

# Reference Data as a Basis for National Spatial Data Infrastructure

Tomáš Mildorf and Václav Čada

Department of Mathematics - Section of Geomatics  
Faculty of Applied Sciences, University of West Bohemia in Pilsen  
Univerzitní 22, 306 14 Pilsen, Czech Republic

mildorf@kma.zcu.cz cada@kma.zcu.cz

## Abstract

*Spatial data are increasingly being used for a range of applications beyond their traditional uses. Collection of such data and their update constitute a substantial part of the total costs for their maintenance. In order to ensure sustainable development in the area of geographic information systems, efficient data custody and coordination mechanisms for data sharing must be put in place. This paper shows the importance of reference data as a basis for national spatial data infrastructure that serves as a platform for decision making processes in society. There are several European initiatives supporting the wider use of spatial data. An example is the INSPIRE Directive. Its principles and the main world trends in data integration pave the way to successful SDI driven by stakeholders and coordinated by national mapping agencies.*

**Keywords:** reference data, INSPIRE, spatial data infrastructure, data integration

## 1. Introduction

The role of spatial data in current society is increasingly important. Spatial data help us to shape the environment we live in, to manage the resources we possess and to preserve our cultural heritage. The importance of spatial data is being recognised by decision makers, whose support is essential for further development of spatial information technologies and the wider use of spatial data in practice. Funding schemes at various levels of administration aim to support projects and initiatives dealing with access to heterogeneous spatial data through innovative technologies. The potential of spatial data is also given by the economic value of spatial information within public sector information in the EU. An analysis dated back to 1999 (see Figure 1) is underpinned by recent studies including ACIL Tasman (2008) and Fornefeld et al. (2009). The range of applications where spatial data play an important role is growing alongside the demand for sustainable spatial data management.

Despite the importance of spatial data for society, there are certain questions which need to be addressed in order to achieve sustainable management and efficient use of spatial data. The Geographic Information Panel (2008) declares that “current users of spatial information spend 80 per cent of their time collating and managing the information and only 20 per cent analysing it to solve problems and generate benefits.” How can we overcome this imbalance?

Recent activities by the European Commission provided a European Interoperability Framework (EIF) aiming to support the interoperability of European public sector information and

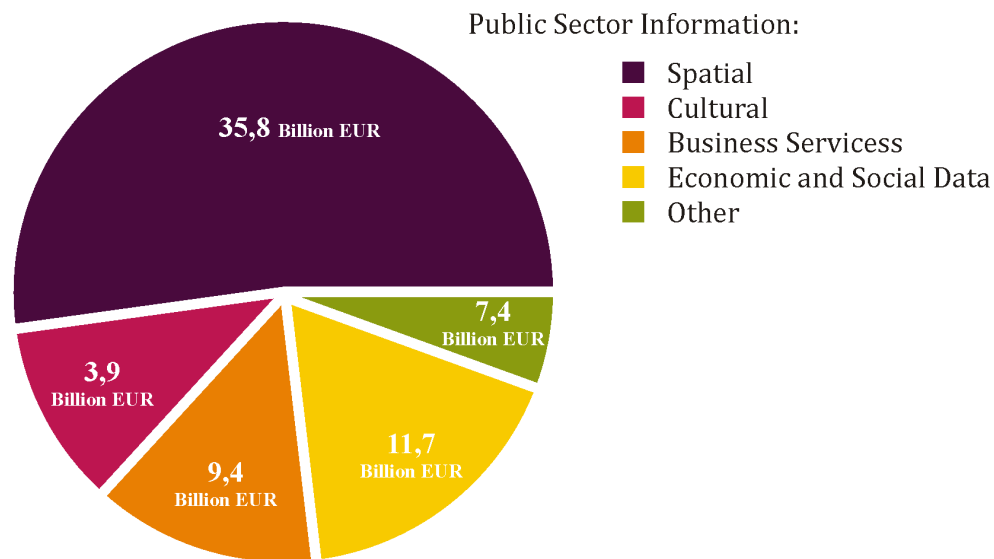


Figure 1: Economic value of public sector information in the European Union in 1999 (Pira International Ltd. & University of East Anglia and KnowledgeView Ltd. 2000).

