

Interaktivna karta slovenskih narečnih besedil

Alenka Kavčič,* Ivan Lovrić,* Vera Smole**

* Fakulteta za računalništvo in informatiko, Univerza v Ljubljani
Večna pot 113, 1000 Ljubljana
alenka.kavcic@fri.uni-lj.si, ivan@lovric.si
** Filozofska fakulteta, Univerza v Ljubljani
Aškerčeva 2, 1000 Ljubljana
vera.smole@ff.uni-lj.si

Povzetek

Sedem narečnih skupin, ki združujejo sedemintrideset narečij, uvršča slovenski jezik med narečno najbolj razčlenjene evropske jezike. Da to bogastvo jezika ne bi ostalo skrito in dostopno le jezikoslovcem in narečjeslovcem, smo izdelali spletno interaktivno karto za slovenska narečna besedila. Aplikacija omogoča prikaz posameznih krajev s posnetim narečnim govorom na interaktivnem zemljevidu, z izbiro kraja pa lahko poslušamo posnetek narečnega govora, si ogledamo njegovo fonetično transkripcijo in poknjžitev, dodana pa je tudi diahrona analiza govora. Aplikacija vključuje tudi skrbniški del, ki prijaviteljem uporabnikom omogoča urejanje obstoječih vsebin in dodajanje novih vnosov. Z željo po uporabniku prijaznem in enostavnem uporabniškem vmesniku smo pri izdelavi aplikacije uporabili uveljavljena spletna orodja in rešitve.

Interactive map of Slovenian dialectal texts

Seven dialect groups combining 37 dialects, rank Slovenian language among the European languages with the greatest dialectal diversity. To offer this richness of the language, now available mostly to linguists and dialectologists, to wider public, an interactive web-based map of Slovenian dialectal texts has been developed. The application shows individual places with the recorded dialectal texts on the interactive map, while selecting a specific place enables listening to the sound recording of the dialectal speech, displays its phonetic transcription and translation to literary language as well as the diachronic analysis of the speech. The application includes an administrative part that enables logged-in users editing of the existing contents as well as adding new entries. Aiming to achieve user friendly and easy to use user interface, the application was developed using conventional web tools and solutions.

1. Uvod

Slovenščina je narečno zelo bogat jezik. S 37 narečji, ki jih združujemo v 7 narečnih skupin, se slovenski jezik uvršča med najbolj razčlenjene evropske jezike.

Čprav se narečjeslovje kot samostojna veda razvija že od druge polovice 19. stoletja (Toporišič, 1987), so raziskovalni rezultati še vedno dostopni pretežno v tiskanih izdajah monografij, (zbirk) člankov, narečnih slovarjev in drugih tematskih publikacij, pri čemer so izjema Slovenski lingvistični atlas (1 in 2) in nekateri narečni slovarji, vključeni v spletni portal Fran.¹ Za večjo prepoznavnost in približanje tematike tudi mlajšim, digitalnim generacijam je zato nujno potrebno izkoristiti možnosti sodobnih informacijsko-komunikacijskih tehnologij, predvsem interneta in svetovnega spleta kot glavnega medija za dostop do informacij.

V tem prispevku bomo opisali spletno aplikacijo za prikaz interaktivne karte slovenskih narečnih besedil, ki je nastala v okviru diplomskega dela na Fakulteti za računalništvo in informatiko Univerze v Ljubljani (Lovrić, 2018). Aplikacija za podane kraje, označene na karti, omogoča predvajanje zvočnih posnetkov narečnih besedil ter prikaz njihove fonetične transkripcije, poknjžitve (zapis z grafemi in fonetiko knjižnega jezika ter oblikoslovnimi in leksičnimi knjižnimi ustreznici v pomenskih oklepajih) in analize večine besedil. Gradivo in analize so rezultat raziskovalnega dela pri dveh predmetih

na Oddelku za slovenistiko Filozofske fakultete Univerze v Ljubljani (Smole in Horvat, 2016; Smole, 2016), vsa besedila pa so na temo stara kmečka hiša (prostori in oprema v njih). Aplikacija vključuje tudi administrativni del, ki omogoča dodajanje novih vsebin in urejanje obstoječih.

