FOUNDATIONS OF COMPUTING AND DECISION SCIENCES
Vol. 40 (2015) No. 3
ISSN 0867-6356
e-ISSN 2300-3405
DE GRUYTER OPEN
DOI: 10.1515/fcds-2015-0012

THE APPLICABILITY OF SWE IN POLISH SPATIAL DATA INFRASTRUCTURES - THE EXAMPLE OF THE SENSORML LANGUAGE

Maciej ROSSA, Mariusz ROGULSKI[1]

Abstract. Mobile and stationary sensors currently used to measure various environmental parameters, functioning independently or as part of monitoring networks and measurement stations, provide vast amounts of data on the state and quality of the environment on the Earth. If the data is to be used effectively, they must be exchanged and shared among IT systems. Systems which offer services of searching, exchange, sharing, visualisation and analysis of dispersed and varied data resources on the widely understood environment are, for example, spatial data infrastructures.

The article presents an overview of IT technologies and standards which offer interoperability in spatial data infrastructures. It first defines interoperability and then describes the most important issues connected with spatial data infrastructures on the example of INSPIRE. An example standard which facilitates interoperability in INSPIRE is the SensorML language, a component of Sensor Web Enablement (SWE). Its practical application is proposed – for description of processes of air monitoring in a spatial data infrastructure that is an element of the Polish national environmental monitoring plan.

Keywords: Spatial Data Infrastructures, SensorML, INSPIRE

1. Introduction

At present, globally, there are hundreds of thousands mobile and stationary sensors, which measure various environmental parameters and function independently or are part of thousands of monitoring networks and measurement stations that monitor all elements of the environment. Additionally, over 50 environment satellites carry out their research missions. Various observations and measurements are also made all the time, both in the field and in laboratories. These activities generate an unimaginable amount of data on the

[1] Mariusz.Rogulski@is.pw.edu.pl; Maciej.Rossa@is.pw.edu.pl, Warsaw University of Technology, Faculty of Environmental Engineering, Department of Information Science and Environment Quality Research, Nowowiejska 20, 00-653 Warsaw

- technical, which includes three more specific levels (Fig. 1):
  - systemic – tools, operation systems, transmission protocols, service security, data services;
  - syntactic – formats and standards of data exchange on the abstract level (languages of data exchange which only have syntactic standardisation);
  - schematic – schemata of data exchange and presentation (domain languages of data exchange);
- semantic – correct, unambiguous understanding of the exchanged and shared information by all its users-agents (both IT systems and people).

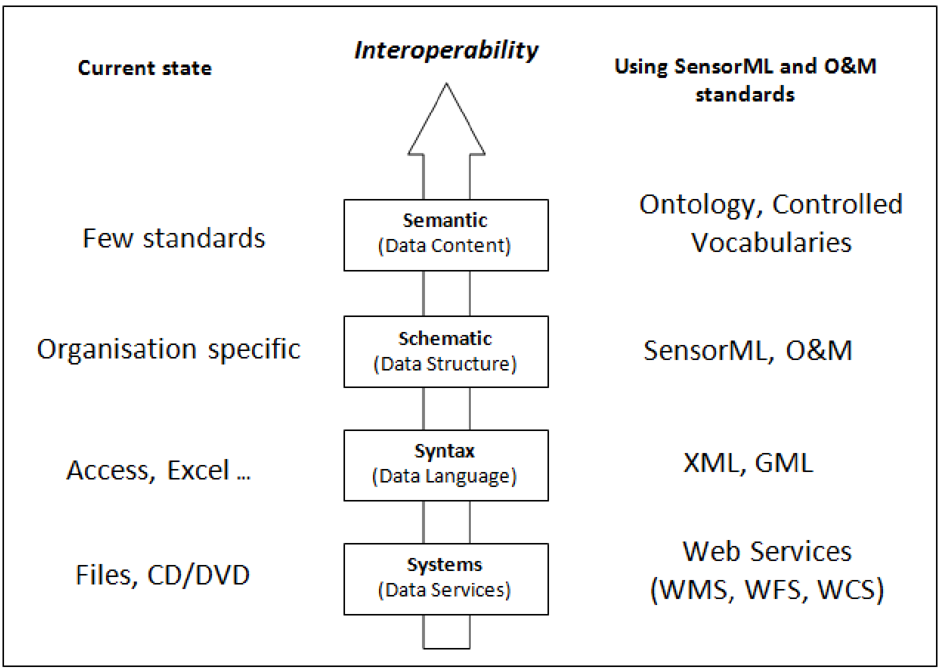

Figure 1. Levels of technical interoperability (according to GeoSciML – modified and complemented)

Activities aimed at achieving interoperability at the formal and organisational and technical levels have for a long time been effectively solved, implemented and realised. Great advances in this area are also visible in the fields connected with environmental studies, i.e. geology, hydrogeology, oceanography and meteorology [24, 29, 30, 31]. However, semantic interoperability is still being developed. Despite constant development of the capabilities of network resources to communicate, especially the development of web services, the issue is still unresolved and needs further research and analyses [21]. An example of the solutions suggested so far may be the use of methods based on ontologies and inference in the project meanInGs [23].

