

Original scientific paper

UDC 528.9:[007:912]:004]
<https://doi.org/10.2298/GSGD2402427M>

Received: October 19, 2024

Corrected: November 28, 2024

Accepted: December 10, 2024

Viktor Marković¹*

** Military Geographical Institute - "General Stevan Bošković", Belgrade, Serbia*

DATA MODEL FOR THE CARTOGRAPHIC REPRESENTATION OF TOPONYMS ON TOPOGRAPHIC MAPS PUBLISHED BY THE MGI

Abstract: An essential aspect of creating topographic maps involves standardization and data modelling for properly visualised toponyms. The significance of this aspect arises from the fact that maps, aside from representing a graphical depiction of geographical space, serve as a medium for conveying specific information and toponyms that must be presented. A significant effort has been directed towards developing a methodology for the standardization of toponyms on the maps of the Military Geographical Institute "General Stevan Bošković" (MGI), which meets the strict criteria of cartographic representation. This paper discusses the methodology and process of creating a database of settlement names for cartographic representation at a scale of 1:25,000 (TM25), as well as in derived scales utilizing a ladder approach system. The focus is on the standardization and enhancement of the cartographic depiction of toponyms using digital techniques. The created database has broad applications not only in cartography but also within the GIS context, enabling complex spatial analyses. Furthermore, this paper explores the advantages and disadvantages of the created toponym data model, its implementation in the production system for creating topographic maps at MGI, and potential opportunities for improvement. The methodology presented in this paper has allowed a relatively smooth transition from analogue to digital representation of toponyms while maintaining the visual identity and functionality of classical topographic maps.

Key words: toponyms, database, GIS, data model, topographic maps

¹ viktor_bre@yahoo.com (corresponding author)
Viktor Marković (<https://orcid.org/0000-0002-1085-8575>)

Introduction

Geographical names represent a key component of any cartographic depiction of geographic space. They are not merely labels for settlements but often play a crucial role in spatial orientation, causal analysis, and the understanding of relationships between humans and the natural environment (see Jordan 2020). In classical cartography, geographical names were depicted following various criteria, such as spatial extent, quantitative characteristics, administrative significance, and historical and cultural context. The transition from analogue to digital cartography has introduced new challenges. While cartographers had freedom of interpretation on analogue maps, GIS (Geographical Information Systems) requires precisely defined rules and standards for data modelling and visualization. This transition not only enabled more precise spatial analyses but also paved the way for more efficient standardization and uniformity in the representation of toponyms. Toponyms are now part of complex databases allowing advanced spatial queries and analyses, making them a significant resource for geographic systems and cartographic applications.

The standardization of toponyms, particularly in the context of their representation in GIS, has greatly contributed to the accuracy and consistency of cartographic depictions. National efforts to standardize toponyms, often in line with international guidelines promoted by organizations like the UNGEGN (UNGEEN, 2015), are crucial for ensuring greater interoperability and accuracy in various cartographic systems (e.g., Perdana & Ostermann, 2018). On the other hand, some researchers highlight the significance of toponyms as cultural symbols, complicating their depiction on digital maps, where they serve both as spatial objects and symbols of cultural identity (e.g. Rose-Redwood et al., 2010).

In Serbia, a significant contribution to the research of toponyms in the GIS context has been made by Borisov et al. (2010) and Tatomirović et al. (2016) who analyzed the development and application of toponym standardization in the national cartographic system. Their works emphasize the importance of a unified representation of toponyms, especially in the context of state and military geographic information systems, such as the one used by the MGI.

This paper presents a toponym modelling method employed by the MGI, focusing on the standardization of toponyms at a scale of 1:25,000 and their integration into the ladder approach of topographic map creation, a practice in use at the MGI since 2012. The presented method allows for a uniform representation of toponyms across different scales, maintaining visual continuity with older analogue maps created during the 20th and early 21st centuries.

Toponym modelling in the GIS environment, as implemented by the MGI, requires strict standardization to achieve visual consistency and efficiency in spatial analyses. This standardization is necessary to retain the traditional appearance of maps while enabling their use in GIS applications.

Materials and methods

To consolidate criteria used in analogue cartography for the representation of toponyms, multiple data sources were utilized. The primary data source consisted of analogue map

editions at a 1:25,000 scale, while analogue maps at scales 1:50,000 and 1:100,000 were employed as auxiliary sources. Based on these maps, along with written instructions for creating geographical names on topographic maps (MGI, 1982), criteria for the depiction of toponyms were synthesized. These maps were scanned, georeferenced, and used within the topographic map production system as services. From these maps, the initial content of toponyms is generated and stored in a database.

The database used in the production system of topographic maps at the MGI is an SQL Enterprise database. This database was created with a database management system (DBMS) integrated into the software package ArcMAP 10.3.1, which was utilized in the research process (ESRI, 2015). The primary function of this database is cartographic production, and it represents a digital cartographic model (DCM) used to create vector maps at a 1:25,000 scale, which serves as a visualized representation of a portion of the database. In addition to its primary function, the database also contains numerous other alphanumeric data that may not need to be visualized on the map but play a crucial role in spatial analyses. These alphanumeric data are tied to individual objects in the database, serving to further describe them or represent their attributes.

When assigning names to certain features on the map from the attributes stored in the database, whether they are geographic names or numerical-alphabetical data that describe or designate a feature, annotations are used. Unlike labels, annotations allow a high degree of manual positional manipulation. In this case, they are created as dependent object classes whose syntactic and semantic context is derived from the vectorized content that needs to be designated or named. As they are dependent but separate object classes, the connection between them and the original vector data is realized through various types of relations or links (ESRI, 2015).

The mentioned cartographic database at a 1:25,000 scale represents a structured, taxonomic GIS database built using ArcGIS software. In this system, individual object classes are defined based on their semantic affiliation to corresponding data sets. Within these object classes, data are further organized into subtypes, facilitating the easy selection of data by type, category, geometry, and internally adopted unique alphanumeric identifiers down to the level of entities.

Technically, the database can create domains, which play a key role in defining individual elements at the lowest level of the subtypes. Within the utilized software package, two types of domains can be created: range domains and coded value domains (ESRI, 2015). The creation and assignment of these domains also contribute to quality control of the created content. By using range domains, the possibility of gross errors in the entry of numerical values is reduced. Coded value domains ensure that only predefined values conditioned by the data model can be entered, ensuring uniformity of the created content while preventing free interpretation, which could complicate or undermine the accuracy of spatial analyses or lead to misinterpretation of geographical phenomena.

Data within the database are organized into datasets. Datasets represent the highest level of classification for related elements based on their nature and semantic interconnectedness. Ten datasets were created, relevant for generating vector-based TM25 maps, and their names indicate the types of phenomena they store. These datasets act as containers for homogeneous content that shares common properties and characteristics defined within the parent dataset.

The third level in the data organization is represented by object classes created within the datasets. These object classes are not only organized by their semantic affiliation but also by their geometry and how they will be represented at a 1:25,000 scale. The point, line, and polygon object classes are present within the datasets.

The fourth level in the data organization is represented by subtypes. This level of taxonomic data organization allows for the differentiation of individual elements at the entity level, which is defined by a unique numerical identifier within the corresponding object class, created according to the defined data model. All subtypes share the same attribute structure defined within the corresponding object class. This hierarchical, taxonomic data organization, from database to subtypes, is a functionality provided by the ArcGIS software platform, enabling efficient data management and retrieval within the GIS environment.

Geographical names are stored within a dataset named „Geografski nazivi“. The elements that contain information about geographical names are represented by three semantic units corresponding to three object classes within the database. In this case, the traditional classification of geographical names was respected: names of settlements - toponyms, names of regions and areas - choronyms, and names of mountains, mountain ranges, hills, ridges, etc. - oronyms (Peterca et al., 1974).

Toponyms represent anthropogenic features and are of significant importance in visualizing geographic space, particularly in cartographic visualization. Specifically, toponyms designate inhabited areas that can be represented by polygons, depicting named regions at a 1:25,000 scale. Accurately defining the boundaries of these areas, considering the data sources used, is often impractical. Therefore, generalization is applied, with polygon elements being represented by a point, which, according to internally established rules, should represent the geometric centre of the feature it defines. Points representing toponyms have attributes that describe the features they geometrically represent. These points are visually invisible, representing points of specific dimensions without colour. This approach was adopted to facilitate spatial analyses and provide greater flexibility in creating dependent annotations, which in this context are crucial for the spatial interpretation of the represented feature. Another reason for this approach is interoperability. The software used in this research supports direct work with annotations, which were employed for visualizing toponyms and other categories of geographical names. To ensure that the created data related to toponyms could be utilized on other software platforms that do not support direct work with annotations, or for which cartographic production is not the primary function, all necessary information about the toponym category is stored within a point-type object class. This allows for the rapid generation of labels based solely on attributes within the point-type object class, without the need for annotations. This approach can be seen as a form of data redundancy since the same data is present both in the point-type object class as an attribute and in the dependent annotation class as a visualized attribute. Although both identical attributes refer to the same feature, the difference lies in their creation, visualization, and usage.

Data modelling

In comparison to the analogue cartographic representation of toponyms, the primary visual difference in this specific case is the choice of font used for the visualization of toponyms within the database environment. For interoperability purposes of vector-

generated content, standard Windows fonts were selected, as they most closely match the characters used on analogue maps.

The next step in the modelling of toponyms involved the categorization within the toponym object class. During the research, it was determined that there are multiple categories of toponyms, each possessing certain specificities that need to be separately considered during the data modelling process.

The first group of toponyms consists of the names of populated places, which are generally represented on the map as urban areas or settlements of a nucleated type with partially defined or visible boundaries. The names of these settlements are positioned horizontally concerning the map frame, without spacing between letters. The modelling of this category of toponyms was carried out exclusively according to quantitative characteristics, i.e., based on the population size.

