DE GRUYTER OPEN

Management Systems in Production Engineering

2016, No 3 (23), pp 178-182

Date of submission of the article to the Editor: 03/2016
Date of acceptance of the article by the Editor: 06/2016

DOI 10.12914/MSPE-06-03-2016

THE UNIFICATION OF THE CODE LISTS PROVIDED WITHIN THE DATA MODEL ORIGINATING FROM THE INSPIRE TECHNICAL GUIDELINES AND THE ONES PROVIDED FOR GESUT DATABASES IN THE CONTEXT OF POTENTIAL EXPLOITATION IN THE MINING INDUSTRY

Andrzej ZYGMUNIAK, Violetta SOKOŁA-SZEWIOŁA
Silesian University of Technology

Abstract:
This study is aimed at exposing differences between two data models in case of code lists values provided there. The first of them is an obligatory one for managing Geodesic Register of Utility Networks databases in Poland [9] and the second is the model originating from the Technical Guidelines issued to the INSPIRE Directive. Since the second one mentioned is the basis for managing spatial databases among European parties, correlating these two data models has an effect in easing the way of harmonizing and, in consequence, exchanging spatial data. Therefore, the study presents the possibilities of increasing compatibility between the values of the code lists concerning attributes for objects provided in both models. In practice, it could lead to an increase of the competitiveness of entities managing or processing such databases and to greater involvement in scientific or research projects when it comes to the mining industry. Moreover, since utility networks located on mining areas are under particular protection, the ability of making them more fitted to their own needs will make it possible for mining plants to exchange spatial data in a more efficient way.

***Key words**: GESUT, INSPIRE, database, interoperability, data harmonization*

INTRODUCTION

Since digital data are commonly used nowadays, there is a special role for those of a spatial character. The aim of fulfilling an idea of precise description provided for a surrounding space including all its elements as well as gathering very accurate information about the natural environment is of high importance, especially in the last decades. Chiefly problems concerning the task of protecting the environment are emphasized. This issue is so important that the European Commission devoted to it a Directive, essential for further considerations, which is called in the shortened way the INSPIRE Directive [2]. The Directive implements a number of regulations with respect to obtain and subsequent processing thematically grouped data related to all objects present in the surroundings. All these actions in principle should be aimed at full and virtual providing environment related data sets in a way which will enable their uniformed and successful protection as well as management in the context of sustainable development. That last idea is an integral part of any mining activity, especially when it comes to the coal mine industry, which makes the INSPIRE Directive [2] as well as its Polish transposition in form of corresponding legal acts considered as those of high importance for this branch of industry.

It is important to note that the INSPIRE Directive was transposed to the Polish legal system in form of *Infrastructure of Spatial Information Act (pol. ustawa o infrastrukturze informacji przestrzennej* [10]), which in a direct way defines only questions concerning the interoperability of newly created data sets and those already managed. In Poland, such databases are intertwined to surveyor measures and surveyors themselves via a number of decrees, which describe a wide range of data to be acquired, processed and gathered in a specific way. Some of the main areas of interest could be highlighted and one of them are utility networks (*pol. sieci uzbrojenia terenu*). All regulations applying to them are presented within the Decree of Administration and Digitization Minister for the local and national Geodesic Register of Utility Networks[1] databases issued on the 21st of October 2015 [9]. Data on utility networks should be considered as very important for mining plants as they are under special protection if located within the boundaries of an area endangered by the negative impacts caused by mining activities. The abovementioned protection is dictated by the need of assuring they will be operated in proper and safe conditions [11, 12]. This is why databases related to GESUT are frequently acquired by mining plants, which are obliged to take into account information stored in them when planning and prosecuting every kind of works.

Unfortunately, it seems that even though there were some inaccuracies signalized formerly [5] between the data

[1] pol. Geodezyjna Ewidencja Sieci Uzbrojenia Terenu – GESUT

models presented within previously binding Decree of Administration and Digitization Minister on geodesic register of utility networks database, topographic objects database and the detailed map issued on the 12th of February 2013 (*pol. Rozporządzenie Ministra Administracji i Cyfryzacji z dnia 12 lutego 2013 r. w sprawie bazy danych geodezyjnej ewidencji sieci uzbrojenia terenu, bazy danych obiektów topograficznych oraz mapy zasadniczej*) [8] and the models for utility networks presented in the Technical Guidelines [1] issued to the INSPIRE Directive, the newly published decree did not solve all the problems. One of the issues still to be solved are code lists used in both models, as they are significantly different in certain parts. This paper presents what are the possibilities of increasing their compatibility as it can facilitate the complicated process of data harmonization when it comes to passing spatial information to foreign partners [6]. The approach described during this study has been used successfully by Geobid Ltd. which thanks to this could give more detailed databases to the foreign partners cooperating in the frames of the GeoSmartCity.

