

# RECEPCIJA VIROV STAREJŠE SLOVENŠČINE IMP

**Tomaž Erjavec**

Institut »Jožef Stefan«, Ljubljana

**Darja Fišer**

Filozofska fakulteta, Ljubljana

UDK 801.8=163.6"1584/1918":004.7

V prispevku predstavljamo medsebojno povezane jezikovne vire starejše slovenščine, ki obsegajo digitalno knjižnico, korpus in besedišče, nato pa se osredotočimo na uporabnike spletne podobe teh virov. Zanimalo nas je, koliko jih je, kaj najpogosteje uporabljajo in kakšna je njihova percepcija teh virov: kaj jim je všeč, kaj jih moti in kakšne predloge imajo. Analizo smo opravili na osnovi dnevnika spletnega strežnika in spletnega vprašalnika.

jezikovni viri, starejša slovenščina, digitalna knjižnica, korpus, leksikon

This paper presents the results of a user-oriented analysis of a collection of language resources of historical Slovene consisting of a digital library, corpus and dictionary. We present the resources and then focus on the user analysis: how many users there are, which resources they use most often and what is their opinion of these resources (what they like, dislike and would like to see improved in the future). The user analysis was based on server logs and an on-line questionnaire.

language resources, historical Slovene, digital library, corpus, lexicon

## 1 Uvod

Na spletu je dostopna že velika količina jezikovnih virov za slovenščino, od digitalnih knjižnic (Hladnik 2009, Krstulović, Šetinc 2005, Šorn, Hadalin 2010) do korpusov (Arhar, Gorjanc 2007, Erjavec, Krek 2008) in slovarjev, kot sta SSKJ in Pleteršnikov Slovensko-nemški slovar. Zato je zanimivo, da je bilo doslej opravljenih zelo malo raziskav recepcije teh virov s strani uporabnikov: kako jih uporabljajo, kako uporabni se jim zdijo in kaj jim pri uporabi povzroča največ težav. Posledično nadaljnji razvoj teh virov večinoma tudi ni izrecno usmerjen v izboljšave, ki si jih uporabniki najbolj želijo oz. jih potrebujejo. Redka izjema je obsežna in temeljita analiza uporabe konkordančnika korpusa Gigafida (Arhar Holdt 2009), na podlagi katere so razvili novo, poenostavljeno različico konkordančnika, ki je brskanje po korpusu približala številnim uporabnikom.

V tem prispevku se posvetimo digitalni postavitvi zbirke starejših jezikovnih virov IMP (Erjavec 2012), ki vsebuje zbirko besedil, jezikoslovno označeni korpus in besedišče. Projekti, ki so podprli izdelavo teh virov, so sicer večinoma predvidevali njihovo uporabo v jezikovnotehnološke namene kot podatkovne baze pri razvoju

boljšega optičnega razpoznavanja znakov in boljšega iskanja po besedilu digitalnih knjižnic naše kulturne dediščine, vendar smo kmalu ugotovili, da bi bili lahko uporabni tudi na področju digitalne humanistike, kjer bi jezikoslovcem, učiteljem, učencem, študentom in drugim zainteresiranim omogočili spletni dostop do starejših slovenskih del in njihovega besedišča. Čeprav so bili naši naporji usmerjeni predvsem v izdelavo virov v pomensko usmerjenem zapisu XML, ki sledi Smernicam za kodiranje besedil TEI<sup>1</sup> in omogočanju njihovega prevzema po licenci Creative Commons, smo za namene digitalne humanistike vse tri vire ponudili tudi na spletu v ljudem prijaznejših oblikah.

Po dobrih dveh letih od postavitve virov starejše slovenščine IMP nas je zanimalo, kakšna je njihova uporaba in percepcija uporabnikov s področja digitalne humanistike v Sloveniji. Članek v razdelku 2 na kratko predstavi vire IMP, v razdelku 3 podamo analizo dnevnika spletnih poizvedb, v osrednjem razdelku 4 pa se osredotočimo na anketo, ki smo jo izvedli, da bi izvedeli mnenja uporabnikov o virih oz. njihovi spletni postavitvi. V razdelku 5 podamo zaključke in smernice za nadaljnje delo.