related services, taking into account legal, organisational, semantic and technical issues. One of the most important projects with a focus on harmonisation of spatial data and services is the Infrastructure for Spatial Information in the European Community (INSPIRE) Directive. INSPIRE entered into force in 2007 as a European directive and over the next two years was transposed into the national legislation of all member states of the European Union. The main objective of the INSPIRE Directive is to establish an infrastructure for spatial information in Europe to “assist policy-making in relation to policies and activities that may have direct or indirect impact on the environment.” (European Parliament 2007). Is fulfilling the INSPIRE Directive sufficient to secure the sustainability of spatial data infrastructures (SDI) on a national level? Is there anything that national mapping agencies must beware of?

The spatial data analyses that are needed for decision making processes in society require spatial data of appropriate quality in terms of completeness, logical consistency and positional, temporal and thematic accuracy. One aspect to which insufficient attention is paid in SDI building, and which is considered by the authors as a basis for national SDI, is the delimitation of reference data. What is understood by the authors under the term reference data? What benefits for decision making processes and the sustainability of national SDIs do they present?

In order to address the above mentioned questions the authors analysed selected data sources in the Czech Republic within the context of the INSPIRE Directive. The next chapter reviews the scope of INSPIRE and its main principles. Chapter 3 presents the results of the analysis of the selected data sources from the Czech Republic. Chapter 4 describes the role of reference data within the context of an SDI. The need for reference data is underlined by global trends in data integration and cohesion of cadastral and topographic data in SDI building in Chapter 5. Cases from the Netherlands and Great Britain give the context for the issues tackled in this paper and provide best practice in the national SDI implementation with regard to reference data and maintenance of the INSPIRE principles. The paper aims to start a discussion about these topics.

## 2. The scope of INSPIRE

The INSPIRE Directive lays down the rules that enable the sharing and reuse of pre-existing data. Heterogeneous spatial data originating from various sources are harmonised according to the common INSPIRE data specifications. A single access point enables users to search the right data for their purposes, to seamlessly view the data and to download them or to perform other spatial services. INSPIRE is a good basis not only for decision makers but also for planners, businesses, emergency management and others.

The success of INSPIRE is based on principles that are crucial for achieving the sustainability of the infrastructure. The INSPIRE principles include:

- Data should be collected once and maintained at the level where this can be done most effectively;
- It should be possible to combine seamlessly spatial data from different sources and share them between many users and applications;
- Spatial data should be collected at one level of government and shared between all levels;
- Spatial data needed for good governance should be available on conditions that are not restricting their extensive use;
- It should be easy to discover which spatial data are available, to evaluate their fitness for purpose and to know which conditions apply for their use. (INSPIRE Website 2012)

All of these principles can be achieved by implementing the INSPIRE mechanisms for data sharing. However, in some cases this is only true to a certain degree. The heterogeneity of spatial data harmonised by INSPIRE can cause certain inconsistencies in the target data, for example when two datasets of different levels of detail are harmonised (see Figure 2). The data provided through the INSPIRE infrastructure for applications requiring data of a high level of detail and high accuracy may not be sufficient. The legislation put in place by the INSPIRE Directive does not affect the collection and processing of data which are considered by the authors as the main sources of inconsistencies. The primary scope of INSPIRE is therefore on data at European, national and regional level where the inconsistencies are diminished by the level of generalisation and the expected level of quality.

