

Metainformation system of geodata for study and practice

Zdena Dobešová

Department of Geoinformatics, Faculty of Science, Palacký University,
Tr. Svobody 26, 771 46 Olomouc, Czech Republic
zdena.dobesova@upol.cz

Abstract: Metainformation systems are pivotal when carrying out operations with geodata, for example, sharing and exchange. This paper presents the metainformation system MICKA, which is running at Palacký University in Olomouc.

Metainformation system is introduced at the department administration board requests. Own metadata department profile *UPOL* is used. This profile complies with the ISO 19115 standard. Metainformation system is a system designed to store three categories of metadata: bought (acquired) data for educational purposes, data created during the scientific and research projects, and students' data that are created in the bachelor, master or doctoral theses. The student diploma works produce the largest number of data sets.

Our own system of coded identifiers was designed to distinguish between the groups of data acquisition. Moreover, user groups, as well as individual users, were defined. Groups and users have specific rights that allow them to search, create and edit metadata records. Metadata records are divided into public and private (non public). Public records are open for anyone in the world to browse through. This is achieved by including the Metainformation system in the Czech National Geoportal INSPIRE [7]. Currently, the system holds over 370 records, of which 118 are public. The non public records are mostly the geodata that were purchased or licensed and, therefore, are not to be shared further.

Metainformation system is in compliance with the requirements of the INSPIRE directive. Information about the spatial extent, data format, contact person, data lineage, and data quality is important from the interoperability point of view. All these data are recorded in metainformation system MICKA. Every new metadata entry is checked to meet the requirements for mandatory

information. The MICKA metadata catalogue will certainly contribute to increase the availability of geographic data not only within the university department. The implementation of the system at Palacký University is so far a unique project of the Czech education system. The metainformation system should immensely contribute to higher availability of data not only at the university but within the scope of all users in geoinformatics profession.

Key words: INSPIRE, maintenance, metadata, MICKA.

INTRODUCTION

According to the 2007/2/ES directive about setting up an Infrastructure for spatial information inside the European association INSPIRE [3, 4], which came to power as a part of the law 123/1998 Sb., the Palacký University is one of the institutions obliged to provide spatial data into the infrastructure. Nevertheless, the necessity to introduce a metainformation system was felt at the department since the year 2004. Every year the numbers of bought, acquired and created data were increasing. In 2004 metadata from students' works were collected for the first time. At that time, the MIDAS Lite was used for that purpose. MIDAS was designed at VSB Technical University in Ostrava, Institute of Geoinformatics [1, 6].

However, with the INSPIRE directive coming into power, it was necessary to migrate to another metainformation system. Metainformation system MICKA was chosen as a new metadata maintenance tool in 2009. The MICKA system is a product of HR-RS, Help Service Remote Sensing company [2]. This system was selected mainly for the fact that it was best complying with the INSPIRE requirements. Moreover, it was the most widely used metainformation system in the Czech Republic. Metainformation system MICKA runs directly on our department server and uses the PostgreSQL database server. Since 2010, the system is fully functional and is regularly filled with new metadata. The system mainly serves the employees of Palacký University and students of Geoinformatics department; it is, however, also open to public.

PRACTICAL PURPOSE

The metainformation system runs on the following URL address: <http://gislib.upol.cz/metadata> [5]. Searching for metadata is either performed by lexical data input (word, keyword, thematic category) or by selecting a bounding box directly inside a map on the right side of the window. It is also possible to combine both. Detailed description of the search process can be found in the user's guide (button Help on the left menu). When browsing the records, only public entries are displayed to a user which is not logged in (user

guest). The search screen of the system is on Fig. 1. The result of a search in “elevation” category is shown in a well-arranged table (Fig. 2). The table includes the names of the metadata entries (bold blue letters), and corresponding abstracts. More details are displayed after clicking the record name.