2. Sorodni pristopi

Za slovenski jezik nismo našli nobene podobne aplikacije, tematsko še najbližja je elektronska oblika Slovenskega lingvističnega atlasa (e-SLA²), ki pa je namenjen bolj narečjeslovcem kot širši publiki.

Nekoliko bližji naši ideji spletne aplikacije je Zemljevid narečij bolgarskega jezika,³ ki ga je izdelal Oddelek za dialektologijo in jezikovno geografijo Inštituta za bolgarski jezik in je prvi elektronski interaktivni zemljevid bolgarskih narečij. Zemljevid prikazuje glavne narečne skupine celotnega jezikovnega ozemlja in predstavlja vsa glavna narečja in njihove meje. Po zemljevidu so razporejene ikone, ki ob kliku prikažejo krajšo analizo govora (ikona knjige) ali omogočijo predvajanje zvočnega zapisa narečnega govora (ikona zvočnika). Aplikacija omogoča tudi dopolnjevanje karte z novimi zvočnimi zapisi.

Ena bolj zanimivih obstoječih rešitev je tudi Zvočna karta naglasov in narečij Velike Britanije,⁴ ki je dostopna na spletnih straneh The British Library. Za prikaz podatkov uporablja zemljevide Google Maps, a na karti so dostopni

¹ Fran, Slovarji Inštituta za slovenski jezik Frana Ramovša ZRC SAZU, dostopno na <http://www.fran.si/>.

² Slovenski lingvistični atlas v obliki html, dostopno na http://sla.zrc-sazu.si/eSLA/Zavihki_na_eSLA_JS.html.

³ Zemljevid narečij bolgarskega jezika, dostopno na http://ibl.bas.bg/bulgarian_dialects/.

⁴ Zvočna karta naglasov in narečij Velike Britanije, dostopno na <https://sounds.bl.uk/Sound-Maps/Accents-and-Dialects>.

le zvočni zapisi brez besedil ali analiz govora. Uporabniška izkušnja ni najboljša, saj se predvajalnik zvoka ne odpre v modalnem oknu, ampak uporabnika ob kliku na zvočni zapis preusmeri na drugo spletno stran.

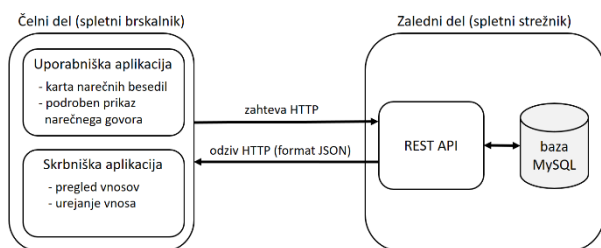
Podobno temelji na Google Maps tudi Zvočni atlas hrvaških govorov,⁵ a se njegova ideja nekoliko razlikuje od naše aplikacije, saj gre za glasoslovni atlas, katerega namen je predstaviti določene foneme v kratkem besedilnem okolju. Na zemljevidu prikazane ikone za zvočne posnetke primerov narečnih govorov tako omogočajo poslušanje vedno le vnaprej izbranega stavka z določeno fonološko strukturo.

3. Interaktivna karta slovenskih narečnih besedil

V tem poglavju bomo opisali izdelano interaktivno aplikacijo, začenši z zasnovo in zgradbo aplikacije, navedbo uporabljenih spletnih tehnologij ter opisom priprave karte narečnih besedil v ustreznem formatu.

3.1. Zasnova in zgradba aplikacije

Spletno aplikacijo sestavljata čelni in zaledni del, kot je to shematično prikazano na sliki 1.



Slika 1: Zgradba aplikacije.

Čelni del se izvaja v spletnem brskalniku in predstavlja tisti del aplikacije, ki ga vidi uporabnik (neprijavljen uporabnik pri ogledu narečne karte in govora posameznega kraja ali prijavljen skrbnik pri pregledu in urejanju vnosov). Zajema tako uporabniško kot tudi skrbniško aplikacijo (obe sta opisani v nadaljevanju). Izdelava tega dela aplikacije je temeljila na sodobnih spletnih tehnologijah in ogrodjih, kot so HTML5, CSS, JavaScript, AngularJS, Bootstrap, za prikaz karte pa smo uporabili knjižnico Leaflet, ki temelji na prosto dostopnih zemljevidih OpenStreetMap.