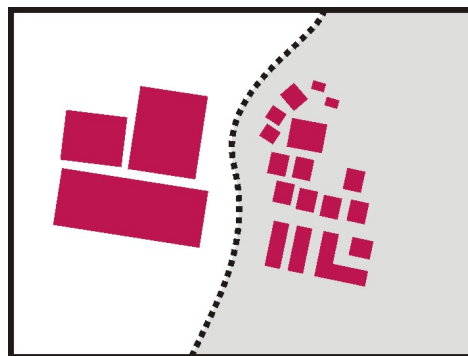


Figure 2: Inconsistency of harmonised datasets of different levels of detail.

INSPIRE represents a solid foundation for the European SDI. National SDIs should benefit from the provisions of INSPIRE. It should be the responsibility of national mapping agencies (NMAs) to combine the national requirements and priorities with INSPIRE and to secure the sustainability of the overall infrastructure. NMAs should take advantage of the INSPIRE implementation and trigger the creation of national SDI.

### 3. Analysis of selected data sources in the Czech Republic

The authors analysed the situation of spatial data management in public administration in the Czech Republic in relation to the INSPIRE Directive and its principles. The focus was mainly on semantic aspects of selected digital data sources of higher level of detail including:

- Cadastral map (KM);
- Technical maps (TMO);
- Planning Analytical Materials (UAP);
- Fundamental Base of Geographic Data (ZABAGED).

All the geographic features from these data sources were compared in terms of their definition and an overview of relations between the features was drawn. Due to complexity of the overview including the definitions of all the features only an indicative table showing the similarities between the analysed data sources is presented (see Table 1).

The results of this analysis show that many geographic features from the selected data sources are duplicated, they are not maintained at the most appropriate level and it is not easy to combine them with other data sources. Three of five INSPIRE principles are not maintained and the sustainability of the data sources forming the national SDI is questionable.

### 4. Reference data

Data collection and their update make a substantial part of the total costs of data maintenance. Sharing of spatial data between different applications enables sharing of the costs for data management. The historical development of spatial data of public administration in many countries, and the lack of coordination between data producers and data users, has led to duplication in data collection. An example is the situation in the Czech Republic analysed by the authors. Real world phenomena are independently captured, processed, stored and updated by several organisations.

Based on the findings of the performed analysis, the authors propose the delimitation of reference data on the highest level of detail that will be shared between several applications of public administration and other users.

Great achievement was done by establishing the basic registers in the Czech Republic including the Register of Territorial Identification, Addresses and Real Estates (RUIAN). Geographic features included in RUIAN thus create a reference base for all applications of public administration and other users. The system of basic registers uses the term reference data as data maintained in basic registers and given by law that are up-to-date, valid and unambiguous for every application of public administration.

