
Дигитализација грађе Одбора за ономастику САНУ – значај, циљеви и први кораци*

Научни рад

DOI: 10.18485/judig.2025.1.ch13

*Снежана Петровић*¹,  0000-0003-1549-7294

*Мирјана Петровић-Савић*²,  0000-0001-6129-4741

*Ана Шпановић*³,  0000-0001-7368-9660

*Ленка Бајчетић*⁴,  0000-0002-2005-5426

*Матија Нешовић*⁵,  0009-0004-5822-6675

*Јована Тодорић*⁶  0009-0002-0706-6312

Апстракт

У раду су приказани значај, циљеви и први кораци у процесу дигитализације ономастичке грађе Одбора за ономастику САНУ, која садржи око 700 000 листића сакупљених са целокупног српског језичког простора у периоду од 1975. године до данас. Процес дигитализације укључује скенирање, документовање и моделирање различитих података записаних на листићима ономастичке грађе за потребе израде структуриране базе података и креирања корисничког интерфејса за приступ бази и уношење података. Приказан је поступак израде и имплементације интерфејса, а указано је и на проблеме и недоумице који су се јавили у овом процесу, као и на примењена

* Овај рад финансирало је Министарство науке, технолошког развоја и иновација Републике Србије према Уговору број 451-03-136/2025-03/ 200174, који је склопљен са Институтом за српски језик САНУ.

¹ Институт за српски језик САНУ, snezzanaa@gmail.com

² Институт за српски језик САНУ mirjana.petrovic@isj.sanu.ac.rs

³ Институт за српски језик САНУ tesicana@gmail.com

⁴ Иновациони центар Електротехничког факултета у Београду д.о.о., lenka.bajcetic@gmail.com

⁵ Институт за српски језик САНУ nesovic1998@gmail.com

⁶ Институт за српски језик САНУ jovanatodoric080@gmail.com

решења. Крајњи циљ дигитализације ономастичких записа јесте објављивање грађе у форми вишеструко претраживе платформе отвореног типа, доступне свима, како истраживачима тако и широј јавности. Платформа ће омогућити да се ономастички подаци организују, повежу и прикажу на више начина, укључујући различите врсте претрага и визуелизацију података у виду дигиталних мапа.

Кључне речи: српски језик, дигитализација, ономастика, израда интерфејса, визуелизација података.

1. Увод

У раду ће бити представљен пројекат дигитализације ономастичке грађе која припада Одбору за ономастику САНУ, као и први кораци у његовој реализацији. Оснивањем Одбора 1974. године створени су услови за систематска ономастичка истраживања на нашем језичком простору. Милан Будимир је био први председник Одбора, а након његове смрти (крајем 1975. године) па до 1981. године Одбором руководи Фанула Папазоглу, коју на том месту 1981. године наслеђује Павле Ивић (Стијовић С., 1997: 203). Рад на теренском бележењу српске ономастичке грађе иницирао је и осмислио академик Павле Ивић. У њима се први ангажовао и надаље имао главну улогу Светозар Стијовић сопственим сакупљачким радом и регрутацијом других сарадника (Петровић-Савић, 2023а: 604). Од 2004. на челу Одбора налази се академик Александар Лома, под чијим руководством су започети послови на дигитализацији ономастичке грађе.

Теренска истраживања започета су 1975. године на подручју Косова и Метохије. Прикупљени подаци из 652 насељена места, смештена у 25 општина, записани су на око 250 000 листића (Петровић-Савић, 2023а: 609).⁷ Истраживања су се истовремено одвијала и у другим деловима Србије, затим у Црној Гори, Хрватској и Босни и Херцеговини (Стијовић С., 1997: 204). Подаци сакупљени из тих области исписани су на око 500 000 листића, тако да данас картотека Одбора за ономастику САНУ укупно садржи око 700 000 листића.

На терену су прикупљане следеће врсте ономастичких података:

- а) ојконими (имена насељених места);

⁷ Део грађе са Косова и Метохије објављивана је у часопису *Ономастолошки прилози*: Божовић, 1983; Букумирић, 1981, 1982, 1984, 1988, 1997, 2011; Пумпаловић, 2015, 2016; Стијовић Д., 1997, 1998, 2002; Стијовић С., 1979, 1981, 1982, 1984; Ћирић, 1989; Цветановић, 1988; Џоговић, 1983, 1985, 1986, 1987, 1996.

