

An Overview of the Use of GML in Modern Spatial Data Infrastructures

Viktors Skoks, *Riga Technical University, EURAC, Bolzano*, Christian Steurer, *EURAC, Bolzano*

Abstract – This paper introduces an overview of the use of Geography Markup Language in modern Spatial Data Infrastructures. The goal of the paper was to indicate some of the main consequences of the use of Geography Markup Language in the important geospatial data harmonisation processes, both search and access, which are in current use. In order to show a practical example of the use of Geography Markup Language, the system for Earth observation data processing and distribution at the Institute for Applied Remote Sensing at EURAC, Bolzano was studied. The results of the paper set out how Geography Markup Language is used in modern Spatial Data Infrastructures, and the degree to which the Geography Markup Language standard is helpful in achieving data harmonisation and interoperability.

Keywords – GML, HMA, INSPIRE, metadata, SDI

I. INTRODUCTION

One commonly faces the challenge of searching and obtaining some necessary information, because when problems arise, data have to be acquired, decisions made, and responses initiated as rapidly as technology permits. Even with the help of the Internet it is sometimes difficult, time-consuming or, at worst, impossible to find the information one wants, especially if the information is heterogeneous, not-harmonised and only available from multiple, inconsistent sources. In general information can be described as ‘harmonised’, when it is both clear and standardised, although when we speak about harmonisation within the context of an information system, this term is much wider. Harmonisation in the context of information systems involves the adjustment of differences and inconsistencies between different measurements, methods, procedures, schedules, specifications, or systems in order to make all of these uniform or mutually compatible, and thus improve the whole system’s productivity. One might think that one can avoid the need for harmonisation by using one of the popular search engines such as Google. However, search engines are of limited use in providing specific information, when the information concerned is only used by a small number of specialists in order to create new products, make decisions or solve specific problems. Thus, the main requirements for information in modern information systems must be its harmonisation; interoperability, in other words its capacity to operate across diverse systems, organisations or data formats; and its capacity to be shared easily.

By and large the same issues and requirements which can be seen to affect information in general, also affect geographic information, by which we mean all the information associated with a location related to the Earth. This kind of information is

essential to many aspects of geographic and environmental research, including general territorial knowledge, the prevention of natural catastrophes and the monitoring of climate change. If the data sources from different providers can be used, experts and scientists are able to create new geospatial products. Typically, two kinds of geospatial data are held – raster data and vector data. These can now be distributed and made available via geo-information systems (GIS) which capture, store, analyse, manage and present geodata.

Nowadays information technology in general and geo-informatics in particular can offer just-in-time access to the distributed, heterogeneous spatial data with the help of a Spatial Data Infrastructure (SDI). From the many definitions of an SDI the following was chosen:

“An information communication system providing users with access to shared spatial data resources as well as providing for the dissemination and exchange via the Internet, or some other publicly accessible global network, in an efficient and flexible way in order to enhance their research and productivity” [1].

Additionally, an SDI can be formed by interconnecting several systems that, in turn, could be seen as SDIs themselves [2]. An SDI consists of sources of spatial data, databases and metadata which can be searched, visualised and evaluated using data networks and data collection, management and representation technologies, which are consistent with institutional arrangements, policies and standards and requirements of end-users [3]. An SDI should enable the discovery and delivery of geospatial data from a data repository, ideally via one or more web-services. Additionally, it is often desirable to create and update spatial data: for this purpose GIS software is required. At present, SDIs generally are based on Service-Oriented Architecture (SOA), in which web-services and XML technologies each play a relevant part [2]. An SOA provides a universally distributable, platform-independent computing platform. Applications can thus be built easily from services running on heterogeneous platforms in multiple locations. This approach is consistent with modern international trends in information technology and greatly enhances the possibility to work on geographical data. More detailed information concerning SDIs can be found in the article by Steiniger and Hunter [4].

As new technologies become available, so too do new standards for encoding the geospatial data and infrastructures. The main organisations operating in this area are ISO, W3C, OGC (The Open Geospatial Consortium) and GSDI (The Global Spatial Data Infrastructure Association). In 2007 the

ISO adopted the Geography Markup Language (GML) as an International Standard (ISO 19136). GML is based on the XML grammar and can be used in the context of SOA and geospatial web-services, and therefore is relevant for the creation of an SDI [1], [5].

Within this overall framework, the specific object of our investigation is GML, which probably represents one of the key steps taken by the geospatial community towards the goal of data interoperability and harmonisation. This markup language, which is later described in detail, was specifically designed for the modelling, dissemination and storage of geographic information [5]. The following sections will introduce the concept of GML and the study of its use within two major initiatives designed to establish harmonised access for spatial information in Europe. The final part of the article is dedicated to an investigation of the use of GML in the information system of the Institute for Applied Remote Sensing at EURAC, Bolzano. Since GML is relatively new and there is not much experience available of its usage in systems of this type – despite the fact that the ideas of geographical data harmonisation and interoperability are spreading around the globe and their implementation appears urgent – this work has proved valuable. Additionally, the attempt to propose some ideas for a new catalogue for the geospatial data stored on the Institute's current system, where GML technology could be used, will be useful also for those who are working on similar problems, because – although GML has many advantages and potential uses – this language is not used globally yet.

