in the National Infrastructure for Spatial Information and their better exploitation for various activities including spatial planning.

In the Czech Republic cadastral data could be used as one of the base datasets forming the National Infrastructure for Spatial Information. This would ensure consistency and improve integration with datasets of lower levels of detail. It is important to note that it is the only database containing cadastral parcels together with ownership data which plays a crucial role in decision making on all governmental levels (local, regional and national). Cadastral data are captured and maintained at the highest level of detail, covering the entire territory of the Czech Republic. It is a unique data set in terms of completeness, logical consistency and positional, temporal and thematic accuracy.

The exploitation of this unique large scale dataset should be supported by model generalisation to enable integration with other datasets, update of datasets of lower level-of-

[1] Geomatics, as defined by the *Canadian Institute of Geomatics,* is a field of activity which, using a systematic approach, integrates all the means used to acquire and manage spatial data required as part of scientific, administrative, legal and technical operations involved in the process of production and management of spatial information. These activities include, but are not limited to, cartography, control surveying, digital mapping, geodesy, geographic information systems, hydrography, land information management, land surveying, mining surveying, photogrammetry and remote sensing.

detail and using cadastral layer as a reference layer for various purposes. Several attempts were already made to delimitate spatial objects from cadastral datasets. As a result a simple feature catalogue was created (see Fig. 2). The catalogue forms simple features out of the subjects of content of digital cadastral map.

CODE								Content of Digital Cadastral Map	Geographic Feature
Map Symbol	Description	Type of land	Land Use	Building Type	Building Use	Type of the Real Estate Protection	Form of the Real Estate Protection		
2.18								inner line	
4.02				1-4	6 7 8 9 10 11 12 13			mansonry, concrete or iron building	building
4.03					15 16 17 18 19			wooden building	
2.19								parcel border	
2.01								state border	territory of the Czech Republic
2.03								regional border	regional territory
2.04								disctrict border	district territory
2.05								municipal border	municipal district
2.06								cadastral disctrict border	cadastral district

Fig. 2 Example of simple feature catalogue for cadastral dataset (Čada & Mildorf 2005)

Further steps are to define semantic, topology and geometry criteria for data model in lower level-of-detail. The use of model generalisation defining the relations between the two different levels-of-detail is essential.

An example of the use of cadastral data as reference layer is spatial planning. Many issues that are being solved by spatial planners include the legal aspects of the territory being discussed. In particular, the issues of ownership of land and other rights and restrictions.

The INSPIRE Directive addresses spatial planning through spatial data themes, primarily through the spatial data theme Land Use. This theme is defined as territory characterised according to its current and future planned functional dimension or socio-economic purpose (e.g. residential, industrial, commercial, agricultural, forestry, recreational). Spatial planning

is becoming part of the INSPIRE and interoperability on all levels is required.
Spatial planning acts in bottom-up and top-down directions between all levels of government. The situation is complicated by the diversity and overall complexity of spatial planning. Spatial planning is a holistic activity. All the tasks and processes must be solved comprehensively with input from many various sources. Several authorities are in charge of single spatially relevant topics (e.g. water management, transport, cadastre, geology, etc.). There is a big diversity in data collection, storing, processing and provision. To combine these sources, to perform an analysis and to ensure valuable results are big challenges in spatial planning (Mildorf 2009). Cadastral data should become a reference layer for the creation of spatial plans and other input data for spatial planning processes. It would ensure the compliance of the legal aspects of land and the spatial plan itself. The schema in Fig. 3 shows the role of model generalisation.

