# The role of IE methodology in SDIs

**Agnieszka ZWIROWICZ, Poland**

**Key words**: Information Engineering (IE), data structure, data integration and exchange, Unified Modeling Language **(**UML), ISO series 19100 standards, Spatial Data Infrastructure (SDI)

## SUMMARY

Information Engineering is an approach to designing and developing information systems. It has many purposes, including organization planning, business re-engineering, application development, information systems planning and systems re-engineering. IE is also an approach to SDIs. To identify needs for SDIs, explore and discuss the new areas of their usage, business modeling can be used. The aspects of spatial information, including spatial information for sustainable management of urban areas, can be expressed in the process of information modeling. In this paper, there are discussed aspects of structure and integration, and exchange of community based data collections in the frame of conceptual modeling and IE tools. Furthermore, there are references to Polish SDI and examples from one of branch SDI – Polish Infrastructure of Geodetic and Cartographic Information, showing data integration and exchange problems, and also proposals to solve these in terms of IE methodology.

# The role of IE methodology in SDIs

**Agnieszka ZWIROWICZ, Poland**

## 1. INTRODUCTION

There are set high hopes on SDI. It can be used in multiple ways and purposes in unlimited areas and communities. It can be a useful tool in, among the others, managing the space – both urban and rural, environment and good governance. The possibilities of its applying have been still under cognition so far. But regardless of assignment of data, there are some characteristic which determine value of the data, e.g.: quality, accuracy, origin, exchange format, distribution, storage. Datasets and bases located in different institutions and companies, at different levels of administrative division are the base for SDI. One data reference the other ones and the data flow is required. There is need for data adjustment and harmonization to support interoperability among systems and databases.
Information Engineering methodology is an architectural approach to planning, analysing, designing, and implementing applications within an enterprise. IE can enable SDI as enterprise, to improve the management of its resources, including capital, people and information systems and support the achievement of SDI's business vision. For example, exploring new applications of SDI could be expressed in frame of business modeling and use cases. Data structure can be described in domain and class models. The harmonization and structure and integration aspects of community based data are discussed in this paper.

## 2. DATA INTEGRATION

The purpose of data integration is data adjusting to each other. It is preceded with data transmission and conversion from different systems and databases, and consists of datasets consolidation and harmonization (geometric and descriptive ones) to become one coherent whole. There can be distinguished both semantic and spatial aspect of data integration, and also the realization one. The semantic aspect of data integration in frame of IE methodology is discussed in the following clauses.

### 2.1 Data structure

Each system can organize data in the free way to reach the most effectiveness of data storage and manipulation. Data model that is represented as data structure, describes rules for data definitions, organization and management. Data definitions describe among others, spatial definition, attributes, cartographic portrayal, metadata. Data structure's aspects play the crucial role in data integration and are the basis of interoperable data interchange.

### 2.2 Application schema

To achieve interoperability between heterogeneous systems two fundamental issues need to be determined. The first issue is to define the semantics of the content and logical structures of geographic data. This shall be done in an application schema. An overview of a data interchange is shown in Fig. 1. System A wants to send a dataset to system B. To ensure a successful interchange A and B must decide on three things, i.e. a common application schema, which encoding rule to apply and what kind of transfer protocol to use. The application schema is the basis of a successful data transfer and defines the possible content and structure of the transferred data, whereas the encoding rule defines the conversion rules for how to code the data into a system independent data structure.