Мотила
 Село се налази у богатској м-
 вотицај равници откопањој велики
 високих планина. Од Витице је уда-
 љено око 5 км. Село је веома старо, на
 селу је староседелачки, старобитни-
 цови. Једно је од највећих села у овом
 крају.
 Село има веома старо гробље,
 са мноштвом старих надгробних

Слика 1а

плота. У њему се више не сахра-
 њују. Зато гробље је Новијат гачу-
 ма.
 Црква св. Ђорђе је веома стара.
 Како је реновирана и обновљена
 се чува. Налази се на локалном
 путу бр. 1. Поред цркве је велика
 катапа и висока збонџа. Веома
 славе су Ђорђевица и Марковдан.
 Село има Пешчаразредну

Слика 1б

Основну школу. Леда да не остани
 школу уде у суседном селу Врбовцу.
 Из Мотила се од 1960. године одсе-
 лило 150 становника. Значај их,
 има 118. Двадесет њихових породица своје
 чорке везује за Турчиново, 7 за Тим-
 лине, 40 једно за село Трпаче и Вр-
 бовац, остали кажу да у чорак-
 лачи бивају.
 Прелевица бр. 1. Значови у чорковица
 је око 4,50.

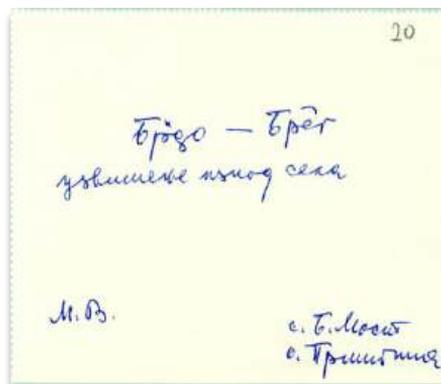
Слика 1в

б) етници (називи за становнике насељених места – мушки и женски) и ктетици (присвојни придеви изведени од имена места);

4
 Етници: Мотилџи; Мотилџица,
 ктетици: Мотилџа.
 Информатори: Петар Мотилџ (54),
 Радомир Мотилџ (68), Златко
 Мотилџ (68).
 с. Мотила
 о. Витица

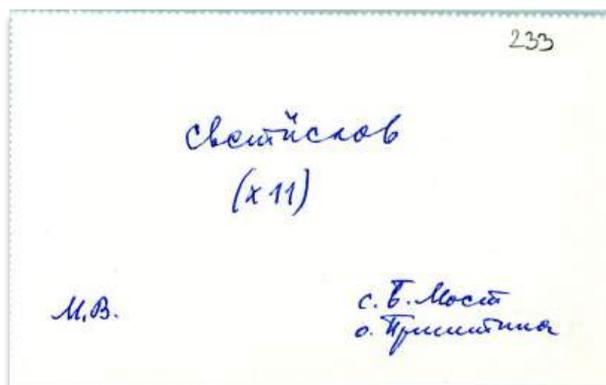
Слика 2

в) (микро)топоними;



Слика 3

г) антропоними (презимена, секундарна презимена, породични надимци, мушка лична имена, женска лична имена, мушки хипокористици, женски хипокористици, мушки лични надимци и женски лични надимци).



Слика 4

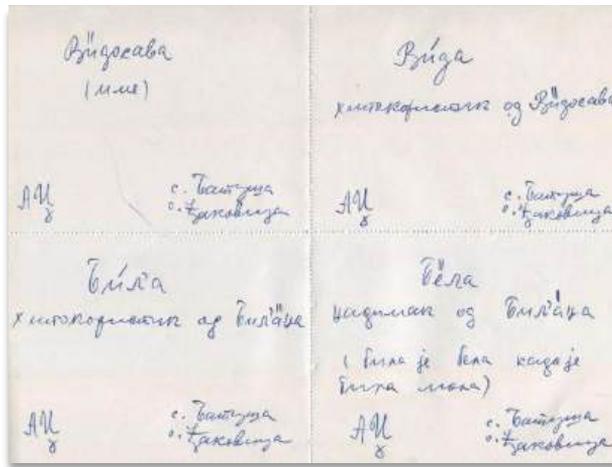
Битно је нагласити да нису у сваком пункту прикупљане све врсте података.⁸

⁸ Приликом бележења ономастичке грађе на терену у сваком посећеном пункту се, по правилу, бележе ојконими, етници, ктетици, микротопоними и патроними (уп. Стијовић & Гочанин, 2021: 294). Лична имена се не бележе у сваком селу. То је најчешће чињено у сваком петом селу једне области.