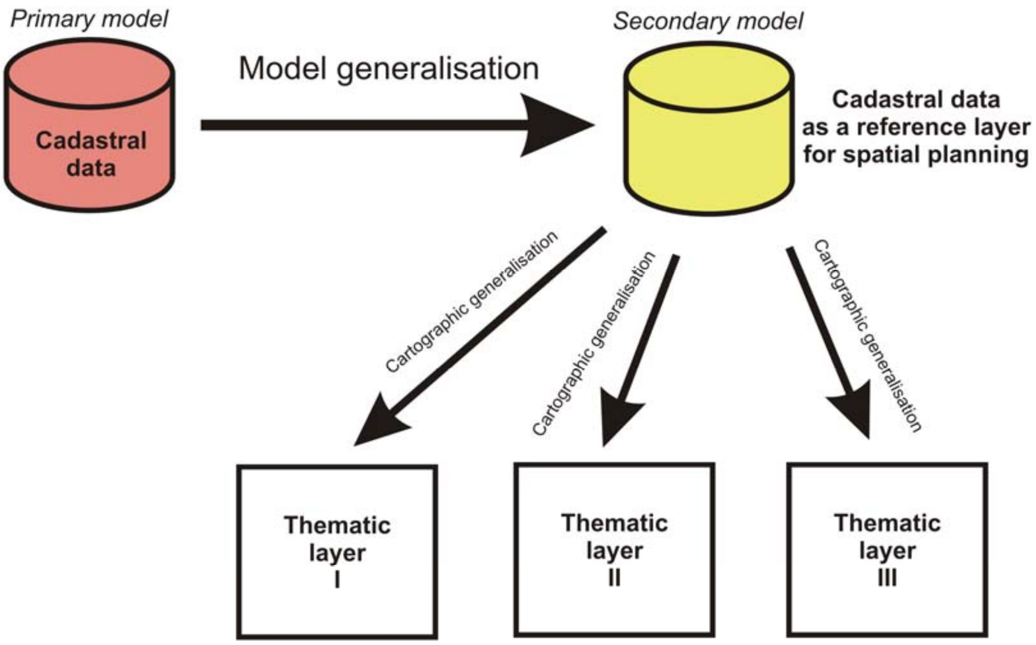

Fig. 3 Simplified schema and the role of model generalisation in the overall process.

8. CONCLUSIONS AND FUTURE PLANS

This contribution introduces use case from the Czech Republic where model generalisation plays a crucial role – cadastral data model. The context of the INSPIRE and related initiatives is to stress that generalisation plays important role in building Infrastructure for Spatial Information. INSPIRE Directive is a European initiative which has mainly top-down approach. It is a legislative which is transposed and then implemented in Member States on national, regional and local level. The use cases on which the INSPIRE Directive is based on are not always from real situation. The involvement of users and stakeholders and their feedback are therefore essential for further developments.
The Czech use case takes a bottom-up approach and studying the Czech use case contributes to the generalisation problem by enabling us to understand the user requirements in this real situation (van Oosterom 2009). Cadastral data are a unique source of spatial information for

National Infrastructure for Spatial Information. Cadastral data are captured and maintained at the highest level of detail and they cover the entire territory of the Czech Republic. The quality of this large-scale dataset has no other equivalent in terms of completeness, logical consistency and positional, temporal and thematic accuracy.
Further developments include the definition of semantic, topology and geometry criteria for the cadastral data model in a lower level of detail. Selected methods for model generalisation will be applied in order to get the target model. The generalisation techniques will be applied to, and tested on, sample cadastral data. The use of model generalisation defining the relations between the two different levels of detail is essential.
The author Vaclav Cada is supported by the Research Plan MSM 4977751301.

REFERENCES

ACIL Tasman, 2008. *The Value of Spatial Information*, Available at: http://www.crcsi.com.au/uploads/publications/PUBLICATION_323.pdf.

Basaraner, M., 2002. Model Generalization in GIS. In International Symposium on GIS. Istanbul, Turkey.

Čada, V. & Mildorf, T., 2005. Delimitation of reference geodata from land data model. In *Proceedings of GIS Ostrava 2005*. GIS Ostrava 2005. Ostrava: VŠB - TUO, p. 13. Available at: http://gis.vsb.cz/GIS_Ostrava/GIS_Ova_2005/Sbornik/cz/Referaty/cada.pdf.