CHARACTER OF THE CODE LISTS

The code lists are orderly set of values which can be used by a user of database organized with accordance to a specific data model. In case of GESUT [9] or utility networks described in the Technical Guidelines [1] the model is based on a GML file structure, which directly originates from an Extensible Markup Language (XML) [7] and stands as an IT way of writing down the structure of a geospatial database.

The code lists themselves are important in a process of describing stored objects since they provide the values possible to use as attributes. These are defined uniquely in every specific data model. In case of utility networks organized in accordance to the Polish Decree (Rozporządzenie] examples of attributes can be *additional information (pol. informacja dodatkowa) or kind of conduit (pol. rodzaj przewodu)* [9]. These two attributes are of a different character since the first one is defined as *character string*, which means that this attribute is to be provided by the user as a written description. On the contrary, the second one is the kind of attribute which obtains its values from a defined code list – *RodzPrzewodu* in this particular case. It means that the only values intended to serve as information given within the attribute are those presented on the list [3].

It is very clear of how great importance is for the code list values to be defined properly. Other question is the problem of making decision if a specific attribute shall be provided with the help of a code list or in other way, which also can lead to some inaccuracies between two data models.

Noteworthy is a fact, that the code lists themselves are subject to some regulations. The most important for further consideration is the one stating that in a creation phase of a specific data model it is possible to define a type of individual list and determine if it will be extensible or not [3]. In a consequence, a person utilizing such model will or will not be able to add some values to the code list. Furthermore, these can be of two different characters and serve as a *CodeList* type or an *Enumeration* one. The only distinction between them is that the CodeList type sets a group of values which can be provided as an attribute and is potentially extensible, as a user of database is allowed to add a value on his own if he states there is a lack of needed option on the list. In case of the *Enumeration* type such action is excluded and the user shall fill in the attribute value only using this one he can obtain from a relevant code list [4].

Having in mind the first option, thus potential ability of adding values to a code list within the very model, it has to be stated that this can serve as an opportunity to make it better adjusted for a specific user. Since the data model originating from the Technical Guidelines [1] posits such procedure at the implementation stage and allows fitting it to the need of individual users, there is also a chance to utilize this feature the other way around. Assuming that the data model provided for the GESUT databases is not, in the overall meaning, fully compatible with its counterpart presented in the Technical Guidelines [5] and one of the areas where the differences are visible are exactly the code lists, then at least this aspect ought to be considered as one which could be eliminated if the Technical Guidelines' model [1] will be expanded with as much code lists values originating from the GESUT model [9] as it is possible. Such extension leads to a situation when a user operating the INSPIRE extended data model will be able to obtain databases managed in accordance to the Polish Decree [9] and these data will not be void of some former not compatible attributes, so the risk they will be incomplete is reduced.

CODE LISTS IN THE GESUT DATA MODEL

The data model provided in the currently binding Polish Decree [9] offers eighteen different code lists defining the values possible to be used when entering information within a specific, relevant attribute. Not all of them have a counterpart presented in the Technical Guidelines data model [1], but in those cases when there is one, it is possible to make an attempt to bring the two to an increased compatibility.

Such correlation can be observed between the following code lists presented altogether with its values [9]:

1. Existance (*GES_Istnienie*):
   - existing,
   - projected,
   - under construction.
2. Exploitation (*GES_Eksploatacja*):
   - used,
   - disused.
3. Function (*GES_Funkcja*):
   - transport,
   - distribution,
   - house connection,
   - other.
4. Water network type (*GES_TypWodoc*):
   - general,
   - local.
5. Sewer network type (*GES_TypKanal*):
   - storm,
   - local,
   - combined,
   - industrial,
   - sanitary.
6. Gas network type (*GES_TypGaz*):
   - high pressure,
   - increased medium pressure,
   - medium pressure,
   - low pressure.
7. Electricity network type (*GES_TypElektr*):
   - highest voltage,
   - high voltage,

- medium voltage,
- low voltage.

