THEORETICAL ASPECTS AND OPEN ISSUES IN SDI ESTABLISHMENT WITHIN GEODETIC AND CADASTRAL DOMAIN AT THE SLOVAK UNIVERSITY OF TECHNOLOGY

Tomas, Kliment[1], Renata, Duraciova[1], Linda, Galova[1], Dusan, Cibulka[1], Martin, Tuchyna[2]

[1]Department of Theoretical Geodesy, Faculty of Civil Engineering, Slovak University of Technology, Radlinskeho 11, 813 68, Bratislava, Slovakia
tomas.kliment@stuba.sk¨

[2]Slovak Environmental Agency, Tajovskeho 28, 975 90, Banska Bystrica, Slovakia
martin.tuchyna@sazp.sk

Abstract

The initiative to establish Spatial Data Infrastructure (SDI) within the Geodetic and Cadastral domain at the Slovak University of Technology aims to implement all principles defined by INSPIRE directive. INSPIRE may be taken into account either as a legal act to force member states to follow its paragraphs or as a "best-practices" in establishing SDI in wider context. Geodetic and Cadastral domain includes geospatial data with a different scale and thematic content and therefore the application of SDI principles is more then appropriate. Therefore, research workers of the Department of Theoretical Geodesy at the Slovak University of Technology propose the project "*Portal of reference data and services for domain of Geodesy and Cartography*". This project is focused on analysis of existing dataset resources that are used within the Geodetic and Cadastral domain and optimization of their workflows (in order to avoid duplication and increase coherency). Realization of the proposed project should ensure interoperable geospatial data discovery, evaluation and consumption in desired quality and quantity. The main objective of the project is data harmonization within selected domain and further implementation of network services that ensure discovery, view, download and processing of these harmonized data. Project has been submitted for approval to the VEGA, a scientific grant agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic and of the Slovak Academy of Sciences, for funding of its realization. There are many opened issues and questions that will most probably arise during the processing of the particular project tasks. All of them are mentioned and described within the grand application, which has been sent for evaluation and review. This paper does not want to find and describe all the answers of those questions, however is trying to outline the actual situation with open issues and consequently propose appropriate solutions based mainly on most economically available applications (cost & benefit considerations), which will be possible to deploy also by the other stakeholders within Geodetic and Cadastral domain, perhaps even in further domains.

Keywords: Spatial Data Infrastructure, INSPIRE, Geodetic and Cadastral domain, Initial investigation, Initial metadata profile, Geospatial data

INTRODUCTION

Many times activities focusing on more effective execution of the tasks and better decisions support within the certain domain requires except good will and know-how also strong drivers behind. These drivers can be represented either by legislation force or by strong and accurate identified user needs (Masser, 2005). In the world, where word "spatial" plays significant role we can identify a lot of such drivers in various shapes of networks, initiatives, projects and so on. One of them was recently born in Europe, and its name is INSPIRE.

The proposed project "*Portal of reference data and services for domain of Geodesy and Cartography*" (thereinafter only project) is focused on data flows optimizing. Scope of the project is within the organization level (among the departments, sections, actions...) and thus reducing duplication and improving consistency. It includes design and testing an appropriate structure of the data from the field of geodesy and cadastre. Data flow in this context refers to the chain of operations with data from their production (acquisition) to their use. Within whole process we can identify operations representing inputs into the system for processing and operations acting as distribution of outputs from processing through web services. The Department of

Theoretical Geodesy, workplace of project promoter, holds data relating to the region of Slovakia and Central Europe. This data repository contain models of geoid and quasigeoid, data from permanent GNSS stations, geodynamic data from long-term points monitoring, digital elevation model, data from troposphere monitoring and atmospheric moisture, spatial data for prediction of archaeological sites and so on. All data will be shared by access rights for scientific research, educational and other purposes.

The paper describes an approach used to investigate and summarize the current situation within particular departments covering *geodetic and cartographic domain at the Slovak University of Technology in Bratislava* (thereinafter only domain) as well as discuss the open issues and propose the most appropriate solutions considering the personal and financial resources of the public sector: The first section in methods describes INSPIRE and other related realisations of SDI that shall be taken into the consideration within the project proceeding, The second one deals with an approach used to perform *an initial investigation* within the domain to obtain primary information about using data resources. The third section provides a graphical overview of organizational structure of stakeholders within the domain. The fourth section proposes the *initial metadata profile* for existing data resources for catalogue purposes in central domain geo-catalogue. The section with results summarizes the current situation within the domain and shows the examples of data, which are currently used. In the discussion part we describe open issues and raise the questions that should be answered during the project processing. Conclusion and future work provides the proposal how to reach the aims based on investigation results that are described within this paper.