II. A BRIEF DESCRIPTION OF THE GEOGRAPHY MARKUP LANGUAGE

GML is a markup language that is used to encode information about real-world objects. In GML, these real-world objects are called features, and these can have geometrical and non-geometrical properties. The elements of GML are used to describe features and also the relationships between them. Features and properties constitute the raw data, which GML stores in something called an 'instance document'. Apart from instance documents, a second type of GML encoding exists, which is called 'GML schema'. GML schemas are 'vocabularies' that define the structure of GML data. Since it is impossible to describe all features and predict their usage and properties *a priori*, two kinds of GML schemas exist – 'GML core schemas' and 'GML application schemas'. GML core schemas contain all the basic elements for describing the geographical information and provide the framework that can be used to construct GML application schemas. GML application schemas, and their corresponding data instances, are always domain-specific and should be created and managed by professionals. The ability to create GML application schemas gives the user the opportunity to describe and model unlimited amounts of diverse geospatial data.

Software can read GML application schemas and determine which elements are features, and then use this information when encountering data instances that conform to those

application schemas, using 'abstract types'. Software uses these abstract types to determine the nature of elements in the data stream [5], [6]. GML data elements can have metadata properties, which can be optionally attached to the GML element types. The metadata information itself contains its own application schema (or several GML schema documents imported into that application schema), which is similar in structure to a GML application schema for features [7]. In fact, XML schema can be thought of as metadata for an XML document, and in order to build applicable XML schema for an object's metadata, several GML application schema documents could be imported into the object's metadata application schema [8].

GML can be used to create both metadata for automation (i.e. machine-readable, in order for example to be used by web-services) and, equally, human readable metadata which can be used directly by an end-user [8]. Of prime importance is the fact that if the contextual and background material of the metadata is provided, the process of discovery of information resources can be automated. As such, GML encoded metadata is a key enabler of data and system interoperability [9].

GML is only concerned with the representation of geographic data content and does not specify how GML data should be presented. To represent the geospatial features in the form of maps, GML data have to be transformed through XSLT (Extensible Style sheet Language Transformations) in order to 'style' the GML geographical contents into one of the common graphical formats (e.g. SVG, VML, X3D), which specify how each element in the GML data should be displayed [5], [10].

Web-services can be briefly defined as software components, which provide functionality to a web user. However, the term web-services has a wide variety of definitions and web-services are not necessarily restricted to particular protocols (such as SOAP) or payloads. Although most web-services involve the transmission of messages encoded in XML, this is not a necessary condition of web-services. Web-services which process 'request messages' (sent to the service) and 'response messages' (sent from the service) that contain the geographic information are called Geospatial Web-Services. GML can be used to describe these web-services as well as to describe the geospatial parts of their messages. With GML, geospatial web-services and supporting technologies are used to distribute GML-encoded information over the Geo-Web (which includes also SDIs) [5].

The OGC has proposed a number of different geospatial web-services standard specifications, including OGC WFS (Web Feature Service), WCS (Web Coverage Service), WMS (Web Map Service) and CSW (Catalogue Service for Web). GML and other OGC standards will be used to support the geospatial aspects of these web-services [11].

Web Feature Service

GML was defined by the OGC in order to represent the data retrieved by the WFS which provides access to geographic feature data and may be used to access repositories of spatial objects (features with geometry in OGC terminology) on the

web. WFS returns original geographic data semantics (e.g. GML data, and not a visualised map). Thus, WFS can be used as client-server data retrieval, server-to-server data synchronisation or as a common access interface on top of heterogeneous geospatial data stores [5], [12].

Web Coverage Service

Also, GML can work with WCS, which supports the electronic retrieval of geographic data as discrete coverages. Coverage is a collection of mappings from geometry (e.g. grid points) into parameter (e.g. temperature or pressure) values, which can be represented as grid data or other geometries (curves, triangles, etc.) [13], [14].

Web Map Service

The GML data, retrieved from WFSs and WCSs can be visualised in the form of static maps (images in formats such as JPEG, GIF, PNG, SVG, etc.) by assigning visualisation styles (colours, line widths, etc.), and WMS can be used to work with this kind of data. WMS server returns to the HTTP query pre-rendered images (layers), projections, or other parameters from different sources, to be layered on top of each other in the client application [5], [12], [24].

Catalogue Service for Web

A Catalogue Service (CSW) is an OGC web-service that supports the storage, retrieval and management of metadata, including GML application schemas, which can be stored in schema registries. The schema registry provides a mechanism for managing application schemas and is the source for all GML application schemas that are supported by various geospatial web-services [5].

Thus, GML can be used to encode geospatial information and its metadata, to describe geospatial parts of web-services and to discover, transfer and retrieve geospatial information using the web-services technology. It is an XML-based standard, compatible with the main geospatial web-services, it is easy to transform and transport. The use of GML facilitates the main requirements of modern SDIs – geospatial data sharing, harmonisation and interoperability.

To explain the feasibility of the general use of GML, a SWOT analysis was performed, which is a method to evaluate the Strengths, Weaknesses, Opportunities, and Threats of a particular subject [15].

The most significant outcomes of the analysis can be observed below.

Strengths:

- XML-based nature;
- Feature-property logic;
- Separation of GML data and schema encodings;
- GML is supported by various OGC W*S standards.

Weaknesses:

- Unsuitability for encoding raster data;
- Difficulties with the visualization;
- Problems with storage of the data;
- Verboseness and redundancy (different users may use different names to represent the same feature and need to

know the schema created by other users in order to achieve real data interoperability).