DKM	ZABAGED	TMO	UAP
hrana koruny a střední dělicí pás silnice nebo dálnice	silnice, dálnice		dálnice včetně ochranného pásma rychlostní silnice včetně ochranného pásma silnice I. třídy včetně ochranného pásma silnice II. třídy včetně ochranného pásma silnice III. třídy včetně ochranného pásma
osa kolejí železniční tratě mimo železniční stanici a průmyslové závody	železniční trať	osa železničních a tramvajových kolejí	železniční dráha celostátní včetně ochranného pásma dráhy celostátní vybudované pro rychlost větší než 160 km/h včetně ochranného pásma
	lanová dráha, lyžařský vlek stožár lanové dráhy	lanová dráha, dopravník	lanová dráha včetně ochranného pásma
nadzemní vedení vysokého a velmi vysokého napětí včetně stožárů	elektrické vedení	elektrické vedení	nadzemní a podzemní vedení elektrizační soustavy včetně ochranného pásma
stavba k využití vodní energie (vodní elektrárna)	elektrárna	elektrárna, spínací stanice nebo měnič, transformovna, transformační stanice	výroba elektřiny včetně ochranného pásma
	dálkový produktovod vodojem věžový	potrubí produktovodu vodojem	produktovod včetně ochranného pásma
ostatní plocha - skládka	skládka		skládka včetně ochranného pásma
stavba odkaliště	usazovací nádrž, odkaliště	odkalovací nádrž, kaliště	odval, výsypka, odkaliště, halda
vodní plocha - rybník			
vodní plocha - vodní nádrž přírodní	vodní plocha	vodní nádrž	vodní útvar povrchových, podzemních vod
vodní plocha - vodní nádrž umělá			
vodní plocha - koryto vodního toku přirozené nebo upravené	vodní tok	vodní tok	
vodní plocha - koryto vodního toku umělé	břehová čára	vodní tok občasný, vysychající, odpadová stoka, suchý příkop	
vodní plocha - zamokřená plocha	bažina, močál	močál	
národní park - I. zóna			
národní park - II. zóna			
národní park - III. zóna	velkoplošné zvláště chráněné území	hranice zvláště chráněného území	národní park včetně zón a ochranného pásma
ochranné pásmo národního parku			
chráněná krajinná oblast - I. zóna			
chráněná krajinná oblast - II.-IV. zóna		hranice ochranného pásma	chráněná krajinná oblast včetně zón
pam. rezervace - budova, pozemek v památkové rezervaci		hranice památkové rezervace	památková rezervace včetně ochranného pásma
pam. zóna - budova, pozemek v památkové zóně		hranice památkové zóny	památková zóna včetně ochranného pásma
chr.lož.území,dob.prostor,chr.území pro zvl.zásahy do z.kůry		hranice chráněného ložiskového území	chráněné ložiskové území

Table 1: The comparison of geographic features from the selected data sources.

Reference data were used as the basis for documents forming the current INSPIRE Directive. The Chapter on Reference Data of the European Territorial Management Information Infrastructure (ETeMII) White Paper (2001) defined the following functional requirements for reference data:

- to provide an unambiguous location for a user's information;
- to enable the merging of data from various sources;
- to provide a context to allow others to better understand the information that is being presented. (p. 5)

The INSPIRE TWG Cadastral Parcels (2009) defines reference data as data that constitute the spatial frame for linking and pointing at other information that belongs to specific thematic fields; e.g. land use, land cover, agriculture and demography. These are considered as application data, which is a complementary term to reference data (INSPIRE Drafting

Team Data Specifications, 2008). Reference data provide a common link between various applications and provide mechanisms for sharing information in society. In the initial phase of the INSPIRE development the following spatial data themes were defined as reference data (RDM Working Group 2002):

- geodetic reference system;
- units of administration;
- units of property rights;
- addresses;
- selected topographic themes;
- orthoimagery;
- geographical names. (p. 11)

It is necessary to note that these are spatial data themes covering a wide range of geographic features. The vision of the authors goes a step beyond the spatial data themes into a selection of particular geographic features that fulfil the above mentioned functional requirements for reference data. The selection of geographic features playing the role of reference data should be available for any application. The consensus on the features, their quality and the sources of updates as well as the forms of exchange should be agreed.

The next chapter introduces examples from Great Britain, the Netherlands and other countries to give context to the above mentioned ideas, especially in relation to reference data and their integration with application data.

## 5. Trends in data integration

### 5.1. Great Britain

Great Britain realised the importance of INSPIRE in time and in 2008 the UK Location Programme<sup>1</sup> was created. The programme combines the national priorities (implementation of the UK Location Strategy) with the requirements of INSPIRE and provides a complex solution for the sharing and reuse of spatial information of public administration.

One of the operational challenges of the UK Location Programme is, according to the Geographic Information Panel (2008), to develop a set of core reference geographies captured at the highest level of detail. In the initial stage, the core reference geographies should include a geodetic framework, topographic mapping (at different resolutions and including ground height information), geographic names, addresses, streets, land and property ownership, hydrology/hydrography, statistical boundaries and administrative boundaries.