8. Vertical position (*GES_Przebieg*):
   - elevated,
   - on ground surface,
   - underground.

   Pole type (*GES_RodzSlup*):
   - street light,
   - lighting mast,
   - telecommunication mast,
   - standard,
   - A-frame pole,
   - pylon,
   - railway overhead line pole,
   - tram overhead line pole,
   - trolleybus overhead line pole,
   - wind turbine,
   - telecommunication tower,
   - other.

   Appurtenance shape (*GES_KsztaltUrzadz*):
   - square,
   - circle,
   - rectangle,
   - oval,
   - other.

It has to be pointed out that in case of the code list concerning vertical position of an object there is no need of further study since the values presented within are exactly the same as those presented in the data model in Technical Guidelines [1]. At the same time, there is a possibility of proper binding the code lists related to existence, exploitation, function and water network type as they are, even though their values are slightly different or the specific value does not have any counterpart in the INSPIRE data model. Moreover, a subject of further study is a code list concerning appurtenances linked to utility networks (*GES_Urzadz*) [9]. It will be described later, when discussing the specific values added from the source code lists to the target ones in the process of increasing their compatibility.

CODE LISTS IN THE TECHNICAL GUIDELINES' DATA MODEL

The code lists presented within the Technical Guidelines [1] can be examined in a similar way. In case of a basic data model the subject of interest will be several code lists describing kinds of appurtenances linked to a specific type of network[2] as well as the code list concerning sewer water type. For the extended data model there are also code lists describing pole types and manhole cover shapes [1].

When it comes to the sewer water type code list, its values are as follows [1]:

- combined,
- reclaimed,
- sanitary,
- storm.

In turn, the types of poles and manhole cover shapes the extended data model proposes such values to be used [1]:

1. For pole types:
   - H-frame pole[3],
   - other,
   - standard,
   - street light,
   - tower,
   - very high voltage.
2. For manhole cover shapes:
   - circle,
   - composite,
   - grid,
   - other,
   - rectangle,
   - square.

At the same time, the basic data model leaves the ability to expand the code lists, especially those related to the appurtenances. That is because the proposed values could not be comprehensive when compared to the real needs dictated by a specific appurtenance type actual used within a network [1].

PROCESS OF DEFINING THE DIFFERENCES

The process of defining the differences between the code lists in both data models shall be considered in two steps. The first one lies in analyzing and comparing the values provided within the GESUT data model [9] and the basic data model presented in the Technical Guidelines [1]. If a specific source value does not have any counterpart, it is needed to repeat the process of comparing it, but with a reference to the INSPIRE extended data model [1]. After realizing that there is also a lack of a corresponding value, it may be accepted that it is necessary to add a new appropriate one to the code list.

As said in the introduction chapter, such attempt has been used with success by Geobid Ltd. when fulfilling tasks related to their engagement in a European project called GeoSmartCity. In the frames of the project the INSPIRE data model [1] was modified by any given means in order to make it better adjusted for each of the project partners. A part of this activity was to increase the compatibility between the source code lists taken from the Polish Decree [9] and the target ones used within the GeoSmartCity database. It is worth mentioning, because of the fact that in that case there is also a third step of the whole process which should be applied after the previous two were accomplished.

Since there are partners in the project who come from cities all across Europe, there has been worked out a way in which they were able to simultaneously express their needs related to the code lists potential expansion. To this end, an appropriate shared Excel spreadsheet was prepared, which can be edited by every one of the engaged parties. After the comparison of both models was made and it was stated

[2] When it comes to the data model provided in the Technical Guidelines [1] every type of network has its own code list of appurtenances, different from others

[3] A pole of this type is linked by a crossbar perpendicular to two vertical masts, whilst the A-frame pole in Polish pole type code list is combined of two masts joined at the top.

that there is a need of adding new value to a code list, there was an obligation to check if none of the other partners has already provided exactly the same value to the spreadsheet as the one demanded by the Polish company. If it has, the process was ended by inserting the Geobid's signature below such entry as of another company which expects this very value to be added. Otherwise, one should provide the value in the spreadsheet with respect to the appropriate rules. In a consequence the spreadsheet fulfilled with the proposed values will serve as a uniform source intended to help in extending all the code lists mentioned by the project engaged parties. The described attempt to evening the differences ensures that the whole process will be conducted simultaneously, without any obstacles and in a precise way what can be considered as beneficial and worth of reusing in the future.