Opportunities:

- Creation of GML-Native database;
- Geo-enabled web (the usage of the GML in interconnected SDIs).

Threats:

- Other languages exist to model geoinformation (e.g. G-XML, SOSI – Systematic Organization of Spatial Information).

The OGC, ISO, GSDI and other organisations are currently working on the complete documentation for GML and all its accompanying technologies in order to reduce verboseness and redundancy problems. Geoinformation, which is encoded in GML, requires more space and is often normalised and mapped into relational databases. Since these processes create difficulties, experts are working to design GML-Native databases to solve the current problems with the storage of GML-encoded geodata. As an XML-based standard, the GML is used in web-technologies, and is supported by the majority of modern GIS. Although other standards exist to encode feature-based geographical information, GML is the preferred standard in the geospatial community.

The results of the analysis confirmed that GML is sufficient for expressing, storing and exchanging feature-based geoinformation, although the fact that GML standard is relatively new, is not used globally yet, and developers do not have much experience with it, can all be considered impediments for its wide adoption.

III. THE STUDY OF THE USE OF GML FOR ONGOING SDI INITIATIVES

In this part GML is examined, not as a standalone technology, but within the context of real examples of modern SDIs. Two main European attempts to establish harmonised access for spatial information were studied: the INSPIRE and GMES-HMA. These two initiatives, which clearly emphasise the need for portal frameworks to provide simple and transparent access to geospatial data and metadata, aim to establish an SDI throughout the EU.

A. INSPIRE and HMA

The development of the Infrastructure for Spatial Information in the European Community (INSPIRE) was started in May 2007 and its full implementation is mandatory by 2019. The INSPIRE project also involved the publication of an EU Directive(2007/2/EC) aimed at creating an SDI throughout the EU, which must be implemented in every European Union Member State; it addresses the fragmentation of datasets and sources, gaps in availability, lack of harmonisation between datasets at different geographic scales and duplication of information collection. INSPIRE will enable the sharing of environmental spatial information among public sector organisations and facilitate public access to spatial information across Europe [16].

In parallel with the INSPIRE project, the GMES (Global Monitoring for Environment and Security) program, initiated by the EC (European Commission) and the ESA (European Space Agency), is also being developed in Europe. GMES is based on observation data received from satellites as well as information from ground based stations. While the INSPIRE project is more oriented to the data harmonisation and global standardisation with the final goal to create a European SDI, the GMES main objective is to rationalise the use of multiple data sources in order to provide the necessary information in a timely manner, and to provide autonomous and independent access to information in relation to environment and security [17], [18]. Both initiatives are complementary, and while INSPIRE provides the sharing of existing spatial data, GMES is an investment in additional European content, which also integrates existing data and ensures the continuous availability of observation data and services, creating new geospatial datasets [18].

In September 2005 the ESA initiated the Heterogeneous Missions Accessibility project (HMA) in order to harmonise ground segment services required to access the Earth observation (EO) data offering access to ground segments for many observation projects or missions [19]. The HMA is an ambitious attempt to provide a portal entry-point to multiple missions in the context of the GMES [17]. GMES-HMA operates on a partially connected net of systems and services, which allows the possibility of further extension, not only to additional HMA mission ground segments, but also to “third-party” missions outside GMES [20], [21].

The prototype of the INSPIRE geoportal can be found at <http://www.inspire-geoportal.eu/index.cfm>, and the HMA geoportal, which is still in the development stage, can be accessed through <http://hma.eoportal.org>.

The architecture of both INSPIRE and HMA is based on SOA, which is the main component of a modern SDI, and in which the definition of ‘Web-services’ is a requirement (it should be noted that in SOA, web-services and XML-related technologies are both GML-compatible). Web-services operate using a layered architecture, which separates the presentation layer, the service (or business logic) layer and the data source layer.

Although HMA, compared to INSPIRE, has a more complicated architecture, both projects’ main components can be outlined as follows [16], [17]:

- A portal or an application – the presentation layer;
- Network services and service execution environment – the business logic layer;
- Registers and user profile repository – a data source layer;
- Database for service metadata and for dataset metadata – a data source layer;
- Spatial data storage – a data source layer.

The Network Services are the core of both INSPIRE and HMA architecture. Although the names of some of INSPIRE and HMA web-services are different, they have the same purpose: to provide the possibility to search for geospatial data, to view data and to access data online or via data-

download. This describes the main services in any modern SDI.

View Service – to search for data (request service capabilities on the service itself and the data it serves) and request for layers to be returned as a map, and display, navigate, zoom, pan and overlay spatial datasets, as well as to display any relevant content of metadata; WMS can be used to view maps [17], [22], [23], [24].

Discovery Service – to search or browse spatial datasets and services based on the metadata content and to display the metadata content; the CSW can be used as a standard basis [25].

Download (Data Access) Service – to download or make available via an interface spatial datasets; the most convenient way is to create a direct link for downloading; the WFS can be used together with GML to download geographical feature-based objects [17], [26].

Ordering Service – to order discovered spatial datasets if they are not available directly from the Download Service. For example, at the HMA geoportal, after an order is processed, the link to download a product is received by the user via e-mail [17].

The possibility to have a Registry Service should also be taken into account because it allows one to ascertain information about all the spatial data and to keep track of any change in the system, and the possibility of having a Transformation Service which enables spatial datasets to be transformed and thus achieve better interoperability. Both Registry and Transformation Services could be internal services of the system, included in the work of other services [22], [23], [27].