METHODS

INSPIRE and other related works

Main purpose of this initiative driven by the Directive 2007/2/EC of the European Parliament and of the Council is to establish an Infrastructure for Spatial Information in the European Community (INSPIRE) for the purposes of Community environmental policies and policies or activities which may have an impact on the environment. To fulfil this ambitious aim, almost all major domains need to be taken into the consideration as environment is affected by human as well as natural related activities. INSPIRE therefore serves unique opportunity to improve aspects, processes and phenomena, where it was difficult to do it by now.

INSPIRE Directive itself is very generic legal framework defining the five main components, which shall be taken into the consideration by affected stakeholders:

- ***Metadata*** *are descriptive information helping users to navigate and discover, inventory and use the most appropriate spatial data or services. In context of this project proposal plays important role as with the properly structured metadata profile can be spatial data easier described and later on discovered and used according user requirements.*
- ***Data specifications*** *defining structure and content of spatial data according the themes identified in 3 annexes of INSPIRE directive. INSPIRE models can serve important reference base for spatial data identified within the project. When proper mapping and transformation will be executed, compliancy tests can be executed to identify level of conformity with target schema/s. This will enlarge possibility to exchange data in interoperable manner across the various domains.*
- ***Network services*** *specifying the five main type of services used within the INSPIRE architecture (discovery, view, download, transformation and invoke), including definition of spatial data services concept. These services within the project should help make metadata easier discoverable as well as allow to display and properly transform content of spatial data via web browser without need for specific software. They should also ensure access to the raw data for further analysis via desktop GIS tools and where possible deploy advanced desktop GIS functionality to the web interface via invoke (for example Web Processing) services,*
- ***Data Sharing*** *is dealing with setting up the rules for spatial data and services access and use.*
- ***Monitoring and reporting*** *focused on evaluation of spatial data infrastructures evolvement.*

Each of these components comes, or in near future will come with specific set of rules defined in legally binding “Implementing rules” which are accompanied with non legally binding “Guidelines”. This set of documentation together with additional INSPIRE documents and registers provide the main base for the implementation within the EU member states. This process just started and according the INSPIRE Roadmap implementation phase fulfilment is foreseen by the 2019, so still long way to go.

Despite of this long term run, there is already a lot of effort and pressure via various pilot implementations and use cases aiming to set up new procedures or adjust existing processes in the various domains utilising spatial data. Their scope is wide from global activities (GSDI, GEOSS), continental (State of play reports, AuScope, Nature SDIplus, oneGeology), through national (géoportail FR, INSPIRE CZ, SK), domain/thematic (Geonet.SK, Geoportal UGKK SR, ENIPI), crossborder (CENTROPE, eSDI-NET+, HUMBOLDT) or regional and local ones (Turistický portál Posázaví, GIS portál Mesta Banská Bystrica).

Examples mentioned above shows only fragments of huge amount and variety of spatial data infrastructures implementations. Each of them are in different level of maturation but when following same principles chance for real interoperability achievement is getting more real.

Approach to investigate the current situation within the domain

The first step of investigation is a general study of data and data workflows in the particular departments, sections and actions within the domain. This study analyzes actual data workflows (which data exists, for what purposes they are used, how they are managed, updated, provided, visualized and shared). From this study there are identified user requirements and then defined future architecture of the system. The result of this step is a metadata inventory describing existing data resources (main identification of all data sources that will be included within the project).

An investigation of the existing data resources (as-is analysis) is performed by interview between the person from research project team and particular person responsible for each data resource (stakeholder). This interview analyzes data resources from several points of view (data theme, format, etc.) to gain information about current state. An interviewer asks individual stakeholders from particular departments questions according to structure, specifically:

- *What sections is a department divided in?*
- *What kinds of actions are realized by those processes (inputs and outputs), data resources (education, research projects, etc.)?*
- *What are the types, themes of datasets used within these actions?*

Furthermore we will get information about current approaches in issues like data search and discovery, maintenance and visualization. Further issues are related to proposals of user requirements for the new system. Received information will be documented for following steps of research project.

