SPATIAL DATA QUALITY CONTROL PROCESS BASED ON ISO 19113 AND ISO 19114

Aleš Veršič
The Environmental Agency of the Republic of Slovenia
Vojkova 1b, 1000 Ljubljana
aversic@gov.si

Abstract
Introduction
The Environmental Agency of the Republic of Slovenia (hereinafter: Environmental Agency) performs expert, analytical, regulatory and administrative tasks related to the environment at the national level. Its mission is to monitor, analyse and forecast natural phenomena and processes in the environment, and to reduce natural threats to people and property. These tasks are performed by the national services for meteorology, hydrology and seismology. The Agency's mission is also to monitor environmental contamination and to provide reliable public environmental data; to this end, the Agency has the appropriate measuring network and laboratories in place. Its mission of exceptional importance is, moreover, to meet the requirements regarding environmental protection deriving from the regulations in force, to preserve natural resources and the biotic diversity and to ensure sustainable development of the country.
Environmental Agency. About the Agency. Available at: http://www.arso.gov.si/en/ about%20 the% 20agency/ [Accessed 10 January 2009].

Objectives
Thanks to the development of modern technologies, data capture has become easier, but the risk of poor data quality has increased. Every data producer should, as part of the work process, establish mechanisms of quality control of data and describe their quality in a standard manner. In this paper we will try to describe the data quality process for the spatial data, for which is the Agency responsible. In this process two international standards where used. The International Organization for Standardization – ISO has published ISO 19113 and ISO 19114 that help data producers to objectively describe the quality of data and quality evaluation procedures for geographic information.

Methodology
In the paper will be described the data quality process which is set-up in the Environmental Agency. In description of process and quality test application we will highlight some part of them.

Results
The agency obtains some data from professional institutions performing certain tasks for the purposes of the Ministry of the Environment and Spatial Planning, whilst some data are generated as part of work processes at the agency. Since these data are the basis

for decision-making, it is even more important that they are of good quality. For this reason, the Environmental Agency has adopted a decision on carrying out data quality control. Accordingly, all spatial data must be checked for quality.

Conclusion

The ISO 19113 and ISO 19114 standards, which are applicable to the area of geographic information systems, help spatial data producers and administrators to describe spatial data quality in a standard manner. Since data quality control is a continuous process, a collection of documents had to be set up, describing the process of quality control for particular data. Control is carried out each time data is modified; a new control document is then drawn up and stored into the collection of documents. The Environmental Agency, in partnership with the company SRC.SI d.o.o., has developed an application that enables the agency to prepare a control document and store it in the collection of documents. With this application Environmental Agency also set-up the quality control process which enable overview of database quality. The main problem is maintaining the process which is continuous.

SPATIAL DATA QUALITY

Spatial data quality plays an important role in making decisions related to spatial issues. Thanks to the use of modern technologies and simplified data capture, the amount of generated and accessible data has increased. But the risk of poor quality of data and metadata has increased as well. In the services industry, the costs of meeting the quality requirements and standards account for up to 30% of total revenue [1].

The term **quality** is generally used for marking an outstanding achievement or confirming the excellence of a product. It is the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs. In creating and setting up a database, **quality** is the desired goal, which can be achieved only through prudent management and control of the course of production. The reliability of the implementation procedure has a direct impact on the usefulness of a database and indirectly affects its value. Spatial data quality is the totality of features and characteristics of a database that bear on its ability to satisfy the set of stated requirements [1].

Accuracy describes how close the measured value is to the actual value. Accuracy relates to the relationship between observations and actual, real-world phenomena. In practice, accuracy can be termed as the discrepancy between the actual value of the selected data and a comparable, more accurate value of the same data, which is usually an estimated mean of the values measured. Accuracy is only one of the data quality elements or indicators. [1]

In quantitative attributes**, precision** is indicated by the number of decimal places [1].

In general, we could say that high quality data are data whose quality is suitable for their intended use in decision-making and planning. For example, spatial data with

metre precision, which are complete for the entire area and are of sufficient quality for our needs. We could have more precise data, with centimetre precision, but this would not mean that our database is complete. We would have very precise, but incomplete data and consequently, a database of poor quality. For the purposes of the Environmental Agency, it is better to have complete data rather than very precise, but incomplete data.