Both INSPIRE and GMES-HMA have an SOA, which is based on international standards as far as possible and helps to achieve the interoperability, harmonisation and cross-border accessibility of geospatial datasets and services. The following part is dedicated to the description of these datasets and the role of GML in their encoding.

B. The Role of GML in INSPIRE and HMA

INSPIRE and HMA services aim to work with geospatial information: geospatial data or metadata.

In accordance with the INSPIRE directive, three different types or levels of metadata are distinguished: metadata ‘for discovery’, metadata ‘for evaluation’, and metadata ‘for use’. Due to its extensibility and flexibility, GML is a recommended encoding for metadata ‘for use’ (as this kind of metadata can be quite rich and different from the metadata for discovery or evaluation which, within INSPIRE, are less rich and more common). For other metadata encoding the ISO/TS 19139 (and information models of ISO 19115/19119) and Dublin Core (ISO 15836) standards should be used. It should be noted that according to the INSPIRE harmonisation requirements the creation of the metadata schemas is one of the highest priorities [28].

An important principle is that the scope of INSPIRE is the handling of spatial data, i.e., most non-spatial aspects of the themes are outside the scope of INSPIRE data specifications.

INSPIRE has a broad thematic scope consisting of 34 spatial data themes. Spatial data can be accessible via Download and View services (where WMS and WFS technologies are used). The default encoding for geographic information is GML – it covers encoding rules for large parts of the INSPIRE application schemas and is compatible with the technical architecture of INSPIRE. Thanks to the use of GML in INSPIRE, the user is able to work with the data using digital maps and coverage representations. However, it is expected that further additional standards, other than GML, for the encoding of geospatial data will become recommended in the future [23], [29].

Since HMA was introduced to provide access to the multi-mission EO data from ground segments and to harmonise ground segment services, HMA is more oriented to the integration of existing data and providing for the continuous availability of observation data and services, which are more heterogeneous than in INSPIRE. The data itself can be stored in different formats, depending on the data provider [18]. The main task of HMA is to provide access to the data, to provide searching and ordering procedures and to offer the ability to make a request for future acquisitions across multiple missions. To achieve this task the use of metadata is essential, and HMA defines the “GML Application schema for EO products” as the main schema for product metadata description [17]. However, for the collection metadata (the metadata for ‘Discovery’ using the INSPIRE terminology) the ISO 19115 is used, which is GML-compatible. Since the technical architecture of HMA is service-oriented, it uses XML-based technology which is also GML-compliant. In short, the use of GML for metadata encoding serves one of the main objectives of HMA – data searching and data access.

GML is well integrated with the current candidate standards of the INSPIRE and HMA network services: WMS, WFS, WCS, CSW. However, to use GML in WMS, which is widely used in INSPIRE, GML data must be transformed using XSLT into one of the geographic data visualisation formats. Also, GML can be used on every level of the systems’ architecture: on the presentation layer for the encoding of metadata and in WMS-styled maps, on the network service layer for the description of web-services and their information encoding, and on the data source layer for the encoding of metadata and feature-based data. Summarising this chapter we can say that GML serves the main objectives of INSPIRE and HMA. INSPIRE is the first attempt to use GML globally, and this is bound to make the standard more widely used.

IV. THE APPLICATION OF GML TECHNOLOGY FOR GEOSPATIAL DATA AT EURAC

There are two main groups of geospatial data at EURAC’s Institute for Applied Remote Sensing: archived data and collected data. These two groups hold many different kinds of geo-information, e.g. “raw data” and value-added products of different levels of processing. In this chapter we will analyse how to apply GML technology to both these kinds of geodata, as well as make some suggestions about the creation of a catalogue for the Institute’s geospatial data.

A. Geospatial data and their management at EURAC

The archived data held at the Institute for Applied Remote Sensing is comprised of the satellite images, received from the Moderate Resolution Imaging Spectroradiometer (MODIS), captured by antenna, and then transferred to the Compact Station (CS) at EURAC. An archive which is part of CS stores all received MODIS data. Data files are stored in HDF format, a kind of container for large and complex datasets, providing very fast data access. Although these raster satellite images are not designed to be converted into GML, their metadata can be extracted in a text format and therefore also in GML. In fact, every HDF file in the CS has a corresponding .tar (derived from ‘tape archive’) file, which contains the quick-look information (JPGfile) and metadata (XML file) related to the original data file [31], [32]. This .tar file is processed by a special Metadata Manager which sends the quick-look information and metadata to the database of a catalogue (called the Sinbad Catalogue), where the metadata information is parsed from the XML file into the database tables. It is known that the structure of metadata information in XML files complies with the “OGC 06-080 GML 3.1.1 Application schema for EO products”, which is similar to the GML metadata application schema used in GMES-HMA. However, the fact that EURAC has a catalogue, which provides data searching, evaluating and downloading functionality (regarding the MODIS data), and that GML technology is used, does not make the system interoperable and easy-to-integrate with existing SDIs. This catalogue cannot be modified without the permission of a third party and does not provide the same functionality for the geospatial datasets, which are formatted differently from the MODIS satellite images [33]. This inconsistency is one of the main reasons to suggest the creation of a new catalogue for geospatial data, which would provide the functionality and usability that is currently missing in the Sinbad Catalogue.