One of the main building blocks of the UK Location Programme is the Digital National Framework<sup>2</sup> (DNF) that includes the interoperability components such as feature catalogue, terminology, metadata, standards and reference model. In relation to reference data, the feature catalogue of the DNF Base Reference Objects is of most importance and provides an

---

<sup>1</sup> <http://location.defra.gov.uk/programme/>

<sup>2</sup> <http://www.dnf.org/>

agreed definition and a set of attributes for every geographic feature. The feature catalogue is not definite and can be further extended.

The basis for the current feature catalogues represents the feature catalogue of the Ordnance Survey (OS) MasterMap. OS MasterMap is a common reference base containing a variety of information in four different product layers: Address Layer, Imagery Layer, Integrated Transport Network Layer and Topography Layer. The OS MasterMap database contains over 450 million geographic features. “Every feature within the OS MasterMap database has a unique common reference (a TOID<sup>®</sup>) which enables the layers to be used together, including the layer of your own information.“ (Ordnance Survey 2012). The example of the OS MasterMap Topography Layer is depicted in Figure 3.

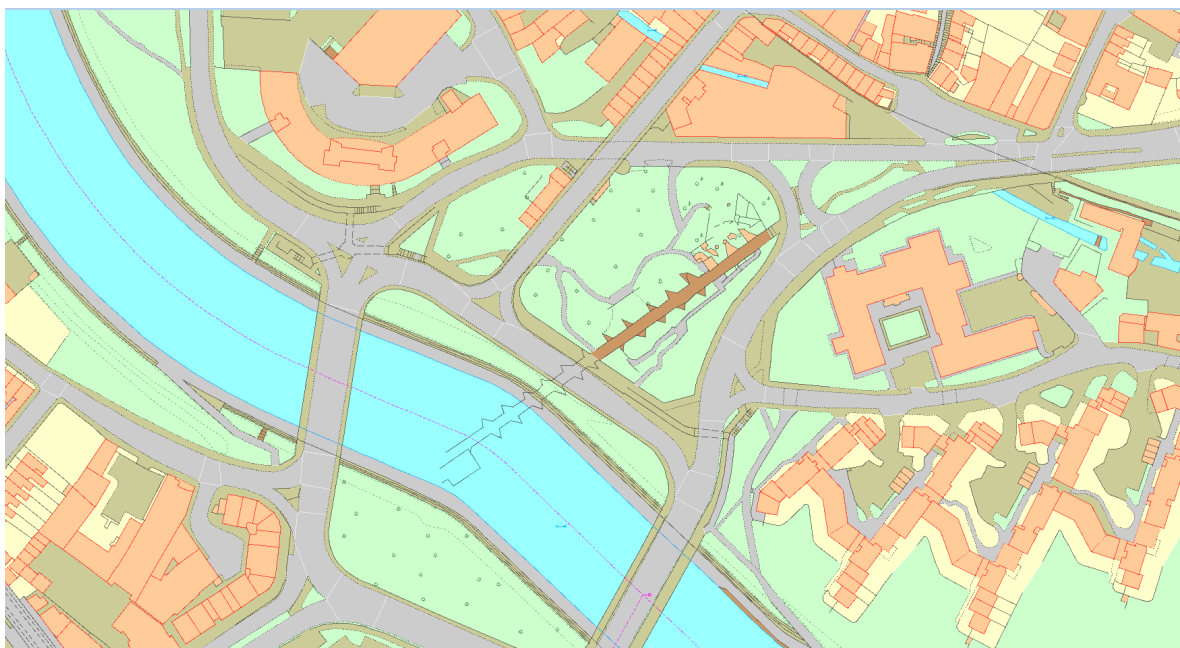


Figure 3: OS Master Map – Topography Layer (Ordnance Survey 2012)

The use of the common reference data aims mainly to improve interoperability, data harmonisation and spatial data quality and increase cross-sector collaboration. The end benefits of reference data encompass (Jones & Wilks 2012):

- reduction of costs for public, private and 3rd sector users of data;
- improvement of quality, efficiency and delivery of services;
- improvement of evidence base for informed policy development and decision making;
- increase of research, innovation and commercial exploitation of location data to benefit UK economy;
- facilitation of other Government initiatives using location based information and tools.

