

ORIGINAL ARTICLE

Considerations about the quality of UML and GML application schemas

Agnieszka Chojka  ^{1*}

¹ Department of Geoinformation and Cartography, Institute of Geodesy and Civil Engineering, Faculty of Geoengineering, University of Warmia and Mazury, 15 Prawochenskiego Str., 10–720, Olsztyn, Poland

*agnieszka.chojka@uwm.edu.pl

Abstract

Exchange of and access to spatial data is the principal goal of any Spatial Data Infrastructure, therefore, one of the key concepts of SDI is interoperability, especially semantic and syntactic. Whereas application schemas and quality issues are one of the aspects that have to be considered to ensure a successful data interchange in SDI.

Two types of application schema are widely used in the European SDI as well as in the Polish SDI. They cover both semantic and syntactic interoperability and are an integral parts of spatial data specifications and relevant regulations in the form of data models. However, working out accurate and correct application schemas may be a challenge. Additionally, faulty or too complex application schemas can influence the ability for valid data interchange, and consequently, prevent achieving interoperability within SDI. Therefore, the capability to examine and estimate the UML and GML application schemas quality seems to be a worthwhile and important issue in the context of semantic and syntactic interoperability in SDI.

The main subject of this article it to set out the context of performed studies, among others, the role of application schema in the interoperable data exchange, issues related to the concept of quality and its measures.

Key words: application schema, quality, interoperability, UML, GML

1 Introduction

Interoperability, in particular semantic and syntactic, is one of the core concepts of Spatial Data Infrastructure (SDI) due to the fact that exchange of and access to spatial data is the foremost aim of any SDI. The main SDI initiative in the European Union (EU) is INSPIRE (Directive, 2007). It is established at the supranational level to support environmental policies and policies or activities that may have a direct or indirect impact on the environment.

According to the CEN/TR 15449–3 (CEN/TC 287, 2012b), there is a lot of data interoperability aspects that have to be taken into account, among others, application schema, quality issues, data transfer, consistency and conformity checks.

Spatial data exchange via SDI, in an interoperable manner, assumed using two types of application schema. The first one is expressed in the UML (Unified Modelling Language) and the

second one in the GML (Geography Markup Language). They comprise semantic and syntactic interoperability. However, working out accurate and correct application schemas may be a challenging task. Many issues should be considered, for example, the recommendations of ISO 19100 series of International Standards in the geographic information domain, relevant regulations for given problem or theme, production opportunities and limitations. In addition, faulty or too complex application schemas can influence the ability to valid data interchange.

Therefore, the capability to examine and estimate the UML and GML application schemas quality, including also exploring their complexity, seems to be a worthwhile, very interesting and important issue in the context of interoperability in SDI, especially semantic and syntactic.

The main subject of research is developing the full and complex methodology for examining and evaluating the UML and GML application schemas quality, mainly used within the Eu-

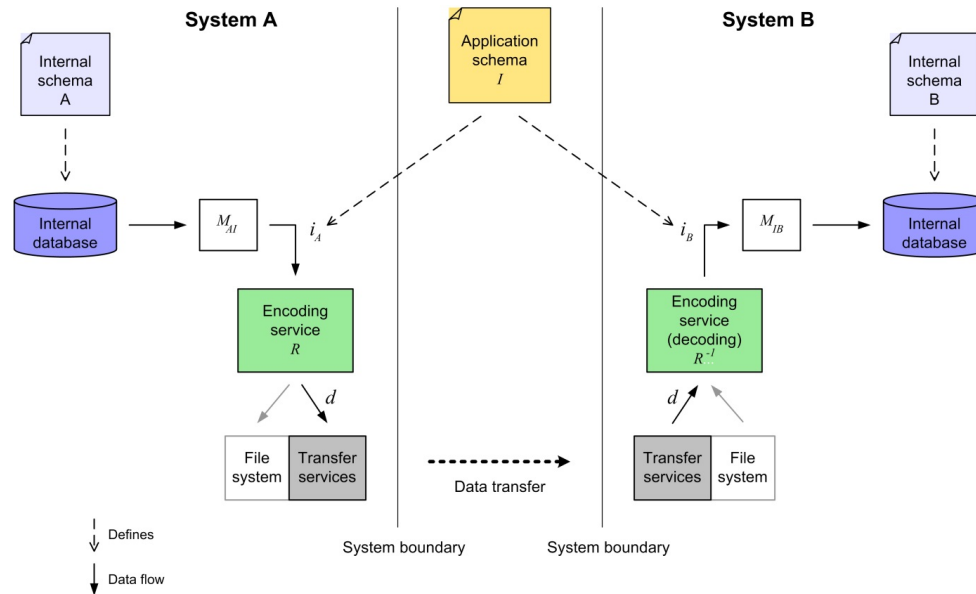


Figure 1. Overview of the interoperable data exchange (source: own elaboration on the basis of ISO 19118 (ISO/TC 211, 2011) and the CEN/TR 15449-3 (CEN/TC 287, 2012b))