## 2 Sestava in spletna predstavitev virov

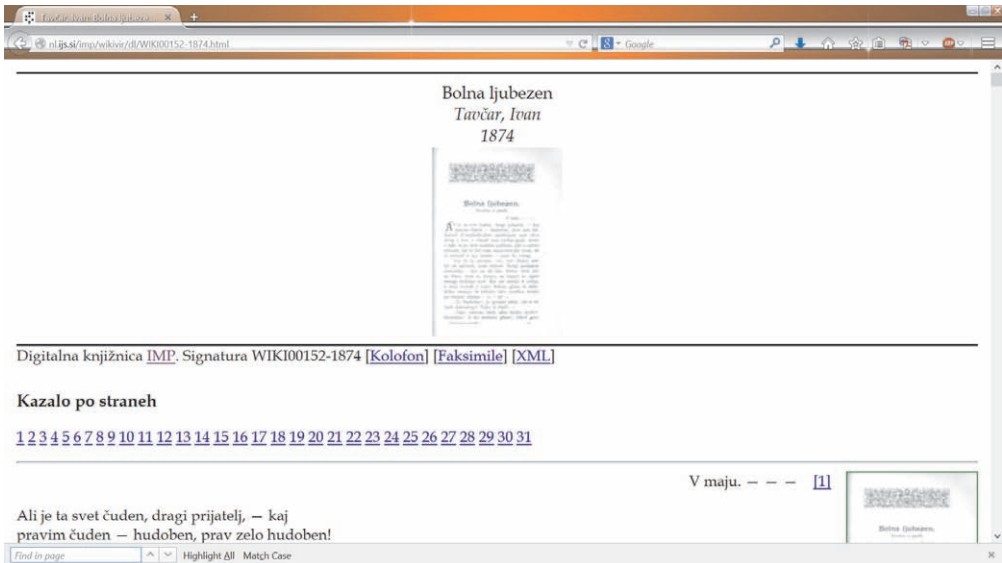
V tem razdelku na kratko predstavimo vire IMP, tj. digitalno knjižnico, besedišče in korpus, pri čemer se osredotočimo predvsem na njihovo spletno predstavitev.

### 2.1 Digitalna knjižnica

Zbirka besedil IMP, na osnovi katere je bila narejena digitalna knjižnica, vsebuje faksimile posameznih del in njihove pregledane prepise. Zbirka trenutno obsega 658 del s 46.645 stranmi, kar je okoli 15 milijonov besed. Besedila zajemajo obdobje 1584–1918, vendar z zelo neenakomerno porazdelitvijo, saj je večina besedil iz 19. stoletja (po številu besed 57,4 % celotne zbirke) in začetka 20. stoletja (39,8 %), medtem ko je iz 18. stoletja samo 2,6 % besedil. Za 16. (0,4 %) in 17. stoletje (0,6 %) vsebuje zbirka samo vzorca dveh knjig, za ostala obdobja pa celotna dela, predvsem knjige, vendar so vključeni tudi članki iz revij, izvodi časopisa *Kmetijske in rokodelske novice* ter posamezni rokopisi. Vsebovana dela so verska, leposlovna in stvarna, tako izvorna slovenska kot prevodna.

Zbirka besedil je na splet postavljena kot digitalna knjižnica s katalogi, urejenimi po avtorju, naslovu, letu izdaje in signaturi. Katalogi so razčlenjeni tudi glede na zvrst besedila. Vsako delo je sestavljeno iz kolofona in celotnega prepisa v eni datoteki HTML, ki je bila avtomatsko izvedena iz izvornega XML/TEI. Spletna oblika vsebuje kazali po straneh in po poglavjih, pri vsakem prelomu strani pa je prikazana tudi sličica faksimila strani, ki se s klikom nanjo prikaže v polni velikosti. Stopnja oblikovanja, torej kako verno prepis odslikava faksimile, se razlikuje glede na vir besedila, vendar vsebuje vsaj naslove, odstavke, opombe in poudarjeno besedilo, ponekod pa tudi popravke tiskarskih škratov v izvorniku, prelome vrstic, tiskarska znamenja itn. Začetek enega od del v digitalni knjižnici je prikazan na Sliki 1.

<sup>1</sup> [www.tei-c.org](http://www.tei-c.org).