Organizational structure of the stakeholders within the domain

The geodetic and cadastral domain at the Faculty of Civil Engineering consists of three departments: Department of Theoretical Geodesy, Department of Surveying and Department of Mapping and Land Consolidation. The results of interview (investigation of data resources) have been modelled in freeware mind mapping program FreeMind (http://freemind.sourceforge.net). FreeMind allows the user to edit a hierarchical structure of any ideas around a central concept as well as provides extensive export capabilities for distribution purposes. Using of this tool brings non-linear approach to map brainstorming outlines and investigation results as ideas are added around the mind map (http://en.wikipedia.org/wiki/FreeMind). Basic structure (Department – Sections - Actions) of individual departments within the domain is mapped using this tool and an example from Department of Theoretical Geodesy is illustrated in Fig 1.

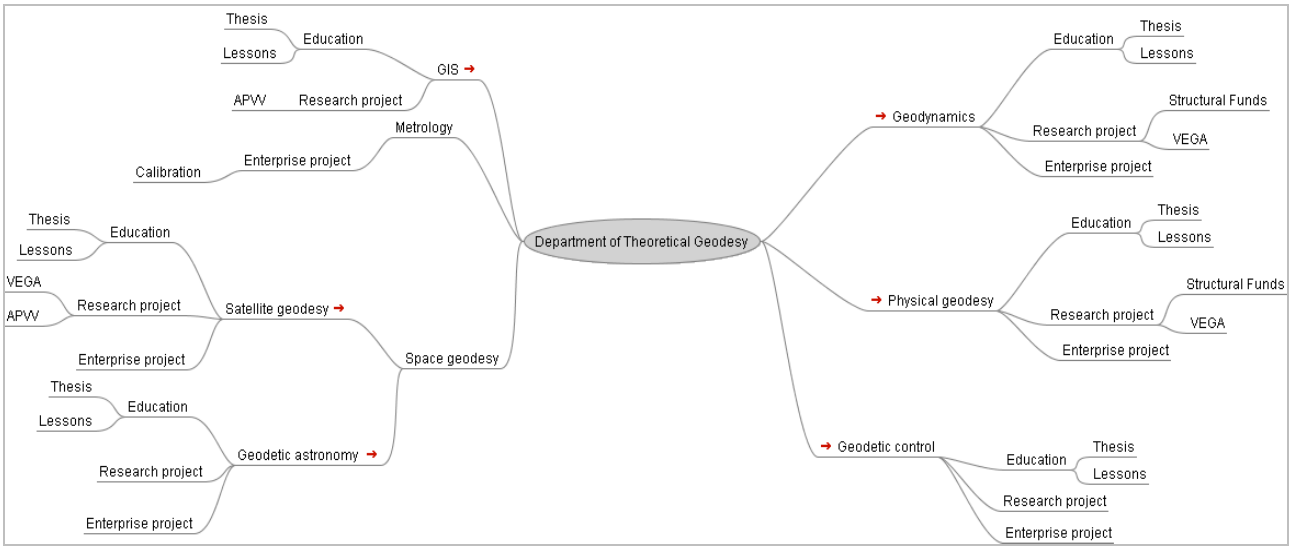

Fig. 1 Department of Theoretical Geodesy – Mapping to Sections and Actions

The data resources mapping from the Department of Theoretical Geodesy is illustrated in Fig. 2 (as an example). Other two departments are mapped in the same way, but the mappings are not included into the paper contents due to extent of the paper.