2. Процес дигитализације⁹

Први кораци у правцу дигитализације ономастичке грађе Одбора за ономастику САНУ начињени су пре петнаестак година, када је израђена база података и десктоп апликација за њихов унос.¹⁰

Наставак дигитализације састојао се из више фаза, од којих је прва скенирање целокупне грађе. Током 2020. године отпочеле су припреме листића за скенирање у Аудиовизуелном архиву и центру за дигитализацију САНУ, а приоритет је дат грађи са Косова и Метохије.¹¹ Будући да су ономастички подаци преко индига исписивани у блокове А4 формата са страницама подељеним на четири дела (Слика 5), листиће је најпре требало физички издвојити из ономастичких свезака, узбучити их и нумерисати. За ове послове у Одбору за ономастику САНУ било је ангажовано двоје хонорарних сарадника.¹²



Слика 5

⁹ Цео процес дигитализације обавља се уз финансијску и институционалну подршку САНУ, а у сарадњи са Аудиовизуелним архивом и центром за дигитализацију САНУ и Одсеком за дигитализацију Института за српски језик САНУ.

¹⁰ Сличан пројекат урађен је на Етимолошко-ономастичком одсеку Института за словеначки језик у Љубљани. Приређено је интернет издање рукописне збирке географских имена Макса Плетершника (Pleteršnikova rokopisna zbirka zemljepisnih imen, <https://przzi.zrc-sazu.si/>), која представља изузетно богато сведочанство о словеначкој топонимији на прелазу из 19. у 20. век. Грађа из те збирке унета је у базу у виду скенираних листића, засад без прекуцавања ономастичких података у њој (Петровић-Савић, 2023б: 203).

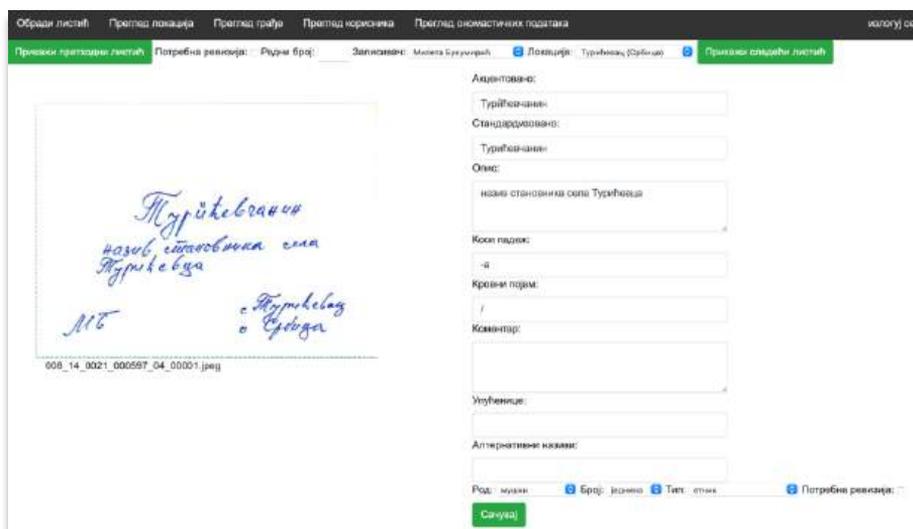
¹¹ У међувремену је комплетан материјал припремљен за скенирање.

¹² Почетком 2022. године окончано је скенирање ономастичке грађе са терена Косова и Метохије. Очекује се да овај посао за грађу из остатка Србије и појединих делова бивше Југославије буде завршен током 2026. године.

Током скенирања су одређени подаци који се налазе на листићу – име села и општине којој оно припада, као и име и презиме сакупљача (на листићу исписано иницијалима) – везани за појединачне скенове и унети у Ексел табелу.

Стара база и интерфејс за унос података, иако у основи добро осмишљени, због ограничених техничких капацитета и немогућности превођења у веб-окружење, нису могли да испуне захтеве пројекта.¹³ Стога је следећа фаза дигитализације подразумевала израду нове базе и веб-апликације за унос ономастичких података са модерним корисничким интерфејсом који омогућава вишекорисничко окружење за унос, преглед и кориговање ономастичких података, као и података везаних за локације (географске координате, измењени називи места и општина и др.). Корисничка апликација је имплементирана помоћу програмског језика *Python* и библиотеке *Django*. У базу су иницијално унети метаподаци из већ поменутог Ексел документа.