3. Spatial data infrastructures

A way of supporting interoperability is processing data in unambiguously defined, strict schemata published in specialised web services which have individual communication protocols. This concept for spatial data is used in spatial data infrastructures (SDI), which are built on the basis of SOA (Service Oriented Architecture) using web service technologies [26]. To some extent, using SDI allows to automate the use of processed metadata and spatial data [17, 18, 19]. An example of practical realisation of this model is INSPIRE - Infrastructure for Spatial Information in the European Community, which comprises SDIs of EU member countries, also the Polish national spatial data infrastructure (IIP). An example of SDI implementation using SOA is the EU system BRISEIDE (BRIdging SErvices, Information and Data for Europe) [27].

Essential for interoperability in SDI, also in INSPIRE and IIP, are geostandards and norms provided by two standardisation organisations: OGC (Open Geospatial Consortium, an organisation that creates and implements open standards for spatial data and services, geographical data systems, for processing data) and ISO (International Organisation for Standardisation). They form the basis for construction and functioning of SDI, supporting full technical interoperability, both in terms of service communication and data exchange [28]. A growing number of data and services entails the necessity to use more and more advanced IT tools and technologies. The possibilities and perspectives of development in this field using i.a. cloud services, as well as grid technology are described in [12] on the example of a project of the European Commission ”enviroGRIDS”. In [11] there is a suggestion of enhancing standard cataloguing services in SDI with semantic elements, while in [9] – a semantic development of cataloguing services in SDI for a region in Italy.

4. Interoperability in SDI on the example of INSPIRE

4.1. Formal and organisational interoperability

Interoperability at the formal and organisational level in INSPIRE is realised through a wide range of formal and legal solutions (procedural and organisational ones), the aim of which is to ensure cooperation of all those involved in creating and processing spatial data included in INSPIRE. These issues are regulated by appropriate legislation (the INSPIRE Directive implemented through the Polish SDI act) and executive regulations (EU regulations – which are binding and do not need to be implemented in national legislation), as well as technical implementation specifications (http://inspire.ec.europe.eu).

4.2. Technical interoperability

Technical interoperability in spatial data infrastructures, both in terms of services and data, is effected through implementation of OGC geostandards and ISO 19100 norms; in the case of INSPIRE also through dedicated technical guidelines.

The foundation of technical syntactic implementation is the GML language (Geography Markup Language), which is an application of the XML language. It is an XML grammar for expressing geographical features and their relations. In simple terms, GML describes only their geometry and topology but not their meaning. It contains generic geographical data that contain points, curves, polygons etc., but not for example, stream gauges, transport routes or borders of cadastre plots [24].

For the purposes of data exchange within a field of knowledge and for support of interoperability at the technical schematic level, on the basis of the GML language, field dedicated languages of data exchange (not formats) are developed and they are its profiles (simplification) or applications (extension) [25]. So far, over 20 different field data exchange languages have been developed, which are application schema or GML profiles (http://www.ogcnetwork.net/gmlprofiles), and work is still being done on new ones.

The above standards are the basis for creation of other languages of data exchange dedicated to various fields. At present, there are over twenty languages of this type in the fields of: Earth sciences (geology, tectonics, hydrogeology, hydrology, soil science), spatial planning, geohazards, agriculture, meteorology and climatology, hydrography and oceanography, and biology. Moreover, there is a group of application schema and GML profiles predefined for 34 subjects of INSPIRE spatial data. In some cases (geology, soil science, mineral resources, hydrology), INSPIRE data exchange languages are equivalent to OGC standards. Examples of this type of language include GeoSciML (GeoScience Markup Language) – a more advanced language of this type, GWML (Ground Water Markup Language), AgriXchange (a GML application scheme for description of spatial data on INSPIRE subjects connected with agriculture) [29].

It is worth noting that all the above mentioned data exchange languages share general and field classes of objects, e.g. some classes of objects of the GeoSciML language, describing geological age, mineral composition of rocks or geological structures, inherited to the ResourceML language predefined to the INSPIRE subject “natural resources” [30, 31].