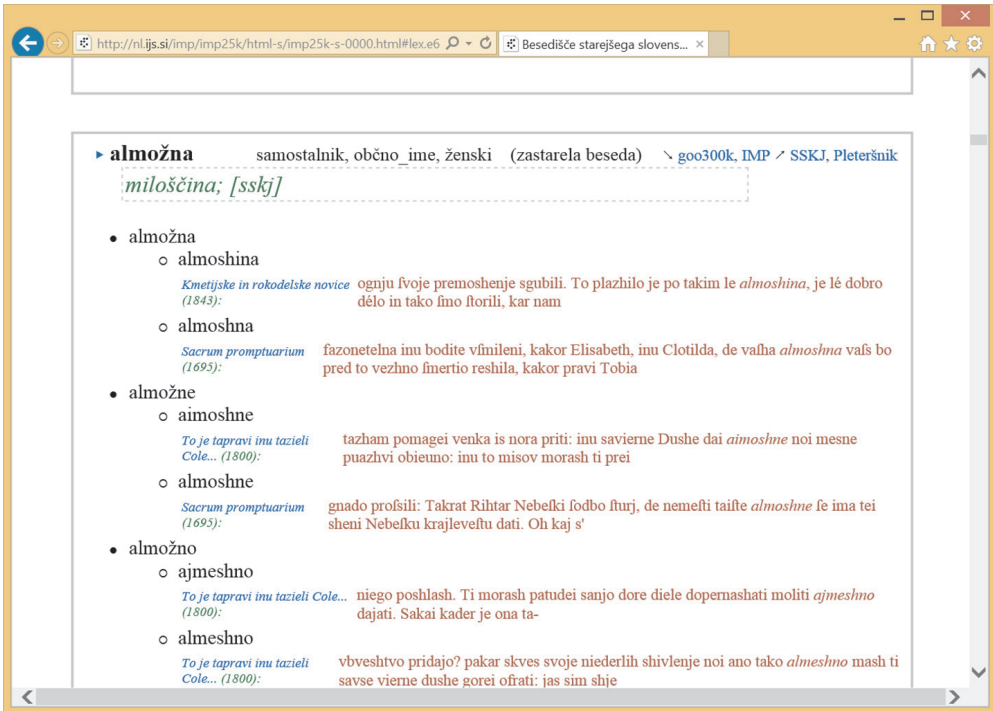


Slika 1: Primer predstavitve dela v digitalni knjižnici

## 2.2 Besedišče IMP

Besedišče je bilo avtomatsko izdelano iz pojavnice, ki so bile ročno označene v korpusu. Vsaki taki besedni pojavnici je bila v korpusu pripisana sodobna besedna oblika, njena lema in oblikoskladenjska oznaka, za zastarele besede pa tudi sodobna ustreznica oz. ustreznice ali krajša razlaga skupaj z virom te informacije. Iz takšnega postopka izdelave sledi, da besedišče vsebuje samo korpusno izpričane oblike. Če zanemarimo številke, zatipkane in tuje besede ter lastna imena, besedišče trenutno vsebuje 21.653 lem, 51.156 besednih in 73.263 zgodovinskih oblik. Seveda je veliko besednih oblik enakih kot v sodobni slovenščini – če se omejimo samo na gesla, ki vsebujejo vsaj eno besedno obliko drugačno, kot je sodobna, je takih lem 12.255, če pa samo na zastarele, ki imajo torej pripisano razlago, je takih lem 4156.

Tudi besedišče je bilo pretvorjeno iz izvirnega XML v HTML in postavljeno na splet kot digitalni slovar, pri čemer je trenutna različica narejena kot statična zbirka spletnih strani, torej brez svojega iskalnika. To bi namreč zahtevalo investicijo v razvoj programske opreme, poleg tega pa se nam je zdelo, da je slovar zaenkrat dovolj majhen, da je možno po naboru gesel iskati kar prek iskanja v HTML. Spletni slovar ima krovno stran, ki vsebuje kazalko na kolofon in kazalo vseh gesel (lem), slovar pa je nato razdeljen na posamezne strani, saj bi bil celoten prevelik za eno datoteko. Poleg stavljenja v HTML pretvorba iz XML tudi razveže oblikoskladenjsko oznako v njene lastnosti v slovenščini, lemi doda povezavo na konkordančnik, bibliografske podatke o posameznih primerih uporabe pa poveže na spletno knjižnico IMP. Kjer je to smiselno, so geslom dodane tudi povezave na spletni postavitvi SSKJ in Pleteršnikovega slovarja. Obliko spletnega slovarja prikazuje Slika 2.