Интерфејс је унапређен и пружа више могућности за унос и повезивање података, а самим тим и за њихову претрагу. На Слици 6 види се тренутни изглед корисничког интерфејса.



Слика 6

На левој страни екрана налази се скенирани листић. Изнад њега су поља „Редни број”, „Записивач” и „Локација”. У поље „Редни број” уноси се број са листића који је уписан приликом припреме

¹³ Прву верзију базе података и интерфејса израдио је Брана Томић.

грађе за скенирање.¹⁴ Поља „Записивач” и „Локација” попуњавају се избором из падајућег менија, начињеног на основу метаподатака из Ексел документа.

На десној страни екрана налазе се поља за унос следећих података: акцентовани облик насловне речи, стандардизовани облик насловне речи, опис ономастичког податка, коси падеж, кровни појам¹⁵ и коментар¹⁶. Поља „Род”, „Број” и „Тип” попуњавају се одабиром одговарајућих опција из падајућег менија. Поље „Алтернативни називи” резервисано је за унос фонолошких, морфолошких и, евентуално, лексичких варијанти одређеног ономастичког податка забележеног на терену, док поље „Упућенице” омогућава једноставније повезивање реферишућих листића.

Поред ових поља, имамо још два истог имена: „Потребна ревизија”. Овако названо поље смештено у горњем левом делу екрана омогућава да се назначи да ли је потребна исправка података преузетих из Ексел документа и везаних за скенирани листић.¹⁷ За разлику од њега, поље смештено у доњем десном углу указује на потребу да се провере и/или измене унети ономастички подаци са скенираних листића.

3. Проблеми и решења

Велики број проблема проистиче из недоследног и неуниформног бележења података на терену. Упутства за сакупљање грађе нису могла бити толико детаљна да обухвате све ситуације на које се наилазило у пракси, те су проблеми и недоумице решавани различито од случаја до случаја. Због тога смо се определили да грађу у базу уносимо прекуцавањем са листића уместо коришћења ОЦР (Optical character recognition) технологије. Пречишћавање података, ручна коректура и парсирање би у садашњим околностима одузели више времена него прекуцавање у појединачна поља, посебно имајући у виду да је максимална тачност података за лингвистичка истраживања услов без којег она не могу бити валидна.

Један од изазова представља и то што се један ономастички податак не везује увек само за један листић. Сходно томе број листића неће одговарати броју ономастичких података, већ се очекује већи број података него листића. У наредним примерима представимо неке од

¹⁴ То је једини податак који се налази на скенираном листићу, а није забележен на терену.

¹⁵ В. проблеме 2 и 3 у наставку.

¹⁶ У ово поље истраживачи могу унети текст који се не налази на листићу.

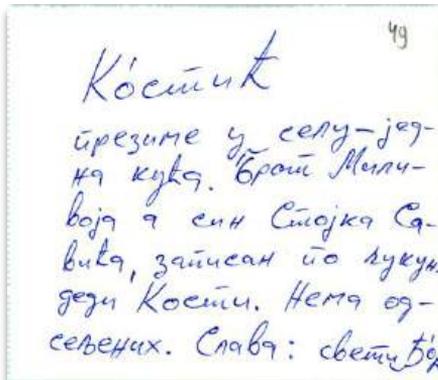
¹⁷ Може се, на пример, десити да је у Ексел документу унето погрешно име записивача теренске грађе, па је то неопходно изменити.

проблема с којима смо се сусрели, као и начине на које смо их разрешили.

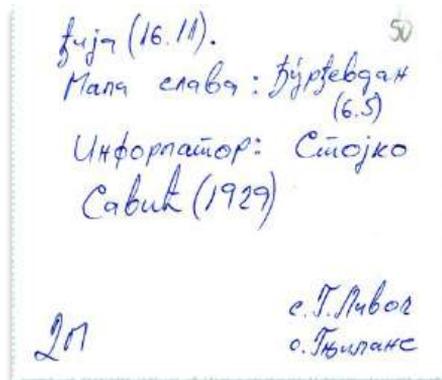
1. Један ономастички податак записан је на више листића (Слика 7а и 7б)

Грађа прикупљена на терену по правилу је уношена у ономастичке блокове тако што је на један листић уписиван један податак. Постоје, међутим, примери у којима су додатне информације уз један податак бележене на више листића (Петровић-Савић, 2023а: 612). То је углавном случај са информацијама о месту у којем се бележи ономастичка грађа (подаци о географском положају насеља у односу на његов административни центар, подаци о броју становника и његовом етничком саставу, о конфигурацији терена, о томе има ли место школу, цркву и слично). Понекад се податак код презимена исписује на више листића и укључује информације о броју породица које то презиме носе, њиховој крсној слави, одакле су досељени и тако даље.