Moreover, each language describing spatial data on the state and quality of the environment, has inheritable classes of objects from two base standards: SensorML and Observations and Measurements (O&M) – data exchange schemata on measurements and observations, their results, procedures and methodologies[2].

In INSPIRE[3], only a few subjects of spatial data include data on the state and quality of the environment which are derived directly from measurements and observations. These subjects include:

- hydrography;
- geology (with hydrogeology);
- soil – soils and subsoil characterised according to depth, texture, structure and content of particles and organic material, stoniness, erosion, where appropriate mean slope and anticipated water storage capacity;
- production and industrial facilities - industrial production sites, including installations covered by Council Directive 96/61/EC of 24 September 1996 concerning integrated pollution prevention and control (1) and water abstraction facilities, mining, storage sites.

[2] http://www.opengeospatial.org/standards/om

[3] http://inspire.ec.europa.eu

SensorML is an approved OGC standard. The latest version (SensorML 2.0) was approved on 2013-12-10.

A basic element of the language is a measurement process for which it is possible to determine input, output, parameters and additional information characterising the process. The user may create many types of processes concerning any environmental component, which as their basis require attributes defined in the base process. Processes may be both physical processes connected with measurements and observations, and processes other than physical ones (e.g. software processes connected with processing measured quantities or modelling).

The most important abstract types of objects offered by SensorML 2.0, which at the same time allow to use a dialect, include:

- ObservableProperty – represents physical parameters of the observed and/or measured phenomenon. Parameters include, for instance, temperature, gravitational force, location, chemical concentration etc. Objects inheriting from this type may also describe quantities connected with phenomena which have been determined in other ways than direct measurement (e.g. as a result of modelling).
- DescribedObject – an abstract class processing metadata for other classes of processes. These include many descriptive properties on general process information (e.g. keywords, classifications), limitations (e.g. timeframe, security limitations, legal limitations), classifications (properties and possibilities), references (contacts and documentation) and history. They are grouped in lists for simpler analysis.
- AbstractProcess – a basic abstract class inheriting from DescribedObject, which additionally offers properties connected with input, output and process parameters, process use, and the possibility of further (descriptive, for instance) development of derived processes.

On the basis of abstract classes, SensorML also has additional classes:

- Simple Process – for indivisible processes, i.e. those the realisation of which is treated as a whole. The class also has properties which enable provision of the methodology used in the process.
- Aggregate Process – for complex, multiple-component processes, with the possibility of mapping flow of data between the components,
- Physical Component – represents real facilities for which it is essential to define spatial and time coordinates,
- Physical System – used to model physical facilities as processes for which location in the real world is known and significant,
- Processes with Advanced Data Types – a class offering support for more advanced types of data than those offered by the abstract class AbstractProcess (e.g. DataArray, Matrix, DataStream and Choice).

SensorML 2.0 is strongly related to GML3.2, because all elements are substitutable for GML AbstractValue and AbstractFeature. For example, DescribedObject is derived from GML AbstractFeature.

Complementation and support of the SensorML language is offered in two other standards – Sensor Observation Service (SOS) and Sensor Planning Service (SPS). Both have been defined as interfaces so they may be used to create web services, for instance.

define criteria of data and when they become available (e.g. as a result of making a measurement), they are instantly sent to the user [8].

In turn, [2] proposes using BPEL and process chains from SensorML, a method of creating workflows for the so-called e-science – for these areas of science that require calculations in heavily dispersed network environments or that use vast amounts of data processed in grid environments.

SensorML is not the only language for standardisation of processes connected with functioning of sensors. An overview of currently developed norms and standards may be found in [32]. The authors present there, for example, standards created by ISO (ISO/IEC 18000 norm), IEEE (IEEE 1451) and the mentioned earlier standards created by OGC which are part of SWE. They also point to cooperation possibilities between these units and perspectives of development of standards. [4] presents SSN ontology created by Semantic Sensor Network Incubator Group, which describes sensors in terms of their possibilities, measurement processes and conducted observations. The ontology includes significant parts of SensorML and O&M languages.

In Poland, the described standards have not been widely used yet. They could be used when observations are made with different devices and methods and the process of collecting and processing a large number of measurements is quite developed. A good example application would be for description of processes of air monitoring. The reason is a big number of varied measuring devices and measured indicators of air quality, as well as quite a developed process of verification, analysis and processing of the data.