## 5.2. The Netherlands

The strategic features of the national SDI in the Netherlands are the key registers (*basisregistratie*) that have been developing since 2008. The system of key registers is coordinated by the Ministry of the Interior and Kingdom Relations. There are 13 key registers planned. 4 of them related to spatial information and maintained by Kadaster, the Dutch Land Registry Office, include:

- Key Register of Topography (*Basisregistratie Topografie*, BRT) – contains the topographic dataset Topo10NL in the level of detail equivalent to scale 1:10 000;
- Key Register of Cadastre (*Basisregistratie Kadaster*, BRK);
- Key Register of Large-scale Topography (*Basisregistratie Grootchalige Topografie*, BGT);
- Key Register of Addresses and Inhabitants (*Basisregistraties Adressen en Gebouwen*, BAG).

The key registers provide reference data for applications of public administrations as well as for private sector. The use of common reference data is obligatory for all public administrations.

The inclusion of the large-scale topography as one of the key registers enables linking non-spatial and spatial information of public administration. Non-spatial data can be analysed in the context of spatial data, e.g. visualisation of average income or violence in certain areas (Peersmann et al. 2009). The current and potential uses of this key register are depicted in Figure 4.

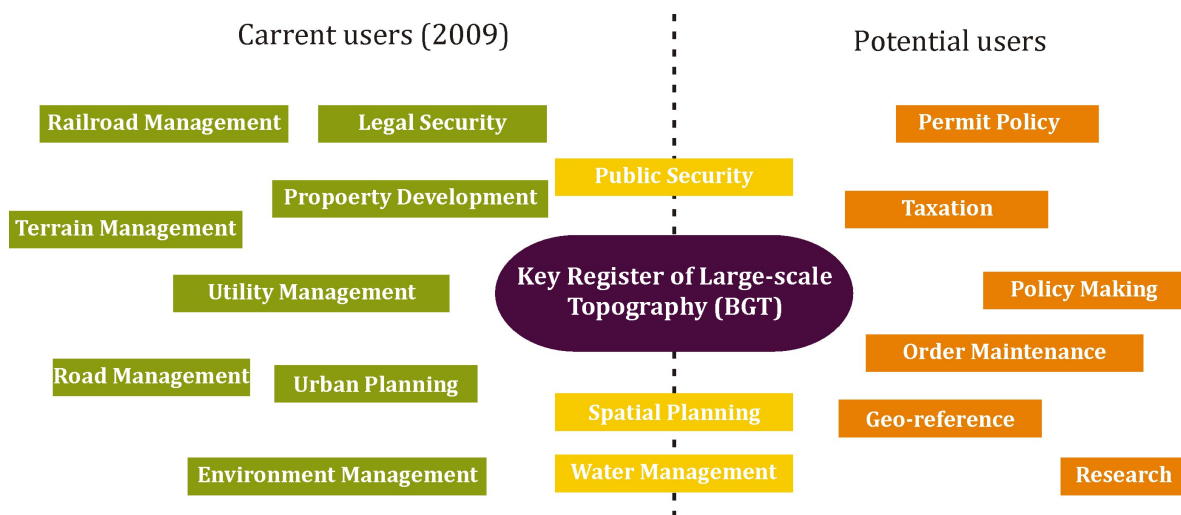


Figure 4: Current and potential uses of BGT (adapted from Peersmann et al., 2009).

An important feature of the Dutch SDI is that cadastral information is maintained together with the large-scale topography in one database. According to Steudler et al. (2009), the layer of buildings is shared between these datasets and cadastral boundaries are aligned to the topographic features.