ropean and Polish SDIs. This article presents the context of studies, including issues related to the interoperability concept, spatial data exchange, the role of application schemas in this respect, as well as the quality concept and its measures. The results of the conducted research will primarily become the contribution base for creating some guidelines and recommendations that will allow to optimise the UML and GML application schemas currently in force in Poland.

2 Application Schema Role

In line with the reference model for SDIs, defined in the CEN/TR 15449-1 (CEN/TC 287, 2012a), one of the general considerations for achieving interoperability is the use of the model-driven approach. This solution, also promoted by the ISO 19100 series of geographic information standards, enables cross-platform interoperability and follows the concepts formulated in the model-driven architecture (MDA) defined by the Object Management Group (2014). The universe of discourse (view of the real or hypothetical world that includes everything of interest (ISO/TC 211, 2014)) is the starting point in this approach. It is expressed in the form of a conceptual model that formally can be represented in one or more conceptual schemas. This schema defines how the model of the real world is described with data and applying a conceptual schema language (CEN/TC 287, 2012b). In turn, the conceptual schema language is a formal language containing the required linguistic constructs to describe the conceptual model in the conceptual schema (ISO/TC 211, 2014). Additionally, conceptual schema can be used by one or more applications and then it is called an application schema. It provides not only a description of the semantic structure of the spatial dataset but can also identify the spatial object types and reference systems required to provide a complete description of geographic (spatial) information in the dataset (CEN/TC 287, 2012b).

2.1 Interoperable Data Exchange

The application schema is the basis of a successful data transfer between different systems as it defines the possible content and structure of the exchanged spatial data, which means it co-

vers both semantic and syntactic interoperability. Applications (software) and users (people) should interpret data and information in the same manner to ensure they are understood as it was planned by the producer of the data.

The general idea of the data interchange between two different systems is shown in Figure 1. System A wants to send a dataset to system B, what follows, system B has to be able to use data from system A. To ensure a successful result of this process, it is necessary for both systems to determine a common application schema I, an encoding rule R and a transfer protocol (ISO/TC 211, 2011).

According to the ISO 19100 suite of standards, the application schema used for encoding should be written in the UML conceptual schema language, in compliance with ISO 19103 (ISO/TC 211, 2015a) and ISO 19109 (ISO/TC 211, 2015c). These documents provide a set of rules for how to properly write the application schema, including the usage of standardized schemas to define feature types. a sender and a receiver of spatial data must have an access to the application schema. It is even recommended to transfer it before data interchange proceeded, to allow both ends of this transaction to prepare their systems by implementing appropriate mappings and data structures corresponding to the application schema (ISO/TC 211, 2011).

The encoding rule is an identifiable collection of conversion rules that defines the encoding for a particular data structure (ISO/TC 211, 2011). In accordance with ISO 19118, it specifies, among others, the syntax and structure of the resulting data structure and is applied to the application schema specific data structures to produce system independent data structures suitable for transport or storage. The conversion rule, by contrast, defines how a data instance in the input data structure is converted to a data instance in the output data structure (ISO/TC 211, 2011).

To conclude, the UML application schema determines the possible content and structure of the interchanged spatial data, whereas the encoding rule defines the conversion rules for how to code the data into a system independent data structure.