In addition, the Institute possesses not only the MODIS data, but also other geospatial data collected over time: raster or vector, satellite based or ground based, original or processed, modelled and other data. To store them, a file system is used, which means that there is no database system for the collected data from which we can extract the meta information. It is also necessary to make the data searchable and easily accessible through a detailed catalogue. Therefore, it is important to understand how the data could be described in a catalogue, i.e. what metadata information should be provided in a catalogue in order to find and choose a particular product. Since the entire collection of data is distributed across the file system and there is no documentation on the data stored at the Institute, it was decided to collect this basic and essential information from every group leader.

The result was that it was found that the Institute for Applied Remote Sensing deals with three groups of data: Air Quality, Snow and Geology. All three groups have some common metadata characteristics (which could be used in the data search process), e.g. geographic location or temporal reference, and specific characteristics, depending on the category of the data source (like satellite images, ground based

data, GIS vector data), e.g. the incidence angle for the radar images or the location of a particulate matter station for ground based data. It turns out that GML is ideally suited to describe all three types of metadata, because GML metadata application schemas can be easily extended (e.g. for satellite information the “OGC 06-080 GML 3.1.1 Application schema for EO products” could be extended) and modified depending on the needs of the Institute.

Thus, the necessary GML metadata application schema should be created, in order to populate the catalogue with new metadata instances. The following chapter provides some ideas about the catalogue and its services, and gives some real examples of existing catalogue software.

B. Suggestions Regarding a New Catalogue for EURAC Geodata

Firstly, it is essential that the new catalogue software should be compatible with the definition of SDI and operate in the context of SOA. Additionally, a definition of web-services' standard is necessary: the main standard supported by the catalogue could be the CSW. Examples of server-based opensource catalogue software are:

GeoNetworkOpensource (<http://geonetwork-opensource.org>), MDweb (<http://www.mdweb-project.org>) and the Deegree framework (<http://wiki.deegree.org>).

The licenses used by free and open source products are recommended because they ensure a low acquisition cost (or no cost) for the software, and ensure that the software is customisable and adaptable in an SDI context. Hence, free and open source software is often compatible with a wide set of OGC standards, and at least permits the addition of components that are OGC/ISO standard compliant. Finally, free and open software products are often discussed in user and developer communities, where one can get the support and maintenance options offered by various individuals and companies [4].

The possibility to upload new products and create new metadata fields would be an important characteristic of the system. This can be done best by uploading new XML-schemas and then executing the transformation into XSLT. Usually server-based open source catalogues support the ISO 19115 standard to describe metadata and also provide the possibility to upload a new metadata schema for user-defined metadata formats. In such a way, the “GML Application schema for EO products”, which is used in the Institute for the MODIS satellite images' meta-information, can be used in a catalogue as well.

In order to locate the necessary information it is important to establish how the search function works. The most essential information needed to find geospatial data is geographical position and temporal reference – these two fields should be included in the metadata description of necessity. To search for other information, the advanced search mode could be used, as it is done in the INSPIRE geoportal. Searching by discipline and browsing catalogues and collections is another type of search for relevant information and this can be implemented in the system as an alternative mode of searching.

In the light of this analysis of EURAC's Institute for Applied Remote Sensing data, the definition of the necessary information for the creation of metadata application schema and suggestions concerning the catalogue for the Institute's geospatial data, it would appear that GML technology can address all the questions raised. After a deeper study of the proposed catalogue software, the analysis of the problem of getting the metadata instances from the Institute's data and the problem of uploading them (and the relevant GML metadata application schema) to the catalogue, it is concluded that the Institute's catalogue system could be used as a part of an SDI, because the use of GML allows the necessary metadata harmonisation, and provides for better interoperability and sharing of the Institute's geospatial data.

V. CONCLUSIONS

This work is intended to give a wide overview of the use of GML in modern SDIs, which are currently the main building blocks of Geo-Web. The paper gives a brief description of GML technology and then gives examples of how this technology is used in current initiatives to construct an SDI. GML technology is used in the web-services of INSPIRE and HMA, for the description of metadata (INSPIRE and HMA) and for feature-based data (INSPIRE). A relevant practical example of the use of GML is the encoding of MODIS satellite images' metadata for the Sinbad catalogue, as well as the use of GML for constructing a new metadata application schema for all geospatial information at EURAC's Institute for Applied Remote Sensing, which has yet to be catalogued. Finally, some suggestions concerning the use of free and open source catalogue software, in which GML technology can be used, are provided.

As a final remark, we would like to summarise the main characteristics of GML when used in a modern SDI, which were identified during this work:

- GML can be used for encoding geospatial data, if the data are feature-based. An example of this usage is the INSPIRE project.
- GML can be used to encode geospatial data for transfer, exchange or storage.
- GML is not suitable for encoding satellite images, aerial photos and other raster images, although it can still be used to describe them.
- GML can be used for metadata encoding, as it is currently used at the Institute, in the INSPIRE and HMA projects.
- GML plays a significant role in the work of geospatial web-services and their information encoding.

Since the GML standard is new, it is not used globally yet, but with time and the gaining of some experience in working with GML, it should become an important and widely adopted standard to be used in modern SDIs, where the principles of data sharing, harmonisation and interoperability are essential.