The second category of toponyms comprises dispersed-type villages. This category is distinctive in that their significance is primarily reflected in the area they occupy, i.e., their spatial position. These are mostly toponyms that are spatially curved relative to the map frame and are placed in such a way that they approximately match the spatial arrangement of the objects within the settlement, thus spatially defining it. The population size is of secondary importance in this case, as it involves a relatively small population spread over a relatively large area. Population data are stored in the database and are available upon request (where available), while the specific geographic name of the toponym was modelled exclusively based on its extent on the analogue TM25.

The third category of toponyms consists of the names of smaller hamlets or parts of settlements that need to be named on the TM25. This category is significant because it enhances the informational value of the TM25, further describing and naming specific areas. This category is complex to model due to its indeterminate nature within the GIS context. These are typically relatively small areas, and it is not practical to classify them based on extent, while also, as parts of settlements or hamlets, they are generally not separately included in population censuses, making it difficult to classify them based on population size. The modelling of this category of toponyms was done arbitrarily to the font size of the settlement name to which these toponyms belong. Consequently, this category of toponyms has the highest degree of conditional indeterminacy. The font size is defined in such a way that, by appropriately selecting the defined category, it does not lead to misinterpretation of the main settlement and its integral parts.

The fourth category of toponyms consists of exit names. Exit names do not represent named areas but indicate the extension of a route to the nearest larger settlement or toponym. These names are not part of the main TM25 content but are located in the map's margins of a printed map sheet. Exit names are standardized and modelled within a single category, meaning that settlements named in this way do not differ in size or font on the TM25.

This method of toponym modelling pertains to vector content, i.e., point-type toponym object classes, while visualization is implemented through the corresponding annotation object class.

Visualization of Toponyms

The visualization of toponyms, along with other categories of geographic names, is accomplished through the use of annotations. These annotations are dependent polygonal-

type object classes that derive their syntactic, semantic, and partial positional context from the attributes of the original point-type object class, which is used for data collection. The syntactic context is created within the annotation object class using a defined data model. This is achieved through the application of subtypes and corresponding symbology on the annotation object class.

The subtypes in the annotations are created based on the subtypes defined in the original point-type object class, meaning they are identical. This data is transferred from the point-type object class as an attribute into the appropriate column within the annotation object class, which is enabled by used software. Concerning these subtypes, textual symbology is created according to the previously mentioned methodology and data model. Symbology in this context refers to defining the method of displaying text stored within the attribute values of a particular column. This includes setting the font, size, style, spacing, etc., according to the defined category determined by a unique subtype identifier.

As they represent dependent yet separate object classes, the relationship between them and the original vector data is realized through a one-to-many (1–M) relationship, meaning that each element in the point-type object class can have multiple corresponding annotations (Berk et al 2003). This type of relationship is necessary because, in classical cartography, toponyms are often represented by more than one name. In addition to the primary name, a secondary name is sometimes used for better understanding by a specific user group. This secondary name is often shown by using phonetic transcription, transliteration, or the use of traditional names with historical significance. The secondary name is stored as an attribute in the column of the point-type object class and is visualized in the same manner as the primary name, although it has certain specificities in the visualization method that directly correspond to the visualization of the primary name. Additionally, exit names, as a category, are stored as attribute values in the corresponding columns of the point-type object class and are also visualized through annotations.

The aforementioned methodology has been used for the visualization of toponyms and other categories of geographic names on maps at a scale of 1:25,000. However, the database that supports this methodology was created to support the generation of a series of maps at various scales, including maps at 1:50,000 (TM50) and 1:100,000 (TM100) directly, and potentially 1:250,000 (OTM250) utilizing the ladder approach system.

While the change in scale does not significantly impact the original point-type object class since it is not visually represented on the printed map sheets, it is crucial for the visualization of annotations. The point-type object class is relatively easy to generalize by applying methods of cartographic reduction and reclassification of content based on the defined categorization of toponyms implemented in the data model. The change in scale affects the size of annotations on the map, as the size values of annotations are linked to the initial 1:25,000 scale. The change in the size of toponym annotations, in line with traditional cartographic practices, does not change linearly with scale but undergoes certain reductions in size, which are not linear across all categories of geographic names. The level of these changes was determined empirically and is numerically defined during the production of previous map editions by the MGI.

The implementation of different annotation display methods of toponyms according to scale required the creation of new symbol sets for each annotation category and the corresponding display scale supported by the database. This problem was solved by creat-

ing coded value domains linked to the annotations. The coded value domain associates the created symbology and display methods for the supported scales with the corresponding field in the attribute structure of the annotation class. The creation and assignment of coded value domains to the annotation object class do not affect the original point-type object class. By changing the scale and display methods according to predefined parameters, the visual identity of toponym representation, as established in traditional cartography on MGI maps, is preserved. However, this does not imply complete automation during the scale change for annotation display on maps. To maintain the spatial orientation of certain categories of toponyms, especially dispersed villages, several formalized activities must be undertaken. Moreover, the alignment of toponyms across derived scales currently requires the involvement of experienced cartographers who ensure proper representation according to internally established cartographic rules. This is mostly caused by the problems addressed by Freeman (Freeman, 2007).

Results

Based on the methodology described, a dataset was created within the cartographic database encompassing all categories of geographic names, with toponyms being an integral part of this dataset.



Fig. 1. Organization of object classes in the dataset GEOGRAFSKI_NAZIVI (screenshot captured in ESRI ArcMap 10.3.1) (MGI, 2024)

As shown in Figure 1, the geographic names classified as toponyms are represented through two object classes (Toponimi and Toponimi_ANNOTACIJE) and the relationship between them, Toponimi_ano. The object class Toponimi is a point-type class used for the collection and categorization of toponyms, based on which content is created in the dependent polygon-type object class, Toponimi_ANNOTACIJE, via the relationship class Toponimi_ano.

The object class Toponimi, besides defining attributes such as a universal identifier, also contains attributes that specify the created feature. Within the Toponimi object class, a primary field is defined for the classification of subtypes, storing unique identifiers for a specific type of feature. The subtypes are determined according to the primary field, wherein for each defined value of the primary field, the smallest common factors representing the characteristics of the feature being entered into the database are defined. These smallest common factors are the attributes of the feature that are unchangeable for the category defined by the data model, and they are entered as initial values in the sub-

type settings. An illustration of some of the created subtypes for the Toponimi object class is shown in Figure 2.

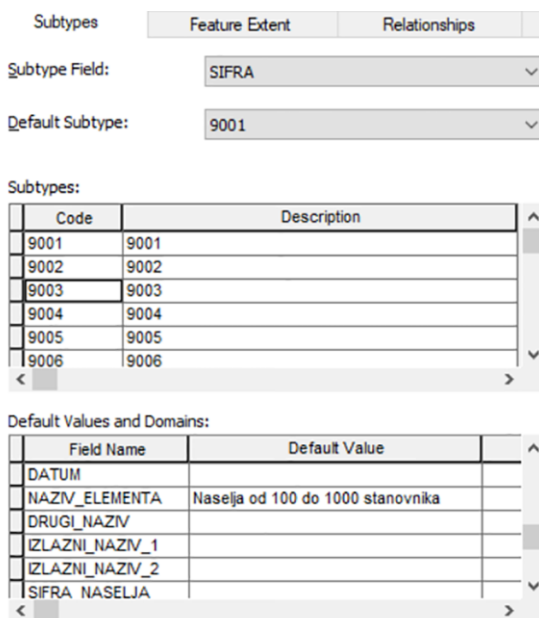


Fig. 2. Display of a portion of created subtypes for the object class Toponimi (screenshot captured in ESRI ArcMap 10.3.1) (MGI, 2024)

According to the stated methodology, data modelling for the object class Toponimi was carried out, with the data model realized through 19 categories that include all four identified types of toponyms.

The first type of toponyms represents the names of urban areas and clustered villages categorized by population size. In this case, the categorization and display methods established in analogue cartography, which also respected population size as a determining factor for the size of toponym labels, were followed. This category is represented by ten subtypes, which include: settlements with up to 100 inhabitants, settlements with 100 to 1,000 inhabitants, settlements with 1,000 to 2,000 inhabitants, settlements with 2,000 to 5,000 inhabitants, settlements with 5,000 to 10,000 inhabitants, settlements with 10,000 to 25,000 inhabitants, settlements with 25,000 to 50,000 inhabitants, settlements with 50,000 to 100,000 inhabitants, settlements with 100,000 to 500,000 inhabitants, and settlements with more than 500,000 inhabitants. Each subtype has a unique identifier and a description of the feature it represents. This categorization is influenced by the characteristics of the area being mapped, specifically the territory of the Republic of Serbia, and is therefore not universal outside this context.

The second type of toponym refers to the names of dispersed villages. The subtypes for these toponyms are modelled based on the length of the linear extension between two residential or industrial buildings at the maximum distance from one another, which are considered part of the named settlement. This distance is then converted into

units of graphic display on a map at a scale of 1:25,000. This data modelling method was primarily implemented for practical reasons. The primary data source for creating toponyms is a scanned, georeferenced map created using methods employed in analogue cartography at MGI. Since toponyms, as a category, are not significantly prone to change over relatively short periods, this data source is considered relevant. It serves as the foundation during the collection of toponyms, and through this data modelling, quicker identification of the relevant toponym category is facilitated, as rather than measuring the length of each toponym, the rectangular coordinate grid, integral to the cartographic source, is utilized. The modelling of dispersed villages is carried out using four subtypes, classified as follows: dispersed villages up to 2 cm on the map, dispersed villages from 2 cm to 10 cm on the map, dispersed villages from 10 cm to 20 cm on the map, and dispersed villages over 20 cm on the map. For this type of toponym, only the label size is defined within the annotations, while the spacing between the characters of the toponym is left for further cartographic processing. This approach was adopted because the spacing between the characters of a given toponym is determined by the number of characters in the name.