2.2 GML Encoding Rule

A good example of encoding rule mentioned above is the XML-based encoding rule for neutral data interchange described in

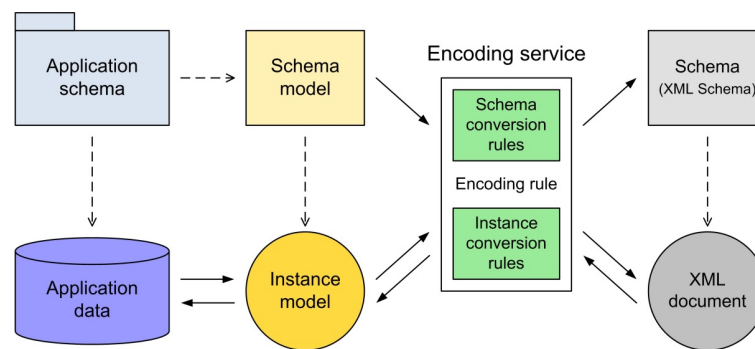


Figure 2. XML-based encoding rule (source: own elaboration on the basis of ISO 19118 (ISO/TC 211, 2011))

detail in ISO 19118 standard (ISO/TC 211, 2011). It is compatible with the UML and defines the encoding rule based on the XML (eXtensible Markup Language). An overview of this encoding rule is illustrated in Figure 2. Within the XML-based encoding rule exists two sets of conversion rules. The first one specifies a mapping from the UML class definitions in the application schema to the type declarations in the XML Schema. The second one defines a mapping from the objects in the instance model to the corresponding element structures in the XML document.

The discussed XML-based encoding rule is applied in the GML standard that is an XML encoding based on principles specified in ISO 19118. The GML provides a common XML encoding for spatial data, along with an open, vendor-neutral framework for the description of geospatial application schemas for the transport and storage of geographic information in the XML (ISO/TC 211, 2007b).

Whereas the GML application schema is an application schema written in the XML Schema in accordance with the rules specified in ISO 19136 (ISO/TC 211, 2007b). Additionally, it has to import the GML schema that comprises XML encodings of a number of the conceptual classes defined in the ISO 19100 series of International Standards.

In the case of INSPIRE, the GML is usually recommended encoding, unless otherwise specified for a specific data theme. By way of illustration, for large volume coverage data such as orthoimagery or computer simulations (e.g., weather forecasts), other, more efficient, file-based encodings (e.g., geoTIFF) may be defined as the default encoding language (Tóth et al., 2012).

2.3 Data Specifications

Both, UML and GML application schemas are commonly used in the European SDI, as well as in the National SDIs (e.g., in Poland). As data models, they are included in the data specifications that usually contain other relevant requirements about data, such as rules for data capture, encoding, and delivery, as well as provisions of data quality and consistency, metadata and so on.

In the broader sense, data specification can refer to both the data product specification and the interoperability target specification in SDI (Tóth et al., 2012).

The data product specification is a detailed description of a dataset or dataset series together with additional information that will enable it to be created, supplied to and used by another party (ISO/TC 211, 2007a). In line with the ISO 19131, it provides a definition of the universe of discourse and a specification for mapping the universe of discourse to a dataset. It may be used for production, sales, end-use or other purposes (ISO/TC 211, 2007a).

One of the key and mandatory items included in the data product specification is the data content and structure informa-

tion, which means an application schema. In case of a feature-based data product, this element is described in terms of an application schema and a feature catalogue (this issue is beyond the scope of this paper and will not be discussed later, for more information, see the ISO 19110 (ISO/TC 211, 2015b)). The application schema provides the formal description of the data structure and content of the data product. It is a conceptual model expressed in the UML (conceptual schema language in terms of the ISO 19100 suite of geographic information standards) in accordance with the ISO 19109. This model should include the representation of feature and property types (including attribute types), feature operations and associations, inheritance relations and constraints, where attribute types cover descriptive, geometric and temporal properties, whereas associations include spatial and temporal relationships such as topological as well as non-spatial relations (e.g., ownership) that occur between feature types (ISO/TC 211, 2007a). In case of coverage-based and imagery data, a coverage is considered a subtype of a feature and behaves like a function, which returns one or more feature attribute values from a direct position within a spatiotemporal domain (ISO/TC 211, 2007a).

The interoperability target specification is used for transforming existing data so that they share common characteristics (Tóth et al., 2012). In the case of INSPIRE, to achieve interoperability in the European SDI, data specifications have been established for the 34 spatial data themes. The Member States of the EU can use these documents to create new datasets or to transform existing datasets according to the specifications by mapping the existing models to the models defined in the data specification guidelines (CEN/TC 287, 2012b).