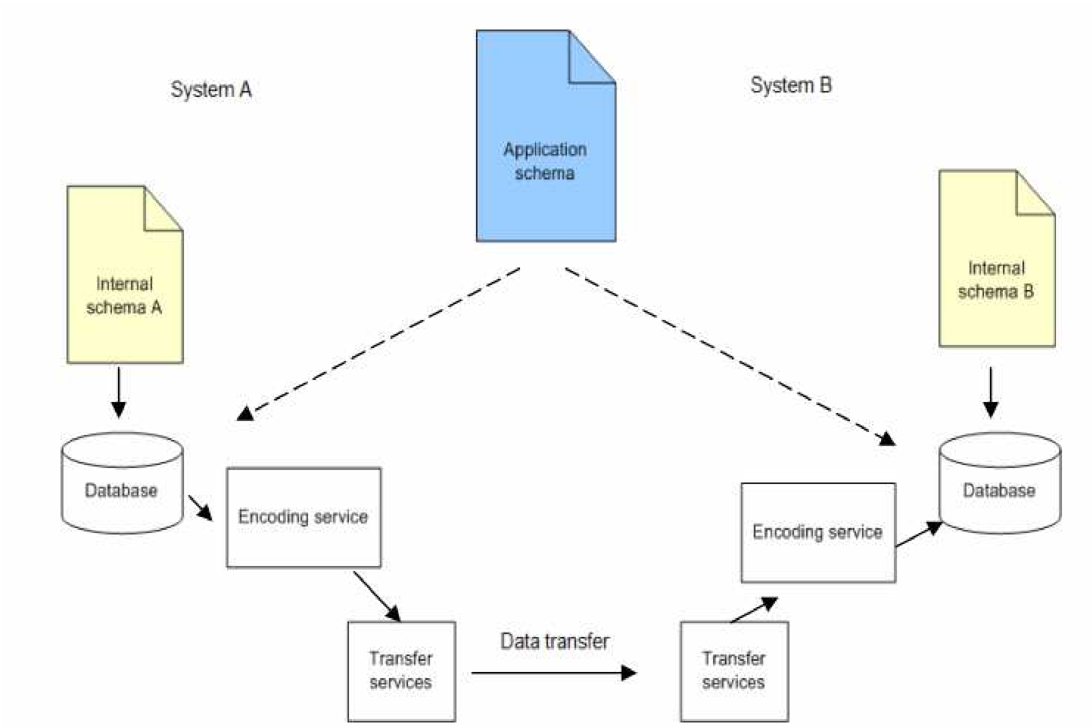


Fig. 1 Interoperable data interchange between systems (ISO Encoding, 2003)

Process of building application schemas (ISO Rules for application schema, 2005) can be divided into following steps: surveying the requirements from the intended field of application, making a conceptual model of the application (identifying feature types, their properties and constraints), describing the application schema in a formal modeling language (for example UML and OCL), integrating the formal application schema with other schemas of ISO series 19100 standards (spatial schema, quality schema, etc.) into a complete application schema.

### 2.3 Case study of Polish Infrastructure of Geodetic and Cartographic Information

To illustrate the aspects and some problems of data integration, there is presented a case study of some elements of Polish Infrastructure of Geodetic and Cartographic Information (PIGCI). Datasets and database of PIGCI, some of them shown in Fig.2, are the reference ones for other datasets and databases and usage (including e.g. spatial planning, registry of property price and value) of Polish SDI.

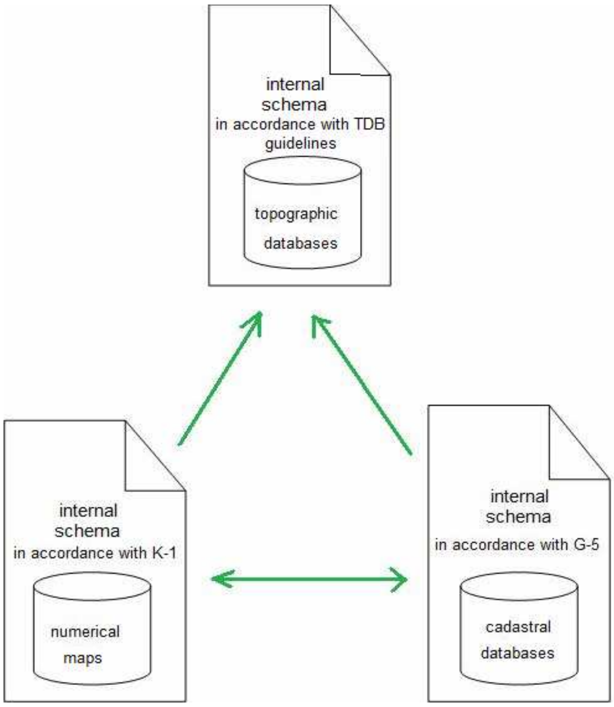


Fig. 2 Data interchange between some of the databases of Polish Infrastructure of Geodetic and Cartographic Information

The quality and reliability and of the reference data has impact on further use of data in different sectors. Unquestionably, PIGCI meets these requirements. But there are some internal problems of data integration between some of the PIGCI databases. In Fig. 2 the arrows shows the direction of data interchange, and so far there have been problems with data integration of datasets from numerical maps and cadastral databases in topographic databases. There are different data structures established in each of case study's databases. The basis of data structure definitions are geodetic technical regulations (e.g. K-1 for numerical map, G-5 for cadastral databases). There are different features classifications and semantics conflicts: names, meanings and schemas. Schemas conflicts include different classes, attributes and relationships.