### 5.3. Other examples

During the international conference ‘Spatial Information for Sustainable Development’ in Nairobi, Ryttersgaard (2001) presented experiences from SDI implementation. One of his visions for further development in this area stated: “Cadastral, topographic and thematic datasets should adopt the same overarching philosophy and data model to achieve multi-purpose data integration, both vertically and horizontally.” (Ryttersgaard 2001, p.7).

One of the main priorities of the Working Group 2 of the Permanent Committee on GIS Infrastructure for Asia and the Pacific is the integration of datasets containing representation of man-made objects and natural objects. This is performed mainly by integrating cadastral and topographical data which are in most cases maintained separately with no existing links between them and which hinder further exploitation (Rajabifard & Williamson 2006).

The National Land Survey of Finland started with providing free access to reference data since May 2012 including topographic maps, imagery and digital elevation model. The publication aims to enable combination of application data from various sectors with common reference data (Ratia 2012). According to Koski (2011), economic growth should be stimulated.

The interoperability of spatial data in Germany is coordinated by the Working Committee of the Surveying Authorities of the German Länder (AdV). The key feature of the interoperability is the AAA concept for modelling of spatial information based on ISO and OGC standards (Working Committee of the Surveying Authorities of the German Länder 2011). AAA stands for AFIS-ALKIS-ATKIS which represent the geodetic control, cadastral and topographical systems of Germany.

## 6. Conclusions

The requirements of the users at a local level go beyond the scope of INSPIRE, especially in terms of data quality. National SDIs supported by NMAs should serve as a basis for INSPIRE. Based on the analysis of the situation in the Czech Republic, we can state that pure implementation of the INSPIRE mechanisms for data sharing without introducing the national context is not sufficient for sustainable national SDI building. NMAs should take the responsibility for coordination of the legal, organisational, technical and semantic aspects; meeting the national priorities and user requirements at local, regional and national level; and supporting the implementation of INSPIRE in connection with a national SDI.

The building of a national SDI should use the experience of already existing infrastructures and best practices in data integration. Several examples were mentioned in the paper in order to provide the underlying information for the authors’ suggested approach for designing and implementing the national SDI in the Czech Republic.

The trends in data integration aim to combine cadastral and topographic data and to provide a reference data for various applications. The INSPIRE principles are taken as key priorities in SDI building. Based on the performed analyses and long-term research activities, the authors are proposing to define reference data at the highest level of detail, so that they can be shared between various organisations of public administration and the private sector. The delimitation of reference data is expected on the level of geographic features. The understanding of the semantics of available data, together with organisational and legal aspects, is

essential for further progress in national SDI building. The Czech NMA should coordinate this process in connection with lead experts from the field of geomatics and geoinformatics. A good progress in this matter started with the establishment of the Register of Territorial Identification, Addresses and Real Estates (RUIAN). However, wider scope and inclusion of other data sources is necessary for better exploitation of spatial data available in public administration and for securing the sustainability of data management.

The definition of reference data mainly represents, but is not limited to, the following benefits:

- sustainable development of the national SDI and support of its extensive use;
- saving costs for data collection and data update of public administration;
- source of guaranteed high quality data;
- unique opportunity for integration of cadastral and topographic information;
- possibility of integration with application data;
- general usability and data availability.

As already mentioned in the introduction, the aim of the paper is to raise a discussion about these topics. What is your opinion?