The third type of toponym represents the names of smaller hamlets or parts of settlements that need to be labelled on TM25. For this type of toponym, it is not practical to define categories using the criteria applied to the previous two types of toponyms. However, as they represent an integral part of a settlement, the label size for these toponyms on the map must be smaller than that of the main toponym, i.e., the toponym that semantically and spatially integrates this type of toponym. The determination of label size for this category of toponyms is carried out relative to the main names in a manner that avoids confusion when distinguishing the main toponym from the name of its part. Four subtypes have been created for this category of toponyms. Two subtypes refer to dispersed villages and are classified as smaller hamlets and larger hamlets. The other two subtypes refer to parts of settlements and are classified as smaller parts of settlements and larger parts of settlements.

The fourth type of toponym is modelled through a single subtype. The purpose of this toponym type is exclusively their graphical representation on the printed map. The visualization method for this subtype is identical across all toponym types, while in GIS applications, it does not hold practical value.

Based on this method of toponym data modelling, the visual representation of settlement names is created through the dependent object class `Toponimi_ANNOTACIJE`. The data in the object class `Toponimi_ANNOTACIJE` is modelled in the same way as in the `Toponimi` object class, using identical identifiers, but with twice as many subtypes. The increased number of subtypes is solely due to the presence of secondary names. As previously mentioned, a vector element from the `Toponimi` object class can possess multiple annotations, which in this case means the visualization of two columns from the `Toponimi` object class through the `Toponimi_ANNOTACIJE` object class.

The annotation object classes in the used software package are predefined to support a wide range of settings related to the display of geographic names, including the previously created attribute structure of the annotation object classes. An illustration of some of the created subtypes for the `Toponimi_ANNOTACIJE` object class is shown in Figure 3.

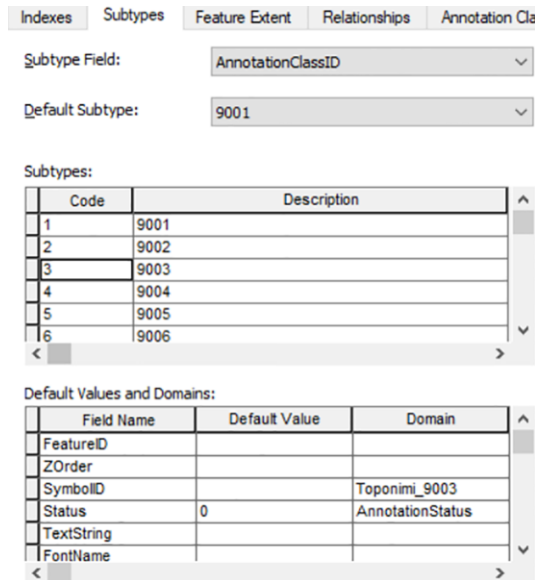


Fig. 3. Display of a portion of created subtypes for the object class Toponimi_ANNOTACIJE (screenshot captured in ESRI ArcMap 10.3.1) (MGI, 2024)

Concerning the dependent object class Toponimi_ANNOTACIJE, the method of creating symbology, i.e., text display on the map for each subtype, is of critical importance. Since the data modelling of toponyms was carried out to enable the visualized part of the database through labelling to achieve a clear differentiation of specific categories of toponyms, the method of display primarily adhered to the determining subtype. Accordingly, a set of textual symbols was created for each subtype, linked through code value domains. This set of symbols includes the defined display method for subtypes across all supported scales intended for cartographic production. In this case, these scales are 1:25,000, 1:50,000, and 1:100,000.

To avoid visualizing the naming column from the Toponimi object class across all created categories within the dependent object class Toponimi_ANNOTACIJE, which would result in multiple visualizations for a single attribute value, a reduction of the visualized categories had to be carried out. This reduction is facilitated by the software package through the writing of simple code for each created category. This code allows the selection of only the desired predefined visualization method based on the defined toponym category, and these codes are executed automatically during the creation of toponym content which is enabled by used software.

By using the presented methodology, the settlement name content was created within the centralized TM25 database following the defined data model and the cartographic display requirements for a 1:25,000 scale map. Figures 4 and 5 provide a depiction of some of the created toponyms for TM25 production according to the defined categories.

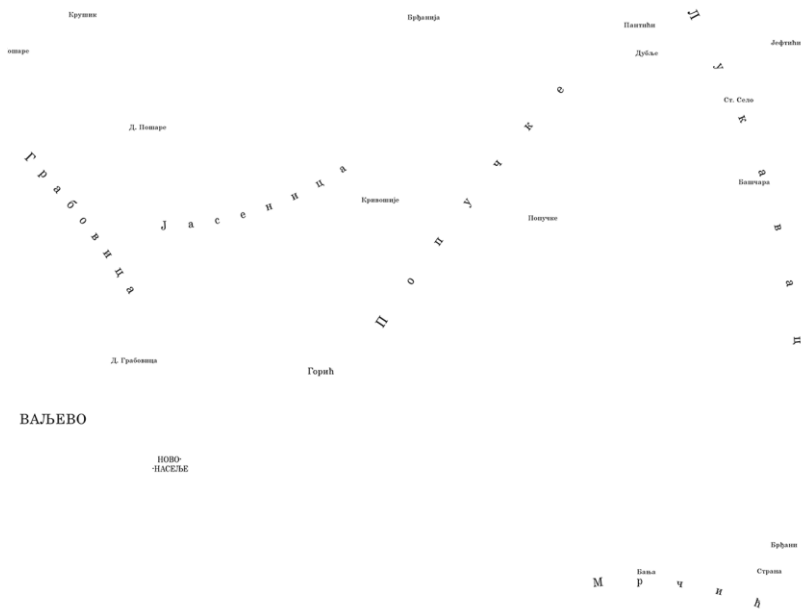


Fig. 4. Display of part of created toponyms for settlement names at a 1:25,000 scale (urban area with settlement parts and surrounding toponyms) (screenshot captured in ESRI ArcMap 10.3.1) (MGI, 2024)



Fig. 5. Display of part of created toponyms for settlement names at a 1:25,000 scale (dispersed settlements, hamlets, and secondary names) (screenshot captured in ESRI ArcMap 10.3.1) (MGI, 2024)

Discussion

Representing names of diverse settlement types on topographic maps created in GIS environment presents a significant challenge. A significant body of research addressing toponyms explores various cultural (Resticka et al., 2023), social (Savage, 2020), historical (Seidl, 2018), linguistic (Radding & Western, 2010) and similar aspects. These studies, to a lesser extent, address the technical and technological aspects of creating toponym content, even though they utilize toponym data that has often been previously collected and visualized on topographic maps. The quality of the analyses conducted depends on the method of visualization, primarily cartographic visualization of toponyms. This holds true when cartographic sources are used for analysis, but not when using tabular data with geographic names lacking a spatial component. This paper presents a comprehensive and structured method for modeling toponyms, addressing this challenge while aligning with MGI's standardization efforts for its published maps. Echoing the principles of hierarchical structured toponym resolution discussed in (Adelfio & Samet, 2013), this approach classifies toponyms into four main types (urban areas, dispersed-type villages, hamlets/parts of settlements, and exit names) with several subtypes based on specific criteria such as area and number of inhabitants. By standardizing the cartographic appearance of toponyms on topographic maps and enabling efficient attribute-based queries and spatial analysis within an ArcGIS environment, a given structured approach facilitates improved data management and reduces redundancy, similar to the specific aspect of categorization explored in (Setyo et al., 2022). This approach to toponym modeling for primarily cartographic purposes offers a significant improvement over traditional methods, enabling more efficient data management, enhanced analytical capabilities, and ultimately, more accurate and informative topographic maps.

This methodology addresses several key challenges in cartographic toponym representation. Firstly, it tackles the issue of data redundancy by implementing a standardized classification system, incorporating hierarchical categories and standardized attributes, which streamlines the process of identifying and managing multiple instances of the same toponym across multiple map sheets. This is particularly relevant in the context of large-scale mapping projects, such as TM25, where managing numerous toponyms can be complex. Secondly, the methodology considers the specific requirements of the Serbian language, which utilizes both Cyrillic and Latin scripts. By incorporating both scripts into the process, the methodology ensures that toponyms can be accurately represented cartographically in Cyrillic while maintaining Latin script attributes which are used for analytic and interoperability purposes. This approach aligns with the broader discussion on the interconnections between toponymy and identity (Woodman, 2014), recognizing the importance of preserving linguistic and cultural diversity in cartographic representation. Thirdly, and crucially in the cartographic context, the applied methodology enables a standardized visual representation of toponyms on MGI cartographic publications, thereby increasing the accuracy and usability of the created maps.

The use of a point-type object class, which can be exported in a standard format like Shapefile, enables interoperability with other software platforms for visualization and spatial analysis. However, more complex analyses involving semantic relationships between toponyms may require further data transformation. Furthermore, it allows for the efficient storage and retrieval of toponym attributes, such as name, type etc. This structured data model enables more sophisticated analyses, including both attribute-based

queries and spatial analyses that incorporate the location of toponyms (Purves et al., 2019). The potential for linking this toponym data to external databases further enhances the analytical capabilities (Mitxelena-Hoyos & Amaro-Mellado, 2023), allowing for a more comprehensive understanding of the relationships between toponyms and other geographic features.

The given methodology represents a significant step towards a more standardized and efficient approach to toponym representation in digital cartography at MGI. The structured data model, the consideration of dual scripts, and the potential for integration with other data sources offer valuable tools for managing and analyzing toponymic information.

