

КАДАСТР І МОНІТОРИНГ ЗЕМЕЛЬ

CADASTRE AND LAND MONITORING

UDC 004.91

R. PERESOLIAK

Land management and cadastre department, Uzghhorod national university, 14, Universytetska, Uzghhorod, 88000, Ukraine, e-mail: roman.peresoliak@gmail.com.

ERRORS IN CADASTRAL DATA EXCHANGE FILES – CASE STUDY

<https://doi.org/10.23939/istcgcap2017.02.066>

Purpose. Given the importance of the land cadastre, the quality of the cadastral exchange file plays an extremely important role, both for the land cadastre and for all other cadastres as a whole. Despite the widespread use of cadastral XML exchange file formats in the field of land management, the lack of a sufficient understanding of the conceptual peculiarities of this thesaurus does not allow us to fully disclose its potential or functionally improve the cadastral exchange file and cadastral system of Ukraine. The materials of this article outline widely known files in the field of land management from a new point of view, such as the markup language. This article gives a detailed analysis and shows the disadvantages of the cadastral exchange file structure elements performed using XML technology. **Result.** Changing the point of view of the cadastral file offered by this article improves the mechanism for making changes to the structure of the cadastral exchange file and to directly identify it. Based on the analysis of the disadvantages of the existing station cadastral file sharing, a new design cadastral file sharing system was developed. In this work, manifestations of unproduction of the structure of the cadastral file and its uncontrollable dynamic changes were demonstrated. At the time of writing this article version 7 was available for the language determinant. Despite the fact that this determinant is developed solely for internal use in the Digital environment, but with the absence of the same determinant specified by law leads to the arbitrary interpretation of the fidelity of the structure of the cadastral file implemented by software tools, such as Digital and others. **Scientific innovation.** The obtained results give the opportunity to use them as a basis for further improvement of cadastral exchange files, as well as to eliminate existing disadvantages and differences with regulatory acts for cadastral file aspects and structure definition.

Key words: cadastral exchange file; xml; markup language; cadastre.

Introduction

Given the importance of the land cadastre, the quality of the cadastral exchange file plays an extremely important role, both for the land cadastre and for all other cadastres as the whole. To evaluate the cadastral exchange file which was made using XML technology and used in the land cadastre, the structural elements of the file have been used in this paper and are determined in the normative field [Vumohy do struktury zmistu, Pro zemleustrii].

The single name (XML) contains simultaneously many related values, which can be confusing. Regardless the name itself, XML is not a markup language: it is a set of rules for creating markup languages [Eryk Rey, 2001].

An important feature of the new cadastral file is that it is the same as other formats executed using XML, from which commonly include:

- SVG (Scalable Vector Graphics) [SVG, 2011];
- GML (Geography Markup Language) [GML, 2013];

- KML (Keyhole Markup Language) [David Burggraf, 2015];

- SLD (Styled Layer Descriptor) [Markus Lupp, 2007].

It is not just the format of the document but rather it is the entire markup language. Consequently, by analogy, it can be clearly stated that the cadastral exchange file is the result of the generated document executed using a cached exchange file markup language. Since any grammar language that uses XML must have a file schema, that is, a set of rules, then the language of the cached file counts must be owned by it. Digital contains a schema file called “IN4XML Schema.xsd”. This is exactly a file for describing the rules of the language of cadastral exchange, since it is precisely by this that we, in our opinion, execute the correctness of filling a cadastral file for the exchange.

Also it should be mentioned that XML represents a whole set of technologies that work closely with each other. In our opinion, it is the

excellent workflow and the number of these technologies that allows storing XML in the first positions, since it is not the only representative of the markup languages, which can be attributed to JSON Schema [JSON Schema]. Checking and generating new, simpler documents such as JSON or YAML, is gaining popularity [RX Schema], but the most advanced set of technologies in XML tools leave it at the forefront.

The work of filling and creation of the cadastral exchange file was performed using Digitals [Digitals] 2016 software.

The use of only one software and its one version, in spite of other analogues available on the market [GIS6] and other versions in particular, is due to the fact that cadastral exchange files generated by Digitals, as well as other environments, are suitable and correct for the Ukrainian automated system of the state land cadastre. Thus disadvantages that can be detected will be inherent to the entire system.

Main part

The system of the cadastral exchange file shortcoming detection is built on the basis of the categories that describe a certain characteristic of the node of cadastral file structure of their group.