REFERENCES

- [1] A.V. Koshkarev, A. N. Antipov, A. R. Batuyev, V. V. Yermoshin, V. P. Karakin, “Geo-portals as Part of Spatial Data Infrastructures:

- Russian Academy-supported resources and geoservices", *Geography and Natural Resources*, vol. 29, pp. 18-27, 2008. Available: Sciencedirect, <http://www.sciencedirect.com>. [Accessed: Sept. 23, 2010].
- [2] **C. Granell, L. Díaz, M. Gould**, "Service-Oriented Applications for Environmental Models: Reusable Geospatial Services", *Environmental Modelling & Sources*, vol. 25, no. 2, pp. 82-198, Feb. 2010. Available: Sciencedirect, <http://www.sciencedirect.com>. [Accessed: Sept. 23, 2010].
- [3] **C. Kiehle**, "Business Logic for Geoprocessing of Distributed Geodata", *Computers & Geosciences*, Vol. 32, Issue 10, Elsevier Ltd, pp. 1746-1757, 2006. Available: Sciencedirect, <http://www.sciencedirect.com>. [Accessed: Sept. 23, 2010].
- [4] **S. Steiniger, A.J.S. Hunter**, Free and Open Source GIS Software for Building a Spatial Data Infrastructure. In E. Bocher and M. Neteler (eds): *Geospatial Free and Open Source Software in the 21st Century: Proceedings of the first Open Source Geospatial Research Symposium*, 2009, LNCS, Springer, Heidelberg, 2010.
- [5] **R. Lake, D. Burggraf, M. Trninic, L. Rae**, *Geography Markup Language*. Chichester: Wiley, 2004.
- [6] D.S. Burggraf, "Geography Markup Language", *Data Science Journal*, vol. 5, pp. 178-204, 2006.
- [7] **R. Lake**, "The Application of Geography Markup Language (GML) to the Geological Sciences", *Computers & Geosciences*, vol. 31, issue 9, pp. 1081-1094, Nov. 2005. Available: EBSCOhost, <http://web.ebscohost.com>. [Accessed: Sept. 23, 2010].
- [8] **A. KalantariOs Kou'ee, M.S.**, "Geo-Metadata Creation Using GML technology", *International Journal of Information Science and Technology*, vol. 5 no. 2, pp. 47-65, July/December 2007. [Online]. Available: Scientific Information Database, <http://www.sid.ir/en/index.asp>. [Accessed: Sept. 23, 2010].
- [9] "National System for Geospatial Intelligence (ngs) Geospatial Metadata Desk Side Reference (NGMDSR)", Version 1.0, Aug. 23, 2007. [Online]. Available: http://www.gwg.nga.mil/documents/FINAL_MDSRG-8-23-07.doc. [Accessed: Sept. 23, 2010].
- [10] **Y. Xi, J. Wu**, "Application of GML and SVG in the Development of WebGIS", *Journal of China University of Mining and Technology*, vol. 18, issue 1, pp. 140-143, 2008. Available: ScienceDirect, <http://www.sciencedirect.com>. [Accessed: May 20, 2010].
- [11] "OGC Standards and Specifications", Open Geospatial Consortium, Inc., 2010. [Online]. Available: <http://www.opengeospatial.org/standards>. [Accessed: Sept. 23, 2010].
- [12] **I. Rinne**, "Introducing OGC Web Services at FMI", 19th EGOWS 2008, 9-12 June, Ljubljana, Slovenia, ARSO, 2008. [Online]. Available: http://www.arso.gov.si/egows2008/FMI_introducing_ogc_web_services_Rinne.pdf. [Accessed: May 20, 2010].
- [13] **T. Hamre, H. Krasemann, S. Groom, D. Dunne, G. Breitbach, B. Hackett, K. Sørensen and S. Sandven**, "Interoperable web GIS services for marine pollution monitoring and forecasting", *Journal of Coastal Conservation*, vol. 13, no. 1, pp. 1-13, 2009. Available: SpringerLink, <http://www.springerlink.com>. [Accessed: May 20, 2010].
- [14] **N. R. Brisaboa, A. Fariña, M. R. Luaces**, "On the Compression of Geography Markup Language", *In Advances in Computer Science and Engineering*, vol. 19, pp. 93-104, 2006.
- [15] "SWOT analysis: a tool for making better business decisions", U.S. Dept. of Agriculture, Risk Management Agency, Issue 1973 of Program aid, 2008. [Online]. Available: http://farm-risk-plans.usda.gov/pdf/swot_brochure_web.pdf. [Accessed: Sept. 23, 2010].
- [16] "The INSPIRE Directive – A Brief Overview", Association for Geographic Information (AGI), Dec. 2009. [Online]. Available: http://www.agi.org.uk/storage/inspire/inspire_intro.pdf. [Accessed: Sept. 23, 2010].