Slika 2: Primer predstavitve gesla v spletnem slovarju

### 2.3 Korpusa in konkordančnik

V sklopu dela na virih IMP smo izdelali tudi manjši, polno ročno označeni korpus goo300k z okoli 300.000 pojavnicami in orodje za avtomatsko označevanje korpusov starejše slovenščine (Erjavec 2013b). S tem programom smo označili celotno digitalno knjižnico in jo pripravili kot korpus besedil IMP, ki vsebuje okoli 15 milijonov besed.

Oba korpusa smo naredili dostopna prek spletnega konkordančnika noSketch-Engine (Rychlý 2007), postavljenega na strežniku nl.ijs.si, ki ponuja že več kot 20 različnih korpusov (Erjavec 2013a). Konkordančnik je hiter in omogoča zelo raznovrstna iskanja po korpusu kot tudi različne načine prikaza podatkov, razvit pa je bil kot orodje za profesionalne leksikografe, zato je za uporabo razmeroma zapleten. Slika 3 prikazuje konkordance besede (oz. leme) »ljubezen« v korpusu goo300k, pri čemer je izpisana samo prva pojavitev te besede v vsakem delu, medtem ko Slika 4 prikazuje uporabo konkordančnika za izpis ključnih besed korpusa goo300k glede na ročno označeni korpus sodobne slovenščine sjs500k.

Iskani niz **ljubezen** 155 > Filter all but first hit in document 46 (128.5 na milijon)

Stran 1 od 3 Pojdi [Naslednja](#) | [Zadnja](#)

Sveti priročnik (vzorec) (1695) vřimite valh fazonetel, ter premislite pamet, **lubesan** /ljubezen / inu poterpeshlivost, katero Socrates je  
 Mali katekizem (1768) so sami na sebi prov ostudni, inu ozhitnu **lufefni** /ljubezni / inu pravizi tega blihnega prutistojee  
 Marianske Kempensar (1769) ta perva nahdati, koker da urafhenja te **lubesne** /ljubezni / se befediti. § 2. En modre lube myr, inu  
 Branja in evangeliji (1777) řposnan. Sedej pak ořtanejo vera , upanje, **lubesen** /ljubezen / lete try: al lubesen je med nymi ta narvezl  
 Županova Micka in Ta vese... (1790) kraja. Al ta nar\_ bøl vashna refsniža je moja **lubesen** /ljubezen / prut Neshki. § Neshka. Kaj mi povejřhl –  
 Pridige za vse nedelje v ... (1794) boshje, inu te řet Zerkovne, Verø, Vupauje, **Lubesen** /ljubezen / , te ředem naglavne grehe, te devet ptuje  
 Kolemonov řegen (1800) noi kervavi puet: jas taba prosim skues **lubesen** /ljubezen / kathero si ti mou sate zhovlesen rued:  
 Zgodbe Svetega pisma za m... (1830) je Joshef per jezharji kmal saupanje ino **ljubezen** /ljubezen / nahfel, ino vie jetnike mu je dal pod oblaf  
 Sedem novih pripovedi za ... (1836) řhtaniřlaj, tak rad ima, ř prineřla; Bog mu je **ljubezen** /ljubezen / do mene, nje\_govo polhtenje, pridnořt ino  
 Genovefa (1841) Tvõjiga Duha, v řposnanju Tēbe in v řvėti **ljubesni** /ljubezni / srediřli, in ga kakor drago, meni isrozhēno  
 Kmetijske in rokodelske n... (1843) , poboshnim ferzam in veřto pridnořt in **ljubesni** /ljubezni / k řvojimj řtanu řdruřli, fe bo kmato řkosi  
 Kmetijske in rokodelske n... (1845) slovensko, de bi njim vsim, ki bi zasliřati, od **ljubezni** /ljubezni / do sloveřšine serce se vneřlomo. – Pa joj  
 Kmetijske in rokodelske n... (1846) domorodecv sije ře jedno sonřice, – oř řlahtna **ljubezen** /ljubezen / do naroda nařegal da domovino ljubijo ,  
 Čujte, řujte, kaj řganje ... (1847) je kje negodno řaljubil, in nedosegiva **ljubezen** /ljubezen / mu řalostno serce napravřija; de je ře nezav  
 Kmetijske in rokodelske n... (1847) tem řlahtnim delu podpiralo in krepřalo? – **Ljubezen** /ljubezen / kristjanska, ali kakor nekteri tudi prav  
 71... (1848) ...