У случајевима овога типа од више листића прави се један ономастички податак.



Слика 7а



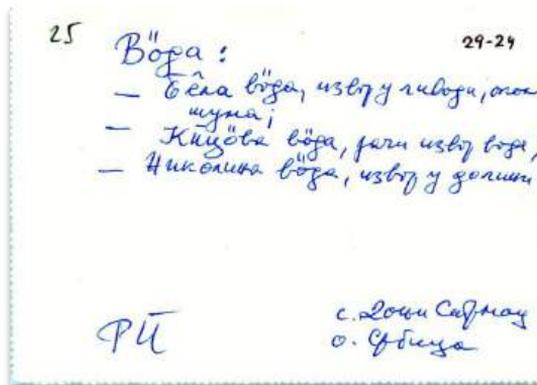
Слика 7б

2. На једном листићу налази се више потпуних ономастичких података (Слика 8)

У примеру на слици насловна реч *Вџда* нема опис, што значи да на терену не постоје ни једночлани хидроним, нити топоним *Вџда*, већ само двочлани називи синтагматског типа са лексемом *вџда* (*Бџла вџда*, *Кџџџа вџда*, *Николина вџда*).¹⁸

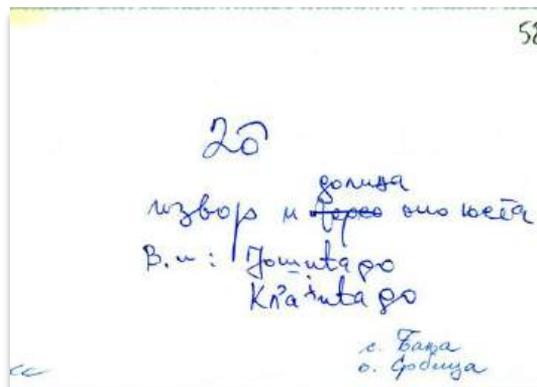
¹⁸ Исто важи за све насловне речи које немају опис, без обзира на то да ли су акцентоване или не.

У овом случају направљена су три ономастичка податка од једног листића: *Бела вода*, *Кљџова вода*, *Николина вода*, а насловно *Вода* уписано је у поље „Кровни појам” уз сваки податак. Да је уз наслов одреднице стајао опис (нпр. *Вода* „извор на одређеном локалитету”), од датог листића био би оформљен и додатни, четврти ономастички податак (уп. Проблем бр. 3).



Слика 8

3. На једном листићу налази се један комплетан ономастички податак и више упућеница (Слика 9)



Слика 9

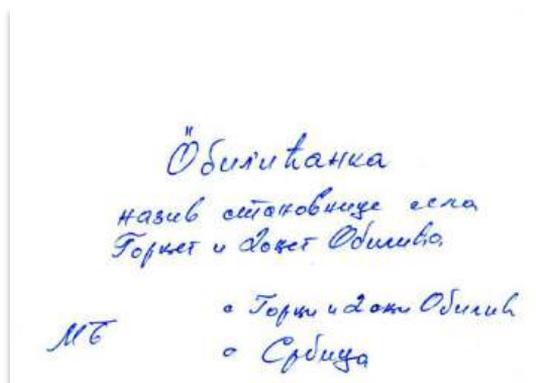
Код овог типа проблема насловна реч не представља кровни појам за вишечлане микротопониме чији је део, већ је самосталан ономастички податак са описом. Уз њега је записивач навео и друге

податке с напоменом „в. и” („види и”), чиме се указује на то да ће се они јавити и самостално, на посебном листићу.

У случају као на слици, микропоним *Дб* обележен је као основни ономастички податак, док су називи *Јошића до* и *Кл'а'ића до* уписани у поље „Упућенице” и биће накнадно обрађени као засебни ономастички подаци када се њихови листићи буду појавили у апликацији.