- [17] **B. Lawrence, M. Pritchard, A. Woolf**, "Review of the Heterogeneous Mission Accessibility Project", BADC, ver. 1.0, 2007 [Online]. Available: <http://home.badc.rl.ac.uk/lawrence/static/2007/03/14/HMA-Review.pdf>. [Accessed: May, 20, 2010].
- [18] **M. Schilcher, A. Fichtinger, K. Jaenicke, V. Kraut, J. Stahl, F. Straub**, "INSPIRE Fundamentals, Examples, Test Results", *Runder Tisch Geoinformationssysteme e.V.*, Munich, pp. 23, 2009. [Online]. Available: http://www.cagi.cz/files/INSPIRE_Broschuere_V4_en_final_web_231109162434.pdf. [Accessed: Sept. 23, 2010].
- [19] HMA, ESA (videoclip). [Online]. Available: http://earth.esa.int/hma/video/HMA_080607.wmv. [Accessed: May, 20, 2010].
- [20] **Е.Б. Кудашев, А.Н. Филонов**, Развитие инфраструктуры распределенных хранилищ спутниковых данных: интегрированная распределенная среда неоднородных информационных ресурсов исследования Земли из Космоса. In *Труды 10-й Всероссийской научной конференции, Электронные библиотеки: перспективные методы и технологии, электронные коллекции*, 2008, RCDL, Dubna, Russia, pp. 299-308, 2008.
- [21] HMA Project organised by the European Space Agency, ESA. [Online]. Available: <http://earth.esa.int/rtd/Projects/HMA/>. [Accessed: May, 20, 2010].
- [22] "Directive 2007/2/EC of the European Parliament and of the Council of March 14, 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)", *Official Journal of the European Union*, 2007, [Online]. Available: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:EN:PDF> [Accessed: May, 20, 2010].
- [23] "GIGAS Technology Watch Report Architecture Technical Note", GIGAS Consortium, Issue 103, 2009. [Online]. Available: <http://www.thegigasforum.eu/cgi-bin/download.pl?f=354.pdf> [Accessed: Sept. 23, 2010].
- [24] Technical Guidance to implement INSPIRE View Services, INSPIRE Drafting Team "Network Services", ver. 2.0, 2009. [Online]. Available: <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/5>. [Accessed: May 20, 2010].
- [25] "Technical Guidance to implement INSPIRE Discovery Services", INSPIRE Drafting Team "Network Services", ver. 2.0, 2009. [Online]. Available: <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/5>. [Accessed: May 20, 2010].
- [26] "Guidelines for the encoding of spatial data", INSPIRE Drafting Team "Data Specifications", ver. 3.1, 2009. [Online]. Available: <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/2>. [Accessed: May 20, 2010].
- [27] "INSPIRE Technical Architecture – Overview", INSPIRE Drafting Teams "Data Specifications", "Network Services", "Metadata", ver. 1.2, 2007. [Online]. Available: http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/network/INSPIRETechnicalArchitectureOverview_v1.2.pdf. [Accessed: Sept. 23, 2010].
- [28] **P. Villa, T. Reitz, M. A. Gomasasca**, "HUMBOLDT Project for Data Harmonisation in the Framework of GMES and ESDI: Introduction and Early Achievements", In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, part B4, pp. 1741 – 1746, Beijing, 2008. Available: ISPRS, <http://www.isprs.org>. [Accessed Sept. 23, 2010].
- [29] **C. Maioli, E. Sánchez Jordán**, A comparative Italian-Spanish analysis of spatial information and right of access, *Proceedings ICTT 3rd International Conference on Information and Communication Technologies: from Theory to Applications*, April 7-11, 2008, Damascus, Syria. IEEE, 2008.
- [30] **S. Cox, P. Daisey, R. Lake, C. Portele, A. Whiteside**, "OpenGIS Geography Markup Language (GML) Implementation Specification", ver. 3.1.0 - OpenGIS Recommendation Paper, Open GIS Consortium, Inc. and ISO. Reference: OGC 03-105r1, pp. 601, 2004. [Online]. Available: http://www.e-cartouche.ch/content_reg/cartouche/webservice/en/literatur/gml_3_1.pdf [Accessed: Sept. 23, 2010].
- [31] "DAPS Ground Segment, ARF Reference Manual", Advanced Computer Systems, DAPS-MA-ACS-EURAC-0107, Issue: 1.0, pp. 99, 2009.
- [32] "DAPS Ground Segment, SDF Reference Manual", Advanced Computer Systems, DAPS-MA-ACS-EURAC-0111, Issue: 1.0, pp. 94, 2009.
- [33] "SINBAD for EURAC, USER Manual Document", Advanced Computer Systems, DAPS-MA-ACS-EURAC-0112, Issue: 1.0, pp. 35, 2009.