Slika 3: Primer izpisa poizvedbe po lemi »ljubezen« v korpusu goo300k

Find: ljubezem Previous Next Options

User: defaults Korpus: goo300k (starejša besedila, ročno označena) Išči v goo300k (starejša be

Konkordance Seznami

Shrani Spremeni nastavitve

**Seznam besed**

Korpus: goo300k (starejša besedila, ročno označena)

Referenčni korpus: ssj500k (ročno označen)

lemma	goo300k (starejša besedila, ročno označena)			ssj500k (ročno označen)		
	Frekvenca	Frekv/milijon		Frekvenca	Frekv/milijon	Rezultat
ljubljana	125.0	347.7	0	0.0	348.7	
preljub	13.0	36.2	0	0.0	37.2	
vkľjub	7.0	19.5	0	0.0	20.5	
ljuboslava	7.0	19.5	0	0.0	20.5	
ljubljanica	7.0	19.5	0	0.0	20.5	
ljubezniv	32.0	89.0	2.0	3.4	20.4	
ljub	211.0	586.9	18.0	30.7	18.5	
rodoljub	7.0	19.5	1.0	1.7	7.6	
ljubiti	113.0	314.3	34.0	58.0	5.3	
obljubiti	64.0	178.0	23.0	39.2	4.4	
poljubiti	16.0	44.5	6.0	10.2	4.1	
obljubljen	6.0	16.7	2.0	3.4	4.0	
ljubeznivo	8.0	22.3	3.0	5.1	3.8	
obljuba	29.0	80.7	13.0	22.2	3.5	
ljubo	5.0	13.9	2.0	3.4	3.4	
ljubezen	155.0	431.2	99.0	168.9	2.5	

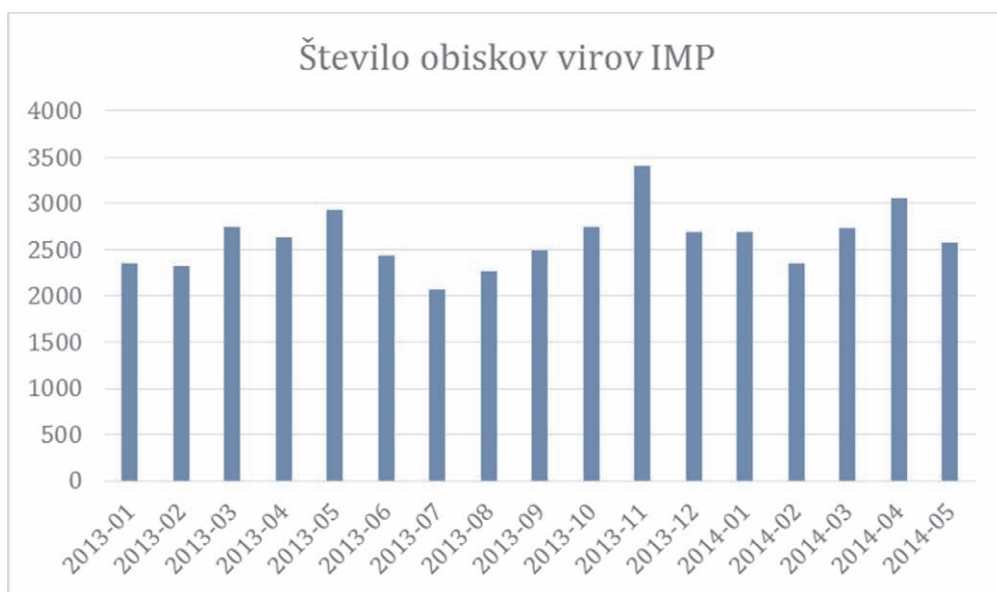
Lexical Computing Ltd. LEX COM  
 Sketch Engine (ver:open-2.59.1-open-2.91.13)  
 Jezik vmesnika: English | řesky | řjednřna řednřna | řesky | řjednřna řednřna | řesky | řjednřna řednřna | řesky | řjednřna řednřna