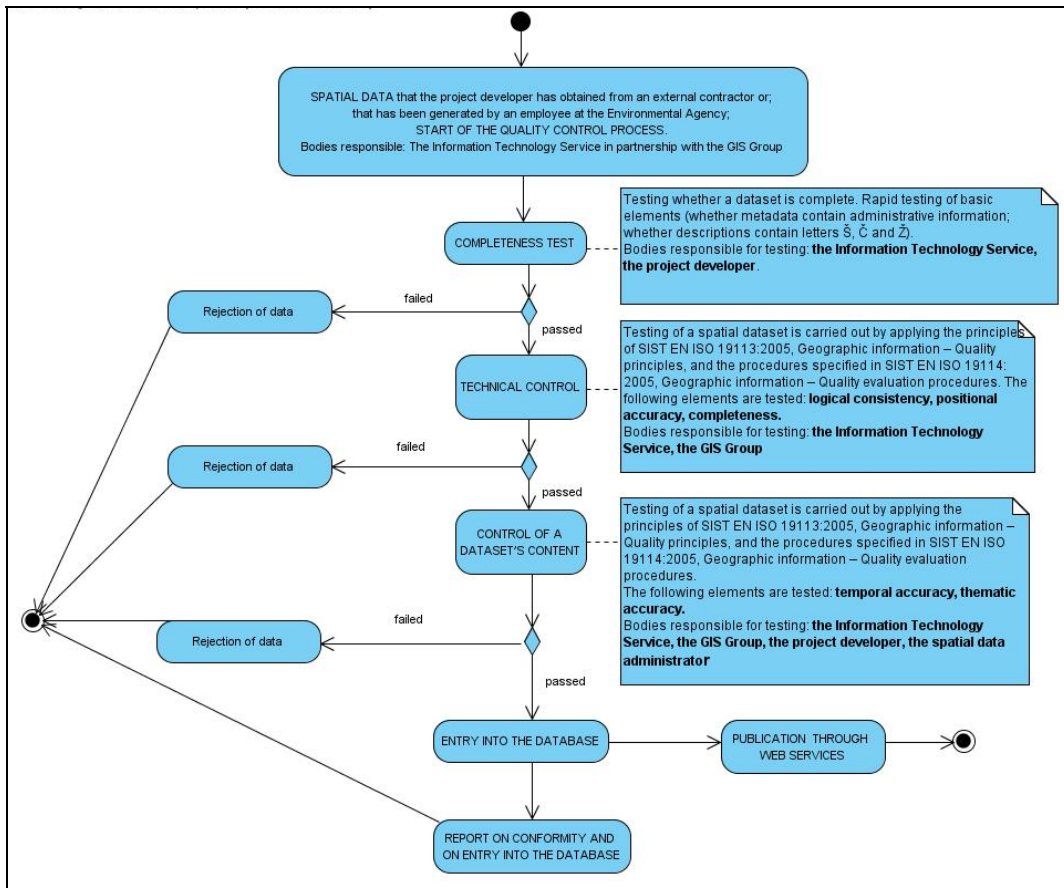**THE QUALITY CONTROL PROCESS AT THE ENVIRONMENTAL AGENCY**
The Environmental Agency is an important administrator of environmental spatial data, which are used in various analyses and spatial plans, and form the basis for adopting decisions. The spatial database is updated in a number of ways. The agency obtains some data from professional institutions performing certain tasks for the purposes of the Ministry of the Environment and Spatial Planning, whilst some data are generated as part of work processes at the agency. Since these data are the basis for decision-making, it is even more important that they are of good quality. For this reason, the Environmental Agency set-up working group for data quality control process. Accordingly, all spatial data must be checked for quality.

In carrying out the quality control of geographic information, we can refer to the following two standards: **SIST EN ISO 19113:2005, Geographic information – Quality principles**, which lays down quality principles, and **SIST EN ISO 19114:2005, Geographic information – Quality evaluation procedures,** which provides a framework for decision-making and quality evaluation and is applicable to all types of digital geographic data. **ISO 19113** provides a quality model enabling spatial data producers to objectively describe the quality of data, and users to determine whether a dataset is suitable for its intended use. **ISO 19114** provides uniform guidelines for determining quality. It enables data users to follow the data producer's quality evaluation procedures if they consider the available information on quality insufficient.

SIST EN ISO 19113:2005 and SIST EN ISO 19114:2005 standards are applicable to geographic data producers providing information on quality to describe how precise a dataset is, and information for users on whether this quality meets their needs. The standard does not define the minimal acceptable level of quality for geographic data. It leaves it to the data producer or the user to decide on the basis of data quality whether they will use the data. [1]

On the basis of internal instructions and taking account of international standards, the Environmental Agency has put in place the quality control process.

As we can see on the Figure 1, spatial data have to be forwarded to the Information Technology Service, which begins quality control. In accordance with ISO 19113 and ISO 19114, the person responsible for data quality control draws up a control document providing detailed instructions on how to carry out quality control. Records of quality control are kept and stored in the Lotus Notes document collection **QUALITY TEST**.

SPATIAL DATA that the project developer has obtained from an external contractor or; that has been generated by an employee at the Environmental Agency; START OF THE QUALITY CONTROL PROCESS. Bodies responsible: The Information Technology Service in partnership with the GIS Group

COMPLETENESS TEST

Testing whether a dataset is complete. Rapid testing of basic elements (whether metadata contain administrative information; whether descriptions contain letters Š, Č and Ž). Bodies responsible for testing: **the Information Technology Service, the project developer**.