Zaledni del aplikacije se izvaja na spletnem strežniku in zajema vmesnik za dostop do podatkovne baze MySQL, v kateri so shranjeni vsi podatki. Vmesnik temelji na arhitekturi REST in je v celoti napisan v jeziku PHP. Vmesnik tako sprejema zahteve HTTP preko metod GET, POST, PUT in DELETE ter vrača odzive operacij v formatu JSON. Dostop do podatkov v bazi je možen le preko vmesnika s pomočjo zahtev HTTP, kar velja tako za uporabniško kot tudi za skrbniško aplikacijo.

3.2. Priprava interaktivne karte narečij

Pri izdelavi aplikacije smo posebno pozornost namenili enostavni in intuitivni uporabi karte, saj je aplikacija namenjena širšemu krogu končnih uporabnikov, ki niso vedno vešči uporabe računalniških programov. Sprva smo nameravali našo karto graditi na dobro poznanem in pogosto uporabljanem zemljevidu Google Maps (z uporabo vmesnika Google Maps API), preko katerega bi izrisali slovenska narečna območja. Vendar smo se zaradi zaprtosti in posledično manjše fleksibilnosti vmesnika Google Maps API na koncu odločili za uporabo odprtokodne Javascriptovne knjižnice Leaflet in prosto dostopnih zemljevidov OpenStreetMap.

Osnova za izdelavo zemljevida z vrisanimi narečnimi območji je bila Karta slovenskih narečij Tineta Logarja in Jakoba Riglerja (1983), ki je nastala na osnovi Ramovševe (1931) *Dialektološke karte slovenskega jezika* in bila še večkrat dopolnjevana. Najprej smo na osnovi Karte slovenskih narečij izdelali vektorski izris zemljevida slovenskih narečij ter ga shranili v formatu svg. Nato smo vektorsko grafiko še geokodirali, in sicer s pretvorbo v format GeoJSON, ki vsebuje geografske podatke za vsak grafični element karte. Na koncu smo geokodirano karto uvozili v aplikacijo s pomočjo knjižnice za izdelavo interaktivnih spletnih zemljevidov Leaflet.

Tako pripravljena karta nam omogoča več interaktivnosti v aplikaciji, saj lahko dinamično spreminjamo oblikovne stile narečnih območij in mej med njimi, odziva pa se tudi na uporabniške vnose in premike z miško, kot je npr. lebdenje z miškinim kazalcem nad posameznim narečjem in podobno.