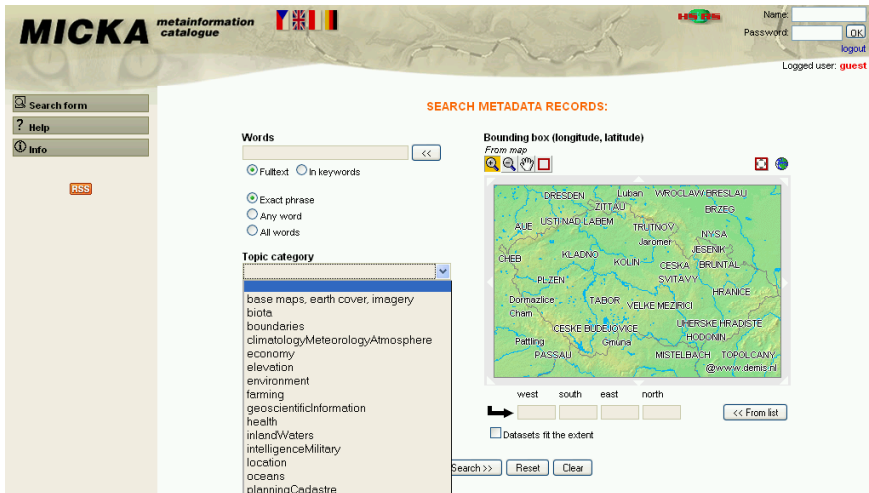


Fig. 1 Interface of Metainformation catalogue MICKA – mode Search

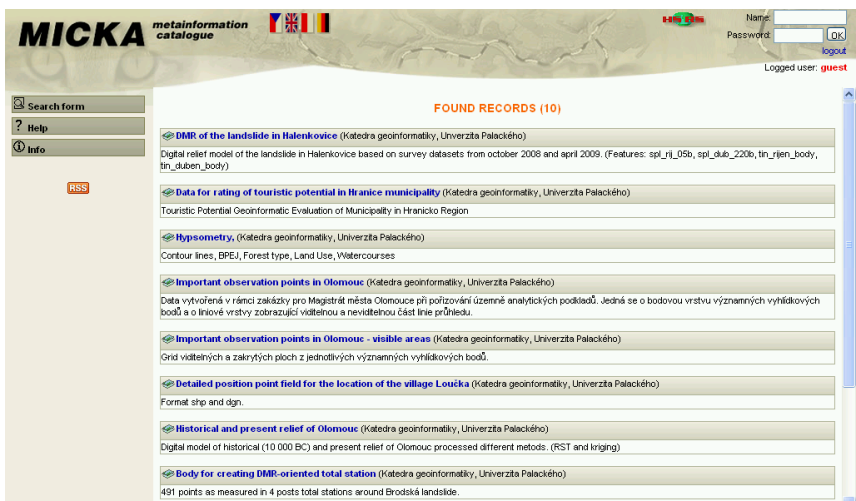


Fig. 2 Resultant metadata records for search in "elevation" category

STRUCTURE OF THE METADATA

All recorded metadata comply with ISO 19115 standard. For the purpose of storing metadata, our own *UPOL* profile was established. Moreover, a system of coded identifiers was created for the purpose of searching and differentiating the records [8]. This code is recorded in the *Identification – Code* element.

Table 1: Coded identifiers for identification of the data origin

Coded identifiers	Meaning
KGIDATA	Purchased (received) data
KGIPROJ	Data of scientific project
KGISBP	Data of student bachelor thesis
KGISDP	Data of student diploma thesis
KGISPP	Data of student doctoral thesis

The next required value is the name of the diploma work and the name of the supervising teacher. These are entered into the *Identification – Purpose* element. Descriptive *Keywords* are chosen from GEMET or INSPIRE thesaurus. The name of the student is stored in the *Metadata Contact* element. Spatial representation, reference system, and data format are another important values saved together with every record. *Data Quality – Lineage* can also be of high significance. Short text, named type of spatial analysis, description of the steps in data creation process, type and technical parameters of measurements (GPS, theodolite) etc. are mentioned in *Lineage* element.