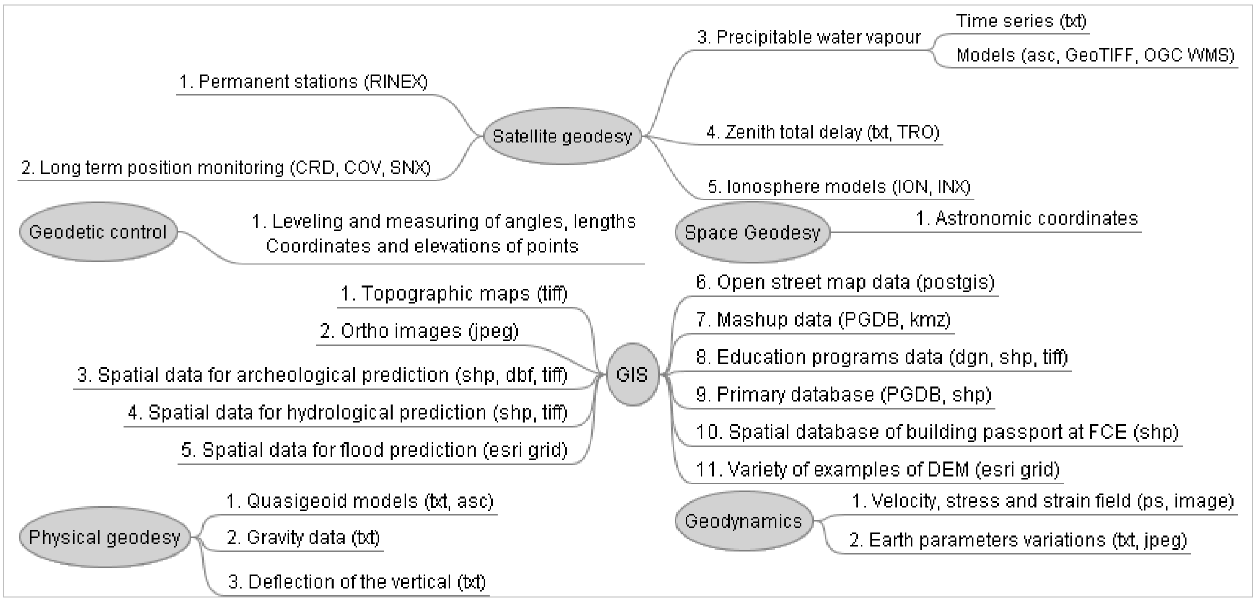

Fig. 2 Department of Theoretical Geodesy – Mapping to data resources

Initial metadata profile

The next investigation step is obtaining the samples of data resources related to particular actions. These samples are further being used to test automatic metadata extraction from them. Software CatMDEdit (http://catmdedit.sourceforge.net/) is used as a tool for extraction. CatMDEditor is an user friendly desktop application for documentation of resources, with focus on geospatial resources. We use this tool to obtain some information from particular data resources (formats) in automatic way as well as to create metadata as XML files defined by ISO (International Organization for Standardization) gmd (ISO, 2007) schema. Content of these metadata will be moreover extended following further interviews with stakeholders to fulfil INSPIRE requirements for metadata defined in metadata regulation (INSPIRE, 2008). Therefore the metadata model defined by INSPIRE regulation is taken into account as the *initial metadata profile* in our work. The initial metadata profile is shown in Fig. 3.

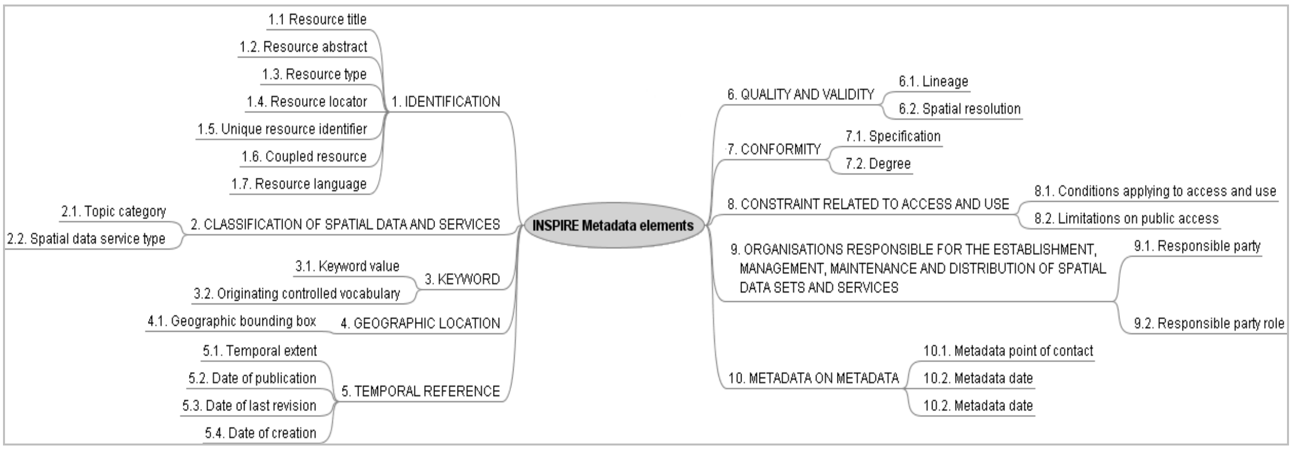

Fig. 3 INSPIRE metadata model (INSPIRE, 2008) – initial metadata profile elements