3.3. Uporabniška aplikacija

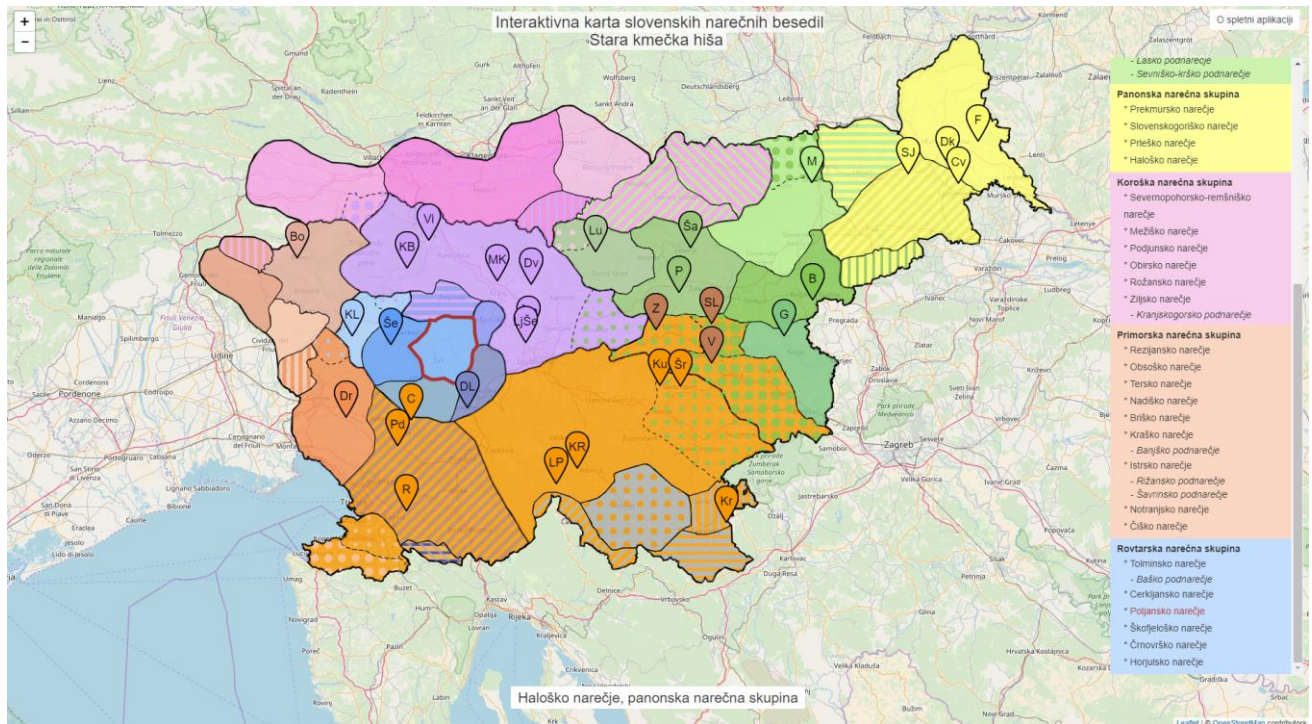
Uporabniški del aplikacije prikazuje spletni zemljevid,⁶ na katerem so barvno označena vsa slovenska narečja, podnarečja in tudi narečne skupine. Ob zemljevidu je tudi legenda z navedenimi narečji in podnarečji za vsako od sedmih narečnih skupin.

Uporabniški pogled na karto je prikazan na sliki 2. Na zemljevidu so označena narečna območja, kjer je vsaka narečna skupina označena z določeno barvo, posamezna narečja in podnarečja pa še z dodatnimi grafičnimi simboli, pikami ali poševnimi črtami, ki ponazarjajo vplive drugih (pod)narečij. Karta namreč z izbranimi barvami, šrafurami in vzorci vsebuje tudi vizualne informacije o tem, kako se na nekaterih območjih prepletajo narečja in podnarečja.