At the end of the study the following features were identified (grouped depending on the type of the characteristic), which have become criteria for assessing the quality of the cadastral exchange file:

- Features of the environment(Digitals)
 - 101 (an element exists when there is at least one element of the first level)
- data that does not apply to the person who creates the file;
 - 102 (not subordinate information);
 - 103 (metadata);
- pointer to the application of the attribute;
 - 104 (the unit name depends on the option of the environment in which the document was generated);
 - 126 (the environment option replace the block that will be present in the document);
- grouping pointer;
 - 121 (block is the element of the list);
- pointer that item name should be changed
 - 105 (the same name for the different destination blocks, elements or nodes);

- 108 (the name of the block does not correspond to its content);
- 109 (the name of the block is too simplistic);
- 110 (the name does not correspond to the unique style of the names);
- 111 (the name contains transliteration);
- 113 (the name contains name of the parent element);
- incompatibility with the regulatory act;
 - 107 (the name does not correspond to the name in the regulatory act);
- complex type pointer;
 - 114 (the block and its components are repeated several times in the document);
 - 106 (block content repeated several times and does not depend on the name of the block);
 - 119 (block contains signs of a typical element);
- additional element characteristics;
 - 112 (block is a link to another object);
 - 116 (the block content depends on the option of the software which generated the document);
 - 122 (block contains an exhaustive list of values);
 - 123 (block value represented by code);
 - 124 (the block value contains spaces);
 - 125 (the block can be represented by list of values);
- document complication;
 - 118 (redundant element);
 - 129 (absence of alternatives).

It should be noted that the above classification does not aim to highlight exclusively negative moments in the formation of a cadastral exchange file. The classification also highlighted the formation features, which make it possible to better describe the peculiarities of the internal structure of the exchange file made with the help of XML technology.

Regarding to the list above, it should be noted that cadastral exchange file use topological model of the geometry representation instead of “spaghetti” and these aspects of cadastral data were not included in the classification. However, importance topology checking goes out of the scope of this publication. Consistency should use system level instead of exchange file level. It is

confirmed by several sources [Zygmunt. M., 2015; Siejka, M, 2014]

In the context of this publication, it is intended to cover exclusively the negative features of the cadastral exchange file structure from the list above, which, in the author's opinion, most strongly demonstrate the imperfection and non-standardization of the electronic document as such.

Structural disadvantages of the cadastral file structure.

The same name is used for the different destination blocks, elements or nodes

This characteristic allows to outline the inability to find some of the information in it using features of the exchange file, since the name of the node is the identifier of the type of data it contains.

```
<!--No. 1-->
<InfoLandWork>
  <Executor/>
</InfoLandWork>

<!--No. 2-->
<Executor>
  <Executor/>
</Executor>

<!--No. 3-->
<ParcelMetricInfo>
  <Area/>
</ParcelMetricInfo>

<!--No. 4-->
<LeaseInfo>
  <Area/>
</LeaseInfo>
```

Block 1. The same name of the different structural elements

For example, a node of Area (Area), from block 1, in the third case is a complex structured element with a number of sub elements and in the case (No. 4) -- an element that contains the numerical value of the area.

The same applies to the block containing information of the performer for works (Executor) in the first case. The element describes the organization of the executor and in case (No. 2) the person-executor who works in the organization-executor.

The name of the block is too simplistic

Given the presence of items names containing full and extended node name value, such as "UkrainianCadastralExchangeFile", the abbreviation shown in block 2, "RegName" does not use

the full name of the element(node), which in turn leads to a lower understanding of the information load of the element.

```
<ProprietorInfo>
  <Privilege>
    <RestrictionInfo>
      <RegName/>
    </RestrictionInfo>
  </Privilege>
</ProprietorInfo>
```

Block 2. Simplistic element name

The name does not correspond to the unique style of the names

This feature, as well as the previous, can be attributed to the same class of features and is associated with names.

The name does not correspond to the unique style of the names

This feature, as well as the previous, can be attributed to the same class of features and is associated with names.

```
<!--№1-->
<Point>
  <UIDP/>
</Point>

<!--№2-->
<PL>
  <ULID/>
</PL>
```

Block 3. Different way to element naming

Block 3 displays the divergence in the naming style "UIDP" (Unique Identifier Point) and "ULID" (Unique Line Identifier).

Block content repeated several times and not depending on the name of the block

This characteristic describes elements of the structure that could be defined as typical, but elements of this class change the name of the highest element, without changing its content.

```
<FullName>
  <LastName>
  <FirstName>
  <MiddleName>
</FullName>

<DKZRHead>
  <LastName>
  <FirstName>
  <MiddleName>
</DKZRHead>
```

Block 4. Typical element with different name

Such application of typical elements leads to the fact that it is impossible to display or find, as an example all persons names regardless of their status. Also, this approach leads to the fact that essentially identical elements appear in the cadastral exchange file with different names, that is, there is no confidence where exactly the name can be written, in this case.