Slika 4: Primer izpisa seznama ključnih besed leme »\*ljub.\*« v korpusu goo300k glede na korpus sskj500k, vzorčen iz korpusa FidaPLUS

### 3 Analiza zabeleženih iskanj uporabnikov

Uporabo spletnih virov je možno kvantificirati z analizo dnevnika poizvedb na strežniku. Ta postopek se sicer umika analizam, pri katerih so že spletne strani opremljene s programsko kodo, ki omogoča beleženje uporabnikov in ima prednost, da je možno slediti tudi identiteti posameznega uporabnika, vendar pa ima to slabost, da uporablja piškotke, zato je v skladu s slovensko zakonodajo uporabnike pri vsakem obisku nove strani treba prositi za privoljenje.

Analizo dnevnikov uporabe smo izvedli za leto 2013 in prvo polovico leta 2014, in sicer z uporabo orodja Piwik.<sup>2</sup> Število dostopov do spletnih virov IMP (kamor štejemo samo dostope do HTML-strani knjižnice in slovarja ter do konkordančnika) je podano na Sliki 5. Kot vidimo, so obiski v celotnem obdobju razmeroma konstantni, pri čemer je na mesec od 2000 do 3400 dostopov, torej okoli 100 na dan, kar povsem upravičuje naše napore za izdelavo spletnih strani. Seveda pa se je treba zavedati, da obiskovalci na strani virov IMP pridejo tudi pomotoma, ko jih nanje napačno usmerijo iskalniki; tako npr. na tem spletišču ne bodo našli odgovorov na svoja iskanja, kot sta npr. »crnogorski peceni prasci« ali »boginja lepote prevedeno na afrikanski«.



Slika 5: Število obiskov do spletnih virov IMP

Analiza držav, iz katerih so prišle poizvedbe, kaže, da je iz Slovenije 66 % obiskovalcev, sledita ZDA s 15 % in Nemčija s 5 %. Največ obiskovalcev pride na strani prek iskalnika Google, temu sledi stran prof. dr. Mirana Hladnika »Slovenska književnost« (lit.ijs.si), tretje mesto si delita Twitter in www.kulinarika.net (verjetno, ker viri vsebujejo tudi dve kuharici), na četrtem mestu pa je Facebook. Zanimivo je še,

<sup>2</sup> <http://piwik.org/>

da največ obiskovalcev uporablja brskalnik Google Chrome, temu sledi Mozilla Firefox, Internet Explorer pa je šele na tretjem mestu.

## 4 Analiza mnenj uporabnikov

Mnenja uporabnikov o virih IMP smo preverili s spletnim vprašalnikom,<sup>3</sup> ki je vseboval 15 vprašanj, razvrščenih v tri sklope: (1) vprašanja, vezana na profil uporabnika, (2) uporabniško izkušnjo ter (3) želje in predloge po nadaljnjem razvoju virov. Vprašalnik smo poslali udeležencem Jezikovnotehnološkega usposabljanja za učitelje, ki ga zadnja tri leta financira Ministrstvo za šolstvo, in na dopisni seznam SlovLit ter ga objavili na spletni strani Oddelka za prevajalstvo Filozofske fakultete ter družbenih omrežjih Facebook in Twitter.

### 4.1 Profil uporabnikov

Vprašalnik je izpolnilo 52 uporabnikov, med katerimi je bilo 81 % žensk in 19 % moških. Največ (40 %) uporabnikov je starih od 41 do 60 let, 38 % pa med 25 in 40 let. 13 % jih je mlajših od 25 let, 4 % pa starejših od 60 let. Vprašanje o poklicu v anketi ni bilo obvezno, zato je nanj odgovorilo le 33 (63 %) vprašanih. Med njimi je največ osnovnošolskih in srednješolskih učiteljev slovenščine (21), ostali so diplomski in podiplomski študenti (5), prevajalci (2), novinarji (2), raziskovalci (1), upokojenci (1) in drugo (1). Večina (44 %) anketiranih je za vire izvedela na internetu (na dopisnem seznamu ali v družbenih medijih) ali na izobraževanju (40 %). Večina (37 %) anketiranih vire uporablja za poučevanje, 31 % ljubiteljsko, 22 % pa za raziskovanje. Med anketiranimi je tudi 9 % uporabnikov, ki vire uporabljajo pri prevajanju, kar kaže na zelo heterogeno ciljno skupino uporabnikov, zato so najbrž tudi njihove želje in potrebe precej različne.