The Head Office of Geodesy and Cartography, the main coordinator of SDI creating and functioning in Poland, adopted a similar approach to implement the INSPIRE Directive and establish the National SDI. The law of the infrastructure for spatial information (Act, 2010) in Poland was approved. It is a transposition of the INSPIRE Directive to the national law. This action involved a necessity of many acts and relative laws' changes, among others, the law on geodesy and cartography (Act, 1989). The existing (very often obsolete) instructions and guidelines were replaced by regulations of the Cabinet or the responsible Minister. These documents became annexes to the law on geodesy and cartography as well as put some of the INSPIRE Directive recommendations into action.

An integral part of these developed regulations are the UML and GML application schemas that define information structures of spatial databases, corresponding to each regulation. These schemas were prepared in accordance with the ISO 19100 series of geographic information standards to ensure interoperability of spatial data sets and GIS applications. Regulations, data product specifications in terms of ISO 19131, cover the whole legal and technical issues regarding the geodetic domain in Poland. This was a very ambitious challenge due to the

methodology of the conceptual modelling and the usage of the UML and GML notations to describe the information content of databases, were applied for the first time in Poland.

3 Problematic Issues

To reach interoperability in SDI, two approaches are possible. The first one is transformation that uses the information and communications technologies and does not impact the original data structures. The second one is harmonisation, which relays on modifying and fine-tuning semantics and data structures to enable compatibility with agreements (specifications, standards, or legal acts) across borders and/or user communities (Tóth et al., 2012). In Tóth et al. (2012) view, when technical arrangements are not sufficient to bridge the interoperability gap between the communicating systems in SDI, then harmonisation is needed. Nevertheless, the combination of these two approaches provides the best solution in SDI.

The process of harmonisation requires either working out new data structures or adjusting existing data structures of spatial databases to INSPIRE guidelines and recommendations. Data structures are specified in the form of UML and GML application schemas. However, working out accurate and correct application schemas is not an easy task. Many issues should be considered, for instance, recommendations of the ISO 19100 series of geographic information standards, appropriate regulations for given problem or topic, production opportunities and limitations (i.e., software, tools).

Moreover, the GML application schema is closely related to the UML application schema. Usually the first one is the result of mapping from the ISO 19109 conformant UML application schema and this process is based on the set of encoding rules specified in the ISO 19136. But not everything that can be written in the UML can be represented in the GML. This can have a significant influence on the spatial data sets and GIS interoperability, and thereby, the ability to validly execute data exchange. In addition, application schema determines the final structure of the database. If it is faulty or too complex, it may influence the ability to generate GML data sets with concrete data (objects), and thereby, can cause various problems and anomalies at the data production stage.

Such problems occurred in Poland. Already during the creation of above discussed application schemas, some technical difficulties were identified. Most of them covered the UML to GML transformation issues. After publishing regulations, that determine data structures for relevant spatial databases, some contractors also reported remarks about application schemas, among others, faults, mistakes or anomalies in their notation. By way of illustration, one of elaborated application schemas included a recursion that prevented from generating the sample of GML data. Another example is the usage of incorrect geometric data type in the UML application schema, by extension in the GML application schema. This fault resulted in problems concerning the proper interpretation of the objects' geometry by the GIS software.

One of the reasons of these situations may be an ambiguity of the UML to GML transformation (Chojka, 2013), while another reason is too much complexity of prepared application schemas. Unfortunately, these problems can influence the possibility of generating GML files with spatial data, as well as the ability of GIS software to process these files and, consequently, make interoperability impossible to achieve.

Fortunately, at the European level, INSPIRE data specifications are revised regularly and some corrigenda or new versions of these documents are published on the INSPIRE website. In Poland, at the Head Office of Geodesy and Cartography, the work is currently underway to detect the most problematic is-

suces related to the existing UML and GML application schemas and to propose some improvements to optimise these structures.

Nevertheless, in connection with the foregoing, examining the quality of application schemas seems to be an extremely important issue in the context of interoperability in SDI, particularly semantic and syntactic. The results of studies could help avoid the above described difficulties in the future and could allow to offer users higher quality application schemas.

However, questions arise as to what does 'the quality of application schema' mean, how to examine or measure this quality, who can be the most interested in the outcomes of the quality evaluation?

4 Quality Concept

The expression 'quality' usually appears in statements like: 'this is good quality/poor quality' or 'sufficient/insufficient level of quality'. Although it is known in theory what this concept actually means, it is still hard to define it precisely and unambiguously. For instance, various dictionaries define quality as: 'the standard of something as measured against other things of a similar kind' (Oxford University Press, nd), 'how good or bad something is' (Cambridge University Press, nd), 'degree or standard of excellence' (HarperCollins Publishers, nd), 'any of the features that make something what it is' (HarperCollins Publishers, nd), 'a distinguishing characteristic, property, or attribute' (HarperCollins Publishers, nd).