The block value contains spaces

In the case of complex descriptive information presentation or data commentary, the presence of spaces is an integral part. However, the application of this approach may also leads to the fact that the information can potentially be chopped into a substructure – which does not use this modification.

```
<Address>
...
  <Settlement/>
...
</Address>
```

Block 5. Node value could contains spaces

In particular, this example demonstrates that the element “Settlement” contains information about the type of settlement and its name. This, in turn, leads to the fact that it becomes impossible to find data exclusively based on the type of settlement. The form of the element value of this type is not defined and determined by the operator and as a result could be presented in a different form (city Uzghorod, c. Uzghorod, Uzghorod).

The name does not correspond to the name in regulatory act

The characteristic describes the phenomenon of the contradiction of the regulatory act with the exchange file. In spite of the obvious simplicity of the features presented in block 3, the neglect of the naming template leads to confusion in the future. A striking example of this is the different names of the same elements in the regulatory act and exchange file, in particular:

- cadastral exchange file contains “UIDP”;
- the decision to approve the procedure for conducting the State Land Cadastre contains the mention “UPID”.

Given the above, it turns out that the cadastral file does not correspond to the regulatory act.

Absence of alternatives

Given the presence of data, the that are fixed and determined by the structure of the cadastral file

in some cases, it should be possible to record data with some variability.

```
<xsd:all>
  <xsd:element name="MoneyRent" type="xsd:double" >
    <xsd:annotation>
      <xsd:documentation>The size of the rent for a land plot
or part thereof in cash</xsd:documentation>
      <xsd:appinfo>Money form</xsd:appinfo>
    </xsd:annotation>
  </xsd:element>
  <xsd:element name="OtherRent" type="xsd:string"
minOccurs="0" maxOccurs="1">
    <xsd:annotation>
      <xsd:documentation>The size of the rent for a land plot
or part thereof, leased in other forms of
payment</xsd:documentation>
      <xsd:appinfo>Another form</xsd:appinfo>
    </xsd:annotation>
  </xsd:element>
</xsd:all>
```

Block 6. Absence of alternatives

As shown in block 6, the cadastral file does not provide for the possibility of absence of information regarding the size of the payment form, in the case of rent payment in a form other than money, that is, the field which must contain information about the payment in cash equivalent must be present, but not contain information.

Block value represented by code

```
<Address>
...
  <Citizenship/>
...
</Address>
```

Block 7. Code representation

The application of this method leads to the fact that the information contained in the cadastral exchange file is not independent, and therefore the file itself is not independent too. This means that without the use of additional documents, interpreting the information inside exchange file is impossible.

Block contains signs of a typical element

The characteristic is an attempt to draw analogies or parallels between existing elements and helps to indicate the duplication of the structure with small modifications.

```

<Executor>
  <ExecutorName>
    <LastName/>
    <FirstName/>
    <MiddleName/>
  </ExecutorName>
  <ExecutorPosition/>
...
</Executor>

<TypeAction>
  <Action>
    <ExecutorName>
      <LastName/>
      <FirstName/>
      <MiddleName/>
    </ExecutorName>
    <ExecutorPosition/>
    <DateApproved/>
  </TypeAction>

```

Block 8. Structurally similar elements

Elements typicality play an important role in data structure. Given this, data should potentially have the same structure. This will make it easier to support markup language.

The name contains transliteration

Characteristics describe the use of the mechanism of transliteration, in contrast to the use of an equivalent in a foreign language.

```

<LegalEntityInfo>
  <EDRPOU/>
</LegalEntityInfo>

```

Block 9. Transliterated elements names

The disadvantage of this method is shown in the example of block 4 where DKZRHead, originating from the Head of the State Land Cadastre (Holova Derzhavnoho Kadastru Zemelnykh Resursiv), when the organization name changes, the name of the block will lose its relevance. Therefore, this approach strongly binds structural blocks of information to the names of organizations.

The name of the block does not correspond to its content

The names of the elements of the cadastral file structure make it possible to better understand the completeness of the block. A name that does not disclose content or misleads understanding of the content and is in a lower quality of work with an exchange file at the file level.

Despite the fact that the item, bearing the name, also carries information on the name of the notary office, and the item directly contains information about the details of the legal entity.

```

<StateActInfo>
  <MarkInfo>
    <NotaryMark>
      <NotaryName/>
    </NotaryMark>
  </MarkInfo>
</StateActInfo>

```

Block 10. Name inconsistency to its content

Block is the element of the list.