Future research could explore the development of automated script conversion tools, the implementation of a seamless toponym national database, and the integration of this methodology with other toponymic research initiatives, such as those focusing on dynamic place naming systems (Mamontova & Klyachko, 2022). By addressing these challenges and building upon the existing framework, the given methodology contributes to a more comprehensive and robust system for representing toponyms. Its flexible and adaptable design, particularly for visualizing toponyms, accommodates evolving cartographic requirements and the addition of new toponym types without requiring significant changes to the underlying data model.

Conclusion

The creation of the content for the names of settlements, or toponyms, represents a database of geographic names tailored for cartographic visualization at an initial scale of 1:25,000 and for derived scales. Additionally, the created database possesses greater usability beyond mere cartographic display, as it contains additional data related to settlement names that enable complex and comprehensive spatial analyses within a GIS context. In this way, the generated geographic names meet the strictest criteria for cartographic visualization, while largely retaining the visual identity of toponym representation compared to the classical, analogue visualization method used by the Military Geographical Institute "General Stevan Bošković" until the early 21st century. This approach facilitates the seamless transition of cartographic data users from traditional maps to new edition maps created using digital methods and techniques, as the fundamental visualization remains unchanged. As a result, additional training for users on utilizing such cartographic publications, particularly regarding the presentation of geographic names, has been avoided, which was imperative when defining the methodology. User training, due to its scope, would have required significant personnel, financial, and time resources.

An additional advantage of this approach to toponym modelling is the potential for higher data usage levels, made possible through analyses that can be either purely attribute-based or that incorporate a spatial component. The created toponym data can also be linked to external databases, potentially expanding the attribute capacity of the conducted analyses.

Internal standardization of settlement name visualization has been achieved, allowing for faster production of cartographic content related to settlement names, improving the quality of the created content, and implementing quality control through domain-coded values.

The use of a point-type object class enables partial interoperability with other software that can visualize certain attribute values stored in the point vector's underlying

database. This approach has already yielded positive results when working with other software platforms.

The names of settlements printed on map sheets were required to be in Cyrillic, while there was also required for Latin script to be able to be represented on the maps. For this reason, the toponym name is entered in Latin script as an attribute in the point-type toponym geometry, while the dependent annotation object class is manually changed to Cyrillic script. This approach remains current but affects the speed of production and increases the risk of errors in the display and quality control of toponyms. Considering the recent development of non-strictly geography-oriented software, this approach may prove unnecessary, and transitioning to a single script would significantly impact the speed and quality of cartographically displayed toponyms. From the perspective of applied methodology and technology, such a transition would not affect the established technological process for creating topographic maps but would lead to inconsistencies with previously created settlement name content.

In terms of territorial coverage, the database used includes data at a scale of 1:25,000 on a national level, covering the territory of the Republic of Serbia with a narrower border zone included in TM25 map sheets. The data within the database are spatially defined according to individual TM25 sheets, leading to redundancy in toponyms that are spatially positioned on the edges of TM25 sheets and in content spanning multiple TM25 map sheets. This indicates that the database used is not seamless and that fact should be considered when conducting spatial analyses that include areas covered by multiple TM25 sheets. To use the created geographic name dataset, specifically the toponym object classes, as a database of geographic names, the strict concept of cartographic production, i.e., the division into 1:25,000 scale sheets, would need to be abandoned to achieve a seamless database without data redundancy.

Changing the methodology used for the representation of settlement names in classical cartography into a methodology for use with databases was not a simple task. This was primarily due to the GIS-incompatible data model used in analogue cartography. Through the application of the presented methodology and the corresponding data model, almost all categories of settlement names have been unified and standardized in a manner that reflects the traditional, analogue cartographic representation of toponyms. The mentioned degree of model uncertainty refers to the categories of hamlets and settlement parts for which quantitative cartographic representation censuses are not defined. Therefore, their cartographic interpretation depends on the assessment of experts who have interpretative freedom defined by the broader categories of the aforementioned toponym types. The number of toponym categories was determined not only based on cartographic rules but also on the need for faster production of cartographic content, resulting in a reduced number of categories, especially for dispersed-type villages, compared to the potential possibilities for representation. The defined number of categories reflects a compromise between the possibilities for representation, production speed, and the risk of errors in toponym classification while maintaining a satisfactory level of accuracy and quality in the representation of settlement names.

During the practical implementation of this task, ESRI software ArcMAP version 10.3.1 was utilized. Since this version is no longer supported by the manufacturer, the migration of the created toponym database to a newer version of the software (ESRI, 2024) may present potential challenges as well as new opportunities.

This paper presents the methodology and process for creating a database of settlement names for the production of a series of topographic maps at various scales at MGI. Beyond cartographic production, settlement names structured in this manner can potentially meet a range of other needs for visualized toponyms, depending on the scale of representation and the applied data model.

Conflicts of Interest: The author declares no conflict of interest.

Publisher's Note: Serbian Geographical Society stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2024 Serbian Geographical Society, Belgrade, Serbia.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Serbia.

References

- Adelfio, M., & Samet, H. (2013). Structured toponym resolution using combined hierarchical place categories. *Proceedings of the 7th Workshop on Geographic Information Retrieval*, 5, 49-56. <https://doi.org/10.1145/2533888.2533931>
- Berk, S. Radovan, D., & Petrovič, D. (2003). *The Problems of Scale and Standardization in the Toponymic Database*. Conference ISPRS WG VI/3 workshop Geoinformation for Practice.
- Borisov, M., Berk, S., Radovan, D., Petrovič, D. (2010). Modeli izgradnje baze geografskih naziva. *Geodetski žurnal*, 8, 20–22.
- ESRI. (2015). ArcMAP (Version 10.3.1.) [Software], Environmental Systems Research Institute. Available at <https://desktop.arcgis.com/en/arcmap/10.3/main/get-started/whats-new-in-arcgis-1031.htm> (Accessed 13 September 2024)
- ESRI. (2015). ArcMAP (Version 10.3.1.) [Software], Environmental Systems Research Institute. Available at <https://desktop.arcgis.com/en/arcmap/10.3/manage-data/geodatabases/an-overview-of-attribute-domains.htm> (Accessed 13 September 2024)
- ESRI. (2024). ArcGIS Pro [Software], Environmental Systems Research Institute. Available at <https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview> (Accessed 13 September 2024)
- Freeman, H. (2007). On the problem of placing names on a map. *Journal of Visual Languages & Computing*, 18(5), 458-469. <https://doi.org/10.1016/j.jvlc.2007.08.006>
- Jordan, P. (2020). *Current Geographical Research Directions in Toponymy*. Proceedings of the 5th International Symposium on Place Names.
- Mamontova, N., & Klyachko, E. (2022). Process Toponymy: A GIS-Based Community-Engaged Approach to Indigenous Dynamic Place Naming Systems and Vernacular Cartography. *University of Toronto Press*, 57(3), 213-225
- Mixelena-Hoyos, O., & Amaro-Mellado, J. (2023). A Comparison of Cartographic and Toponymic Databases in a Multilingual Environment: A Methodology for Detecting Redundancies Using ETL and GIS Tools. *Multidisciplinary Digital Publishing Institute*, 12(2), 70-70. <https://doi.org/10.3390/ijgi12020070>

- Perdana, A. P., & Ostermann, F. O. (2018). A citizen science approach for collecting toponyms. *ISPRS International Journal of Geo-Information*, 7(6), 222. <https://doi.org/10.3390/ijgi7060222>
- Peterca, M., Radošević, N., Milisavljević, S., Racetin, F. (1974): Kartografija, Vojnogeografski institut, Beograd.
- Purves, R. S., Winter, S., & Kuhn, W. (2019). Places in information science. *Journal of the Association for Information Science and Technology*, 70, 452-466. <https://doi.org/10.1002/asi.24194>
- Radding, L., & Western, J. (2010). What's in a name? Linguistics, Geography and Toponyms. *Geographical Review*, 100(3), 394-412. <http://www.jstor.org/stable/25741159>
- Resticka, G. A., Nurdiyanto, E., & Marahayu, N M. (2023). Toponyms in Banyumas as Representation of Banyumas Community Culture. *Jurnal Lingua Idea*, 14(1), 71-71. <https://doi.org/10.20884/1.jli.2023.14.1.8122>
- Rose-Redwood, R., Alderman, D., & Azaryahu, M. (2010). Geographies of toponymic inscription: New directions in critical place name studies. *Progress in Human Geography*, 34(4), 453-470. <https://doi.org/10.1177/0309132509351042>
- Savage, V. R. (2020). Place names. In: A. Kobayashi (Ed.), *International Encyclopedia of Human Geography* (pp. 129-138). Elsevier. <https://doi.org/10.1016/B978-0-08-102295-5.10852-2>
- Setyo, W. F. N., Muryani, C., & Nurhadi, N. (2022). Toponymic and Historiography Influences on Place Naming of Villages in Klego District, Indonesia. *KnE Social Sciences*, 89-100. <https://doi.org/10.18502/kss.v7i16.12155>
- Seidl, N P. (2018). Engraved in the Landscape: The Study of Spatial and Temporal Characteristics of Field Names in the Changing Landscape. *Journal of Onomastics*, 67(1), 16-29
- Tatomirović, S., Regodić, M., Radojičić, S., & Banković, R. (2016). Evaluation of Cartographic Sources for Producing Serbian Military Geographic Names Database. *Journal of Cartography and Geographic Information*, 66(6), 310-15. <https://doi.org/10.1007/BF03545306>
- United Nations Group of Experts on Geographical Names (2015). The Benefits of Geographical Names Standardization. *UNGEGN Information Bulletin*, 49, 1-39.
- Vojnogeografski institut (1982). *Uputstvo za ispisivanje geografskih naziva i drugih natpisa na kartama izdanja Vojnogeografskog instituta*. Vojnogeografski institut.
- Woodman, P. (2014). The interconnections between toponymy and identity. *Review of Historical Geography and Toponomastics*, 9(17-18), 7-20.

