

GRASS GIS Metadata Management Analysis

M. Landa

Department of Mapping and Cartography
Faculty of Civil Engineering
Czech Technical University in Prague
Thákurova 7, 166 29 Prague 6, Czech Republic
martin.landa@fsv.cvut.cz

January 7, 2007

Introduction to Geospatial Metadata

General conception of “metadata” has been studied in several scientific papers since the 1970s. The origin of the word came from the Greek and can be translated as “data about data”. The ISO 211 Technical Committee defines metadata as a structured data describing content, its quality and other (geo)data properties.

Metadata is an essential addition for particular and accurate identification, verification and interpretation of data. Metadata for spatial-orientated data follows the same basic principals as metadata for non-spatial data. On the other hand metadata for spatial data has unique characteristics whose formulation (i.e. description and structure) is still under development. This kind of metadata is oriented rather on spatial data description and should be organised using specialised metainformation systems.

Geodata documentation can be different depending on the regional or national nature of the geospatial infrastructure. Moreover the documentation is very often mutually incompatible. This fact justifies the effort of creating an international standard for metadata description – *ISO 19115, Geographic Information – Metadata*.

Beginning of the standardisation process in this field dates back to the 1990s. Internationally used standards for metadata in geoinformatics are mainly U.S standard *FGDC Content Standard for Digital Geospatial Metadata*, European *CEN ENv 12657, Geographic Information – Data description – Metadata* (its development has been stopped) and finally the *ISO 19115, Geographic Information – Metadata*. There is strong tendency to integrate the FGDC and ISO standards.

One of the prominent initiatives in information technology (IT) field is *The Dublin Core Metadata Initiative* (DCMI). DCMI is an open organisation engaged in the development of interoperable online metadata standards. DCMI defines the Dublin Core Metadata Element Set (DC) which is a vocabulary of fifteen properties for use in resource description (the ISO standard 15836). All metadata standards should follow the DC instructions.

Sophisticated metadata management is going to be an essential feature of all GIS applications. Most of the commercial GIS products reflect this need and support in a given level

the ISO standard 19115.

GRASS GIS and Metadata Management

Capabilities of the world's leading Free Software GIS — GRASS — in metadata handling are strongly delimited and moreover not unified. In the result, GRASS GIS is not able to compete in this field (in contrast with e.g. raster data processing) with comparable applications.

Metadata management in GRASS GIS is relatively poor, considerably out-of-date, and not extensible. Metadata form, user interface and especially API (Application Programming Interface) is different depending on the data type (raster, vector map layers). The core libraries (GIS, raster and vector library) have been created and maintained in different stages of the GRASS development.

Metadata can be an integral part of the data (often without any significant organisation) or stored as a separate unit. Metadata management which is the specific part of the system has more flexible structure. The next stage is an independent metainformation system separated from the data management system. In any case metadata management should be a fundamental part of any Geographic Information System (GIS) and not an extension of functionality.

Metadata can be stored in various formats, which are very often XML-based. XML as a language is highly worth considering metainformation model requirements. Data structure is strictly defined; XML documents can be validated against DTD (Document Type Definition) or XSD (XML Schema Definition). XML is an open and widely used industrial platform independent standard orientated to data publication and distribution. XML documents can be kept as flat files in file system or stored in database management system (DBMS). There are various implementations available based on the relational, object-oriented and native XML DBMS.

GRASS metadata can be represented as a tree structure; the root of the structure is "GRASS location". A GRASS location is defined by its coordinate system, map projection and geographical boundaries. Each GRASS location can have many mapsets. Metadata of the GRASS data layer contains metadata elements of mapset and location metadata set.

Conclusion

The ISO standard 19115 is going to be widely used. In this connection we can also mention harmonisation effort with U.S. standard FGDC-CSDGM or the INSPIRE project in Europe. Although software support of the ISO standards in this field is not so great after all.

Currently GRASS GIS includes very limited metadata management system (no support of any standard). New design and subsequent implementation of metainformation management GRASS subsystem should reflect the prevailing trends, natively support the international ISO standard 19115 including ISO 19139 implementation and other ISO standards.

References

- [1] Martin Landa. GRASS GIS Metadata Management Analysis. In International Symposium GIS Ostrava 2007, January 28-31 2006.
- [2] Martin Landa. GRASS GIS: Importing and processing of cadastral data in Czech cadastral exchange data format. In International Symposium GIS Ostrava 2006, January 23-25 2006.