We will collect all metadata elements illustrated in Fig. 3 on the dataset hierarchy level for discovery purposes. However, the high data diversity within the domain means that INSPIRE metadata may not accommodate all applications. Hence, we will later on extend them again interviewing stakeholders by further elements defined in ISO metadata standard (ISO, 2003) for purposes of evaluation and use. We will create particular community profiles using the rules defined in ISO standard to better serve special stakeholders needs. All metadata records will be created, validated and catalogued using geo-catalogue application to provide web interface for searching, editing and updating of them, as well as to provide discovery service interface endpoint for querying from remote clients.

RESULTS

Overall description of the actual state within the domain

We can say, on the basis of the investigation performed with stakeholders that the data diversity within the domain is in wide range. In the following part we provide summarize tables of actual data resources within the domain with basic metadata like, data format, short abstract and keywords obtained from the stakeholders within the interviews. Table 1 summarizes data resources from the Department of Theoretical Geodesy.

Table 1 Actual data resources of the Department of Theoretical Geodesy (Summary table)

Session	Keywords	Format	Abstract
Geodetic control	control points, coordinates, reference system, reference frame, reference epoch	txt, dgn	Coordinates and elevations of points
Space Geodesy	astronomical coordinates, astronomical catalogue		Astronomical coordinates
Geodynamics	Earth, parameters, polar motions, precession, nutation, lenght of day	txt, jpeg	Earth parameters variations
	velocity, stress field, strain field	ps, image	Velocity, stress and strain fields
GIS	topography, contours, ortho-imagery	tiff jpeg	Topographic maps Ortho images

	hydrological prediction	shp, dbf, tiff	Spatial data for hydrological prediction
	flood prediction	esri grid	Spatial data for flood prediction
	open street map	postgis database	Open street map data
	mashup, information system	esri pgdb, kmz	Mashup data
	education, training	dgn, shp, tiff	Data for education programs
	primary database for the GIS	esri pgdb, shp	Primary database for the GIS in Slovakia
	passport of building, structure of faculty	shp	Spatial data of Faculty of Civil Engineering passport
	DEM	esri grid	Variety of examples of DEM
Satellite Geodesy	global navigation satellite system	RINEX	Permanent stations measuring data
	global navigation satellite system	CRD, COV, SNX	Long term positions monitoring
	precipitable water vapour, meteorology, monitoring	txt, asc, GeoTIFF, OGC WMS	Precipitable water vapour determination. Time series data and models.
	zenith total delay	txt, TRO	Zenith total delay
	Ionosphere	ION, INX	Ionosphere models
Physical geodesy	Quasigeoid	txt, asx	Quasigeoid models
	Gravity	txt	Gravity data
	Deflection	txt	Deflection of the vertical

Table 2 summarizes data resources from the Department of Surveying and the last table (Table 3) summarizes data resources from the Department of Mapping and Land Consolidation:

Table 2 Actual data resources of the Department of Surveying

Session	Keywords	Format	Abstract
Surveying Engineering surveying Photogrammetry	building, current state	txt, dxf	Measurements of current state
Surveying Engineering surveying	control point	txt, dxf	Setting-out networks, points – control
Surveying Engineering surveying	project data	dwg	Project data – data inputs from architects and planner for geodetic works
Surveying Engineering surveying	hydro meteorological parameters, temperature, atmospheric pressure	doc, txt	Hydro meteorological parameters –temperature (measured), atmospheric pressure, humidity.
Surveying Engineering surveying Photogrammetry	cloud of points	txt, dxf	Cloud of points of measured objects
Surveying Engineering surveying	building tilt	xls,	Measurements of tilts
Surveying Engineering surveying	Cadastre	vgi	Cadastral maps
Engineering surveying	displacement	txt, dxf	Displacements measuring data

Photogrammetry	digital images	jpeg, tiff, raw data	Digital images
	photogrammetry, control points, control feature	txt, dxf	Set of control points, set of control features
	photogrammetry	txt, dxf	New feature collections – edges, lines, polylines, polygons, surfaces
	ortho images	tiff, jpeg	Scaled ortho images
	simulations, animations	avi, wmv	Simulations and animations