The preferable, interoperable data interchange needs common application schema that will define semantics shared by all databases (Fig 3). Aspects of thematic harmonization are discussed in Clause 3.

## 3. HARMONIZATION

Harmonization allows to eliminate ambiguity and to reach cohesion of guidelines. Process of harmonization should include analysis of information flow, demands, use cases, data and schemas modeling. This is an iterative process that helps to reach the best solutions and include changes in requirements and new scenario.

There could be two alternative ways of thematic models harmonization (Pachelski and Parzyński, 2007; Pachelski et al., 2007). First proposal is connected with defining and completing feature types, attributes and relationships within the thematic areas, which are schemas of e.g. numerical map, cadastral databases, topographic databases (Fig. 4) and determining the concepts of the features types which are shared by thematic areas, and these can be described in frame of conceptual modeling and UML notation. Each application schema can use structures from other schemas (Fig. 5).

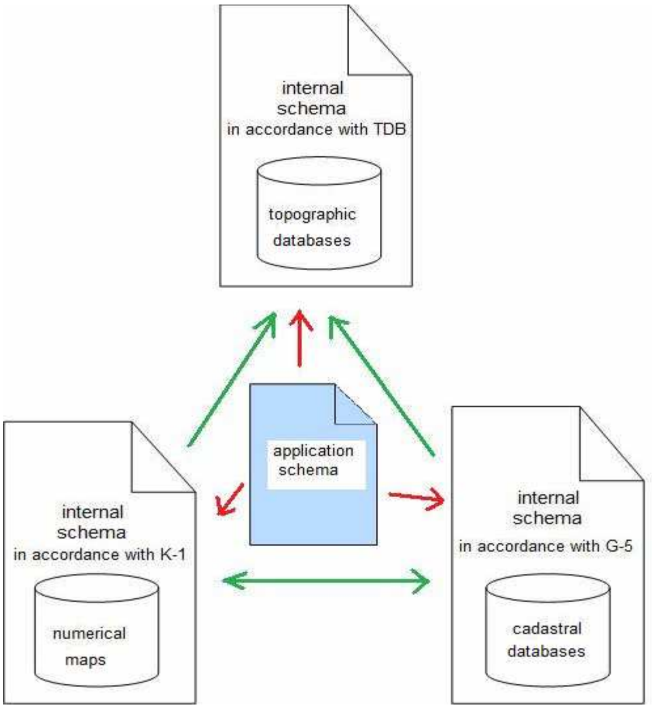


Fig. 3 Preferable data interchange between some of the databases of Polish Infrastructure of Geodetic and Cartographic Information