Air monitoring is a process conducted by voivodship inspectorates for environmental protection as part of the national environmental monitoring plan (NEMP). The obligation and its extent is defined in the act Environmental Protection Law (art. 25-29) [1]. In accordance with the act, the aim of NEMP is support for environmental actions through systematic information of administration units and the society on, e.g., the quality of natural elements, changes in the quality of natural elements and the reasons for them. The monitoring obligation for the Inspection of Environmental Protection is laid down in the Act on National Environmental Inspection.

At present, results of measurements of air pollution concentrations from automatic stations (in some cases, along with accompanying meteorological parameters) are automatically transmitted via a telecommunications network to databases maintained by voivodship inspectorates for environmental protection. These systems also receive measurement results made with non-automatic methods. Verified data is used to: create and update reports from special zones, where there is a risk of exceeding warning levels of air pollution concentrations, to develop short-term forecasts of air quality in special zones, to create reports (e.g. on the state of the environment), etc. The data is also available to the society.

Currently, all 16 voivodship inspectorates present results from stations of automatic air monitoring in the voivodship (on-line, on their websites). In all cases, the data is presented in the form of a static table on the website, sometimes supported with a presentation of the measurement results on a graph (graphs). This is usually effected in two forms – it is either possible to generate a graph for all parameters of a station or for a chosen parameter for all stations (e.g. in the Szczecin inspectorate), or it is possible to choose from a list of parameters on one's own and compare them for different measurement stations. It is impossible to generate graphs in the case of Warmińsko-mazurskie, Lubuskie and Silesian

inspectorates), other inspectorates present the map as jpg graphics where one can click on a chosen station. The map is not interactive and does not allow to change the scale, for example.

No inspectorate offers the possibility to download online data on an ongoing basis or any services connected with interoperability. The Lower Silesia inspectorate allows to download an Excel file with 24-hour mean values from automatic measurements conducted by industrial plants and manual measurements (updated monthly). The Masovia inspectorate, in turn, allows to download results of measurements from automatic stations only for 2012. Other units do not offer even such options.

The Lublin inspectorate, which started creation of a geoportal, is an important example. The portal presents a few layers on air – for example locations of measurement stations, area classification, areas for air quality assessment and data on emissions.

Information on measurement stations (e.g. station address, measured pollution, meteorological parameters) is published on websites by 14 out of 16 inspectorates. The information cannot be found only on Małopolska and Silesian inspectorates' portals. Other units present varied data on measurement stations. Some also publish photos of station surroundings. All information is presented on websites and it is not possible to download it in any format (apart from saving the webpage as HTML). Choice of location may be done from a map or from a list. Available solutions include Google maps (e.g. the Masovia inspectorate) or jpg pictures with added station locations (e.g. the Podlasie inspectorate).

Unfortunately, like in the case of monitoring data, it is not possible to download data for further processing (apart from copying it from the website in HTML). Additionally, practically each inspectorate describes measurement stations using a different set of parameters. It is difficult to see any kind of interoperability here.

This situation needs to be changed in the future. The INSPIRE Directive obliges Poland to implement various themes, among which there is Annex III.7 – Environmental Monitoring Facilities. Its description can be found in "D2.8.III. 7 Specification Data on Environmental Monitoring Facilities – Technical Guidelines" [6].

The above-mentioned document defines mandatory and optional elements of the target model of spatial data and metadata in the following areas:

- database objects including spatial objects,
- descriptive information describing the database objects, including the required IDs, etc.,
- dictionaries, code lists, etc.,
- metadata that extend the INSPIRE general profile with unique elements for environmental monitoring.

In terms of observations and measurements, the specification of environmental monitoring facilities recommends using "Guidelines for the use of Observations & Measurements and Sensor Web Enablement-related standards in INSPIRE Annex II and III data specification development" [7]. According to this document, Observations & Measurements and Sensor Web Enablement-related standards shall be used in INSPIRE to cover requirements in the following thematic domains:

- geology,
- oceanographic geographical features,
- atmospheric conditions and meteorological geographical features,
- environmental monitoring facilities,

- possibility of modelling a simple sensor or complex sensor systems,
- supporting geolocation information about sensors,
- providing accuracy information of sensors measurements,
- possibility of using the standard for various monitoring programs and their integration.

7. Summary

In the times of collecting and processing large amounts of information, users of data are more and more aware and demanding. They want to have access to information, be able to download, process and combine it. This is particularly true for environmental information, which in most cases is spatial data. This makes interoperability a vital aspect which should be considered when designing IT systems for spatial data infrastructures.