Table 3 Actual data resources of the Dept. of Mapping and Land Consolidation

Session	Keywords	Format	Abstract
	Topographic map	cit, jpeg, dgn, dwg, shp	Digital raster and vector topographic maps
	digital elevation model, elevation	dgn, txt	DEM
Cartography	primary database for the GIS	shp	Primary database for the GIS in Slovakia
	thematic map, parcel	jpeg, tiff, shp	Thematic raster and vector cadastral maps
	Imagery	jpeg	Ortho-photographs
	3D model	kmz, SKP	3D models of real objects
Cadastre of Real Estate and Mapping	cadastre, cadastral map, parcel	cit, vgi	Digital raster and vector cadastral maps
	Cadastre	dbf, txt	Descriptive Data File
Land Consolidation	land consolidation, land use, land management	vyk, vgi, dgn, GeoTIFF	Projects of land consolidation
	orto imagery	jpeg	Ortho-photographs

For each obtained data samples is planned to create an Initial metadata record, based on INSPIRE regulation as mentioned above. The example of Initial metadata record was created for one dataset with name **St_Mesto_2010_02_ar_shp**. This metadata records was created by CatMDEdit and online INSPIRE multilingual metadata editor (http://www.inspire-geoportal.eu/EUOSME/userguide/about.html). CatMDEdit was used for automatic metadata extraction from dataset. Automatic metadata extraction procedure does not support full requirements defined by INSPIRE rules. Therefore this metadata record was imported in XML file and extended in INSPIRE multilingual editor. The fragment of XML file is listed in Fig. 4. Dataset described by this metadata record (fragment of XML file) is showed in Fig. 5.


```
<gmd:MD_Metadata xsi:schemaLocation="http://www.isotc211.org/2005/gmd http://schemas.opengis.net/iso/19139/20060504/gmd/gmd.xsd">
 + <gmd:fileIdentifier></gmd:fileIdentifier>
 + <gmd:language></gmd:language>
 + <gmd:characterSet></gmd:characterSet>
 + <gmd:hierarchyLevel></gmd:hierarchyLevel>
 + <gmd:contact></gmd:contact>
 + <gmd:dateStamp></gmd:dateStamp>
 + <gmd:metadataStandardName></gmd:metadataStandardName>
 + <gmd:metadataStandardVersion></gmd:metadataStandardVersion>
 − <gmd:identificationInfo>
   − <gmd:MD_DataIdentification>
     − <gmd:citation>
       − <gmd:CI_Citation>
         − <gmd:title>
             <gco:CharacterString>St_Mesto_2010_02_ar_shp.</gco:CharacterString>
           </gmd:title>
         + <gmd:date></gmd:date>
         − <gmd:identifier>
           − <gmd:RS_Identifier>
             − <gmd:code>
               − <gco:CharacterString>
                   KMPU_Cartography_Thesis_ThematicCadastralMaps_poly_shp
                 </gco:CharacterString>
               </gmd:code>
             </gmd:RS_Identifier>
           </gmd:identifier>
         </gmd:CI_Citation>
       </gmd:citation>
```


