DESIGN OF AN ISO 19115 COMPLIANT PROFILE FOR DOCUMENTING SPATIAL DATASETS AND SERIES OF THE HELLENIC CADASTRE

D. A. Sarafidis[a], O. D. Mavrantza[b], I. G. Paraschakis[a]

[a] Aristotle University of Thessaloniki Greece, Department of Rural and Surveying Engineering, (sarafid, jpar) @topo.auth.gr
[b] KTIMATOLOGIO S.A, Quality Control Executive omavratz@ktimatologio.gr

KEYWORDS: Hellenic Cadastre, vector spatial dataset, metadata, ISO 19115, raster spatial data.

ABSTRACT:

The role of Metadata is to describe and document spatial data. The implementation of metadata serves several goals, such as archiving spatial data, for data managers and searching spatial data, for data users. Spatial metadata is one of the key-components of a geo-information infrastructure. In the present study, an innovative, ISO 19115 - compliant metadata profile is created, in order to cover the additional metadata information needed for documenting spatial datasets and dataset series of Hellenic Cadastre (HC).
Current state of the art (at International and National level) concerning spatial data Infrastructures and metadata standards for spatial data is also presented. Certain issues have been arisen, investigated and resolved concerning the manner in which ISO 19115 is implemented for the creation of the HC metadata profile. These issues are referred to the following: (a) in which manner is the hierarchy of series of (raster / vector) spatial data documented using the ISO 19115 standard, (b) how ISO 19115 can be extended in order to add elements needed for proper documentation of Hellenic Cadastre spatial data and (c) how data quality information is documented. An additional important issue preliminary investigated in this work, concerns the documentation of raster spatial data of the HC. The ISO TC-211 Committee is currently working on an extension of ISO 19115, namely the ISO 19115 - Part 2, in order to cover the documentation of raster datasets. In the present work ISO 19115-2 is being under preliminary research (in its present form) in order to meet the needs concerning the documentation of raster spatial data (aerial photographs and orthophotos).

1. INTRODUCTION

As technology provides faster and more efficient ways to transmit and process geographical data, data producers and users are realizing the potential value of exchanging and sharing spatial datasets. Metadata, or “data about data”, provide information such as the characteristics of a dataset, its history, and the organizations to contact to obtain a copy of it. Standardized metadata elements provide a means to document datasets within an organization, to contribute to catalogues of data that help individuals find and use existing data, and to help users understand the contents of datasets, which they receive from others (Crompvoets et al, 2004).
The evolution of GIS usage in Greece has also increased awareness of the importance of the documentation of spatial data according to a widely acceptable metadata standard, especially as the enormous project of the Hellenic Cadastre (HC) has become one of the largest archives of spatial information in the country (Hatzilakos et al, 2000; Halaris, 2004).
In order to cover additional metadata information needed for the documentation of spatial datasets and dataset series of HC, a new metadata profile is created following ISO 19115 metadata standard rules for creating a compliant profile. This profile is briefly presented in this work.
An additional important issue, which is preliminary investigated in this work, concerns the documentation of raster spatial data of the HC. In the present work, ISO 19115-2 is being under preliminary study (in its present form) in order to meet the needs concerning the documentation of raster spatial data (aerial photographs and orthophotos).

2. METADATA FOR SPATIAL DATA IN INTERNATIONAL AND NATIONAL LEVEL

2.1. Data clearinghouses and spatial data infrastructures

The search of geographic information requires a service for their exploration, a search engine of Geographic Information. Occasionally The Geographic Information Community has given a variety of names for this efford, e.g. the terms "catalogue services" (OpenGeospatial Consortium), "spatial data catalogue" (Australian Spatial Data Infrastructure) and "Clearinghouse" (FGDC USA). Even if different names are used, the common objective requires the exploration of spatial data via their metadata. These services are generally reported as "catalogue services". Many countries have spent considerable resources over the past few years debating optimal national spatial data infrastructures. One of the main elements of these infrastructures is the national catalogue services for accessing required spatial data and provision of complementary services (Crompvoets et al, 2004; NAGII, 2004).
The EU Directive INSPIRE (Infrastructure for Spatial Information in Europe) envisions a panEuropean Spatial Data Infrastructure (pan-ESDI) comprising interlinkage among national and / or regional SDIs of EU Member States. This interlinkage will be achieved only if each individual SDI that comprises the ESDI can communicate via a common global infrastructure. For purposes of geospatial data discovery, (eventual) access and use (exploitation), standards become of paramount importance. The INSPIRE Directive is a legislative document stating that international standards will be used to record and publish metadata about geospatial data holdings at all levels of government, from local to national. The actual standard used is ISO 19115, which is confirmed by the formed INSPIRE Team on metadata (Longhorn, 2005).
In 2000 the Hellenic Mapping and Cadastral Organization (HEMCO), which belongs to the Ministry of Environment, Physical Planning and Public Works, proposed the development of the Hellenic SDI, called NaGi2 (National

Geographic Information Infrastructure), so as economic, social, environmental and planning activities to be facilitated in Greece. NaGi2 will operate as a distributed network of databases based on a set of interoperable standards. The databases will be electronically connected, and they will provide data from various resources, such as ministries, government organizations, private companies and data from the Hellenic Cadastre, to the widest possible group of users. (Alexiadou and Rajabifard, 2006).