Več kot polovica (60 %) vseh vprašanih vire IMP uporablja enkrat mesečno, 17 % jih po virih posega enkrat tedensko, 13 % enkrat letno, 6 % vsak dan, 4 % pa nikoli. Uporabniki največ uporabljajo besedišče (občasno ali pogosto 77 %) in korpus (73 %), medtem ko digitalno knjižnico občasno ali pogosto uporablja le 59 % anketiranih. Med podpornimi viri anketirani najpogosteje uporabljajo domačo stran projekta (41 %) in oblikoskladenjske specifikacije (41 %), bibliografijo pa precej manj (19 %).

### 4.2 Uporabniška izkušnja

Praktično vsi (94 %) anketirani uporabniki se (zelo) strinjajo, da so viri IMP pomembni in koristni za jezikoslovne raziskave, 89 % jih meni, da so pomembni in koristni za celotno slovensko družbo, 87 % pa se zdi, da so pomembni in koristni tudi za pouk slovenščine. Največ kritike sta prejela premajhen obseg virov in prekompleksno iskanje (po 41 %), slaba vizualizacija moti 37 % anketiranih, samo 17 % vprašanih pa meni, da viri vsebujejo preveč napak. Praktično vsem anketiranim (95 %) se viri zdijo priporočila vredni. Kar 35 % anketiranih je vire že priporočilo drugim, 60 % pa bi jih priporočilo sodelavcem, študentom, prijateljem.

<sup>3</sup> Rezultati ankete so dostopni na povezavi <http://tinyurl.com/pg6fq8a>.

### 4.3 Mnenja in predlogi o nadaljnjem razvoju virov

Anketirani si želijo, da bi v nadaljnjem razvoju vire prilagodili predvsem jezikoslovcem (14 %), študentom, dijakom in učencem (po 13 %), srednješolskim in osnovnošolskim učiteljem (po 12 %) ter leksikografom (9 %). Glede dodajanja vsebin se jim zdi najbolj nujno dodati več leposlovja (26 %), manj pa periodike (16 %), rokopisov (16 %) in neleposlovja (12 %).

Največ izboljšav si anketirani želijo pri slovarju. Po njihovem mnenju je prva prioriteta iskalnik po slovarju, ki se zdi nujen 35 % anketiranim. Izboljšave razlag se zdijo nujne 32 %, razširitev slovarja z novimi gesli pa 27 % vprašanih. Izboljšan prikaz gesel se zdi nujen 24 %, boljšo povezavo slovarja s konkordančnikom si zelo želi 20 % uporabnikov, možnost popravljanja in dodajanja gesel v slovar s strani uporabnikov pa se zdi nujna le 15 % uporabnikov, jih pa kar 60 % meni, da bi bila koristna.

Sledijo želje po izboljšavah digitalne knjižnice. Tu se anketiranim zdi najbolj nujno dodati transkripcije v sodobni slovenščini (32 %) in izdelane vire integrirati v druge digitalne knjižnice, kot sta na primer dlib.si in Wikivir (27 %). Bolj kot razširitve digitalne knjižnice si anketirani želijo omogočenega prenosa celotnega faksimila (26 %),<sup>4</sup> vizualno privlačnejšega branja na zaslonu (22 %), dostopa do knjižnice z različnih mobilnih naprav (20 %) in omogočenega prenosa celotne transkripcije v PDF-formatu (19 %). Med novimi gradivi se jim zdi največja prioriteta leposlovje (26 %), medtem ko se jim zdi vključevanje dodatne periodike (16 %), rokopisov (16 %) in neleposlovja (12 %) manj pomembno.

Pri nadaljnjem razvoju korpusa se zdi anketiranim še najbolj nujno ponuditi uporabniku prijaznejši konkordančnik (29 %). Dodajanje besedil v korpus se podobno kot pri digitalni knjižnici ne zdi ključno večini anketiranih, saj jih le 18 % meni, da bi bilo nujno dodajanje novejših besedil (po letu 1918), in le 14 %, da bi bilo to nujno za starejša besedila (pred letom 1800).