Оригинални научни рад

UDC 528.9:[007:912]:004]
<https://doi.org/10.2298/GSGD2402427M>

Примљено: 19. октобра 2024.

Исправљено: 28. новембра 2024.

Прихваћено: 10. децембра 2024.

Виктор Марковић^{1*}

* Војногеографски институт – „Генерал Стеван Бошковић“, Београд, Србија

МОДЕЛ ПОДАТАКА ЗА КАРТОГРАФСКИ ПРИКАЗ ТОПОНИМА НА ТОПОГРАФСКИМ КАРТАМА ИЗДАЊА ВГИ

Апстракт: Приликом креирања топографских карата, веома важан аспект представља стандардизација и моделовање података у сврси правилне визуализације топонима. Важност овог аспекта произилази из чињенице да карте, осим што представљају графички приказ географског простора, представљају и средство за комуникацију одређених информација и топонима који морају бити јасно представљени. У складу са тим, посебна пажња је усмерена на развој методологије за стандардизацију топонима на картама Војногеографског института „Генерал Стеван Бошковић“ (ВГИ), који задовољава строге критеријуме картографског приказа. Овај рад разматра методологију и поступак креирања базе назива насељених места за картографски приказ у размери 1:25.000 (ТК25), као и у изведеним размерама применом мердевинастог система. Фокус рада је на стандардизацији и унапређењу картографског приказа топонима путем дигиталних техника. Креирана база података поседује широку примену не само у картографији, већ и у ГИС контексту, омогућавајући комплексне просторне анализе. Додатно, овај рад истражује предности и недостатке креираног модела података топонима, начине његове имплементације у производни систем за креирање топографских карата у ВГИ, као и потенцијалне могућности за унапређење. Методологија приказана у овом раду је омогућила релативно лаку транзицију са аналогног на дигитални приказ топонима уз задржавање визуелног идентитета и функционалности класичних топографских карата.

Кључне речи: топоними, база података, ГИС, модел података, топографске карте

¹ viktor_bre@yahoo.com (corresponding author)
Виктор Марковић (<https://orcid.org/0000-0002-1085-8575>)

Увод

Географски називи представљају кључну компоненту сваког картографског приказа географског простора. Они нису само ознаке за места, већ често играју пресудну улогу у просторној оријентацији, анализи узрочно-последичних веза и разумевању односа између човека и природног окружења (види Jordan, 2020). У класичној картографији, географски називи су приказивани уз поштовање различитих критеријума као што су простирање, квантитативне карактеристике, административни значај, историјски и културни контекст. Прелазак са аналогне на дигиталну картографију поставио је нове изазове. Док су на аналогним картама картографи имали слободу у интерпретацији, ГИС (Географски информациони системи) захтевају прецизно дефинисана правила и стандарде за моделовање и визуализацију података. Овај прелазак није само омогућио прецизније просторне анализе, већ је и отворио пут за ефикаснију стандардизацију и униформност у приказу топонима. Топоними су сада део сложених база података које омогућавају напредне просторне упите и анализе, чинећи их значајним ресурсом за географске системе и картографске примене.

Стандардизација топонима, нарочито у контексту њиховог приказа у ГИС, у великој мери је допринела тачности и доследности картографских приказа. Национални напори за стандардизацију топонима, често у складу са међународним смерницама које промовишу организације попут UNGEGN-а (UNGEKN, 2015), су кључни за обезбеђивање веће интероперабилности и тачности у различитим картографским системима (нпр. Perdana & Ostermann, 2018). С друге стране, поједини истраживачи истичу значај топонима као културних симбола, што додатно компликује њихов приказ на дигиталним картама, где су они истовремено просторни објекти и симболи културног идентитета (нпр. Rose-Redwood et al., 2010).

У Србији, значајан допринос истраживању топонима у ГИС контексту дали су Borisov et al. (2010) и Tatomirović et al. (2016), који су анализирали развој и примену стандардизације топонима у националном картографском систему. Њихови радови наглашавају важност јединственог приказа топонима, посебно у контексту државних и војних географских информационих система, као што је онај који користи ВГИ.

Овај рад приказује метод моделовања топонима који се користи у ВГИ, са нагласком на стандардизацију топонима у размери 1:25.000 и њихову интеграцију у мердевинасти приступ креирања топографских карата, који је у употреби у ВГИ од 2012. године. Приказан метод омогућава униформан приказ топонима кроз различите размере, задржавајући визуелни континуитет са старијим аналогним картама које су креиране током 20. и почетком 21. века.

Моделовање топонима у ГИС окружењу, како је имплементирано у ВГИ, захтева строгу стандардизацију како би се постигла визуална доследност и ефикасност у просторним анализама. Ова стандардизација неопходна је како би се задржао традиционални изглед карата и истовремено омогућила њихова употреба у ГИС апликацијама.

Материјал и методе

У сврси сублимације критеријума употребљених у аналогној картографији за приказ топонима, коришћено је више извора података. Као примарни извор података су коришћена аналогна издања карата у размери 1:25000 док се као помоћни извори

користе аналогне карте у размерама 1:50000 и 1:100000. На основу наведених карата, као и на основу писаних упутстава за израду географских назива на топографским картама (ВГИ, 1982), извршена је сублимација критеријума за приказ топонима. Поменуће карте су скениране и геореференциране и користе се у продукционом систему топографских карата у виду сервиса. Употребом наведених картографских извора креиран је иницијални садржај топонима и похрањен у базу података.

База података која се користи у продукционом систему топографских карата у ВГИ је SQL Enterprise база података. Сама база података је креирана са системом за управљање базама података (DBMS – database management system) који је саставни део софтверског пакета ArcMAP 10.3.1. који је коришћен у реализацији истраживања (ESRI, 2015). Примарна функција ове базе података је картографска продукција и она представља дигитални картографски модел (ДКМ) на основу кога се креира векторска карта у размери 1:25000 која представља визуализован приказ дела базе података. Поред своје примарне функције, база података располаже и мноштвом других алфанумеричких података који није непходно да буду визуализовани на карти али имају изузетну улогу у просторним анализама. Поменути алфанумерички подаци су у бази података везани за припадајуће појединачне објекте које додатно описују односно представљају њихове атрибуте.

Приликом именовања одређене појаве на карти атрибутима похрањеним у базу података, било да су у питању географски називи или бројчано-словни подаци који описују или означавају неку појаву, коришћене су анотације. Анотације, за разлику од лабела, омогућавају висок степен ручне манипулације у позиционом смислу. У конкретном случају, оне представљају зависне објектне класе које свој синтаксни и семантички контекст добијају из векторски креираног садржаја који је потребно означити односно именовати. Како представљају зависне, а опет засебне објектне класе, веза између њих и изворног векторског податка се реализује путем различитих типова релација односно веза (ESRI, 2015).

Поменућа картографска база података у размери 1:25000 представља структурирану, таксономску ГИС базу података креирану употребом ArcGIS софтвера. У креираном систему, појединачне објектне класе су дефинисане на основу њихове семантичке припадности одговарајућем сету података. У оквиру ових објектних класа, подаци су даље организовани у подтипове, што омогућава једноставну селекцију података по типу, категорији, геометрији и интерно усвојеним јединственим алфанумеричким идентификаторима до нивоа ентитета.

База података у техничком смислу, поседује и могућност креирања домена који имају кључну улогу на најнижем нивоу дефинисања појединачних елемената у оквиру подтипова. У оквиру коришћеног софтверског пакета, могуће је креирати два типа домена: домен распона вредности и домен кодних вредности (ESRI, 2015). Креирањем и додељивањем ових домена постигнута је и својеврсна контрола квалитета креираног садржаја. Употребом домена распона вредности, смањена је могућност грубе грешке код уноса нумеричких вредности. Домен кодних вредности омогућава унос само предефинисаних вредности условљених моделом података, чиме се постиже униформност креираног садржаја, док је онемогућена слободна интерпретација која би могла да отежа односно наруши тачност просторних анализа или доведе до погрешне интерпретације географских појава.

Подаци у оквиру базе података су организовани у сетове података. Сетови података представљају највиши ниво класификације припадајућих елемената на основу њиховог карактера и семантичке повезаности. Креирано је 10 сетова података који су од значаја за креирање векторске ТК25 и самим њиховим називима је очигледно који типови појава су похрањени у оквиру њих. Наведени сетови података представљају својеврсне коверте за смештај истородног садржаја који дели заједничке особине и својства која су дефинисана у оквиру матичног сета података.

Трећи ниво у организацији података представљају објектне класе које су креиране у оквиру сетова података. Поменуте објектне класе, осим према семантичкој припадности креиране су и организоване и према геометрији односно начину на који ће бити представљене у размери 1:25000. Објектне класе тачкастог, линијског или површинског типа фигуришу у оквиру сетова података.

Четврти ниво у организацији података представљају подтипови. Овај ниво таксономске организације података омогућава диференцијацију појединачних елемената на нивоу ентитета и он је дефинисан јединственим нумеричким идентификатором у оквиру припадајуће објектне класе а креиран је на основу дефинисаног модела података. Сви подтипови поседују идентичну атрибутску структуру која је дефинисана у оквиру припадајуће објектне класе. Ова хијерархијска, таксономска организација података, од базе података до подтипова, представља функционалност ArcGIS софтверске платформе, омогућавајући ефикасно управљање подацима и њихово проналажење у GIS окружењу.