Viktors Skoks graduated from the Riga Technical University in 2010 as Master of Engineering in Information Technologies. During his studies he also had an internship in the Free University of Bozen-Bolzano, Italy and at the European Academy of Bolzano (EURAC), where he has produced his final thesis. His main fields of study are large DBMs and the geospatial data harmonisation.

Since 2008 V. Skoks has worked for Accenture Latvia as a Programmer. He specialises in developing IT solutions for Oracle products (TIA – the Insurance Application, Oracle CRM On-Demand) and has experience in application development and testing, as well as data migration.

Christian Steurer graduated from the Politecnico of Milan in 1997 as a telecommunication engineer and obtained his professional accreditation in 1997 after the Public Examination. He specialises in the development of IT and communication systems including radio communication and their respective infrastructures.

C. Steurer has worked as a teacher of physics, mathematics, informatics and

electrical installations and as head of department in a private regional telecommunication company. Since 2002 he has worked at the European Academy of Bolzano (EURAC) as the Technical Head of EURAC's Institute for Applied Remote Sensing and Project Manager of EURAC's service department for Information and Communication Technology, with major expertise on IT-security, networking and IT-infrastructure.

Viktors Skoks, Kristiāns Steurers. Pārskats par GML izmantošanu mūsdienu telpisko datu infrastruktūrās

Ģeogrāfiskā Iezīmēšanas Valoda (GML) ir balstīta uz XML gramatiku, un to var pielietot serviss-orientētas arhitektūras un ģeotelpisko tīmekļa pakalpojumu ietvaros. Raksts sniedz īsu aprakstu par GML tehnoloģiju un tās izmantošanu mūsdienu ģeotelpisko datu infrastruktūrās (SDI – Spatial Data Infrastructure). Darba mērķis – noteikt dažas GML izmantošanas pamatpriekšrocības svarīgākajos mūsdienu SDI – ģeotelpisko datu piekļuves, meklēšanas un harmonizācijas procesos. Bez tam, raksts sniedz piemērus par GML izmantošanu divos lielākajos mēģinājumos izveidot harmonizēto piekļuvi telpiskajai informācijai – INSPIRE (the European Infrastructure for Spatial Information in the European Community) un HMA (Heterogeneous Missions Accessibility). INSPIRE mērķis ir izveidot SDI visai Eiropas Savienībai, savukārt HMA mērķis ir nodrošināt portāla izveidošanu vairākiem projektiem vai misijām GMES (Global Monitoring for Environment and Security) programmas ietvaros. Ievērojama raksta daļa ir veltīta GML izmantošanas iespējai Zemes novērošanas datu apstrādes un izplatīšanās informācijas sistēmas izpētei. Šī sistēma pieder Lietišķās Attālās Uzrādes institūtam EURAC, Bolcāno. Šo sistēmu ir nepieciešams uzlabot, izveidojot jaunu institūta ģeotelpisko datu katalogu, ko ilgtermiņā varētu izmantot kā SDI sastāvdaļu.

Mūsu pētījumā tiek atzīmēts, ka GML izmanto INSPIRE un HMA tīmekļa pakalpojumos, kas veicina labāku sadarbību starp SDI un citām sistēmām. Bez tam, GML izmanto metadatu (INSPIRE un HMA) un telpisko objektu datu aprakstīšanai (INSPIRE). EURAC GML tiek izmantota, lai kodētu satelītattēlu metadatus. Jāatzīmē, ka institūta katalogs izmanto Lietišķo GML metadatu shēmu, kas ir līdzīga tai, kuru izmanto HMA. GML tehnoloģija var tikt izmantota, lai izveidotu jaunu Lietišķo metadatu shēmu visiem ģeotelpiskajiem datiem, kuri ir uzkrāti institūtā. Šī Lietišķā shēma var tikt padota katalogam, lai veiktu informācijas meklēšanu un nodrošinātu institūta ģeotelpiskās informācijas piekļuvi.

Nobeigumā, GML izmantošana katalogos veicina labāku ģeotelpisko datu meklēšanu un piekļuvi. GML izmantošana ģeodatu vai to metadatu kodēšanai veicina labāku ģeoinformācijas mijiedarbību. Savukārt, kā vienotais un savstarpēji saderīgais ar citiem standart, GML veicina ģeogrāfiskās informācijas integrāciju un harmonizāciju.

Виктор Скок, Кристиан Штойпер. Обзор по использованию GML в современных инфраструктурах пространственных данных

Географический язык разметки (GML) основан на XML-грамматике и может быть использован в рамках сервис-ориентированной архитектуры и геопространственных веб-сервисов. В статье дается краткое описание технологии GML и её использования в современных инфраструктурах геопространственных данных (SDI – SpatialDataInfrastructure). Цель работы – обозначить некоторые основные преимущества использования GML в наиболее важных процессах в современных SDI – процессах обеспечения доступа, поиска и гармонизации геопространственных данных. Кроме того, в статье приводятся примеры по использованию GML в двух основных попытках установления гармонизированного доступа к пространственной информации – в INSPIRE (the European Infrastructure for Spatial Information in the European Community) и HMA (Heterogeneous Missions Accessibility). INSPIRE направлена на создание SDI для всего Европейского Союза, в то время как HMA направлена на создание портала для многочисленных проектов или миссий в контексте программы GMES (GlobalMonitoringforEnvironmentandSecurity). Значительная часть статьи посвящена исследованию использования GML в информационной системе для обработки и распространения данных дистанционного зондирования Земли, принадлежащей Институту Прикладного Дистанционного Зондирования в EURAC, Больцано. Данную систему необходимо улучшить, создав для нее новый каталог геопространственных данных Института, который в перспективе может быть использован как часть SDI.

В нашем исследовании заявляется, что GML используется для веб-сервисов как в INSPIRE, так и в HMA, что способствует лучшему взаимодействию между этими SDI и другими системами. Кроме того, GML используется для описания метаданных (в INSPIRE и HMA) и данных пространственных объектов (в INSPIRE). В EURAC GML используется для кодирования метаданных спутниковых изображений. Следует отметить, что каталог Института использует прикладную GML схему метаданных, аналогичную той, которая используется в HMA. Технология GML может быть использована для построения новой прикладной схемы метаданных для всех геоданных, собранных в Институте. Данная прикладная схема может быть передана каталогам в целях осуществления поиска информации и обеспечения доступа к геопространственным данным Института.

В заключение, использование GML в каталогах способствует лучшему поиску геоданных и предоставлению доступа к ним. Использование GML для кодирования геоданных или их метаданных способствует лучшему взаимодействию геоинформации. Более того, как единый и взаимно совместимый с другими стандарт, GML способствует интеграции и гармонизации географической информации.