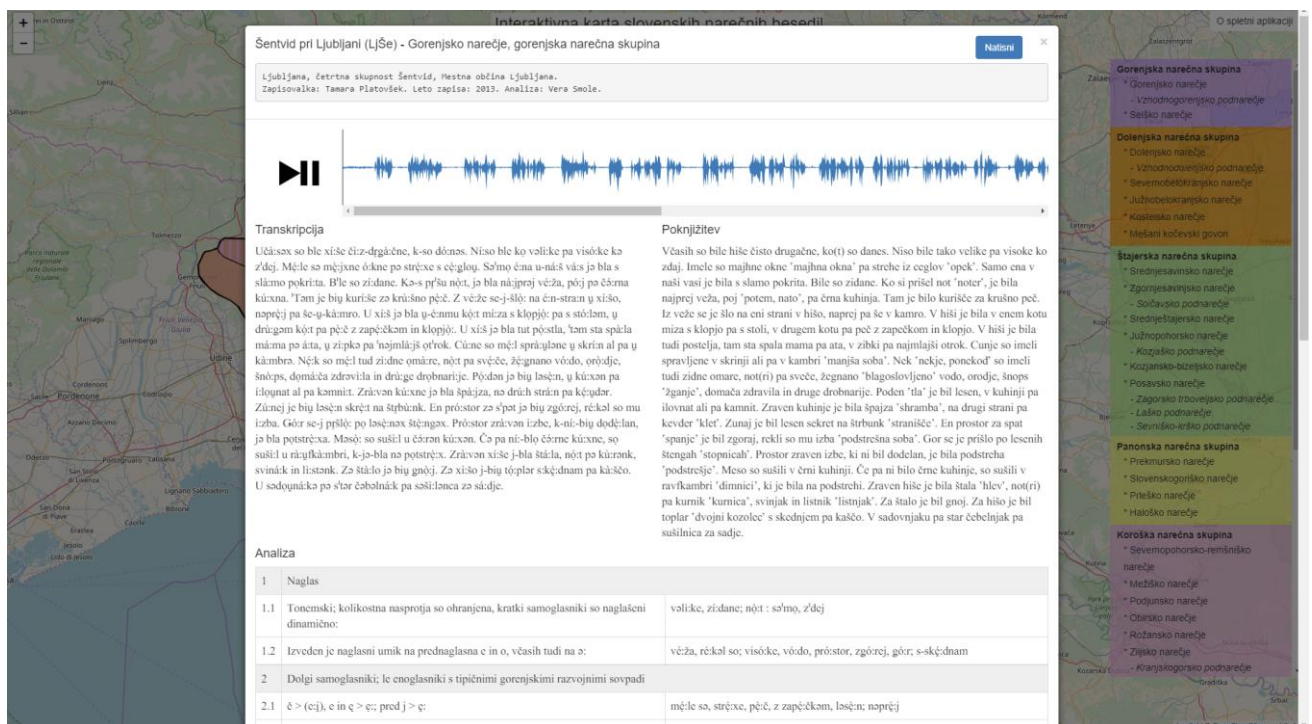
Na desni strani je prikazana tudi legenda z izpisanimi vsemi narečji in podnarečji (poševni tisk), ki so združena v narečne skupine (krepki tisk). S prehodom z miško čez legendo se na karti označi ustrezno narečno območje (primer na sliki 2 prikazuje označeno poljansko narečje, ki je v legendi izpisano rdeče, na karti pa označeno z debelejšo rdečo obrobo). Aplikacija omogoča tudi poljubno približevanje in premikanje po zemljevidu. To je posebej praktično v primeru, ko imamo na manjšem področju označenih več krajev z narečnimi posnetki in s približevanjem lažje ločimo med posameznimi posnetki.

⁵ Zvočni atlas hrvaških govorov, dostopno na <http://hrvatski-zvucni-atlas.com/>.

⁶ Interaktivna karta slovenskih narečnih besedil je dostopna na <http://narecja.si>.



Slika 2: Uporabniški del interaktivne karte.



Slika 3: Pojavno okno z narečnim besedilom za izbrano lokacijo.

Na zemljevidu so dodane še oznake v krajih, kjer so bili posneti zvočni zapisi. Posamezni kraj oz. postavitev oznake je določena z geografskimi koordinatami (geografsko dolžino in širino) kraja. S klikom na to oznako se odpre pojavno okno, prikazano na sliki 3, v katerem lahko

uporabnik predvaja zvočni zapis narečnega govora, prebere fonetično transkripcijo tega zapisa, njegovo poknjížitev na glasoslovni ravni ter velikokrat tudi narečjeslovno analizo govora z vidika značilnosti na sedmih jezikovnih ravneh narečnega govora (naglas, dolgi samoglasniki,

kratki naglašeni samoglasniki, kratki nenaglašeni samoglasniki, soglasniki, oblikoslovni pojavi in leksika). Zaradi boljše uporabniške izkušnje se pri predvajanju zvočnega posnetka nismo omejili le na funkciji *predvajaj* in *ustavi*, ampak smo zvok tudi vizualizirali z valovnimi oblikami in omogočili premikanje naprej in nazaj po posnetku.

V pojavnem oknu je dodana tudi možnost za tiskanje vsebine okna, to je transkripcije, poknjžitve in analize, skupaj s podatki o kraju in narečju (privzeto se vsi ti podatki shranijo v datoteko pdf na lokalnem računalniku). Le zvočnega posnetka narečnega govora ne moremo izvoziti iz aplikacije in lokalno shraniti.

3.4. Skrbniška aplikacija

Skrbniška aplikacija zahteva najprej avtentikacijo uporabnika, saj je dostop do urejevalnika vsebine omejen. Ob uspešni prijavi se prikaže seznam vseh krajev, urejen abecedno (slika 4, levo okno). Vsak kraj lahko uredimo (spremenimo podatke) ali izbrisemo, lahko pa vnesemo tudi povsem nov kraj.