Географски називи су грписани у оквиру истоименог сета података – „Географски називи“. Елементи који садрже информације о географским називима представљени су помоћу три семантичке целине које су представљене у оквиру базе података са три кореспондирајуће објектне класе. У овом случају, поштована је традиционална класификација географских назива на називе насељених места – топониме, називе предела и крајева – хорониме и називе планина, планинских масива, брда, коса и сл. – орониме (Peterca et al., 1974).

Топоними представљају појаве антропогеног карактера и од изузетне су важности приликом визуализације геопростора, а нарочито у картографској визуализацији. Конкретно, топоними означавају насељена места која у размери 1:25000 могу бити представљена полигонима, односно представљају одређени именовани простор. Прецизно одређивање граница тог простора, узимајући у обзир изворе података који се користе, најчешће није сврсисходно. Из тог разлога примењује се генерализација јер се, по природи полигони елементи, приказују тачком која по правилима треба да представља геометријски центар појаве коју дефинише. Тачке које представљају топониме поседују атрибуте који описују појаву коју геометријски представљају. Поменуте тачке су у визуалном смислу – невидљиве, односно представљају тачке одређених димензија без боје. Ово решење је усвојено како би се омогућиле просторне анализе и добила већа флексибилност приликом креирања зависних анотација које су у овом случају детерминишуће у просторном контексту појаве која се приказује. Такође, додатни разлог представља и интероперабилност. Софтвер који се користио у реализацији овог истраживања подржава директан рад са анотацијама и оне су употребљене за визуализацију топонима и осталих категорија географских назива. Како би се креирани подаци везани за топониме могли користити на другим софтверским платформама које не подржавају директан рад са

анотацијама, или којима картографска продукција није примарни задатак, све неопходне информације о категорији топонима се налазе у оквиру објектне класе тачкастог типа. На овај начин је омогућено брзо креирање лабела само на основу атрибута који се налазе у оквиру објектне класе тачкастог типа без неопходног присуства анотација. Овакав приступ може да се интерпретира као вид редундансе података јер се исти податак налази у објектној класи тачкастог типа у виду атрибута као и у зависној објектој класи анотација, у виду визуализованог атрибута. Иако се оба идентична атрибута односе на исту појаву, разлика је у начину креирања и визуализације као и саме употребе.

Моделовање података

У односу на аналогни картографски приказ топонима, основну разлику у визуалном смислу у конкретном случају представља избор фонта који се користи за визуализацију топонима у окружењу базе података. Из разлога интероперабилности векторски креираног садржаја, одређени су Windows стандардни фонтови који по својим карактеристикама у највећој мери одговарају карактерима који су коришћени на аналогним картама.

Следећи корак у моделовању топонима је представљала категоризација у оквиру објектне класе топоними. Током реализације истраживања, утврђено је да постоји више категорија топонима, од којих сваки поседује одређене специфичности које је неопходно посебно разматрати приликом моделовања података.

Прву групу топонима представљају називи насељених места која су генерално на карти представљена као урбана подручја као и насеља ушореног типа са делимично дефинисаним или уочљивим границама. Називи ових насељених места се постављају хоризонтално у односу на оквир листа карте, без размака између слова. Моделовање ове категорије топонима је реализовано искључиво према квантитативним карактеристикама односно према броју становника.

Другу категорију топонима представљају села разбијеног типа. Ова категорија је специфична по томе што се њихов значај огледа првенствено у простору који заузимају, односно њиховој просторној позицији. Најчешће су у питању топоними који су положајно закривљени у односу на оквир листа карте а постављени су тако да се просторно приближно подударају са објектима који се налазе у оквиру насеља на тај начин га просторно дефинишући. Број становника је у овом случају од секундарног значаја јер је у питању релативно мали број становника настањен на релативно великом подручју. Подаци о броју становника фигуришу у бази података и доступни су на упит (тамо где су доступни) док је конкретан географски назив топонима моделован искључиво у односу на дужину простирања на аналогној ТК25.

Трећу категорију топонима представљају називи мањих засеока или делова насеља које је неопходно именовати на ТК25. Ова категорија је значајна јер употпуњује информативни карактер ТК25 односно додатно описује и именује одређени простор. Основни проблем овде представља немогућност јасне категоризације и моделовања ове врсте топонима у ГИС контексту. Најчешће су у питању релативно мала подручја и није их сврсисходно класификовати према простирању док такође, као делови насеља или засеоци, углавном нису засебно обухваћени пописом становништва тако да их најчешће није могуће класификовати према броју становника. Моделовање ове категорије топонима је извршено произвољно у односу на величи-

ну исписа назив насеља чији саставни део чине топоними који припадају овој категорији. Као последица наведеног, ова категорија топонима поседује највиши степен условне неодређености. Величина исписа је дефинисана на такав начин да адекватним избором дефинисане категорије, не доводи до погрешне интерпретације главног насеља и његовог интегралног дела.

Четвру категорију топонима чине излазни називи. Излазни називи не представљају именовани простор, већ указују на протезање комуникације у односу на најближе, веће насељено место, односно топоним. Они нису саставни део експонираниог садржаја ТК25 већ се позиционирају у међуоквирни простор ТК25 на штампаном отиску листа карте. Излазни називи су стандардизовани и моделовани преко једне категорије, што значи да се насеља која су именована на овај начин не разликују у величини и начину исписа на ТК25.

Наведен метод моделовања топонима се односи на векторски садржај, односно објектну класу тачкастог типа топонима док је визуализација реализована преко кореспондирајуће објектне класе анотација.

Визуализација топонима

Визуализација топонима као и других категорија географских назива је реализована употребом анотација. Оне представљају зависне објектне класе полигоног типа које свој синтаксни, семантички и делимично положајни контекст добијају из атрибута изворне објектне класе тачкастог типа преко које се врши прикупљање података. Синтаксни контекст се креира у оквиру објектне класе анотација, употребом дефинисаног модела података. Ово је постигнуто употребом подтипова и одговарајуће симболије над објектном класом анотација.

Подтипови у анотацијама се креирају у односу на подтипове креиране у изворној објектној класи тачкастог типа, односно идентични су. Овај податак се преноси из објектне класе тачкастог типа као атрибут у адекватну колону у оквиру објектне класе анотација, што је омогућено коришћеним софтвером. У односу на ове подтипове, креирана је текстуална симболија у складу са претходно наведеном методологијом и моделом података. Симболија у овом случају представља дефинисање начина исписа текста похрањеног у оквиру атрибутских вредности одређене колоне. То подразумева подешавање фонта, величине, стила, шпациона и сл. у односу на дефинисану категорију назива одређену јединственим идентификатором подтипа.

Како представљају зависне, а опет засебне објектне класе, веза између њих и изворног векторског податка се реализује путем релације један на више (1 – М) што практично значи да сваки елемент који се налази у објектној класи тачкастог типа геометрије може да поседује више кореспондирајућих анотација (Berk et al., 2003). Овај тип везе је неопходан јер су топоними у класичној картографији према потреби приказивани са више назива. Као додаток основном називу, понекад се користи и други назив ради лакшег разумевања код одређеног круга корисника. Други назив се често приказује применом фонетске транскрипције, транслитерације или употребом традиционалних назива који поседују историјски карактер. Други назив је похрањен у виду атрибута у колону објектне класе тачкастог типа и он се визуализује на идентичан начин као и основни назив са тим да поседује одређене специфичности у начину визуализације који је у директној вези са начином визуализације основног назива. Такође, излазни називи као категорија су похрањени као атрибут-

ска вредност у одговарајуће колоне у оквиру објектне класе тачкастог типа и оне се такође визуализују путем анотација.

Поменута методологија је коришћена у визуализацији топонима и осталих категорија географских назива на карти у размери 1:25000. Међутим, база података која подржава наведену методологију је креирана у сврси подршке за генерисање размерног низа карата који укључује карте у размерама 1:50000 (ТК50) и 1:100000 (ТК100) директно и потенцијално 1:250000 (ПТК250) употребом мердевинастог приступа креирању картографског система.

Док промена размере не утиче у значајној мери на изворну објектну класу тачкастог типа, будући да визуално није присутна на штампаним листовима карата, од пресудног је значаја за визуализацију анотација. Такође, објектну класу тачкастог типа је релативно лако генералисати применом метода картографске редукције и прекласификације садржаја у односу на дефинисану категоризацију топонима имплементирану моделом података. Променом размере мења се величина приказа анотација на карти, будући да су вредности величина анотација везани за иницијалну размеру 1:25000. Промена величина исписа анотација топонима, у складу са маниром традиционалне картографије, се не мења линеарно у односу на промену размере већ трпи одређене промене у смислу редукције величине исписа, која опет није линеарна за све категорије географских назива. Ниво промена је установљен емпиријски и дефинисане су нумеричке вредности током израде претходних издања карата у ВГИ.

Имплементација различитих начина исписа анотација топонима у зависности од размере захтевала је креирање нових сетова симбола за сваку категорију анотација према свакој размери приказа коју подржава база података. Овај проблем је решен креирањем домена кодних вредности везаних за анотације. Домен кодних вредности повезује креирану симбологију и начине исписа за подржане размере са кореспондирајућим пољем у атрибутској структури класе анотација. Креирање и додељивање домена кодних вредности објектној класи анотација не утиче на изворну објектну класу тачкастог типа. Променом размере и начина исписа према предефинисаним параметрима, задржава се визуелни идентитет приказа топонима утврђен у традиционалној картографији на картама издања ВГИ. Међутим, ово не подразумева потпуну аутоматизацију приликом промене размере за испис анотација на картама. Да би се задржала просторна оријентација појединих категорија топонима, посебно села разбијеног типа, неопходно је предузети неколико формализованих активности. Такође, међусобно усклађивање топонима на изведеним размерама тренутно захтева ангажовање искусних картографа, који ће осигурати исправан приказ у складу са интерно дефинисаним картографским правилима. Ово је углавном узроковано проблемима које је обрадио Фриман (2007).