## References

- [1] ACIL Tasman, 2008. *The Value of Spatial Information*, Available at: [http://www.crcsi.com.au/Documents/ACILTasmanReport\\_full.aspx](http://www.crcsi.com.au/Documents/ACILTasmanReport_full.aspx) [Accessed May 31, 2012].
- [2] INSPIRE Website, 2012. Available at: <http://inspire.ec.europa.eu/>.
- [3] 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> [Accessed May 31, 2012].
- [4] European Territorial Management Information Infrastructure, 2001. ETeMII White Paper: Chapter on Reference Data. Available at: <http://www.ec-gis.org/etemii/reports/chapter1.pdf> [Accessed March 30, 2012].
- [5] Fornefeld, M. et al., 2009. *Assessment of the Re-use of Public Sector Information (PSI) in the Geographical Information, Meteorological Information and Legal Information Sectors*, Düsseldorf, Germany: MICUS Management Consulting GmbH.
- [6] Geographic Information Panel, 2008. *Place matters: the Location Strategy for the United Kingdom*, Great Britain.
- [7] INSPIRE Drafting Team Data Specifications, 2008. *D2.5: Generic Conceptual Model, Version 3.0*, Available at: [http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.5\\_v3.0.pdf](http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.5_v3.0.pdf) [Accessed May 31, 2012].
- [8] INSPIRE TWG Cadastral Parcels, 2009. D2.8.I.6 INSPIRE Data Specification on Cadastral Parcels – Guidelines, v3.0.

- [9] Jones, G. & Wilks, P., 2012. UK Location Programme, Benefits Realisation Strategy. Available at: <http://data.gov.uk/sites/default/files/Benefits%20Realisation%20Strategy%20v2.0%20Final.pdf>.
- [10] Koski, H., 2011. Does Marginal Cost Pricing of Public Sector Information Spur Firm Growth?
- [11] Ordnance Survey, 2012. OS MasterMap - definitive geographical information of Britain. Available at: <http://www.ordnancesurvey.co.uk/oswebsite/products/os-mastermap/index.html> [Accessed April 17, 2012].
- [12] Peersmann, M., Eekelen, H. & Meijer, M., 2009. The Large Scale Topographic Base Map of the Netherlands (GBKN): The Transition from a Public-Private Partnership (PPP) to a Legally Mandated Key Registry (BGT). In GSDI World Conference. Rotterdam, The Netherlands. Available at: [www.gsdi.org/gsdiconf/gsdii1/papers/pdf/267.pdf](http://www.gsdi.org/gsdiconf/gsdii1/papers/pdf/267.pdf) [Accessed March 21, 2012].
- [13] Pira International Ltd. & University of East Anglia and KnowledgeView Ltd., 2000. *Commercial exploitation of Europe's public sector information*, Luxembourg: Office for Official Publications of the European Communities. Available at: <http://www.ec-gis.org/docs/F15363/PIRA.PDF> [Accessed August 3, 2011].
- [14] Rajabifard, A. & Williamson, I., 2006. Integration of Built and Natural Environmental Datasets within National SDI Initiatives. In Seventeenth United Nations Regional Cartographic Conference for Asia and the Pacific. Bangkok, Thailand: United Nations.
- [15] Ratia, J., 2012. SDI Interviews Jarmo Ratia of National Land Survey of Finland on Open Data. Available at: <http://www.sdimag.com/20120302584/SDI-Interviews-Jarmo-Ratia-of-National-Land-Survey-of-Finland-on-Open-Data.html> [Accessed May 31, 2012].
- [16] RDM Working Group, 2002. Reference Data and Metadata Position Paper. Available at: [http://inspire.jrc.ec.europa.eu/reports/position\\_papers/inspire\\_rdm\\_pp\\_v4\\_3\\_en.pdf](http://inspire.jrc.ec.europa.eu/reports/position_papers/inspire_rdm_pp_v4_3_en.pdf) [Accessed May 31, 2012].
- [17] Ryttersgaard, J., 2001. SPATIAL DATA INFRASTRUCTURE, Experiences and Visions.
- [18] Steudler, D. et al., 2009. Cadastral Template, A Worldwide Comparison of Cadastral Systems. *Cadastral Template, A Worldwide Comparison of Cadastral Systems*. Available at: <http://www.fig.net/cadastraltemplate/index.htm> [Accessed May 31, 2012].
- [19] Working Committee of the Surveying Authorities of the German Länder, 2011. *National Report 2010/2011*, Working Committee of the Surveying Authorities of the German Länder. Available at: <http://www.adv-online.de> [Accessed May 31, 2012].

