

Ana MIHALJEVIĆ  
Josip MIHALJEVIĆ  
Staroslavenski institut  
Zagreb  
*amihaljevic@stin.hr*  
*jmihaljevic@stin.hr*

UDK: 811.163.42'374:004  
DOI: <https://doi.org/10.31745/s.74.10>  
izvorni znanstveni članak  
primljen: 18. travnja 2023.  
prihvaćen: 7. studenoga 2023.

## MREŽNA INAČICA RJEČNIKA CRKVENOSLAVENSKOGA JEZIKA HRVATSKE REDAKCIJE\*

U radu se predstavlja proces razvoja mrežne inačice *Rječnika crkvenoslavenskoga jezika hrvatske redakcije* (RCJHR). Izrada toga rječnika dugogodišnji je projekt Staroslavenskoga instituta u Zagrebu. Projekt je započeo 1958., a prvi je sveščić izišao 1991. Dosad su objavljena 23 sveščića s riječima od *a do izgnilstvo*. Rječnik se izrađuje u MS Wordu. U radu se analiziraju i prikazuju koraci i problemi u retrodigitalizaciji već izdanih sveščića. Trenutačno se izrađuje mrežni portal na kojemu će rječnik biti dostupan. S obzirom na to da rječnik ima složenu mikrostrukturu, a rječnički članci sadržavaju pet jezika (hrvatski, hrvatski crkvenoslavenski, engleski, grčki i latinski) i četiri pisma (glagoljica, stara ćirilica, latinica, grčki alfabet), proces je izrade takvoga portala složen. Nakon provedenoga iscrpnog istraživanja raznih mogućih rješenja i modela izrađen je vlastiti model za retrodigitalizaciju RCJHR-a. Nakon detaljne analize autori su odlučili da se mrežna inačica rječnika izradi u sustavu *WordPress*, a istražuju se i mogućnosti da se nove natuknice također obrađuju u tome sustavu te da se tim sustavom koristi kao sustavom za izradu rječnika (engl. *dictionary writing system* – DWS). Za sveščiće koji postoje i u formatu DOCX omogućeno je pretraživanje teksta napisanoga svim pismima i na svim jezicima. U sustav su implementirani glagoljični i ćirilčni fontovi kako bi autori i urednici mogli učinkovito raditi na tekstu koristeći se jednostavnim korisničkim sučeljem. Kako bi se korisnicima omogućila lakša pretraga, omogućen je prikaz virtualne glagoljične tipkovnice, koja olakšava upis glagoljičnih slova. Riječi koje upućuju jedna na drugu povezuju se hiperpo-

\* Ovaj je rad nastao u okviru projekata *Razvoj modela digitalne infrastrukture Staroslavenskoga instituta – DigiSTIN* i *Rječnik crkvenoslavenskoga jezika hrvatske redakcije*, koje financira Europska unija – NextGenerationEU. Za iznesene stavove i mišljenja odgovorni su samo autori te ti stavovi ne odražavaju nužno službena stajališta Europske unije ili Europske komisije. Ni Europska unija ni Europska komisija ne mogu se smatrati odgovornima za njih.

veznicama. U radu se prikazuju sve faze istraživanja i tehnički aspekti izrade mrežne inačice RCJHR-a, problemi u radu i budući planovi.

Ključne riječi: *Rječnik crkvenoslavenskoga jezika hrvatske redakcije*, e-rječnik, mrežni rječnik, retrodigitalizacija, hrvatski crkvenoslavenski jezik

## 1. O RJEČNIKU CRKVENOSLAVENSKOGA JEZIKA HRVATSKE REDAKCIJE

Izrada *Rječnika crkvenoslavenskoga jezika hrvatske redakcije* (RCJHR) najdugovječniji je i temeljni projekt Staroslavenskoga instituta u Zagrebu. Konceptija RCJHR-a i rad na njemu izloženi su u uvodu samoga rječnika (RCJHR 1991: I–XLVIII; također 2000: I–XLVIII), a dodatne informacije o konceptu i postupku izrade rječnika dostupne su u nekim radovima koji su se bavili ovom temom.<sup>2</sup>

Ideja za stvaranje RCJHR-a potječe iz 1958. godine s Četvrtoga međunarodnog slavističkog kongresa u Moskvi.<sup>3</sup> Prvi je sveščić objavljen 1991. (RCJHR 1991), a dosad (travanj 2023.) su objavljena 23 sveščića s natuknicama od *a* do *izgnilstvo* (RCJHR 2000; 2015; 2016; 2018; 2021). Po deset se sveščića uvezuje u tvrdo ukoričene sveske, što znači da su zasada oblikovana dva takva sveska (RCJHR 2000; 2015).

Rječnik je korpusno vođen<sup>4</sup> te se njegova obrada temelji na dvama fizičkim korpusima u kojima je ekscerpirano više od 60 hrvatskoglagoljičnih izvora – tzv. azbučnome, u kojemu su riječi poredane prema azbučnome redu, i mjesnome, u kojemu su riječi poredane prema spomeniku u kojemu su potvrđene. Turkalj i Vukoja taj korpus opisuju kao referencijalan, reprezentativan, usporedan i parsiran (TURKALJ; VUKOJA 2016: 170–171). Osim tih dvaju

<sup>2</sup> V. MAREŠ 2007; NAZOR 1997; ŠIMIĆ 2017.

<sup>3</sup> O inicijativi za pokretanje rječnika i okolnostima koji su do toga dovele više v. NAZOR 2008; NAZOR 2017.