Pogled za urejanje vnosa prikazuje slika 4 (desno okno). Sestavljajo ga trije deli: Osnovni podatki, Besedilo in Analiza. Prvi del zajema podatke o lokaciji, kjer je bil posnet zvočni zapis narečnega govora: ime kraja, njegove

zemljepisne koordinate (širino in dolžino), v katerih se na karti prikaže oznaka tega kraja, in oznako, to je kratico, ki se prikaže v oznaki kraja na karti. S spustnega seznama izberemo narečje oziroma podnarečje, kamor govor umeščamo, ter naložimo zvočni zapis govora v formatu mp3, ki se prenese in shrani na strežniku. Dodamo lahko tudi poljubne metapodatke, ki navadno vključujejo dodatne opise kraja, podatke o zapisovalcu in informatorju, leto zapisa ali katerekoli druge pomembne podatke.

V drugem delu sta dve polji za vnos besedila: transkripcija in poknjžitev. Transkribirano besedilo je vedno zapisano v pisavi ZRCola (Weiss, 2004) in ga lahko v aplikacijo kopiramo iz urejevalnika besedila, ki podpira vnašalni sistem ZRCola. Zaradi grafične doslednosti smo isto pisavo uporabili tudi v polju za poknjžitev, čeprav ta ne uporablja posebnih fonetičnih znakov. Poknjženo besedilo lahko vnesemo neposredno v tekstovno polje v aplikaciji ali pa ga kopiramo iz urejevalnika besedila, v katerem smo ga predhodno pripravili.

Tretji del pa je namenjen vnašanju narečjeslovne analize narečnega govora. Analiza je sestavljena iz sedmih sekcij; vsaka opisuje značilnosti narečnih govorov na eni od jezikovnih ravnin in vsaka vsebuje poljubno število vnosov s primeri iz besedila, ki jih lahko dodajamo sproti.

The screenshot shows the 'Urejevalnik narečnih besedil' application. On the left is a list of regions with columns for '#', 'Kratnica', 'Kraj', and 'Urejanje'. The main area is divided into three sections: 'Osnovni podatki', 'Besedilo', and 'Analiza'.

Osnovni podatki: Includes fields for 'Oznaka' (Ljše), 'Lokacija' (Šentvid pri Ljubljani), 'Narečje oz. podnarečje' (Gorenjsko narečje), 'Zemljepisna širina in dolžina' (46.097049, 14.465839), 'Metapodatki' (Ljubljana, četrtna skupnost Šentvid, Hestna občina Ljubljana), and 'Zvočni zapis (priporočeno: mp3)' (Naložena datoteka: /resources/audio/24.mp3).

Besedilo: Contains 'Transkripcija' (Učásax so ble xiše čiz-drgáčne, k-so dómas. Níso ble kó valíke pa visóke ka z'dej. Měle sa mějxne ókne pa strěxse s cěglou. Sa 'mo éna u-náš vás jo bla s slámo pokrita. B'le so zídane. Ka-s pršú nóť, jo bla nájpraj véža, pój pa čárna kúxna. 'Tam je biú kuríše za krušúno peč. Z véže so-j-sló: na é-n-strán y xišo, naprěj pa še-y-kámro. U xiš jo bla y-čnúmu kóť míza s klopjé: pa s stólam, y drúgam kóť pa peč z zapéčkom in klopjé. U xiš jo bla tut póstla, 'tam sta spála máma pa áta, y zítka pa 'najmlájši ótrok. Cúne so méj sprájlone y skrín al pa y kámbrá. Něk so méj tud zídne ómáre, nóť pa svéčce, žé-gnano vódo, óródjje, šnóps, domáča zdróvítla in drúge dróbnarije. Pódan jo biú lašén, y kúxan pa ílqnan al pa kómmit. Zrávon kúxne jo bla

Analiza: A table with columns 'Sekcija', 'Opis', and 'Primeri'. It shows a section for 'Naglas' with a red 'x' icon and a description: 'Tonemski; kolikostna nasprotja so ohranjena, kratki samoglasniki so naglašeni dinamično:'. The example text is 'valíke, zídane; nóť : so'mo, z'dej'.