THE ANALYSIS OF POSSIBILITIES IN CONFORMING THE CODE LISTS VALUES

The analysis of possibilities in conforming the values of individual code lists comes down to determining which one of them present in the GESUT data model can be added to the relevant code lists existent in the INSPIRE one. It has to be taken care of that, whenever possible, a short description of each newly added value should be provided so it will be easier to understand.

First of all, this task should have been done concerning the basic data model. After its analysis has been completed, it was stated that there is a possibility of adding an *industrial* value to the sewer water type code list. Due to this fact, every type of sewer water provided within the data model originating from the Polish decree [9] will have direct counterparts.

Moreover, there were demonstrated a number of appurtenances which should be added to the specific code lists in the target model. In this case, the proposed addition is related not to one, but five kinds of networks.

Concerning sewer networks, the related code list should be extended by a *septic tank* value along with its description.

In case of oil, gas and chemicals networks, inasmuch the INSPIRE data model provides such kind of network [1] instead of three different types presented in the relevant Polish legal act [9], the value intended to be added is a *valve.*

For electricity networks it was pointed out that the relevant code list should be expanded by the following values:

- cable joint,
- traffic signal,
- pole-mount transformer.

In the instance of water networks such values are proposed in a similar way:

- well,
- drilled well,
- valve,
- drinking fountain.

Lastly, for thermal networks, following appurtenances should be added:

- pump station,
- vent,
- valve.

Other appurtenances either have counterparts in values presented within the target model, or there is no option of adding them since they are not only an appurtenance, but rather a topographic object (e.g. gas pump existent in the GESUT data model [9]).

When it comes to the extended data model provided within the Technical Guidelines [1], there can be expanded the code list defining the pole types as well as the one related to manhole cover shapes.

In case of the first of them, following values can be added:

- A-pole,
- pylon.

The second one should be expanded by a value *oval*.

Noteworthy, that in the instance of the manhole cover shape code list originating from the Polish Decree [9] it is not only used to describe manhole covers, but also for some other appurtenances. Whereas the target data model, the INSPIRE one [1], a corresponding code list applies only to them. Nonetheless a correct binding acquired by adding proper values allows to fully harmonize data coming from any GESUT database at least in the area concerning this very type of appurtenance. Since – on average – such database stores numerous objects considered as manholes, the manhole cover shape code list mentioned above is of high importance.

SUMMARY

Despite the overall differences revealed [5] between data models originating from the old binding Polish Decree [8] concerning GESUT as well as BDOT databases and the one presented in the Technical Guidelines, not all of the described problems [5] were eliminated by entering into force the new Decree [9], which relates only to the GESUT databases. As an example can serve some of the code lists values which are still incompatible. This situation can potentially lead to serious hindrances if there is a need of passing a file including data exported out of the GESUT database to a party operating an INSPIRE compatible data model [1]. Geobid Ltd. suffered such problem when cooperating within the frameworks of the European project called GeoSmartCity, since there was a necessity to pass spatial data to other partners.

Disregarding the fact that this task could not be fulfilled directly because of the incompatibility of the very data model [5], a process of harmonizing data [6] to the target model was also interfered by differences between the code lists themselves. However, there is an opportunity of extending the target lists and such extension was made every time when it was possible regarding a character of a specific code list. Ultimately it resulted in an increased ease of harmonizing data since at least some of the code lists values in the Polish and INSPIRE data models were better correlated.

Unfortunately, analysis of the two models indicates that such operation can be conducted only on some of the code lists, especially those related to appurtenances and to a lesser extent to attributes describing pole types or manhole cover shapes.

CONCLUSION

With a view to a still increasing need for spatial data and emphasis on making way they are collected, processed and shared uniformed it has to be stated that all the questions related to the organization of spatial databases are of high importance. Especially in the context of an idea of interoperability expressed in the INSPIRE Directive [2] as well

as in its transposition to the Polish legal system [10]. Moreover, since the Directive [2] derives from a need of gathering spatial data in order to make environment management more efficient, this goal should be realized consequently by individual Member States so the European Union's policy will not be harmed.