It is especially important now when the Chief Inspectorate of Environmental Protection is implementing the act on spatial data infrastructure in the field of "Environmental Monitoring Facilities" through standardisation and harmonisation of its resources (data and metadata) and creating a geoportal which would provide INSPIRE services: searching (CSW – Catalog Service for Web), viewing (WMS – Web Map Service) and downloading (WFS – Web Feature Service, WCS – Web Coverage Service). The use of the technologies (standards) presented in the article could be a basis for development of standards of harmonisation of NEMP and also greatly contribute to automation of data generation processes for environmental monitoring.

The article has presented an overview of IT technologies and standards that support interoperability in spatial data infrastructures. An example standard for this purpose is a component of Sensor Web Enablement – the SensorML specification. Its application for description of processes of air monitoring has been suggested. Implementation of the standard of description of measuring devices and the measurement process (from measuring the parameter to its verification and processing) would greatly facilitate combining and analysing the collected data, which is at present virtually impossible for the user.

References

[1] Act of 27 April 2001 Environmental Protection Law.
[2] Chen N., Hu C., Chen Y., Wang C., Gong J., Using SensorML to construct a geoprocessing e-Science workflow model under a sensor web environment, *Computers & Geosciences*, **47**, 2012, 119–129.
[3] Chen N., Wang X., Yang X., A direct registry service method for sensors and algorithms based on the process model, *Computers & Geosciences*, **56**, 2013, 45–55.

[18] Iwaniak A., Kubik T., Paluszyński W., Tykierko M., Kaczmarek I., *Semantic based extension of search capability of SDI*, Global Geospatial Conference 2012 “Spatially Enabling Government, Industry and Citizens”, Québec City, Canada, May 14-17.
[19] Iwaniak A., Łukowicz J., Strzelecki M., Kaczmarek I., Implementation aspects of an intelligent spatial information infrastructure, *Annals of Geomatics*, **X**, 4, 2012, 103-117.
[20] KRI, 2012 *Rozporządzenie Rady Ministrów z dnia 16 maja 2012 r. w sprawie Krajowych Ram Interoperacyjności, minimalnych wymagań dla rejestrów publicznych i wymiany informacji w postaci elektronicznej oraz minimalnych wymagań dla systemów teleinformatycznych* (Dz. U. z 2012 r. poz. 526).
[21] Kaczmarek I., Iwaniak A., The role of thesaurus in creating semantic interoperability, *Annals of Geomatics*, **IX**, 4, 2011, 61-69.
[22] Liang S.H.L., Croitoru A., Vincent Tao C., A distributed geospatial infrastructure for Sensor Web, *Computers & Geosciences*, **31**, 2005, 221–231.
[23] Lutz M., Sprado J., Klien E., Schubert C., Christ I., Overcoming semantic heterogeneity in spatial data infrastructures, *Computers & Geosciences*, **35**, 2009, 739–752.
[24] Michalak J., Languages derived from and connected with GML, *Annals of Geomatics*, **VI**, 6, 2008, 75-84.
[25] Michalak J., UML geospatial data models and their transformation into GML schemas and database structures, *Annals of Geomatics*, **X**, 1, 2012, 15-34.
[26] Nałęcz T., The use of spatial data models and their transformation (UML, XML, GML) in geology and hydrogeology, *Annals of Geomatics*, **IX**, 4, 2011, 105-115.
[27] Prandi F., De Amicis R., Conti G., Debiasi A., *Use of OGC Web Standard for a Spatio-Temporal Enabled SDI for Civil Protection*, Proceedings of the 17th International Conference on 3D Web Technology, New York, 2012.
[28] Rossa M., Gogołek W., Łukasiewicz A., *Geostandardy, metadane i dyrektywa INSPIRE. Poradnik metodyczny Zintegrowanego Systemu Kartografii Geologicznej IKAR.* Państwowy Instytut Geologiczny, Warszawa, 2009.
[29] Rossa M., *Aplikacje GML dla środowiska na przykładzie GeoSciML i GWML2 – stan aktualny, kierunki rozwoju, zastosowania praktyczne*, Konferencja „GML w praktyce”, 12.04.2012, Kon-Dor, Warszawa, 2013.
[30] Rossa M., 2014a *Zastosowanie języka GML do przetwarzania danych geologicznych, XXIII Szkoła eksploatacji górniczej*, 24-28.02.2014, Kraków.
[31] Rossa M., 2014b *Wykorzystanie języka GML do opisu danych hydrogeologicznych i hydrograficznych*, *XXIII Szkoła eksploatacji górniczej*, 24-28.02.2014, Kraków.
[32] Sánchez López T., RFID and sensor integration standards: State and future prospects, *Computer Standards & Interfaces*, **33**, 2011, 207–213.

Received 11.09.2014, accepted 28.05.2015