Fig. 4 Fragment of XML file for dataset St_Mesto_2010_02_ar_shp

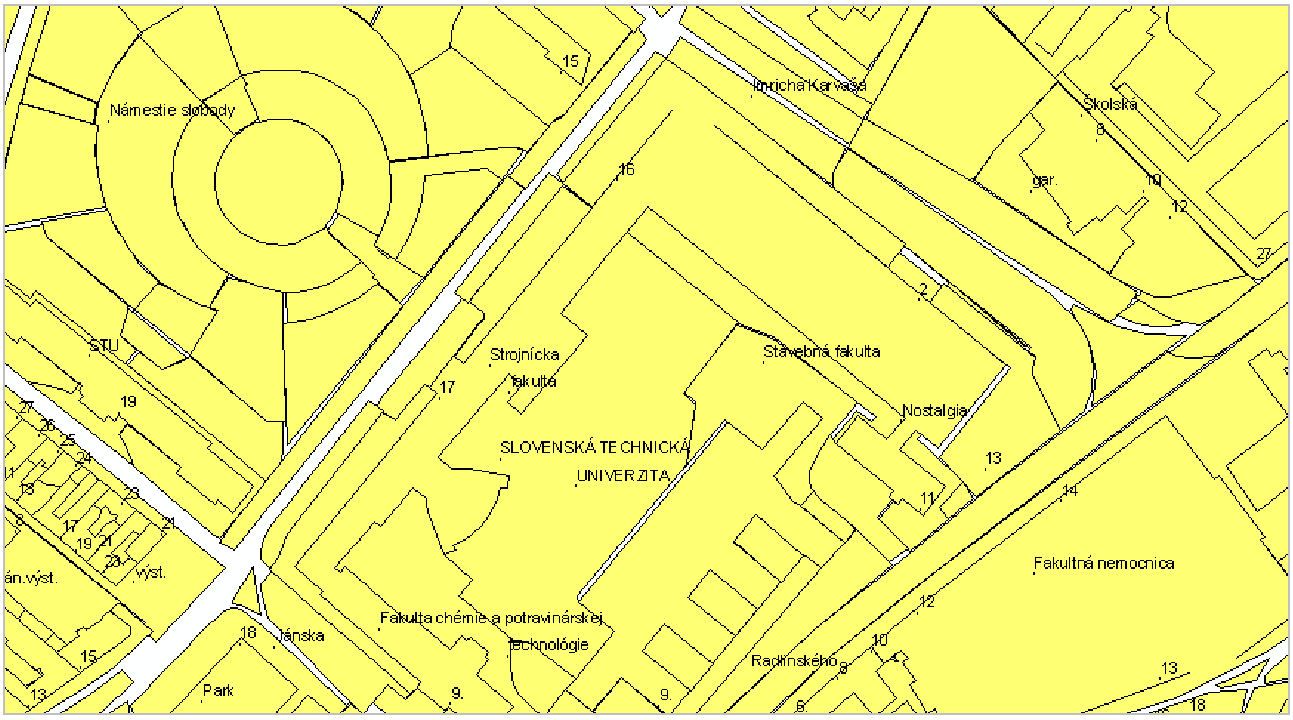

Fig. 5 Sample of dataset portrayal - St_Mesto_2010_02_ar.shp and St_Mesto_2010_02_tx.shp

DISCUSSION

The main objective of the project is data harmonization within selected domain and further implementation of network services that ensure discovery, view, download and processing of these harmonized data. There are many opened issues and questions that will most probably arise during the processing of particular project tasks: (i) Investigation of existing geospatial data resources within the domain – What are the heterogeneities within the data structures? What are the actual data flows? What are the application models used (if there are any)? How to model of existing work flows and propose new? (ii) Modelling of new application models based on reference data specifications (INSPIRE) and user requirements with extensions defined by domain experts – users, data providers and data transformers (producers of value added products based on the data themes of INSPIRE) – Which INSPIRE data specification should be used? How those models should be simplified or extended to ensure user requirements? (iii) Mapping and transformation data resources to the new models and testing the conformance of the mapping process. (iv) Metadata creation – What will be the reference metadata model (INSPIRE)? How to extend this model with specific metadata elements defined in standard ISO 19115 to ensure data resource description according to user requirements for discovery, evaluation and consumption purposes? How to ensure interconnection among related resources (datasets, services) by particular metadata elements? (v) Network services and SDI portal implementation to provide connection between end users and datasets. What types (discovery, view, download, transformation, invoke) of the network services should be implemented regarding to user requirements? How application profiles (OGC, INSPIRE) of particular services should be simplified or extended? What issues need to be taken into account in particular client application (discovery, view, download, transformation) of the final SDI portal development?

CONCLUSION AND FUTURE WORK

To achieve the above mentioned main objective of the project we would propose the following approach of solution:

1. General study (above)

2. Study of user requirements

- Development of highly structured methodology for workflow mapping to aid the final requirements capturing.
- Application of methodology within the organization among the particular data source responsibilities for further advanced user requirements.

On the base of user requirements study (performed within first step) we will model use cases for developing system using graphic notation of UML language (Unified Modeling Language). For modeling tasks we will use available CASE tools (Select Architect or open source tools like StarUML (http://staruml.sourceforge.net/en/) or ArgoUML (http://argouml.tigris.org/)).

3. System architecture design based on studies performed in previous steps.

Based on use case definition and modelling performed within point 2 we will define the most important entities (data, metadata, services) and their relationships. It will serve as a model to help with understanding the nature of the entities and relationships, their purposes, context and how they could interact to realize use cases and workflows. We will use the same tools mentioned in previous point for system architecture design. The second part will be proposals and evaluation of software tools that could be used for practical implementation of particular system components.