It is assumed that quality was defined for the first time by Plato as 'a degree of excellence' (Greek 'poiôtês'). At that time, it was a philosophical term and remained as such to the present days. As a result of numerous disputations, it was only determined that quality have some objective and subjective characteristics. The first features are measurable, such as weight or shape, the second are evaluated differently by everyone, for example, colour or smell. Cicero introduced the word 'qualitas', the Latin translation of the Greek term that became a part of some Romance and Germanic languages, as Italian ('qualità'), French ('qualité'), German ('die Qualität') or English ('quality').

The current definitions of quality place increased emphasis on the social aspects, especially product quality and its value in use. The literature mentions many various definitions of quality, depending on the context and area of application. By way of illustration, in sociological terms, quality is regarded as an expression of consumers' point of view on certain quality attributes, and in humanities, as creating an adequate quality of life and work that increase the level of culture in society (Bielawa, 2011).

Economic sciences, in the area of quality management, specify quality as a degree to which a specific product meets customer needs (marketable quality), a level of compliance of the product with its model, pattern or requirements (conformance quality), an extent of how a certain product has customer's priority over any other product after comparing them (preferences quality), and finally, a distinguishable set of features that are essential for a given product, for example, size, appearance, reliability and so on (Juran, 1993).

In technical terms, quality is defined excluding the recipient. a project or standard play the most important role in this context, for example, predictable degree of uniformity and dependability at low cost with a quality suited to the market (Deming, 1986) or conformance to requirements (Crosby, 1979).

For comparison, according to the International Standards, quality refers to the totality of features and characteristics of a product or service that bears its ability to satisfy stated or implied needs (ISO/TC 176/SC 1, 1986) either a degree to which a set of inherent characteristics of an object fulfils require-

ments (ISO/TC 176/SC 1, 1986).

In the geographic information (geoinformation/geomatics) domain, special attention should be given to the quality concept defined in the ISO 19100 series of standards that are used for creating the European SDI as well as National SDIs by the Member States of the EU.

4.1 Quality According to ISO 19100

The ISO 19100 series of geographic information standards share a common definition of quality, previously defined in ISO 9000, cited above (ISO/TC 176/SC 1, 2015). In general, quality issues are currently discussed in the following normative documents:

- ISO 19157:2013 Geographic information – Data quality (ISO/TC 211, 2013),
- ISO/TS 19157-2:2016 Geographic information – Data quality – Part 2: XML schema implementation (ISO/TC 211, 2016),
- ISO/TS 19158:2012 Geographic information – Quality assurance of data supply (ISO/TC 211, 2012).

ISO 19157 provides the principles for the description of geographic data quality and specifies components for reporting quality information as well as procedures for the evaluation of geographic data quality.

In turn, ISO/TS 19157-2 concerns data quality encoding in the XML. It defines an XML Schema implementation of ISO 19157 and the data quality related concepts from ISO 19115-2 (ISO/TC 211, 2009).

ISO/TS 19158 establishes a quality assurance framework specific to geographic information that is based upon the quality principles and quality evaluation procedures identified in ISO 19157 and the general quality management principles defined in ISO 9000.

Interestingly, this standard introduces the concepts of ‘customer’, ‘supplier’ and ‘product’ derived from ISO 9000. According to both these documents, the customer is defined as an organization or person that receives a product, the supplier is an organization or person that provides a product and the product is a result of a process (defined as a set of interrelated or interacting activities, which transforms inputs into outputs) (ISO/TC 176/SC 1, 2015). Additionally, it was clarified that the supplier has provided the product via the process that can have some impact on the quality (ISO/TC 211, 2012). The ISO/TS 19158 arrangements are applicable to customers and suppliers of all geographic information.

However, the above standards consider quality mainly at the data level and the subject of the research activities is taking into consideration quality at the higher level of abstraction, that is, the level of data structures. It may be assumed that by ensuring the suitable quality of spatial data structures (application schemas), the quality of spatial data, including spatial sets and databases, can be increased significantly.

4.2 Quality Aspects in SDI

SDI relies on standards and specifications in the domain of geographic information and information technology. Therefore, the above discussed ISO standards regarding quality are also relevant in terms of any SDI development and implementation.