2.2. Spatial data standards and profiles

Four main developments have been undertaken to standardize metadata for spatial data (Aalders and Hunter, 2003; NAGII, 2004; Longhorn, 2005):

1. CEN/TC 287 resulting in CEN - ENV 12657 Geographic Information - Data description - Metadata;
2. ISO/TC 211 from 1994 till now resulting in ISO DIS 19115, also used by the OpenGIS Consortium;
3. Dublin Core Metadata Initiative - DCMI
4. FGDC (Federal Geographic Data Committee) resulting in FGDC-STD-001-1998 Metadata Standard

The main metadata standard implemented in the geospatial world comes from the International Organization for Standardization (ISO) Technical Committee 211 – Geomatics / Geographic Information (generally referred to as 'TC 211'), in 19115 standard (NAGII, 2004; Longhorn, 2005).

2.3. The ISO 19115 metadata standard

In May 2003 ISO/TC 211 released the metadata standard ISO 19115 as a result of intensive consultations of organizations pioneering this field, among them in a leading role the US Federal Geographic Data Committee (FGDC), the Technical Committee 287 of the Committée Européenne de Normalisatione (CEN/TC287) and other members of the OpenGIS Consortium. Its development was affected by several standards, including the FGDC and ANZLIC standards, but is more comprehensive than any of them. The FGDC and ANZLIC are establishing "profiles" that consist of metadata elements unique to the ISO 19115 standard. All major standards will be interoperable with this standard (GRDC, 2006).
ISO 19115 defines more than 300 metadata elements (86 classes, 282 attributes, 56 relations), most of which can be applied optionally. At the topmost level, the classes (or entities) are grouped in 14 packages (the "root" "Metadata entity set information" plus 13 dependent packages), which are also available for use in other TC 211 standards. The prerequisite of the various entities or classes is quite variable and flexible. The complex, hierarchical nested structure and relationships between the components are shown using 16 UML diagrams. Additionally, the definitions are listed in a tabulated dictionary. Both the UML diagrams and the dictionary are normative (TC 211, 2003a, GRDC, 2006).

3. A METADATA PROFILE FOR DOCUMENTING SPATIAL METADATA OF HELLENIC CADASTRE

3.1. The need for an HC metadata profile

According to the current state concerning metadata for spatial data stated above, a metadata profile based on ISO 19115 should be created for the documentation of spatial data and series of the HC in order to be incorporated in a national or international SDI.
Due these reasons the design of a metadata profile for the documentation of spatial data of HC based on ISO 19115 standard (HC metadata profile for short in this paper) is under research conducted in the main author's Doctoral Thesis.

3.2. ISO 19115 and the creation of a compliant profile

Several issues have arisen concerning the manner ISO 19115 is implemented for the creation of the HC metadata profile. These issues address the following questions:

1. Which spatial data of HC should be documented?
2. Which is the mechanism that ISO 19115 uses for documenting series of spatial data?
3. How can ISO 19115 be extended in order to add elements needed for proper documentation of HC spatial data?
4. Which ISO 19115-elements are not necessary for the HC metadata profile?

3.2.1. Spatial data of the Hellenic Cadastre: According to the specifications of HC, there are spatial data in both digital and analog form for every survey area (HEMCO, 1997; Arvanitis, 2000; Ktimatologio S.A, 2003).

1. Cadastral spatial data, topographical spatial data, orthophotos and digital elevation models (digital cadastral data)
2. Cadastral, topographic and photogrammetric map sheets (analog form products of the HC).

3.2.2. Documentation of the series of spatial data using ISO 19115: After an extensive research in ISO 19115 contents and in metadata profiles based on ISO 19115 (New Zealand Government Geospatial Metadata Standard, 2004 and EuroRoads profile) it was found that the ISO 19115 standard supports two methods of spatial data documentation:

1. Documentation of many datasets belonging to a series in a single metadata entity
2. Documentation of every dataset in a single metadata entity and a separate metadata entity for the series level metadata.

The method selected for the HC spatial data documentation was the second one. The advantage of this method is that the metadata entity structure is much simpler than the one of the first method.

ISO 19115 standard uses the following elements to document spatial datasets series as well as individual datasets (Sarafidis and Paraschakis, 2006).