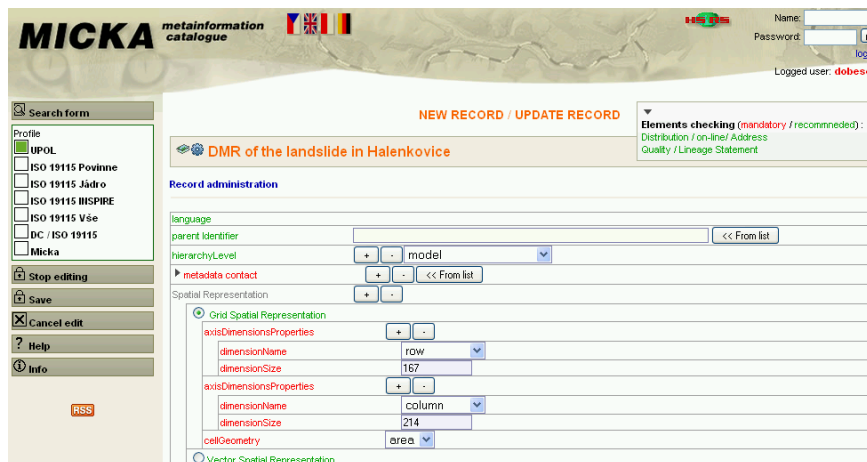


Fig. 3 The interface for "New record /Update record" with checking mandatory/recommended elements (upper right menu)

In order to ensure correct metadata input, a manual with instructions was made and included within the application. In addition, the mandatory fields are checked for errors when new information is being added (Fig. 3.). Red-coloured fields are mandatory, while green fields are recommended. Information is filled both in Czech and English, so that the metadata records can be searched by users all over Europe.

WAYS OF OBTAINING METADATA

Altogether, there are three sources of data whose metadata are stored in the metainformation system. The first group is represented by the data that are purchased or acquired from external distributors. These are vector or raster data sets provided by companies and state organisations such as ČÚZK (Czech Office for Surveying, Mapping and Cadastre), ČSÚ (Czech Statistical Office), ArcDATA Praha, GEODIS etc. These data are of pivotal importance, and serve as a base for education and diploma works. It is also essential to let the members of the department know that these data are available. In total, there are 34 data sets from this group stored in the system. These entries are not public due to the fact that the University is not authorised to share them further on. The data of this sort are added by the department administrator.

The second group comprises of the data created during a scientific and research projects of the department. This group contains the smallest number of 9 entries. The data from this group are always added by the administrator after a consultation with the author of the data.

The third group is the largest, and it is formed by the data produced by students' diploma works. Since 2009, the entries are added by the student themselves. The main reason for this is the fact that the author of the work is the one who can best summarize and describe the data as far as the content, accuracy, quality and the process of formation are concerned. For this purpose, two extra users of the system are defined "bakalar" (bachelor) and "magistr" (master). Both of the two extra users have permission to add new entries. The only restriction is the time period within which the students have to upload their records. They are obliged to do so in the span of two weeks before the final deadline of their works. The final review of the new records is done by the administrator, who is also in charge of setting the entries as public or non public, depending on their type and purpose. Students also include the source data, which they used for the diploma work and which are always non-public. In most cases, these are 10 free map lists from ČÚZK.

Apart from the metadata of the contemporary works (last 3 years), metadata from the works since the year 2000 were also recorded. The precise description was, however, more difficult, as the accuracy and the method of the data formation could not be determined.

Table 2: Numbers of metadata records (15. August 2011)

Type of record	Number (count)
Total count of records	370
Public record	118
Nonpublic (private) record	252
Student bachelor thesis	138
Student diploma thesis	187
Purchase (receive) data	36
Data of scientific project	9

Gradually, the number of the records is growing. Every year, about 40 new student works extend the total amount. Yet not every work introduces absolutely new data. Additionally, latest purchased and research project data contribute to the number of entries.

Metadata (ISO 19115)	
Spatial Representation	
Grid Spatial Representation:	
axisDimensionsProperties:	
dimensionName:	row
dimensionSize:	167
dimensionName:	column
dimensionSize:	214
cellGeometry:	area
Reference System	
MD_ReferenceSystem:	
referenceSystemIdentifier:	
RS_Identifier:	
code:	S-JTSK
codeSpace:	CZ
Identification	
Data Identification:	
citation:	
Citation:	
title:	DMR of the landslide in Halenkovice
date:	
CI_Date:	
dateType:	publication
date:	2009-05-05
identifier:	
RS_Identifier:	
code:	KGISDP-2009
abstract:	Digital relief model of the landslide in Halenkovice based on survey datasets from october 2008 and april 2009. (Features: spl_rji_05b, spl_dub_220b, tin_rjen_body, tin_duben_body)
purpose:	Diploma thesis.
status:	completed
Point of Contact:	
Responsible Party:	
Individual Name:	Zdena DOBEŠOVÁ
Organisation Name:	Katedra geoinformatiky, Univerzita Palackého

Fig. 4 Detail of the metadata record

The metadata information system is managed by a member of the department – the administrator of the metadata (Z. Dobesova). The administrator is also filled in *Point of Contact* element for the purpose of data distribution. More user groups and users were created for the maintenance purposes. The groups are “*administrator*” and “*diplomant*”. When a student needs access to the non public entries the account “*Student*” is to be used.