Slika 4: Skrbniški pogled aplikacije – urejevalnik narečnih besedil in pogled urejanja vnosov.

Ker za vse kraje nimamo vedno pripravljenih vseh podatkov, lahko transkripcijo, poknjžitev ali analizo pustimo tudi neizpolnjeno.

4. Zaključek

Predstavljena aplikacija je prva tovrstna aplikacije za slovenska narečna besedila. Poskušali smo izdelati uporabniku prijazno spletno aplikacijo, ki bi bila dovolj enostavna tudi za širšo uporabo. Rezultat je interaktivna spletna karta slovenskih narečnih besedil, ki omogoča pregled vseh slovenskih narečnih skupin, narečij in podnarečij ter omogoča poslušanje zvočnih zapisov narečnih govorov, ogled njihovih fonetičnih zapisov, prevodov v knjižno slovenščino in analiz narečnih posebnosti. Poleg tega skrbnikom z Oddelka za slovenistiko omogoča dodajanje novih vnosov in tako zagotavlja vsebinsko vedno bogatejši spletni vir ne le za jezikoslovce in narečjeslovce, ampak tudi za učence in dijake, njihove profesorje ter vse, ki jih zanima odkrivanje bogastva in posebnosti slovenske narečne govornice.

Interaktivno karto narečnih besedil nameravamo še nadgraditi, predvsem za izboljšanje uporabniške izkušnje. Aplikacijo bomo prenesli na strežniško infrastrukturo CJVT⁷ in tako poskrbeli tudi za njeno dolgoročno vzdrževanje in delovanje. S tem bomo omogočili tudi trajnostno upravljanje z zbranimi gradivi.

Načrtujemo tudi vsebinsko dopolnitev aplikacije z dodatnimi 30 kraji, za katere imamo že pripravljene zvočne posnetke, transkripcije besedil in pripadajoče poknjžitve, za nekatere govore pa so izdelane tudi diahrone analize.

Zbrane podatke o narečnih govorih, vključenih v aplikacijo, nameravamo ponuditi tudi v obliki spletne zbirke podatkov, saj predstavljena aplikacija ne omogoča enostavnega izvoza in ponovne uporabe teh podatkov. Tako bodo zbrani podatki enostavno dosegljivi tudi drugim raziskovalcem, predvsem pa bosta omogočena njihovo strojno branje in obdelava.

5. Literatura

- Tine Logar in Jakob Rigler. 1983. *Karta slovenskih narečij*. Izdelal Geodetski zavod SRS, kartografski oddelek. Ljubljana, DDU Univerzum (stenski zemljevid).
- Ivan Lovrić. 2018. *Interaktivna spletna aplikacija za slovenska narečna besedila*. Diplomsko delo, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani. <http://eprints.fri.uni-lj.si/4117/>.
- Fran Ramovš. 1931. *Dialektološka karta slovenskega jezika*. Rektorat Univerze kralja Aleksandra I – J. Blasnika nasl. Ljubljana, Univerzitetna tiskarna.
- Vera Smole in Mojca Horvat. 2016. *Stara kmečka hiša, Narečna besedila z analizo I*. Znanstvena založba Filozofske fakultete Univerze v Ljubljani.
- Vera Smole. 2016. Drugi zvezek Slovenskega lingvističnega atlasa (SLA) in narečna besedila o stari kmečki hiši. V: Remnëva Marina Leont'evna, ur., *Slovenskij jazyk, literatura i kul'tura v slavjanskem i evropejskom kontekste: tezisy Meždunarodnoj naučnoj konferenciji*. MGU im. V. Lomonosova, Filologičeskij fakul'tet, 28-29 nojabra 2016 goda, str. 91–94. MAKS Press, Moskva.

⁷ Center za jezikovne vire in tehnologije Univerze v Ljubljani (CJVT), dostopno na <https://www.cjvt.si/>.

- Jože Toporišič. 1987. Slovensko narečjeslovje. V: Toporišič Jože, *Portreti, razgledi, presoje*, str. 217–256. Založba Obzorja, Maribor.
- Peter Weiss. 2004. ZRCola. *Jezikovni zapiski*, 10(1):145–152.