TECHNICAL CONTROL

Testing of a spatial dataset is carried out by applying the principles of SIST EN ISO 19113:2005, Geographic information – Quality principles, and the procedures specified in SIST EN ISO 19114: 2005, Geographic information – Quality evaluation procedures. The following elements are tested: **logical consistency, positional accuracy, completeness.** Bodies responsible for testing: **the Information Technology Service, the GIS Group**

CONTROL OF A DATASET'S CONTENT

Testing of a spatial dataset is carried out by applying the principles of SIST EN ISO 19113:2005, Geographic information – Quality principles, and the procedures specified in SIST EN ISO 19114:2005, Geographic information – Quality evaluation procedures. The following elements are tested: **temporal accuracy, thematic accuracy.** Bodies responsible for testing: **the Information Technology Service, the GIS Group, the project developer, the spatial data administrator**

Rejection of data

failed / passed

ENTRY INTO THE DATABASE

PUBLICATION THROUGH WEB SERVICES

REPORT ON CONFORMITY AND ON ENTRY INTO THE DATABASE

Figure 1: Diagram of the spatial data quality control process at the Environmental Agency

The person responsible for quality control works with the spatial data administrator and provides help in verifying the data content, which takes place after the technical control of the data has been carried out. After a successful control of the data content, the GIS database administrator is informed that the data have successfully passed the quality test; the GIS database administrator then enters the data into the database and completes the conformity report. Depending on the status of the data, the database administrator, in partnership with the spatial data administrator agrees on the publication of the data through web services.

If, at any particular stage, data fail to meet the quality criteria, they will be rejected, and the project developer or the spatial data administrator will be informed of their rejection. The data producer is then invited to correct errors and resubmit the data for quality control. This procedure is repeated until the data meet the basic principles of spatial data quality.

Table 1 show 5 elements and 15 sub-elements that are used for describing spatial data quality and are checked in the control process. [2] [3]

Name of the basic element	Name of the sub-element	Name of the sub-element	Name of the sub-element	Name of the sub-element
Completeness of spatial data or a model.	Data commission	Data omission	/	/
Logical consistency of conceptual and logical rules of the data model and data structure.	Conceptual consistency	Domain consistency	Format consistency	Topological consistency
Positional accuracy of the position of features in a dataset.	Absolute accuracy	Relative accuracy	Gridded data position accuracy	/
Temporal accuracy of attributes and temporal relationships of features in a dataset.	Accuracy of a time measurement	Temporal consistency	Temporal validity	/
Thematic accuracy of the classification of features, of quantitative attributes and of the correctness of non-quantitative attributes in a dataset.	Classification correctness	Quantitative attribute accuracy	Non-quantitative attribute correctness	/

**Table 1: The 5 basic data quality elements and 15 data quality sub-elements**

QUALITY TEST APPLICATION
The Environmental Agency, in partnership with the company SRC.SI d.o.o., has developed an application for keeping records of the quality control process for a particular dataset. The person carrying out quality control uses the "Quality Test" application to draw up a control document containing data quality elements that are to be used in the control process. On the Figure 2 we can see some of the elements that can be used.
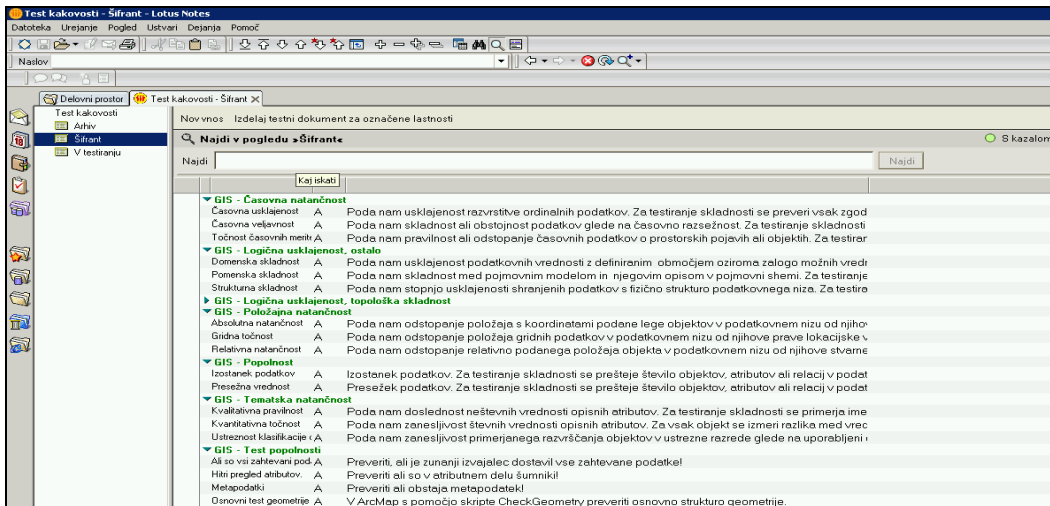
**Figure 2: Preparing a control document to carry out spatial data quality control**