Резултати

На основу наведене методологије, у оквиру картографске базе података, креиран је сет података којим су обухваћене све категорије географских назива чији интегрални део представљају и топоними.

Сл. 1. Организација објектних класа у сету података GEOGRAFSKI_NAZIVI (снимак екрана направљен у софтверу ESRI ArcMap 10.3.1.) (ВГИ, 2024) (стр. 433)

Како се може уочити са слике 1, географски називи који се класификују као топоними, реализовани су путем две објектне класе (Торониџи и Торониџи_АНОТАЦИЈЕ) као и везе између њих, Торониџи_ано. Објектна класа Торониџи представља објектну класу тачкастог типа која служи за прикупљање и категоризацију топонима на основу које се посредством класе везе Торониџи_ано креира садржај у зависној објектној класи полигоног типа Торониџи_АНОТАЦИЈЕ.

Објектна класа Торониџи се састоји у атрибутском смислу осим од детерминишућих атрибута попут универзалног идентификатора, и од атрибута који ближе одређују креирану појаву. У оквиру објектне класе Торониџи одређено је примарно поље за дефинисање подтипова и у оквиру њега су похрањени јединствени идентификатори одређеног типа појаве. Подтипови су дефинисани у односу на примарно поље односно, за сваку дефинисану вредност примарног поља одређени су најмањи заједнички чиниоци који представљају карактеристике појаве која се уноси у базу података. Најмањи заједнички чиниоци представљају атрибуте одређене појаве који су непромењиви за моделом података дефинисану категорију и они су у поставкама подтипова као атрибутске вредности унесени као иницијалне вредности. Приказ дела креираних подтипова за објектну класу Торониџи је представљен на слици 2.

Сл. 2. Приказ дела креираних подтипова за објектну класу Торониџи (снимак екрана направљен у софтверу ESRI ArcMap 10.3.1) (ВГИ, 2024) (стр. 434)

У складу са наведеном методологијом, извршено је моделовање података за објектну класу Торониџи и модел података је реализован кроз 19 категорија које укључују сва четири идентификована типа топонима.

Први тип топонима представљају називи урбаних подручја и села ушореног типа категоризована према броју становника. У овом случају поштована је категоризација и начин приказа установљен у аналогној картографији која је такође поштовала број становника као детерминишући фактор за величину исписа топонима. Ова категорија је реализована преко 10 подтипова који обухватају: насеља до 100 становника, насеља од 100 до 1000 становника, насеља од 1000 до 2000 становника, насеља од 2000 до 5000 становника, насеља од 5000 до 10.000 становника, насеља од 10.000 до 25.000 становника, насеља од 25.000 до 50.000 становника, насеља од 50.000 до 100.000 становника, насеља од 100.000 до 500.000 становника и насеља преко 500.000 становника. Сваки од наведених подтипова поседује јединствени идентификатор као и опис типа појаве која се представља. Наведена категоризација је условљена пре свега карактеристикама подручја које се картира, у конкретном случају простор Републике Србије, тако да она није универзална ван контекста наведеног подручја.

Други тип топонима представљају називи насељених места који означавају села разбијеног типа. Подтипови који именују ову врсту топонима су моделовани према дужини линеарног простирања између два стамбена или индустријска објекта на међусобно максималном растојању за које се сматра да припадају насељу које се именује. Одређено растојање је потом конвертовано у јединице графичког приказа на карти у размери 1:25000. Поменут начин моделовања података је реализован пре свега из практичних разлога. Као примарни извор података у креирању топонима користи се скенирана, геореференцирана карта креирана методама коришћеним у аналогној картографији у ВГИ. С обзиром да топоними као категорија у односу на саме називе, нису у значајној мери подложни проме-

нама у релативно кратком временским периодима, наведени извор података се сматра релевантним. Он се поставља као подлога приликом прикупљања назива топонима и, на овај начин моделованим подацима, омогућава се бржа идентификација припадајуће категорије топонима јер се уместо мерења вредности дужине сваког топонима користи правоугла координатна мрежа која је саставни део коришћеног картографског извора. Моделовање назива села разбијеног типа је реализовано употребом 4 подтипа који су класификовани на следећи начин: села разбијеног типа до 2 cm на карти, села разбијеног типа од 2 cm до 10 cm на карти, села разбијеног типа од 10 cm до 20 cm на карти, села разбијеног типа преко 20 cm на карти. За овај тип топонима, у оквиру анотација дефинисана је само величина исписа, док је растојање између карактера топонима остављено за даљу картографску обраду. Поменуто решење је усвојено јер је размак између карактера одређеног топонима условљен самим бројем карактера у називу.

Трећи тип топонима представљају називи мањих засеока или делова насеља који је неопходно да буду именовани на ТК25. За овај тип топонима није сврсисходно дефинисати категорије према критеријумима који су примењени приликом моделовања претходна два типа топонима. Међутим, будући да представљају интегрални део насеља, величина њиховог исписа на карти мора бити мања од назива главног топонима, односно топонима који семантички и положајно интегрише овај тип топонима. Утврђивање вредности величина исписа ове категорије топонима је реализовано у односу на главне називе на начин који не доводи до конфузије приликом одређивања главног топонима и назива његовог дела. За ову категорију топонима је креирано 4 подтипа. Два подтипа се односе на села разбијеног типа и они су класификовани као засеок мањи и засеок већи. Друга два подтипа се односе на делове насеља и они су класификовани као део насеља-мањи и део насеља-већи.

Четврти тип топонима је реализован моделовањем једног подтипа. Сврха овог типа топонима је искључиво графичко присуство на штампаном отиску карте. Начин визуализације овог подтипа је идентичан за сваки тип топонима док у ГИС применама не поседује употребну вредност.

На основу овог начина моделованих података топонима креира се визуална представа назива насељених места путем зависне објектне класе `Toronimi_ANNOTACIJE`. Подаци су у објектној класи `Toronimi_ANNOTACIJE` моделовани на идентичан начин као и у објектној класи `Toronimi` уз употребу идентичних идентификатора са тим да је број подтипова двостуко већи. Повећан број подтипова је дефинисан искључиво због другог назива. Како је претходно наглашено, један векторски елемент из објектне класе `Toronimi` може поседовати више анотација што у конкретном случају значи вузуализацију две колоне из објектне класе `Toronimi` путем објектне класе `Toronimi_ANNOTACIJE`.

Објектне класе анотација су у коришћеном софтверском пакету предефинисане да подржавају широк распон подешавања везаних за начин исписа географских назива, укључујући и претходно креирану атрибутску структуру објектних класа анотација. Приказ дела подтипова креираних за објектну класу `Toronimi_ANNOTACIJE` је приказан на слици 3.

Сл. 3. Приказ дела креираних подтипова за објектну класу `Toronimi_ANNOTACIJE` (снимак екрана направљен у софтверу ESRI ArcMap 10.3.1) (ВГИ, 2024) (стр. 436)

У односу на зависну објектну класу *Toponimi_ANOTACIJE* од пресудног значаја је начин креирања симболије, односно текстуалног исписа на карти за сваки подтип. Будући да је моделовање података топонима реализовано како би визуализовани део базе података путем исписа довео до јасне диференцијације одређених категорија топонима, приликом креирања начина исписа поштовао се првенствено детерминишући подтип. У складу са тим, за сваки подтип је креиран сет текстуалних симбола који су за њега повезани путем домена кодних вредности. Поменути сет симбола у себи садржи дефинисан начин приказа подтипа у свим подржаним размерама намењеним за картографску продукцију. У конкретном случају то су размере 1:25000, 1:50000 и 1:100000.

Како би се избегла визуализација именујуће колоне из објектне класе *Toponimi* у свим креираним категоријама у оквиру зависне објектне класе *Toponimi_ANOTACIJE*, чиме би се добила вишеструка визуализација за једну атрибутску вредност, морала се извршити редукција визуализованих категорија. Редукција је у коришћеном софтверском пакету омогућена писањем једноставног кода за сваку креирану категорију. Овим кодом је омогућена селекција само жељеног предефинисаног начина визуализације у односу на дефинисану категорију топонима и ови кодови се извршавају аутоматски током креирања садржаја топонима што је омогућено коришћеним софтвером.

Коришћењем приказане методологије, креиран је садржај назива насељених места у оквиру централне базе података *TK25* у складу са дефинисаним моделом података као и захтевима картографског приказа на карти размера 1:25000. На сликама 4 и 5, је дат приказ дела креираних топонима за потребе израде *TK25* према дефинисаним категоријама.