In the CEN/TR 15449-1 view (CEN/TC 287, 2012a), quality issues, particularly checking the quality, are an important consideration for achieving interoperability. However, quality is not considered in absolute terms, but when it comes to user requirements (CEN/TC 287, 2012b). Users, first of all, require information about the quality of datasets to assess whether the datasets are useful for them or not (‘fitness for purpose’). The quality levels for each spatial dataset are defined using the

criteria established by the ISO 19100 suite of standards, including completeness, consistency, currency, accuracy and usability. In turn, quality information concerning individual spatial objects is part of the metadata associated with the respective spatial objects and generally should be described as part of the application schema (CEN/TC 287, 2012b).

According to Tóth et al. (2012), from the point of view of SDI, poor data quality may compromise interoperability. Therefore, it can be generally stated that poor application schema quality also may impact negatively on achieving interoperability within SDI.

5 Methodology Concept for Quality Evaluation

Based on the above facts and considerations, a methodology for evaluating application schemas quality was suggested. First of all, it should consist of procedures and measurements of quality evaluation. Moreover, some abstract test suite (ATS) can also be taken into account. In line with ISO 19105 (ISO/TC 211, 2000), ATS is an abstract test module that specifies all the requirements to be satisfied for conformance and includes a set of related abstract test cases that in turn are generalized tests for a particular requirement. In addition, abstract test cases are a formal basis for deriving executable test cases that generate executable test suite (ETS).

In general, this standard provides the framework, concepts and methodology for testing, as well as the criteria to be achieved to claim conformance to the family of the ISO geographic information standards (ISO/TC 211, 2000). In particular, it provides a framework for specifying ATS and for defining the procedures to be followed during conformance testing. According to ISO 19105 (ISO/TC 211, 2000), conformance may be claimed for data or software products or services or by specifications including any profile or functional standard. Therefore, such conformance can also be requested for application schemas.

The application schema reflects the universe of discourse that can be seen differently from the creator perspective (e.g., software analyst) and the user perspective (e.g., database operator) of such described database structure. For this reason, the quality of UML and GML application schemas can be considered from the marketing point of view that emphasizes the commercial aspects of quality (Fraś, 2000). Thus, quality of application schemas can be described from the producer and consumer perspective, in other words, from the point of view of the entity that ‘experience quality’. Then it will be understood as the capability of the product (work of manufacturer) to meet consumer (user) needs. However, both the producer and consumer can interpret quality differently due to their distinct expectations related to the certain product. Such features of the product as competitive, providing economics benefits or satisfying technological needs are important for the manufacturer. In contrast, the user demands, among others, reliability, high comfort of use or even aesthetic design of the product. In this context, these aspects of application schemas quality are discussed below.

5.1 Producer Perspective

The quality of UML and GML application schemas at the producer level, that is the manufacturer or supplier level (in the considered case, the EU and the Head Office of Geodesy and Cartography), can be defined as a ‘formal quality’ or ‘legal quality’. This means that these application schemas (products) conform to the relevant standards of the ISO 19100 series of geographic information standards, as well as to the appropriate

legal regulations concerning specific thematic issue. Additionally, such product characteristics as profitability (cost, market size) and competitiveness, including among others, technological leadership or an ongoing development of the product and process (Fraś, 2000), can also be taken into account. The concept of quality management, called the Deming cycle (Deming, 1986), known as the PDCA cycle, is linked with the latter feature. The continuous improvement of products and processes is an iterative four-step management method including Plan, Do, Check and Act stage. Similar approach was applied in the methodology for data specification development, presented by Tóth et al. (2012), with the effect that INSPIRE data specifications are revised regularly.

It is recommended that primarily, the methodology of formal quality examination for the UML application schema should consider its verification against the ISO 19103 (ISO/TC 211, 2015a), for example, choosing the proper data types, including geometry types, as well as ISO 19109 (ISO/TC 211, 2015c), generally correctness of the schema design. Conformity with relevant legal provisions in the scope of the given topic also should be taken into consideration.

In respect of evaluating the GML application schema quality, it is necessary to check its conformance with ISO 19136 (ISO/TC 211, 2007b) recommendations, mainly validation of schema structure, and also verification of their range of information.

When it comes to the evaluation of formal quality of application schemas, some previously prepared ATS can be applied. Examples of such ATS can be found in INSPIRE data specifications.