Unfortunately, what was proved by this study, there are some inaccuracies between the Polish regulations, at least in case of the GESUT data model, and the ones presented in form of a data model for utility networks within the Technical Guidelines [1] issued to the INSPIRE Directive [2]. In fact, the latter one is provided just as the one best realizing all the regulations concerning the networks, however one is not obliged to use it instead of these parts where *Implementation Rules* states otherwise [1]. Nonetheless, what is shown by the experience gained when cooperating with foreign partners, European parties extensively implemented the discussed model directly with only slight modifications so it will meet local expectations. At the same time the Polish Decree on GESUT [9] provides a data model developed from scratch on the basis of local needs and merely similar to the INSPIRE one [1]. In a consequence, disregarding other potential discrepancies [5], even the code lists values are largely not compatible.

Therefore, when the data exchange is needed, a party engaged in the process on the Polish site is forced to first conduct a harmonization process of a database [6] which is even harder because of the discrepancies this paper presents and which make it impossible to correlate some of the code lists values directly. This can lead to a distortion or loss of important information about objects stored in the database.

What this study demonstrates is that there is a possibility of some adjustment for the code lists provided by the INSPIRE data model [1], which leads them to an increase of correlation with the Polish counterparts. In a consequence, if there was a need of data harmonization, it would be facilitated to a certain extent as well as decreased would be the risk of losing any attribute information. With a view to the paramount idea connected to this study by means of the INSPIRE Directive, thus the idea of protecting the environment and of its sustainable development, it has to be stated that also for spatial data managed by mining industry or just concerning mining areas there could occur a necessity of harmonizing them to a target data model other than used thus far. The presented solution for expanding the code lists by adding new values in a target model makes it possible to prepare for such task in the most exact way, what can provide, for instance, to make it easier to cooperate with foreign partners.

Furthermore, the possibility of harmonizing data gathered in GIS databases managed by mining plants will affect in a positive way any process of investigating requests for starting new research or scientific projects related directly to the object branch of the industry. What has been shown by the experience gained so far, the question of a straitened data exchange often enough led to rejections in such cases, which has a very negative impact on the potential for development in the mining industry.

REFERENCES

[1] F. Brönnimann et al*. D2.8.III.6 INSPIRE Data Specification on Utility and Government Services – Technical Guidelines*, European Commission Joint Research Centre, 2013.
[2] *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)*. Official Journal of the European Union L108/1, 2007.
[3] D. Mercer. *XML kurs podstawowy*. Kraków: Edition 2000, 2002.
[4] Z. Parzyński and A. Chojka. *Infrastruktura informacji przestrzennej w UML*. Warszawa: Geodeta, 2013.
[5] P. Sikora and A. Zygmuniak. „Porównanie użytkowe modelu baz danych GESUT do modelu przedstawionego w Wytycznych Technicznych wydanych do Dyrektywy INSPIRE", in *MOTROL*, vol. 17(1), 2015, pp. 115-120.
[6] V. Sokoła-Szewioła and A. Zygmuniak. „Proces transformacji struktury plików zawierających dane geoprzestrzenne za pomocą aplikacji HALE", in *Mining of sustainable development Conf.*, Gliwice, Poland, 2015.
[7] M. Śmiałek. *Zrozumieć UML 2.0. Metody modelowania obiektowego*. Gliwice: Helion, 2005.
[8] *Rozporządzenie Ministra Administracji i Cyfryzacji z dnia 12 lutego 2013 r. w sprawie bazy danych geodezyjnej ewidencji sieci uzbrojenia terenu, bazy danych obiektów topograficznych oraz mapy zasadniczej*, Dz. U. 2013 poz. 383, 2013.
[9] *Rozporządzenie Ministra Administracji i Cyfryzacji z dnia 21 października 2015 r. w sprawie powiatowej bazy GESUT i krajowej bazy GESUT,* Dz. U. 2015 poz. 1938, 2015.
[10] *Ustawa z dnia 4 marca 2010 r. o infrastrukturze informacji przestrzennej*, Dz. U. 2010 nr 76 poz. 489, 2010.
[11] *Ustawa z dnia 9 czerwca 2011 r. – Prawo geologiczne i górnicze*, Dz. U. 2011 nr 163 poz. 981, 2011.
[12] *Zasady oceny możliwości prowadzenia podziemnej eksploatacji górniczej z uwagi na ochronę obiektów budowlanych*, GIG Standard 12, 2000.

mgr inż. Andrzej Zygmuniak, dr hab. inż. Violetta Sokoła-Szewioła, prof. Pol. Śl.
Silesian University of Technology, Faculty of Mining and Geology
ul. Akademicka 2A, 44-100 Gliwice, POLAND
e-mail: andrzej.zygmuniak@polsl.pl
violetta.sokola-szewiola@polsl.pl