1. **fileIdentifier** unique identifier for the metadata file,
2. **parentIdentifier** file identifier of the metadata, which this metadata is a subset (child) of,
3. **hierarchyLevel** scope, which the metadata applies to (dataset level, series level etc.)

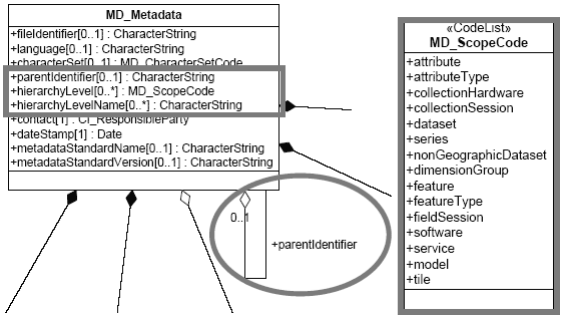

Fig.1 Modeling of metadata hierarchy level (Source: ISO 19115, EuroRoads)

These elements are included in **MD_Metadata** class inside Metadata entity set information package. These elements also are used to HC metadata profile to document the hierarchy of metadata entities between series and spatial data of HC.

3.3.3. ISO 19115 extension for of the HC spatial data documentation: In order to include elements important for the documentation of spatial datasets of the HC, a number of new elements are defined as extension in this metadata profile following the rules stated in ISO 19115 documentation for the creation of metadata extensions.

3.3.4. ISO 19115 unnecessary elements for the HC metadata profile: ISO 19115 elements, which are concerned as unnecessary for the documentation of cadastral spatial data, are not included in the HC metadata profile. These elements are defined as optional in ISO 19115.
Elements contained in Portrayal catalogue information package and in Application schema information package are examples of elements removed from HC metadata profile.

3.3.5. Other changes in HC metadata profile compared to ISO 19115 profile: A number of changes applied to ISO 19115 for the proper documentation of HC spatial data following the rules stated in ISO 19115 documentation. These changes concern the following aspects:

1. The condition or conditions under which certain metadata elements or classes should be documented. An example of such element is **parentIdentifier** (element file identifier of the metadata to which this metadata entity is a subset). This element is defined in ISO 19115 standard as optional. In HC metadata profile is defined as mandatory because is necessary for the documentation of the hierarchy of metadata entities.
2. Restriction or expansion of the domain values of certain elements defined in the ISO 19115 standard. An example of such element is **CI_PresentationFormCode** (mode in which spatial data is represented) which domain values are expanded in order to cover the ability to document cadastral map sheets in digital plot form.

4. DOCUMENTING DATA QUALITY INSIDE HC METADATA PROFILE

4.1 Scope of data quality information in HC metadata profile

Quality elements as well as other elements of the metadata catalogue should not only be documented globally for the entire data set. Instead of that it should be possible to measure and document the data quality also for special spatial and / or thematic aspects (EuroRoads, 2005).
ISO 19115 supports data quality documentation of spatial datasets concerning spatial (planar and vertical), thematic and temporal aspects through the **DQ_Scope** type. which is included in **Data Quality information** package.

Fig.2 Modelling of **DQ_Scope** (Source: ISO 19115)

DQ_Scope data type includes **level** (**MD_ScopeCode** data type) and **levelDescription** (**EX_Extent** data type) elements for the description thematic aspect and extent element for the spatial and temporal aspect. Data quality of HC spatial data are documented for spatial aspect

4.2 Data quality elements documented in HC metadata profile

According to HC specifications, the reference area in data quality measurements in spatial data of HC is the cadastral map sheet. Quality measurements are done in following quality elements for every map sheet

- Positional accuracy
- Attribute accuracy
- Completeness
- Logical consistency

Reference map sheet area can be effectively documented inside metadata documentation using the **EX_GeographicBoundingBox** element, which is part of **EX_Extent** element referred above.

5. DOCUMENTATION OF RASTER SPATIAL DATA

As stated earlier, digital orthophotos and photogrammetric diagrams are also parts of HC spatial data. However, ISO 19115 does not cover the documentation for raster spatial datasets. Due to this reason ISO Technical Committee 211 is currently working on a new standard that extends ISO 19115:2003 Geographic Information – Metadata by defining the schema and additional metadata required for imagery and gridded data. The name of this new standard is "Geographic information — Metadata — Part 2: Metadata for imagery and gridded data" (ISO 19115-2). The ISO 19115-2 Standard extends the existing geographic metadata standard by defining a schema required for describing imagery and gridded data and is still under construction subject to INSPIRE policies This standard provides information about the following subjects:

1. Data acquisition equipment.
2. Geometry of the measuring process of the equipment.
3. Production processes / algorithms used to digitize raw data.