5.2 Consumer Perspective

The quality of UML and GML application schemas at the consumer level, which means their user (in the considered case, all users of the European SDI, as well as surveying and cartographic administration in Poland and also ordinary users of spatial data), can be called a ‘technical quality’. This type of quality includes some functional and non-functional requirements for application schemas. By way of illustration, economy, reliability, maintainability or generally user comfort (Fraś, 2000) can be recognized as functional user needs. On the other hand, aesthetic design or image building (e.g., maintaining database and generating data that conform to obligatory application schema) are examples of non-functional requirements.

It is recommended that technical quality examination for the UML application schema should include, among others, verification of association roles, which are used in the UML for describing relations between classes. Their lack in the UML application schema causes the impossibility of encoding such association in the XML Schema. Besides, checking whether there exists double reference between classes (classes can indicate to each other what leads to recursion), whether applied <<voidable>> stereotype and its correctness and validity of use, occurrence of abstract classes, review of geometry types (e.g., in some cases, defined geometry can be too general or too detailed) and the way of specifying enumeration types (CodeList and Enumeration; in some cases, giving only codes without any additional description can be ambiguous for users).

Regarding technical quality examination for the GML application schema, verification of the correct usage of types with ‘PropertyType’ and ‘Type’ suffix, encoding attributes with assigned <<voidable>> stereotype in the UML (in the GML ‘nilReason’ corresponding attribute), relations between classes (e.g., the lack of association roles in the UML entail the lack of information about relation in the GML code) and the manner of

their encoding, as well as examination how enumerations and code lists are encoded in the XML Schema, should be mainly considered.

5.3 Other Aspects of Quality Examination

In addition to the already discussed aspects of evaluating application schemas quality, the considered methodology should also include some characteristics that can be simply measured or even calculated. An illustrative example is the entropy and complexity of application schema. It is worth investigating how complexity of application schema (e.g., the number of classes and how they are linked) influences its quality.

For this purpose, it is necessary to apply adequate complexity metrics. a good example of such a measure is a cyclomatic complexity, that is, a software metric (measurement) used in computer science to indicate the complexity of a program, developed by McCabe (1976). Mathematically, this approach is based on the graph theory and it can be easy adapt to verify the complexity of the UML application schema. The final results of these research activities will be presented in the following article about the complexity of UML and GML application schemas, focusing on a number of selected application schemas prepared in the Head Office of Geodesy and Cartography in Poland within the INSPIRE Directive implementation works, as well as application schemas from INSPIRE data specifications.

Another important point that must be looked at is the correlation between the quality of application schema and the quality of spatial data sets and services. This raises a question as to whether the faulty application schema results in defective data; in other words, whether the error propagation occurs.

6 Summary and Further Works

Modern science has yet to determine conclusively the term ‘quality’. This is largely due to its abstraction (Bielawa, 2011). The quality itself does not exist, and therefore, it should only be considered in relation to its objective to be achieved (Olejnik and Wieczorek, 1982). In the context of UML and GML application schemas, from the producer’s point of view, this aim is to avoid mistakes and anomalies in elaborated schemas, as well as to prevent an ambiguity of the UML to GML transformation. From the user’s point of view, this goal is an ease of databases use or ease of generating GML files with spatial data.

It should be noted that due to the relativity of quality, there is no ideal quality pattern. Both, the producer and user, can place different demands regarding application schemas. a quality measure may be a degree of meeting their specific needs.

For the above reasons, in order to meet these objectives, the methodology for examining and evaluating the UML and GML application schemas quality is needed. Obviously, development of suitable supporting software tools to automate the whole process of quality verification should also be considered (equivalent of the executable test suites). More thought is needed about opportunities to certify the quality of UML and GML application schemas.

The foregoing deliberations are an attempt to work out in the future full and complex methodology for examining and evaluating the UML and GML application schemas quality, mainly used within the European and Polish SDIs.

The methodology outlined in the previous chapters requires further clarification and systematization. Drawing on the ISO 19137 guidelines, first of all the framework of application schema quality concepts should be defined. Besides, components, measures, evaluation process and methods, and reporting should also be determined.