4. Step by step implementation

- Consolidation and harmonization of existing data inventories to the common application models defined by current directions and standards and extended on the base of results from general study realized within the point 1 of this approach. Reference data will be harmonized to common application model based on INSPIRE data specifications.
- Reference data will be described by metadata (metadata creating) and catalogue (metadata management within searchable metacatalogue) in compliance with directives (INSPIRE), standards and implementation specifications (ISO, OGC). Realization of this step will allow organizations to easily discover data sources that are available and gain information like who are responsible persons and what are the licensing conditions.
- An intranet application will be developed that allows users to visualize reference data in accordance with directive (INSPIRE), standards and implementing specifications (ISO, OGC).
- Development of portal extension for download, transformation and geoprocessing facilities based on the study of advanced user requirements.
- Testing phase – pilot services deployment within reference data and service portal, feedback from users.

5. Analysis and development of methods and procedures for processing of spatial data based on WPS (Web Processing Service)

- To develop each process functionality part of WPS based on user requirements. Following current needs service would implement processes to achieve these tasks:
  - Functions using models of geoid and quasigeoid for vertical deviation and altitude calculations.
  - Geo-kinematics functions and long term monitoring of geodetic points (point stability determination, deformation parameters calculations).
  - Functions based on DEM (determinations of altitude, terrain slope, visibility analysis etc.).
  - Data conversion among various data formats used within geodetic and cadastral domain.
  - Transformations among coordinate systems.

- Methods and procedures for data processing will be developed for facilitate the work with reference data sources for each group of users.

All above mentioned answers and solutions of the proposals probably and even more questions are the main goals of the proposed project. With this paper we would like to point out aspects, which many of the stakeholders will have to consider in near future because of new legislation drivers (INSPIRE) as well as upcoming technology developments and new users expectations.

This paper has been made thanks to VEGA No 1/0894/10.

REFERENCES

INSPIRE (2008) Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata - Metadata Implementing rules, European Commission, Brussels.

ISO (2003) 19115:2003 Geographic information - Metadata, Switzerland.

ISO (2007) 19139:2007 Geographic information – Metadata – XML schema implementation, Switzerland.

Masser, I., (2005) The future of spatial data infrastructures, Proceedings of ISPRS WG IV/1 Workshop on 'Service and Application of Spatial Data Infrastructure' (Hangzhou, China, 2005), http://www.isprs.org/proceedings/XXXVI/4-W6/papers/7-16%20Ian%20Masser-A001.pdf ,

FreeMind, http://freemind.sourceforge.net, October 29, 2010

FreeMind, http://en.wikipedia.org/wiki/FreeMind, October 29, 2010

CatMDEdit, http://catmdedit.sourceforge.net/, October 29, 2010

EUOSME, http://www.inspire-geoportal.eu/EUOSME/userguide/about.html, October 29, 2010

StarUML, http://staruml.sourceforge.net/en/, October 29, 2010

ArgoUML, http://argouml.tigris.org/, October 29, 2010

GSDI, http://www.gsdi.org/, December 30, 2010

GEOSS, http://www.earthobservations.org/geoss.shtml, December 30, 2010

INSPIRE State of play reports, http://inspire.jrc.ec.europa.eu/index.cfm/pageid/6/list/4, December 30, 2010

AuScope, http://www.auscope.org.au/, December 30, 2010

Géoportail, http://www.geoportail.fr/, December 30, 2010

INSPIRE CZ, http://inspire.gov.cz/, December 30, 2010

INSPIRE SK, http://inspire.enviroportal.sk/, December 30, 2010

GEONET.SK, http://www.geonet.sk, December 30, 2010

Geoportal UGKK SR, http://www.geoportal.sk, December 30, 2010

ENIPI, http://isu.enviroportal.sk/index.php/item/enipi, December 30, 2010

CENTROPE, http://pgo.centropemap.org/, December 30, 2010

eSDI-NET+, http://www.esdinetplus.eu/, December 30, 2010

HUMBOLDT , http://www.esdi-humboldt.eu/home.html, December 30, 2010

Turistický portál Posázaví, http://tourist.posazavi.com/cz/mapvylet.asp, December 30, 2010

GIS portál Mesta Banská Bystrica, http://www.uhabb.sk/mapa/htdocs/bystrica.php, December 30, 2010