Izboljšava oznak se zdi nujna le 10 % vprašanih, kar ponovno kaže na to, da so uporabniki s kakovostjo izdelanih virov večinoma zadovoljni in da napake v virih, ki so posledica avtomatske obdelave besedil ali pomanjkljive ročne validacije, ne zmanjšujejo uporabne vrednosti virov za namene, kot jih uporabniki uporabljajo.

### Sklep

V prispevku smo predstavili medsebojno povezane jezikovne vire starejše slovenščine IMP, pri čemer smo se osredotočili na uporabnike in njihovo mnenje o virih. Dnevnik spletnega strežnika kaže konstanten obisk okoli 100 uporabnikov na dan, pri čemer jih je dve tretjini iz Slovenije. Spletni vprašalnik kaže, da vire uporablja zelo raznolik profil uporabnikov za zelo različne namene, zato je zanimivo, da so z njimi in z njihovo kakovostjo v prevladujoči večini vseeno zadovoljni in da se jim zdi, da so viri pomembni tako za jezikoslovje kot za poučevanje slovenščine. Anketirani najpo-

<sup>4</sup> To je razmeroma enostavno doseči, zato je v zadnji postaviti digitalne knjižnice že omogočen prenos celotnega faksimila kot PDF.



gosteje uporabljajo slovar in korpus, manj pa digitalno knjižnico. Med izboljšavami so uporabniki kot najnujnejše izpostavili: razvoj iskalnika po slovarju in izboljšave razlag zastarelih besed, dodajanje transkripcije v sodobni slovenščini in uporabniku prijaznejši konkordančnik, kar bodo tudi naše prioritete pri nadaljnjem razvoju virov IMP.

## Literatura

- ARHAR HOLDT, Špela, 2009: Uporabniška evalvacija korpusa FidaPLUS: zasnova vprašalnika, prvi rezultati. Marko Stabej (ur.): *Infrastruktura slovenščine in slovenistike. Obdobja 28*. Ljubljana: Znanstvena založba Filozofske fakultete. 19–26.
- ERJAVEC, Tomaž, 2012: Jezikoslovni viri starejše slovenščine. *Knjižnica* 56/3. 205–221.
- ERJAVEC, Tomaž, 2013a: Korpusi in konkordančniki na strežniku nl.ijs.si. *Slovenščina 2.0* 1. 24–49. [www.trojina.org/slovenscina2.0/arhiv/2013/1/Slo2.0\\_2013\\_1\\_03.pdf](http://www.trojina.org/slovenscina2.0/arhiv/2013/1/Slo2.0_2013_1_03.pdf)
- ERJAVEC, Tomaž, 2013b: Posodabljanje starejše slovenščine. *Uporabna informatika* 21/4. 186–195.
- HLADNIK, Miran, 2009: Infrastruktura slovenistične literarne vede. Marko Stabej (ur.): *Infrastruktura slovenščine in slovenistike. Obdobja 28*. Ljubljana: Znanstvena založba Filozofske fakultete. 161–169.
- KRSTULOVIĆ, Zoran, ŠETINC, Lenart, 2005: Digitalna knjižnica Slovenije – dLib.si. Aleksander Novaković idr. (ur.): *Informatika kot temelj povezovanja: zbornik posvetovanja*. Ljubljana: Slovensko društvo Informatika. 683–689.
- RYCHLÝ, Pavel, 2007: Manatee/Bonito – A Modular Corpus Manager. *1<sup>st</sup> Workshop on Recent Advances in Slavonic Natural Language Processing*. Brno: Masarykova univerzita. 65–70.
- ŠORN, Mojca, HADALIN, Jurij, 2010: Spletni portal SIstory: prost dostop do dosežkov slovenskega zgodovinopisja. Maja Božič, Igor Zemljič (ur.): *Zbornik prispevkov 4. skupnega posvetovanja Sekcije za specialne knjižnice in Sekcije za visokošolske knjižnice Zveze bibliotekarskih društev Slovenije*. Ljubljana: Zveza bibliotekarskih društev Slovenije. 103–107.