4. Подаци из две локације записани су на једном листићу (Слика 10)

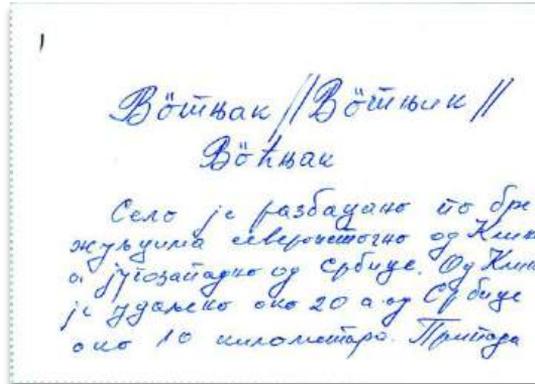
У случају када су на једном листићу записани ономастички подаци из два различита села, формирају се два одвојена ономастичка податка како би се податак потврђен у двама локацијама (у примеру на слици *Обил'ићанка*) при претрази појавио у обе локације, чиме би се добио потпун резултат.



Слика 10

5. Насловна реч има више фонолошких варијанти (Слика 11)

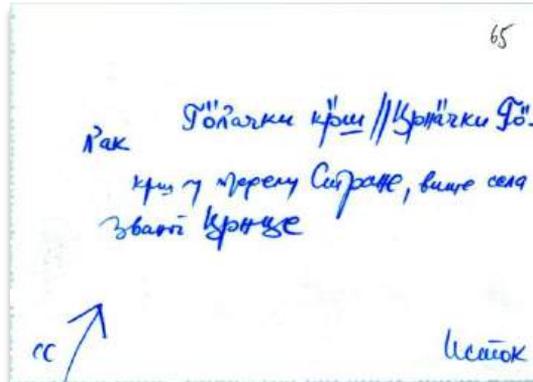
У случају када је на терену забележено више фонолошких варијанти насловне речи, у бази се формира један ономастички податак. Као примарни облик бира се онај којем је записивач дао предност при навођењу (на слици облик *Вџтњак*), док се остале фонолошке варијанте (на слици *Вџтњик* и *Вџћњак*) смештају у поље „Алтернативни називи”. На тај начин ће корисницима бити омогућено да претрагом добију увид у све фонолошке варијанте насловне речи.



Слика 11

6. Локалитет има више различитих назива (Слика 12)

У случају да локалитет на терену има више од једног назива (на пример, микротопоним на слици забележен је као *Гол'ачки крш* и *Црн'ачки Гол'ак*) формирају се два одвојена ономастичка податка. Они се у апликацији обавезно везују и упућују један на други тако што се други податак уписује у поље „Упућенице”.

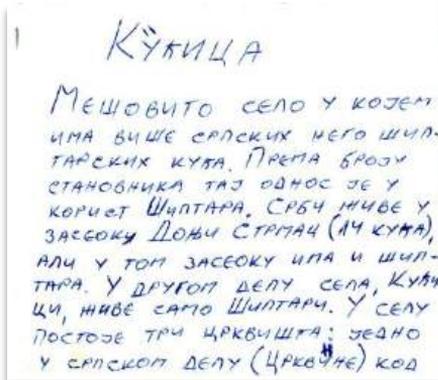


Слика 12

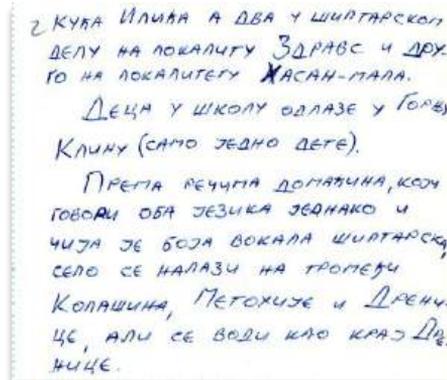
7. Недоследан начин бележења етника и ктетика (Слике 13а, 13б и 13в)

У инструкцијама за сакупљање грађе од записивача је тражено да бележе етнике и ктетике изведене од назива посећеног места, али не и од заселака који му припадају. Међутим, ово правило није увек поштовано, као што је случај у примерима на слици. Записивач је забележио два ктетика: *Кућичанин*, становник села Кућица, и

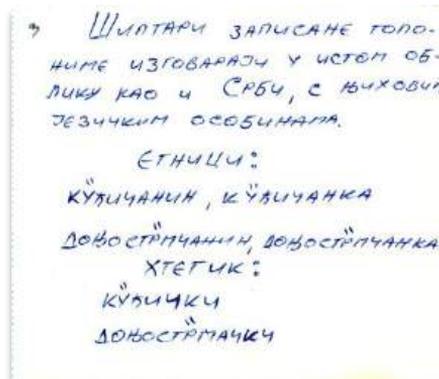
Доњострџмчанин, становник његовог засеока Доњи Стрмац. У оваквим случајевима, у апликацији ће се за сваки од ових назива формирати посебан ономастички податак.