In spite of the fact that this feature, in general, is the norm, since a large amount of data can be represented by the list, however, infologically, this leads to the fact that the data contained in the exchange file can be treated differently.

```

<CadastralQuarterInfo>
  <Parcels>
    <ParcelInfo/>
    <ParcelInfo/>
    ...
  </Parcels>
</CadastralQuarterInfo>

```

Block 11. Several parcels within one quarter

In other words, the record given in block 11 is considered correct and means that within a quarter several parcels may be saved by means of a cadastral file, which in turn leads to the fact that cadastral exchange file is a prototype format for saving general spatial information and not a data format for displaying the state of the object of land cadastre – land parcel.

The unit name depends on the option of the environment in which the document was generated.

Despite the name of the characteristic, this feature leads to the fact that the structure is always subject to change, as an example (Block 11) in the structure of the exchange file, either “EDRPOU” or “TaxNumber” depending if it is an entrepreneur or an enterprise. That is, in the case of looking for data about the executor code, we first need to have information about the type of person who created the file.

The block 11 also demonstrates the fact that the elements may be empty, however, they may carry certain content. That is, there is a conflict as to how the elements should be interpreted in terms of their completeness. Since in one case one should pay attention to the internal content of the node and in the other only to its name.

```

<Executor>
...<EDRPOU/> або <TaxNumber/>
</Executor>

<Metricinfo>
  <HeightSystem>
    <Baltic/> або <Baltic77/> або <Other/>
  </HeightSystem>
</Metricinfo>

```

Block 11. Optional structure variability

The examples presented above aim at demonstrating the problem of the uncertainty of the internal structure of the cadastral exchange file at least on the technical level, not to mention the normative level of certainty.

It is worth knowing that at the time of the research several variants of the language determinant of the cadastral exchange file were available from the Digitals environment:

- In4XmlSchema.xsd version 0.6;
- XmlSchema.xsd version 0.7.

The version information is explicitly indicated in the file schema:

```

<xsd:schema
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified" version="0.7">

```

Block 12. Language determinant version

However, version 0.7 does not possess a large number of necessary elements determined by law, and also implements several new elements, for example:

- LegalEntityInfo2
- NaturalPersonInfo2

Therefore, the basis for the comparison was taken using version 0.6 of the determinant of the markup language. Despite the older version, its generation does not play any role in the research results, as well as software, since it is aimed at demonstrating the change in the structure of the cadastral exchange file in time and its impracticability.

Conclusions

The use of XML technology for land management purposes is, of course, the first step towards effective data use. However, in order to realize the opportunities that are already available in XML technology at an adequate level, the existing cadastral file requires large-scale changes. In this work, manifestations of unproduction of the

structure of the cadastral file and its uncontrollable dynamic changes were demonstrated.

At the time of writing the article version 7 was available as the language determinant. Despite the fact that this whole determinant is developed solely for internal use of the Digitals environment, but the fact of the absence of the same determinant specified by law leads to the arbitrary interpretation of the fidelity of the structure of the cadastral file implemented by software tools, such as Digitals and others.

The authors see the need for the implementation of the mechanism of the publication of the determinant of language. With any change in the legal field that affects the content of the cadastral exchange file (the transition from licenses to certificates has not been fixed by any regulatory document that introduced specific changes in the structure of the cadastral file). In view of this, as well as the inevitability of further changes in the structure of the cadastral exchange file we consider it necessary:

- at the law level to determine no other mechanism is used for changing the structure of the cadastral file sharing, except for the publication of the *.xsd file-determinator;
- to take note of the comments identified by the materials of this publication.

This allows regulation of the structure of the cadastral exchange file and to prevent the unauthorized adaptation of the file at the software level. Also, this mechanism will make it possible prevent collisions in the field as it is now. In spite of the fact that the structure is determined to a certain extent by the law of Ukraine "About the state land cadastre", however, the presentation method proposed in this normative document relates solely to the legal aspect of this document, and the technical side is not considered, which in turn, in our opinion, has resulted in a number of problems.