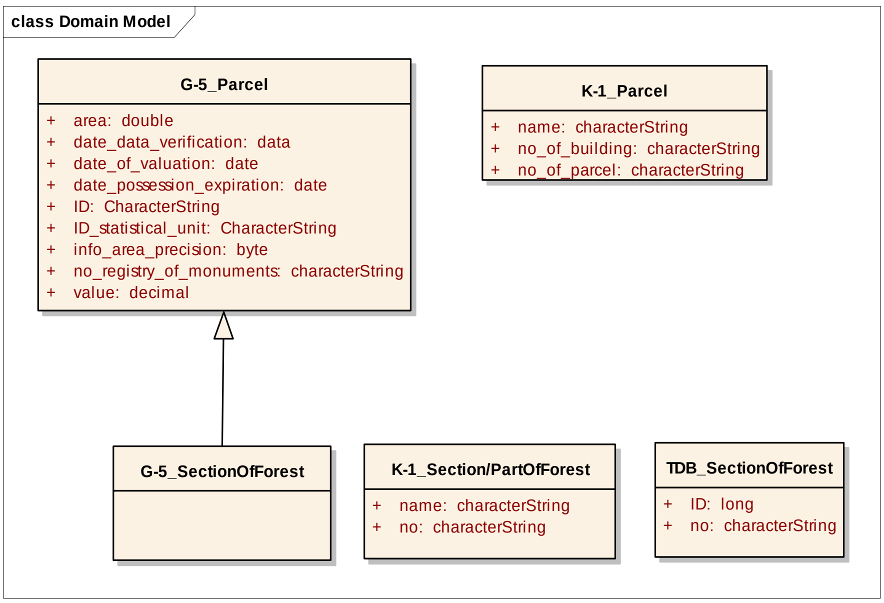


Fig. 4 The example of difference in definitions of feature type SectionOfForest

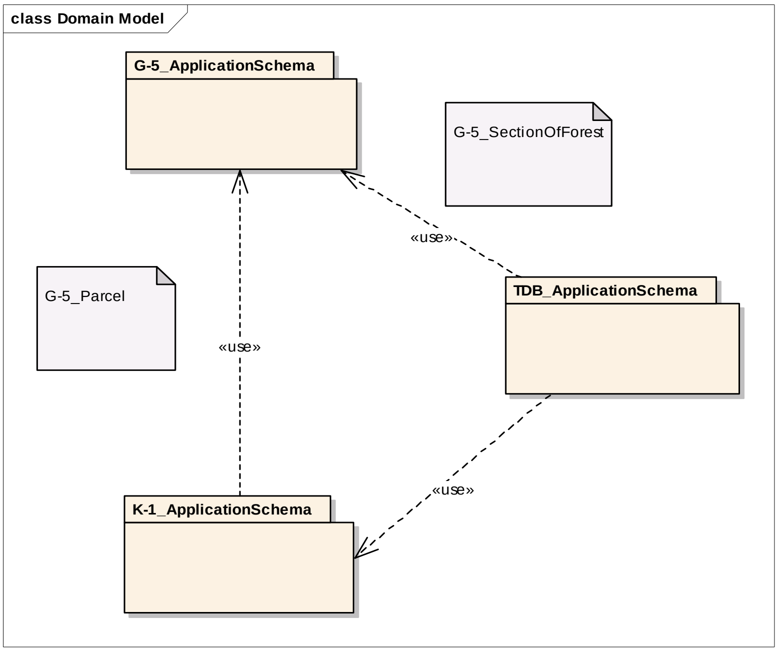


Fig. 5 The examples of shared classes between application schemas

The alternative method of harmonization is defining General Geodetic Model (GGM), which is some abstract supermodel of thematic models based on e.g. G-5, K-1, TDB guidelines (Fig. 6). GGM defines the basic reference data and generalize definitions of classes that are shared by thematic areas. The generalized class is the subtype of GGM class (Fig. 7). The base for GGM should be the Core Cadastral Domain Model (CCDM) (Lemmen and van Oosterom, 2006), which version 1.0 was accepted by FIG in 2002, and included in normalization work schedule of ISO/TC 211 in 2006.