In Figure 3 is sample of quality test document which was drawn up by person carrying out quality control. Document contains ID of test document, title, name of data controller, basic instruction and detailed instructions.
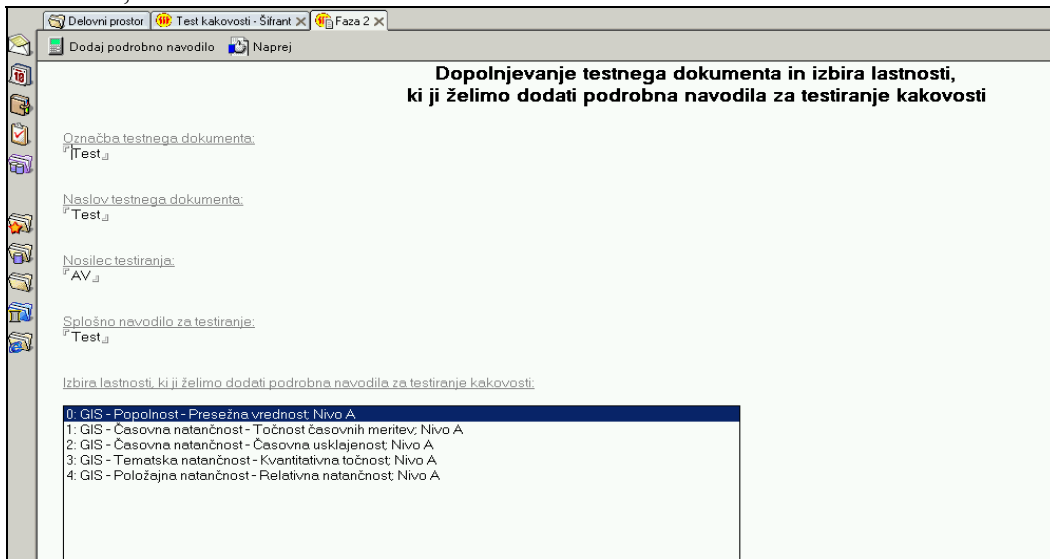
Figure 3: Detailed instructions on quality control

The person in charge of the process then adds to the control document detailed instructions on how to carry out quality control for each particular element.

When a control document is prepared and all detailed instructions provided, the person responsible for quality control begins testing. In so doing, he/she must follow the instructions and, after the process is completed, enter any observations obtained during testing.
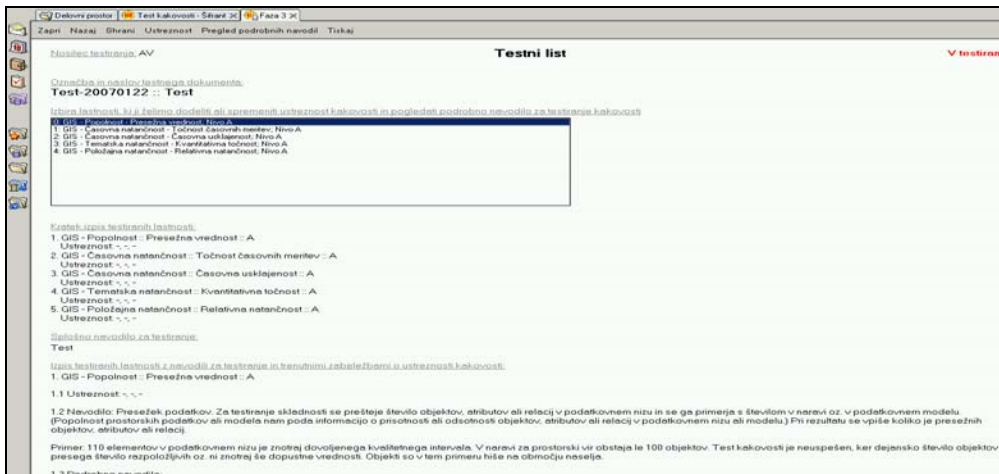
Figure 4: Report document on quality control

SOURCES AND REFERENCES

ŠUMRADA, Radoš: Strukture podatkov in prostorske analize, 2005. University of Ljubljana, Faculty of Civil and Geodetic Engineering

SIST EN ISO 19113:2005; Geographic information – Quality principles (ISO 19113:2002)

SIST EN ISO 19114:2005; Geographic information – Quality evaluation procedures (ISO 19114:2003)