Слика 13а



Слика 13б



Слика 13в

4. Закључак

У раду су приказани први кораци у дигитализацији ономастичке грађе и неки од проблема с којима смо се суочили током рада на њој. Наредне фазе дигитализације обухватиће даље унапређење апликације за унос података и уношење у базу свих података са Косова и Метохије, а затим и скенирање, обраду и унос преостале ономастичке грађе. Један од главних резултата пројекта биће визуелизација података на мапи у веб-апликацији, који ће претходно бити геолоцирани. Апликација ће имати кориснички

интерфејс који ће омогућити да обрађени подаци буду представљени на прегледан начин, као и ефикасну претрагу ономастичких података и листића. Рад на неким од будућих корака одвијаће се истовремено.

Оставља се могућност накнадног увођења ОЦР-а за рашчитавање података са листића, чему би морала претходити озбиљна анализа која би показала да ли би ова технологија уопште допринела ефикасности, брзини и прецизности уноса ономастичких података (имајући у виду различите рукописе бројних сакупљача, недоследности и грешке које су последица људског фактора и сл.).

Крајњи циљ овог пројекта јесте стварање свеобухватног дигиталног ономастикона српских народних говора који ће корисницима омогућити да грађу претражују по више различитих параметара. С обзиром на то да ће овај ресурс бити јавно доступан, послужиће широј публици као поуздан извор информација о нашим географским и личним именима и истовремено представљати драгоцен материјал за научна истраживања стручњацима различитих профила.¹⁹

Литература

- [1] Божовић, М. (1983). Ономастика села Бановог Дола код Звечана. *Ономатолошки прилози, IV*, 453–467.
- [2] Букумирић, М. (1981). Ономастика дела Прекорупља омеђеног рекама Мирушом, Белим Дримом и Клином. *Ономатолошки прилози, II*, 339–404.
- [3] Букумирић, М. (1982). Ономастика Прекорупља (II део). *Ономатолошки прилози, III*, 303–418.
- [4] Букумирић, М. (1984). Из ономастике јужне Метохије. *Ономатолошки прилози, V*, 413–603.
- [5] Букумирић, М. (1988). Ономастика Сиринићке жупе. *Ономатолошки прилози, IX*, 253–470.
- [6] Букумирић, М. (1997). Ономастика јужнога Косова. *Ономатолошки прилози, XIII*, 339–495.
- [7] Букумирић, М. (2011). Из ономастике Липљанске околине. *Ономатолошки прилози, XXI*, 47–190.
- [8] Лома, А. (2013). *Топонимија Бањске хрисовуље*. Београд: Српска академија наука и уметности.

¹⁹ Ономастичка грађа забележена на терену вишеструко је корисна за дијахронијски и синхронијски усмерена лингвистичка истраживања, као и за историјске, етнографске, културолошке и друге студије. У својим истраживањима историјске топонимије Александар Лома (2013) и Митар Пешикан (1981; 1982; 1983; 1984; 1986) обилаго су користили податке сакупљене за Одбор са подручја Косова и Метохије.