Corbin, C., 2008. *Public Private Partnerships for building Spatial Data Infrastructures: Summary and recommendations*, Available at: http://www.epsiplus.net/content/download/13853/173287/file/PPP4SDI_Report_Part1_V1_1_Final.pdf.

DPLI Working Group, 2002. Data Policy & Legal Issues Position Paper. Available at: http://inspire.jrc.ec.europa.eu/reports/position_papers/inspire_dpli_pp_v12_2_en.pdf.

European Parliament, 2007. DIRECTIVE 2007/2/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Available at: http://eurlex.europa.eu/JOHtml.do?uri=OJ:L:2007:108:SOM:EN:HTML.

Gruenreich, D., 1992. ATKIS - a topographic information system as a basis for GIS and digital cartography in Germany. In *Geologisches Jahrbuch Reihe A*. pp. 207-216.

INSPIRE Drafting Team "Data Specifications", 2008a. *D2.5: Generic Conceptual Model, Version 3.0*, Available at: http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.5_v3.0.pdf.

INSPIRE Drafting Team "Data Specifications", 2008b. *D2.6: Methodology for the development of data specifications*, Available at: http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.6_v3.0.pdf.

INSPIRE Thematic Working Group Cadastral Parcels, 2010. D2.8.I.6 INSPIRE Data Specification on Cadastral Parcels – Guidelines.

International Organization for Standardization, 2002. ISO 19113 - Geographic information - Quality principles.

McMaster, R. & Shea, S., 1992. *Generalization in Digital Cartography*, Washington D.C.: Association of American Cartographers.

Mildorf, T., 2009. PLAN4ALL - Interoperability of Spatial Planning Information in the Context of the Infrastructure for Spatial Information in the European Community – INSPIRE. In *INSPIRE, GMES and GEOSS Activities, Methods and Tools towards a Single Information Space in Europe for the Environment*. Riga: Tehnoloģiju attīstības forums, Wirelessinfo. Available at: http://www.earthlook.cz/book/GMES_BOOK.pdf.

van Oosterom, P., 2009. Research and development in geo-information generalisation and multiple representation. *Computers, Environment and Urban Systems*, 33, pp.303–310.

Sester, M. et al., 2009. *Summary report of the Dagsthul seminar 09161 on Generalization of Spatial Information 14-17 April 2009*, Dagsthul. Available at: http://ica.ign.fr/dagsthul2009/Dagsthul2009Report.pdf.

BIOGRAPHICAL NOTES

Tomas Mildorf - MSc. (2004) Geomatics - specialization in cartography - University of West Bohemia, Pilsen. Research activities: Infrastructure for spatial information, model generalisation, metadata for geographic information. Participation in several EU projects (e.g. Humboldt, INSPIRE, COSIN). Traineeship in Joint Research Centre of the European Commission in Ispra (Italy) – Institute for Environment and Sustainability, Spatial Data Infrastructure Unit. At present Ph.D. student at University of West Bohemia. Currently a coordinator of the eContentplus project Plan4all.

Vaclav Cada - MSc. (1981), PhD. (1990), Ass. Prof. (2004) - Czech Technical University, Prague. Research activities: surveying, computer cartography and GIS. Successful leader of many national projects and member of Czech Union for Surveyors and Cartographers, Committee for FIG and member of Czech Association for Geoinformation (CAGI).

CONTACTS

Tomas Mildorf
University of West Bohemia
Univerzitni 22
Pilsen
CZECH REPUBLIC
Tel. +420 377 632 673
Fax +420 377 632 602
Email: mildorf@centrum.cz
Web site: http://www.kma.zcu.cz/mildorf

Assist. Prof. Vaclav Cada
University of West Bohemia
Univerzitni 22
Pilsen
CZECH REPUBLIC
Tel. +420 377 632 678
Fax +420 377 632 602
Email: cada@kma.zcu.cz
Web site: http://www.kma.zcu.cz/vaclav.cada