References

- Act (1989). Act of 17 May 1989 Geodetic And Cartographic Law. Official Journal 1989, No. 30, item 163.
- Act (2010). Act of 4 March 2010 on the Infrastructure for Spatial Information. Official Journal 2010, No. 76, item 489.
- Bielawa, A. (2011). Postrzeganie i rozumienie jakości – przegląd definicji jakości [Understanding and Perception of Quality – Review of the Quality Definition]. *Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania*, 21:144–152.
- Cambridge University Press (n.d.). Quality. In *Cambridge Dictionary*. Retrieved 27 April 2020 from <http://dictionary.cambridge.org/>.
- CEN/TC 287 (2012a). *Geographic information – Spatial data infrastructures – Part 1: Reference model*. Brussels, Belgium. CEN/TR 15449-1.
- CEN/TC 287 (2012b). *Geographic information – Spatial data infrastructures – Part 3: Data centric view*. CEN/TR 15449-3.
- Chojka, A. (2013). Niejednoznaczność transformacji UML–GML [UML–GML Transformation Ambiguity]. *Roczniki Geomatyki*, 11.
- Crosby, P. B. (1979). *Quality is free: The art of making quality certain*, volume 94. McGraw–Hill, New York, USA.
- Deming, W. E. (1986). *Out of the Crisis*. MIT Center for Advanced Engineering Study.
- Directive, I. (2007). Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). *Published in the official Journal on the 25th April*.
- Fraś, J. (2000). *Zarządzanie jakością w instytucjach gospodarczych [Quality management in economic institutions]*. Wydawn. Naukowe Uniwersytetu Szczecińskiego, doi:10.18276/sip.2015.39/1-07.
- HarperCollins Publishers (n.d.). Quality. In *Collins Dictionary*. Retrieved 27 April 2020 from <https://www.collinsdictionary.com/>.
- ISO/TC 176/SC 1 (1986). *Quality – Vocabulary*. (ISO Standard No. 8402).
- ISO/TC 176/SC 1 (2015). *Quality management systems – Fundamentals and vocabulary*. (ISO Technical Standard No. 9000).
- ISO/TC 211 (2000). *Geographic Information – Conformance and testing*. (ISO Standard No. 19105).
- ISO/TC 211 (2007a). *Geographic Information – Data product specifications*. (ISO Standard No. 19131).
- ISO/TC 211 (2007b). *Geographic information – Geography Markup Language (GML)*. (ISO Standard No. 19136).
- ISO/TC 211 (2009). *Geographic information – Metadata – Part 2: Extensions for imagery and gridded data*. (ISO Standard No. 19115-2).
- ISO/TC 211 (2011). *Geographic information – Encoding*. (ISO Standard No. 19118).
- ISO/TC 211 (2012). *Geographic Information – Quality assurance of data supply*. (ISO Technical Standard No. 19158).
- ISO/TC 211 (2013). *Geographic Information – Data quality*. (ISO Technical Standard No. 19157).
- ISO/TC 211 (2014). *Geographic Information – Reference model, Part 1: Fundamentals*. (ISO Standard No. 19157-1).
- ISO/TC 211 (2015a). *Geographic information – Conceptual schema language*. Geneva, Switzerland. (ISO Standard No. 19103).
- ISO/TC 211 (2015b). *Geographic Information – Methodology for feature cataloguing*. (ISO Standard No. 19110).
- ISO/TC 211 (2015c). *Geographic information – Rules for application schema*. (ISO Standard No. 19109).
- ISO/TC 211 (2016). *Geographic Information – Data quality – Part 2: XML schema implementation*. (ISO Technical Standard No. 19157-2).
- Juran, J. M. (1993). *Quality planning and analysis; from product development through use*. Number 04; TS156, J8 1993. in McGraw–Hill Series in Industrial Engineering and Management Science. McGraw–Hill College, 3rd edition.
- McCabe, T. J. (1976). A complexity measure. *IEEE Transactions on software Engineering*, SE-2(4):308–320, doi:10.1109/TSE.1976.233837.
- Object Management Group (2014). *Model Driven Architecture (MDA), MDA Guide rev. 2.0*.
- Olejnik, T. and Wieczorek, R. (1982). *Kontrola i sterowanie jakością: (w przemyśle elektromaszynowym) [Monitoring and Control of Quality]*. PWN, Poznań.
- Oxford University Press (n.d.). Quality. In *Oxford Dictionary*. Retrieved 27 April 2020 from <https://en.oxforddictionaries.com/>.
- Tóth, K., Portele, C., Illert, A., Lutz, M., and Nunes de Lima, V. (2012). *A conceptual model for developing interoperability specifications in spatial data infrastructures*. Office for Official Publications of the European Commission, Ispra, Italy, doi:10.2788/21003.