References

- [1] Horáková, B. MIDAS on-line geodata catalogue. Need for legislation to support technological solutions, GIM International, Volume 17, Issue 7, 2003, pp. 12-13.
- [2] HS-RS. Metadata, HR-RS Help Service Remote Sensing, Benešov URL: <<http://www.bnhelp.cz/produkty/metadata/>> [cit. 15. 8. 2011].
- [3] INSPIRE. Infrastructure for spatial Information in Europe. URL: <<http://inspire.gov.cz>> [cit. 15. 8. 2011].
- [4] INSPIRE, European Commission, Inspire Directive URL: <http://inspire.jrc.ec.europa.eu/> [cit. 15. 8. 2011].
- [5] MICKA. Metainformation catalogue, Department of Geoinformatics, Palacký University in Olomouc, URL: <<http://gislib.upol.cz/metadata>> [cit. 15. 8. 2011].
- [6] MIDAS, URL: <http://gis.vsb.cz/midas/> [cit. 15. 8. 2011].
- [7] Národní geoportál INSPIRE Metadata. [National geoportal INSPIRE] URL <<http://geoportal.gov.cz/web/guest/uvod/>>. [cit. 15. 8. 2011].
- [8] Šrubl, O. Metainformation system of geodata of Department of Geoinformatics, diploma thesis, 51 p., 2010, Olomouc.

Shrnutí

Pro sdílení a výměnu geodat je důležité provozovat metainformační systémy. Řada geodat vzniká v rámci studentských prací na vysoké škole. Na Universitě Palackého v Olomouci je od začátku roku 2010 provozován naplněný metainformační systém MICKA.

Metainformační systém byl upraven pro potřeby pracoviště. Používá se vlastní metadatový katedrální profil UPOL odpovídající normě ISO 19115. Metainformační systém má navržený systém pro ukládání třech skupin metadat: nakoupená (získaná) data pro výuku, data vzniklá v rámci vědeckých a výzkumných projektů a studentská data, která vznikla v rámci bakalářských, magisterských nebo doktorských prací. Největší počet záznamů je o datech vzniklých v rámci studentských prací. Byl navržen vlastní systém kódových identifikátorů pro odlišení skupin dat a roku pořízení. Jsou nadefinovány skupiny uživatelů a jednotliví uživatelé. Skupiny a uživatelé mají nadefinovaná práva, která umožňují vyhledávání a nebo tvorbu a editaci metadatových záznamů. Metadatové záznamy jsou rozděleny na veřejné a neveřejné. Veřejné záznamy jsou přístupné komukoliv v ČR. Informace o datech jsou tak přístupné i pro uživatele z praxe. Metainformační systém je zařazen do

celostátního vyhledávání pod Národní geoportál INSPIRE [7] . Řada dat, které se zpracovávají pro reálné potřeby v praxi se běžně z univerzity předávají na koncová pracoviště mimo UP.

V současné době je v systému evidováno přes 370 záznamů. Z toho je 118 záznamů veřejných. Neveřejné záznamy jsou o datech, která jsou nakoupena nebo účelově vázána a tudíž nemohou být poskytnuta dále.

Metainformační systém je v souladu s požadavky směrnice INSPIRE. Z hlediska interoperability geodat jsou důležité informace o prostorovém rozsahu, formátu dat, kontaktní osobě, rodokmenu dat, kvalitě dat atd. Všechny tyto údaje jsou v metainformačním systému MICKA zaznamenány. Při vstupu údajů do metadatového záznamu je prováděna kontrola na vyplněnost povinných údajů.

Provozovaný metainformační systém určitě přispěje k vyšší dostupnosti geodat to nejen v rámci pracoviště ale i uživatelů v praxi. Článek prezentuje konkrétní realizaci a provozování metainformačního systému na Univerzitě Palackého, které je v oblasti vysokého školství ojedinělá.