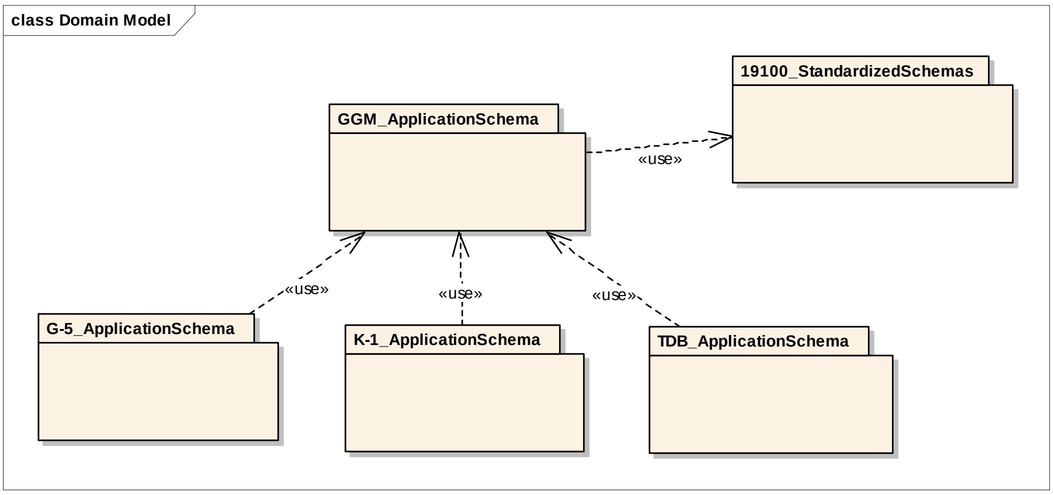


Fig. 6 General Geodetic Model as generalization of thematic models

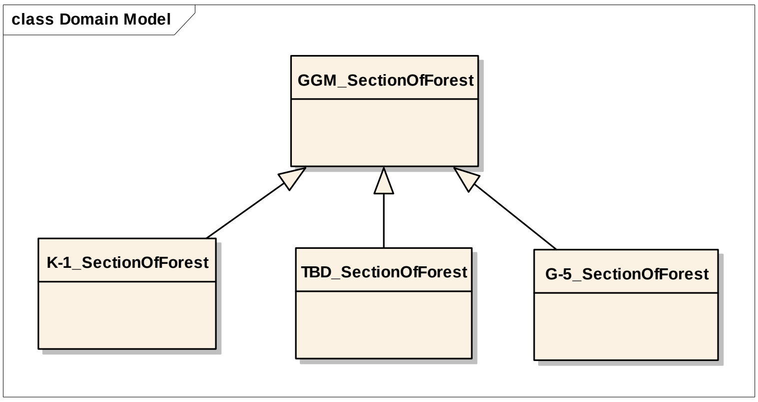

Fig. 7 Subtypes of GGM_SectionOfForest

## 4. CONCLUSIONS

IE can enable SDI as enterprise, to improve the management of its resources, including capital, people and information systems and support the achievement of SDI's business vision. It helps to identify needs for SDIs, explore and discuss the new areas of their usage. The multiple use of spatial data can be supported, if there are databases and dataset of good quality and that they exist at all. Multiple use means also data transmission and conversion from different systems and databases. Data structure's aspects play the crucial role in data integration and are the basis of interoperable data interchange. To achieve interoperability between heterogeneous systems there should be defined the semantics of the content and logical structures of geographic data. This should be done in an application schema. Harmonization allows to eliminate ambiguity and to reach cohesion of guidelines and technical regulations. One harmonization's method is to redefine and complete feature types definitions within thematic areas and use of proper structures from different schemas. Another one is connected with definition of abstract generalization – supermodel of thematic models.

## REFERENCES

ISO:2003: Geographic information — Encoding.
ISO 19109:2005: Geographic information — Rules for application schema. ISO 2005.
ISO,2007: Draft New Work Item Proposal: Geographic information — Core Cadastral Domain Model (CCDM). ISO/TC 211 N 2125.
Lemmen, Ch.; van Oosterom, P., 2006: *FIG Core Cadastral Domain Model Version 1.0.* GIM, Vo. 20, Issue 11, Nov. 2006.
OMG, 2003. Object Management Group, Model Driven Architecture, Guide Version 1.0.1. http://www.omg.org/mda/.
Østensen, O. 2004: Report at 18th plenary meeting of CEN/TC 287 – an update since last CEN/TC 287 plenary. ISO/TC 211.

Pachelski W., Parzyński Z., 2007. Methodological aspects of using standards of 19100 series to develop geodetic components of National Spatial Data Infrastructure. Annals of Geomatics, Vol. 5, z.3, pp. 113 – 122.
Pachelski W., Parzyński Z., Zwirowicz A., 2007. Implementation aspects of geographic information conceptual models. Archives of Photogrammetry, Cartography and Remote Sensing, Vol.17b, pp.591-601
Pachelski W., Parzyński Z., Zwirowicz A., 2008. Problems of integration of national georeference data models with the ISO series 19100 standards. Annals of Geomatics (in print).

## BIOGRAPHICAL NOTES

Agnieszka Zwirowicz – 28 years old, lecturer at the University of Warmia and Mazury in Olsztyn, Poland; MSc in 2004 at UWM in Olsztyn, MSc thesis "Proposal of the mathematical model of the land development"; the reward of the district surveying company in Olsztyn for the best students; PhD in 2007 at UWM in Olsztyn, PhD thesis "Concept of the information system supporting land management at the commune level"; the award of the faculty council for PhD thesis; the author and co – author of more than 14 papers and 13 oral presentations at national and international conferences and 4 posters; co-author and webmaster of e-dictionary for ISO series 19100 standards (http://www.e-przewodnik.gugik.gov.pl).

## CONTACTS

PhD Agnieszka Zwirowicz
University of Warmia and Mazury in Olsztyn
Oczapowskiego 2, 10-957
Olsztyn
POLAND
Tel. +48 601083147
Email: agnieszka.zwirowicz@uwm.edu.pl