- [9] Петровић-Савић, М. (2023а). Косовско-метохијска грађа при Одбору за ономастику САНУ: преглед обрађених и преосталих пунктова, послови скенирања, дигиталног уноса и интернет презентације. У: Д. Војводић (ур.), *Заштита, очување и афирмација српског културног наслеђа на Косову и Метохији, II* (стр. 603–624). Београд: Српска академија наука и уметности.
- [10] Петровић-Савић, М. (2023б). Pleteršnikova rokopisna zbirka zemljepisnih imen. *Зборник Матице српске за филологију и лингвистику, LXVI/1*, 203–205.
- [11] Пешикан, М. (1981). Из историјске топонимије Подримља. *Ономатолошки прилози, II*, 1–92.
- [12] Пешикан, М. (1982). Зетско-хумско-рашка имена на почетку турскога доба. *Ономатолошки прилози, III*, 1–120.
- [13] Пешикан, М. (1983). Зетско-хумско-рашка имена на почетку турскога доба. *Ономатолошки прилози, IV*, 1–135.
- [14] Пешикан, М. (1984). Зетско-хумско-рашка имена на почетку турскога доба. *Ономатолошки прилози, V*, 1–182.
- [15] Пешикан, М. (1986). Стара имена из доњег Подримља. *Ономатолошки прилози, VII*, 1–119.
- [16] Пумпаловић, Р. (2015). Ономастика Бањске и Звечана. *Ономатолошки прилози, XXII*, 599–738.
- [17] Пумпаловић, Р. (2016). Ономастика једног дела Ибарског Колашина. *Ономатолошки прилози, XXIII*, 369–511.
- [18] Стијовић, Д. (1997). Ономастика Суве Реке и околних села. *Ономатолошки прилози, XIII*, 497–573.
- [19] Стијовић, Д. (1998). Ономастика дела северне Метохије. *Ономатолошки прилози, XIV*, 157–256.
- [20] Стијовић, Д. (2002). Ономастика шест села у околини Витине. *Ономатолошки прилози, XV*, 227–263.
- [21] Стијовић, С. (1979). Ономастика источног дела Метохијског (Пећког) Подгора. *Ономатолошки прилози, I*, 237–380.
- [22] Стијовић, С. (1981). Ономастика средишњег дела Метохијског (Пећког) Подгора. *Ономатолошки прилози, II*, 197–338.
- [23] Стијовић, С. (1982). Ономастика западног дела Метохијског (Пећког) Подгора. *Ономатолошки прилози, III*, 193–300.
- [24] Стијовић, С. (1984). Ономастика једног дела поречја Кујавче. *Ономатолошки прилози, V*, 313–411.

- [25] Стијовић, С. (1997). Развој српске ономастичке науке у последњој четвртини века. *Јужнословенски филолог*, *LIII*, 203–210.
- [26] Стијовић, Р., Гочанин, М. (2021). Савремена ономастичка истраживања Косова и Метохије. У: Р. Младеновић, П. Пипер (ур.), *Резултати досадашњих и правци будућих истраживања српских народних говора Косова и Метохије* (стр. 291–309). Београд – Косовска Митровица: Српска академија наука и уметности – Институт за српски језик САНУ – Филозофски факултет.
- [27] Ђирић, Љ. (1989). Ономастика Изморника. *Ономатолошки прилози*, *X*, 365–496.
- [28] Цветановић, В. (1988). Из ономастике Гњилана. *Ономатолошки прилози*, *IX*, 517–542.
- [29] Џоговић, А. (1983). Топонимија југословенског дела Проклетија: (I део). *Ономатолошки прилози*, *IV*, 225–344.
- [30] Џоговић, А. (1985). Ономастика Дечана и околних села. *Ономатолошки прилози*, *VI*, 439–508.
- [31] Џоговић, А. (1986). Ономастика североисточних огранака Проклетија. *Ономатолошки прилози*, *VII*, 247–436.
- [32] Џоговић, А. (1987). Ономастика Ђаковице и суседних села. *Ономатолошки прилози*, *VIII*, 221–304.
- [33] Џоговић, А. (1996). Ономастика Горе. *Ономатолошки прилози*, *XII*, 33–366.

Digitization of the Sanu Onomastic Committee's Onomastic Material—Importance, Goals and First Steps

Snežana Petrović, Mirjana Petrović Savić, Ana Španović, Lenka Bajčetić, Matija Nešović, Jovana Todorčić

Summary

The paper presents the significance, objectives, and initial steps in the process of digitizing the onomastic material of the Committee for Onomastics of the Serbian Academy of Sciences and Arts (SASA). The material comprises approximately 700,000 paper slips with data collected from the Serbian linguistic area, dated from 1975 to the present. The digitization process includes scanning, documenting, and subsequently modelling various data recorded on the paper slips to create a structured database and develop a user-friendly interface for accessing and entering data. The paper also outlines the procedure for designing and implementing the interface, highlighting the challenges encountered in this process and the applied solutions. The ultimate goal of digitizing onomastic records is to publish the material in the form of a multi-searchable, open-access platform available to both researchers and the general public. This platform will enable the presentation of onomastic data in multiple ways, including various search options and data visualization in the form of digital maps.

Keywords: Serbian language, digitization, onomastics, interface development, data visualization.