REFERENCES

- Clemens Portele, ed. Geography Markup Language (GML) – Extended schemas and encoding rules. 2012, p. 91. url: <http://www.opengeospatial.org/standards/gml>
- David Burggraf, ed. OGC KML 2.3. 2015, c. 266. url: <http://docs.opengeospatial.org/is/12-007r2/12-007r2.html>
- Digitals. [Electronic resource]. URL: <http://www.vinmap.net/>.
- Digitals. IN4XmlSchema.xsd
- Digitals. XmlSchema.xsd

- Eryk Rey. Izuchaem XML. Ed. by Halunov A. Sankt-Peterburh: Symvol-Plyus, 2001. P. 408.
- GIS6. [Electronic resource]. URL: <http://www.gis.org.ua/>.
- JSON Schema [online]. url: <http://json-schema.org/>
- Markus Lupp, ed. Styled Layer Descriptor Implementation Specification. 2007, c. 54. url: <http://www.opengeospatial.org/standards/sld>
- Pro zatverdzhennia Vumoh do struktury zmistu ta taformatu oformlennia rezultativ robit iz zemleustroi v elektronni vyhliadi (obminnoho failu). Hakaz derzhavnogo komitetu iz zemelnykh resursiv. [Electronic resource]. URL: <http://zakon3.rada.gov.ua/laws/show/z0157-10>.
- Pro zemleustrii. Zakon Ukrainy [Electronic resource]. URL: <http://zakon3.rada.gov.ua/laws/show/858-15>.
- RX Schema [online]. url: [http://rx.codesimply.com/Scalable Vector Graphics \(SVG\). 1.1 \(Second Edition\). 2011, p. 826. url: \[https://www.w3.org/TR/SVG11/REC-SVG11-20110816.pdf.\\(26\\)\]\(https://www.w3.org/TR/SVG11/REC-SVG11-20110816.pdf.\(26\)](http://rx.codesimply.com/Scalable Vector Graphics (SVG). 1.1 (Second Edition). 2011, p. 826. url: https://www.w3.org/TR/SVG11/REC-SVG11-20110816.pdf.(26)
- Siejka, M, Ślusarski, M., & Zygmunt, M. (2014). Verification technology for topological errors in official databases with case study in Poland. Survey Review, 46(334), 50-57. url: <http://doi.org/10.1179/1752270613Y.0000000054>
- Zygmunt, M, Siejka, M., Slusarski, M., Siejka, Z, Piech, L, & Bacior, S. (2015). Database inconsistency errors correction, on example of LPIS databases in Poland. Survey Review, (47), 256—264 url: <http://doi.org/10.1179/1752270614Y.0000000134>

Р. ПЕРЕСОЛЯК

Кафедра землевпорядкування та кадастру, Ужгородський національний університет, 14, вул. Університетська, Ужгород, 88000, Україна, ел. пошта: roman.peresoliak@gmail.com.

ПОМИЛКИ У КАДАСТРОВИХ ФАЙЛАХ ОБМІНУ — ВИВЧЕННЯ ПРОБЛЕМИ

Мета. З огляду на важливість земельного кадастру якість кадастрового обміну має надзвичайно важливу роль як для земельного кадастру, так і для всіх інших кадастрів загалом. Незважаючи на широкий вжиток кадастрових файлів обміну формату XML у сфері землеустрою, відсутність достатнього розуміння концептуальних особливостей цієї технології не дає змоги повною мірою розкрити її потенціал і, як наслідок, функціонально покращити кадастровий файл обміну та систему кадастрів України. Матеріали цієї статті дають змогу окреслити широко відомі файли в сфері землеустрою із нової точки зору, як мову розмітки. Детально проаналізовано та висвітлено недоліки елементів структури кадастрового файлу обміну, виконаного за допомогою технології XML. **Результат.** Зміна погляду на кадастровий файл обміну, запропонована цією статтею, дає змогу удосконалити механізм внесення зміни до структури кадастрового файлу обміну та безпосередньо визначити його. Базуючись на аналізі недоліків наявного стану кадастрового файлу обміну, розроблено новий дизайн кадастрового файлу обміну. В цій роботі продемонстровано прояви неопрацьованості структури кадастрового файлу та його ніким нерегульовані динамічні зміни. На момент написання статті доступною була версія визначника мови кадастрового файлу обміну 7 версії. Незважаючи на те, що це всього визначник, розроблений винятково для внутрішнього застосування середовищем Digitals, однак, сам факт відсутності такого самого визначника визначеного законом приводить до самовільного трактування вірності будови кадастрового файлу реалізованого засобами програмного забезпечення, як Digitals, так й іншого. **Наукова новизна.** Отримані результати дають змогу використати їх як базис для подальшого вдосконалення кадастрового файлу обміну, а також усунути наявні недоліки та розбіжності із нормативно-правовими аспектами визначення структури кадастрових файлів.

Ключові слова: кадастровий файл обміну; xml; мова розмітки; кадастр.

Received 14.09.2017