Сл. 4. Приказ дела креираних топонима за називе насељених места у размери 1:25000 (Урбано подручје са деловима насеља и околни топоними), (снимак екрана направљен у софтверу ESRI ArcMap 10.3.1) (ВГИ, 2024) (стр. 437)

Сл. 5. Приказ дела креираних топонима за називе насељених места у размери 1:25000 (Села разбијеног типа, засеоци и други назив), (снимак екрана направљен у софтверу ESRI ArcMap 10.3.1) (ВГИ, 2024) (стр. 437)

Дискусија

Представљање назива различитих типова насеља на топографским картама креираним у ГИС окружењу представља значајан изазов. Значајан број истраживања која се баве топонимима истражују разне културолошке (Resticka et al., 2023), социјалне (Savage, 2020), историјске (Seidl, 2018), лингвистичке (Radding & Western, 2010) и сличне аспекте. Ова истраживања у мањој мери разматрају техничке и технолошке аспекте креирања садржаја топонима, иако често користе податке о топонимима који су претходно прикупљени и визуализовани на топографским картама. Квалитет спроведених анализа зависи од методе визуализације, првенствено картографске визуализације топонима. Ово је нарочито значајно када се картографски извори користе за анализу, али не и када се користе табеларни подаци са географским именима без просторне компоненте. Овај рад представља свеобухватан и структурисан метод моделовања топонима, којим се решава поменути изазов у складу са напорима ВГИ-а за стандардизацијом назива на публикованим картама. У складу са принципима хијерархијског структурирања

топонима описаним у раду (Adelfio & Samet, 2013), овај приступ класификује топониме у четири главна типа (урбана подручја, села разбијеног типа, засеоци/делови насеља и излазни називи) са неколико подтипова одређених на основу специфичних критеријума као што су површина и број становника. Стандардизацијом картографског приказа топонима на топографским картама и омогућавањем ефикасних упита на основу атрибута и просторних анализа у оквиру ArcGIS окружења, приказан структурирани приступ поједностављује управљање подацима и смањује редундансу, слично специфичним аспектима категоризације приказаним у (Setyo et al., 2022). Овај приступ моделовању топонима, примарно за картографске сврхе, представља значајно унапређење у односу на традиционалне методе, омогућавајући ефикасније управљање подацима, побољшане аналитичке могућности и, коначно, тачније и информативније топографске карте.

Приказана методологија третира неколико кључних изазова у картографском представљању топонима. Прво, решава проблем редундансе података применом стандардизованог система класификације, укључујући хијерархијске категорије и стандардизоване атрибуте, што поједностављује процес идентификације и управљања вишеструким приказима истог топонима на више картографских листова. Ово је посебно релевантно у контексту пројеката великог обима, као што је ТК25, где управљање бројним топонимима се може показати као сложено. Друго, методологија узима у обзир специфичне захтеве српског језика, који користи два писма, ћирилицу и латиницу. Интеграцијом оба писма у процес, методологија обезбеђује да се топоними могу тачно картографски приказати на ћирилици, уз задржавање латиничних атрибута који се користе у аналитичке и интероперабилне сврхе. Овај приступ је у складу са широм дискусијом о повезаности топонимије и идентитета (Woodman, 2014), препознајући важност очувања језичке и културне разноликости у картографском представљању. Као треће и пресудно у картографском контексту, примењена методологија омогућава стандардизован визуални приказ топонима на картографским издањима ВГИ чиме се повећава тачност и употребна вредност креираних карата.

Употреба објектне класе тачкастог типа геометрије, која се може експортирати у стандардном формату као што је Shapefile, омогућава се интероперабилност са другим софтверским платформама за визуализацију и просторну анализу. Међутим, за сложеније анализе које укључују семантичке односе између топонима, може бити неопходна даља трансформација података. Осим тога, овај приступ омогућава ефикасно складиштење и преузимање атрибута топонима, као што су назив, тип и слично. Овакав структурирани модел података омогућава сложеније анализе, укључујући и упите засноване на атрибутима, као и просторне анализе које укључују локацију топонима (Purves et al., 2019). Потенцијал за повезивање ових података о топонимима са спољашњим базама података додатно побољшава аналитичке могућности (Mitxelena-Hoyos & Amaro-Mellado, 2023), омогућавајући свеобухватније разумевање односа између топонима и других географских појава.

Приказана методологија представља значајан корак ка стандардизованијем и ефикаснијем приступу репрезентацији топонима у дигиталној картографији у оквиру ВГИ. Структурирани модел података, разматрање двојног писма и потенцијал за интеграцију са другим изворима података представљају вредне алате за управљање и анализу топонимских информација.

Будућа истраживања би могла да истраже развој аутоматизованих алата за конверзију писма, имплементацију јединствене националне базе података топонима и интеграцију ове методологије са другим иницијативама у истраживању топонимије, као што су оне које се фокусирају на динамичке системе именовања места (Mamontova & Klyachko, 2022). Решавањем ових изазова и ослањањем на постојећи оквир, дата методологија доприноси свеобухватнијем и поузданијем систему за представљање топонима. Њен флексибилан и прилагодљив дизајн, нарочито у визуелизацији топонима, омогућава прилагођавање нарастајућим картографским захтевима и додавање нових типова топонима без значајних измена основног модела података.

Закључак

Креиран садржаја назива насељених места односно топонима представља својеврсну базу података географских назива прилагођену за потребе картографског приказа у размери 1:25000, као и изведених размера. Такође, креирана база података поседује већу употребну вредност од искључиво картографског приказа јер у себи садржи додатне податке везане за саме називе насељених места који омогућавају комплексне и обухватније просторне анализе у ГИС контексту. На овај начин креирани географски називи испуњавају најстроже критеријуме картографског приказа и у великој мери задржавају визуални идентитет приказа топонима у односу на класичан односно аналогни начин визуелизације који се налазио у употреби у Војногеографском институту – „Генерал Стеван Бошковић“ све до почетка 21. века. Оваквим приступом је омогућена лака транзиција корисника картографских података са употребе традиционалних карата на карте новог издања које су креиране дигиталним методама и техникама, јер се у основи за њих ништа не мења. Избегнута је додатна обука корисника у коришћењу на овај начин креираних картографских публикација са аспекта приказа географских назива, што је само по себи био императив приликом дефинисања методологије. Обука корисника би због масовности захтевала изузетно велике кадровске, финансијске и временске ресурсе.

Додатни квалитет овог приступа моделовању географских назива представља и могућност вишег нивоа употребе података који је омогућен путем анализа које могу бити чисто атрибутске али које могу поседовати и просторну компоненту. Такође је омогућено повезивање креираних података који представљају географске називе са екстерним базама података чиме се потенцијално може проширити атрибутски капацитет спроведених анализа.

Постигнута је интерна стандардизација визуелизације назива насељених места, што је омогућило бржу израду картографског садржаја назива насељених места, подигнут је квалитет креираног садржаја и спроведена контрола квалитета путем домена кодних вредности.

Употребом објектне класе тачкастог типа, омогућена је делимична интероперабилност са другим софтверима који поседују могућност визуелизације одређене атрибутске вредности похрањене у бази података која се налази у основи тачкастог вектора. Овакав приступ је већ дао позитивне резултате у раду са другим софтверским платформама.

У претходном периоду се захтевало да приказ назива насељених места на штампаном отиску карте буде ћириличан док је била неопходна и могућност латиничног исписа на карти. Из овог разлога се у тачкастом типу геометрије топонима, као атрибут, уноси назив топонима у латиници док се у зависној објектној класи аногација тај испис ручно мења у ћирилични начин исписа. Овај приступ остаје актуелан, али утиче на брзину израде и повећава ризик од грешака у приказу и контроли квалитета топонима. Обзиром на развој, не строго географски оријентисаног софтвера у претходном периоду, овакав приступ би могао да се покаже непотребним и прелазак на једно писмо би имао значајан утицај на брзину израде и квалитет картографски приказаних топонима. С аспекта примењене методологије и технологије, овакав приступ не би имао утицај на успостављен технолошки процес креирања топографских карата али би довео до недоследности у односу на претходно креиран садржај назива насељених места.

Према територијалном захвату, база података која се користи обухвата податке у размери 1:25000 на националном нивоу и захвата простор Републике Србије са ужим појасом заграничне територије који су обухваћени листовима ТК25. Подаци су у оквиру базе података просторно дефинисани према појединачним листовима ТК25, што доводи до редунадансе топонима који се просторно позиционирају на ивицама листова ТК25 као и садржаја који просторно обухвата више листова карте ТК25. Ова чињеница говори да коришћена база података није у потпуности једнозначна и треба је узети у обзир приликом просторних анализа које обухватају простор који је обухваћен са више листова ТК25. Да би се креирани сет података географских назива, у конкретном случају објектних класа топонима, могао користити као својеврсна база података географских назива требало би да се напусти концепт строго картографске продукције односно, поделе на листове у размери 1:25000 како би се добила једнозначна база података без редунадансе података.

Превођење методологије коришћене за приказ назива насељених места у класичној картографији на методологију која ће се користити са базама података није био једноставан задатак. Ово је пре свега условљено, са ГИС аспекта, некомпатибилним моделом података коришћеним у аналогној картографији. Применом приказане методологије и према томе креираним моделом података, готово све категорије назива насељених места су унифициране и стандардизоване на начин који осликава манир традиционалног - аналогног картографског приказа назива насељених места. Поменути степен моделске неодређености се односи на категорије засеока и делова насеља за које нису квантитативно дефинисани цензуси картографског приказа. Њихова картографска интерпретација зависи од процене стручњака који опет има слободу интерпретације дефинисану ширим категоријама поменутих типова топонима. Број категорија топонима је одређен не само на основу картографских правила већ и на основу захтева за већом брзином израде картографског садржаја, услед чега је број категорија, поготово категорија села разбијеног типа, смањен у односу на потенцијалне могућности приказа. Дефинисан број категорија осликава компромис између могућности приказа, брзине израде и могућности грешке у класификацији топонима док се истовремено задржава задовољавајући ниво тачности и квалитета приказа назива насељеног места.

Током практичне реализације коришћен је ESRI софтвер ArcMAP у верзији 10.3.1. Пошто произвођач више не подржава ову верзију софтвера, миграција креи-

ране базе топонима на новију верзију софтвера (ESRI, ArcGIS Pro) може представљати потенцијалне изазове, али и нове могућности.

У овом раду је приказана методологија и начин креирања базе назива насељених места за потребе израде размерног низа топографских карата. Поред картографске продукције, на овај начин структурирани називи насељених места потенцијално могу да задовоље низ других потреба за визуализованим топонимима са обзиром на размеру приказа као и примењен модел података.

© 2024 Serbian Geographical Society, Belgrade, Serbia.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Serbia.

Литература (погледати у енглеској верзији текста)