Furthermore the new extended standard includes added the metadata packages (MEs) relative to image and gridded data as follows:

1. *Identification information metadata*: This package contains sub-classes and aggregations concerning the *instrument identification* as well as *information relative to platforms* and / or *missions* for data acquisition.
2. *Data quality information metadata*: This package contains additional sub-classes and aggregations concerning the *sources and production processes* in producing a raster dataset.
3. *Spatial representation information*: The current metadata information package contains sub-classes and aggregations concerning the Georeferencing description (**ME_GeoreferencingDescription)** that contains additional information used to support geo-rectification of the raster / gridded data. **ME_GeoreferencingDescription** is an aggregation of the entities concerning *instrument parameters* (instrumental position according to the ground coordinate reference system), *Control Point Information* (relation of GCPs in the dataset to geographic coordinates) and *Cell Description* (properties of the swath track on the ground).
4. *Content Information*: Provides information concerning (a) Coverage Description (**ME_CoverageDescription)** that in turn is referred to *sensor information* (frame camera / Scanning sensors), (b) Image Description (**ME_ImageDescription)** that provides metadata that determine the suitability of an image for use, and (c) Band (**ME_Band)** that defines additional attributes for specifying properties of individual wavelength bands in the dataset.

HC metadata raster profile schema is still under thorough investigation. Our future work is orientated towards the needs originated from the raster data used in HC and includes the inventory of raster metadata profile parameters that should be taken into account concerning: image data origin (from digital cameras), geo-rectification models (orthorectification), GCP collection (ground / LIDAR), etc

7. CONCLUSIONS

According to the current state concerning catalogue services, the creation of national and international spatial data infrastructures, a metadata profile for the documentation of HC vector spatial data based on International metadata standard ISO 19115 is created. This is done in order HC spatial data to be incorporated HC spatial data in a national or international SDI. A brief description of issues concerning ISO 19115 implementation and how data quality is documented in HC spatial data.
For the documentation of raster spatial data of HC, a brief description of ISO 19115 –2 draft version is made.

8. REFERENCES

Aalders, H. and Hunter, G, 2003. Spatial Data Standards, in *Advanced Geographic Information Systems*, edited by Claudia Maria Bauzer Medeiros. Encyclopedia of Life Support Systems (EOLSS) 26pp.

Alexiadou S. and Rajabifard A., 2006. Developing national SDI platform for Greece. *Coordinates* cGIT Publication, ISSN 0973-2136, pp. 28-34.

Crompvoets J., Rajabifard A., Williamson I., 2004. Assessing the worldwide developments of national spatial data clearinghouses. *INT. J. GEOGRAPHICAL INFORMATION SCIENCE* VOL. 18, NO. 7, pp 665–689.

Global Runoff Data Centre (GRDC), 2006. "ISO 19115", http://grdc.bafg.de/servlet/is/2376/?lang=de&highlight=iso,19115 (accessed on 13 Mai, 2007)

EuroRoadS, 2005. Report WP 6.8 "D 6.8 Metadata Catalogue", EuroRoads. http://www.euroroads.org/php/Reports/D6.8%20Metadata%20catalogue.pdf (accessed on 13 Mai, 2007)

Hatzilakos A., Halaris G., Kavouras M.,Kokla M., Panopoulos G., Paraschakis I., Sellis T., Tsoulos L., Zervakis M., 2000. Interoperability and the definition of a National standard for geospatial data. *International Journal of Applied Earth Observations and Geoinformation (JAG)*, Vol 2, Issue 2, pp. 34-42.

Halaris G., 2004. Metadata and their role in National Geograhic Information Infrastructure. Proceedings of the 3th conference of HellasGI (in Greek), Athens, Greece. http://www.hellasgi.gr/papers3/Xalaris_2004.pdf (accessed on 13 Mai, 2007)

ISO TC 211, 2003. ISO 19115. Geographic Information – Metadata. ISO/TC 211 Geographic Information / Geomatics.

ISO TC 211, 2003. Metadata – Part 2: Extension for imagery and gridded data. ISO/TC 211 Geographic Information / Geomatics.

Longhorn R., 2005, "Geospatial Standards, Interoperability, Metadata Semantics and Spatial Data Infrastructure", paper for NIEeS Workshop on Activating Metadata, 6-7 July 2005, Cambridge, UK.

NAGII, 2004, "Organization of system and data of NAGII – Final report", NAGII, Greece. http://ontogeo.ntua.gr/nagii/FINAL%20Deliverable%202.2.htm (accessed on 1 Feb. 2007).

Sarafidis D., Paraschakis I., 2006. A tool for managing ISO 19115 compliant metadata for the spatial dataset series of Hellenic Cadastre in ArcGIS 9, 21st European Conference for ESRI Users, 2006 Athens Greece.