<sup>4</sup> O korpusno vođenome pristupu vidi TOGNINI-BONELLI 2001: 84–100. U bazi *Hrvatskoga strukovnog nazivlja* – *Struni* definicija je korpusno vođenoga rječnika: »rječnik u kojemu se obrađivač služi isključivo korpusom, pa se u rječniku nalazi samo ono što se nalazi u korpusu« (<http://struna.ihj.hr/naziv/korpusom-vodjen-rjecnik/51783/#naziv>), za razliku od korpusno utemeljenoga rječnika, koji je definiran ovako: »rječnik u kojemu se obrađivač služi korpusom, ali može slobodno procijeniti što treba unijeti u rječnik te rječnik može po potrebi dopunjavati i riječima iz drugih izvora te kolokacijama i značenjima koji nisu potvrđeni u korpusu« (<http://struna.ihj.hr/naziv/korpusno-utemeljen-rjecnik/51782/#naziv>). Više o *Struni* v. npr. ČAVAR; JAZBEC; NAHOD 2009.

korpusa postoje i tri pomoćna geslara – hrvatski crkvenoslavenski s oko 18 100 kartica, grčki s oko 60 000 kartica i latinski s oko 200 000 kartica (v. VUKOJA 2012: 214–216; VUKOJA 2014: 1223–1227; TURKALJ; VUKOJA 2016: 171). Budući da je riječ o rječniku u kojemu su svi primjeri iz korpusa, on je i korpusom oprimgjeren.<sup>5</sup>

Rječnik se izrađuje u sustavu MS Word i dosad je bio objavljivao isključivo u tiskanome obliku.

## 2. DIGITALIZACIJA KORPUSA ZA IZRADU RCJHR-A

Kao dio pokušaja modernizacije RCJHR-a u tijeku su pokušaji digitalizacije dvaju osnovnih korpusa na temelju kojih RCJHR nastaje jer bi njihova digitalizacija, lematizacija te napredne mogućnosti pretrage koje bi iz toga proizšle omogućile učinkovitiji i brži rad.<sup>6</sup> Kartice su osnovnih korpusa skenirane i računalno dostupne članovima projekta.<sup>7</sup> Kao motive za digitalizaciju korpusa Antonio Magdić navodi sljedeće želje:

- »1. da izvorna datoteka mjesnoga kataloga više ne bude i radna
2. da se napravi kopija izvornih kartica koja se radi sigurnosti može premjestiti
3. da se znanstvenicima olakša pretraživanje rječničke građe, tako da ne moraju pretraživati metalne kutije od kojih svaka sadrži oko 3.000 kartica
4. da se znanstvenici mogu služiti rječničkom građom i izvan Instituta.« (MAGDIĆ 2015: 172).

Kartice su očitavane s pomoću programa ABBYY FlexiCapture<sup>8</sup>, međutim, proces njihova očitavanja još nije dovršen. Na skeniranim su karticama autori ovoga rada također isprobavali modele za strojno očitavanje teksta koji su dostupni u Transkribusu<sup>9</sup>, međutim, s obzirom na to da se na svakoj kartici

<sup>5</sup> U *Struni* je korpusom oprimgjeren rječnik definiran ovako: »rječnik u kojemu se obrađivač isključivo služi korpusom za navođenje primjera« (<http://struna.ihjj.hr/naziv/korpusom-oprimgjeren-rjecnik/53367/#naziv>).

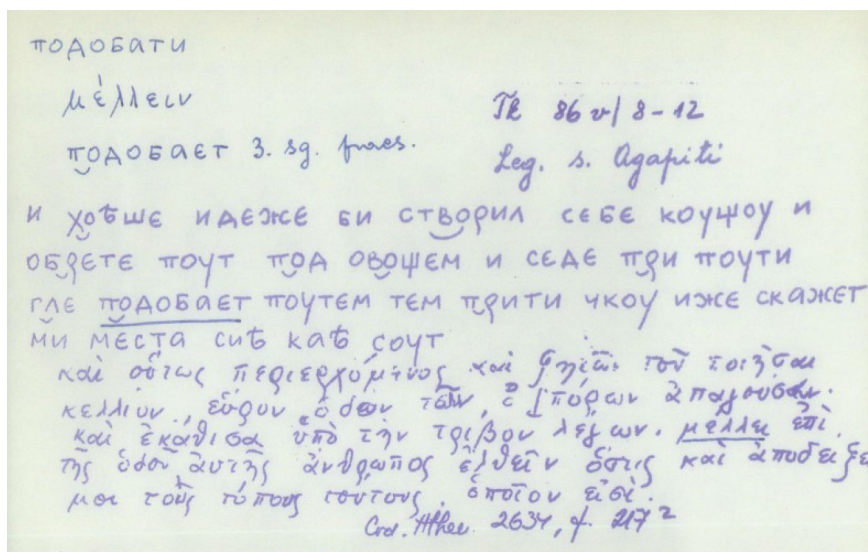
<sup>6</sup> Više o korpusu v. VUKOJA 2012; VUKOJA 2014. O sličnim procesima digitalizacije v. RIBAROV; CAMUGLIA 2003.

<sup>7</sup> O procesu njihova skeniranja v. MAGDIĆ 2015: 171–176.

<sup>8</sup> O tome programu v. <https://www.abbyy.com/flexicapture/>. Više o procesu digitalizacije korpusa za RCJHR v. VUKOJA; TURKALJ 2016.

<sup>9</sup> O Transkribusu v. <https://readcoop.eu/transkribus/>. Transkribusom se u posljednje vrijeme vrlo učinkovito koristi za transliteraciju hrvatskoglagoljičnih tekstova. Tako je s pomoću

pojavljuju dva ili tri pisma (stara ćirilica, latinica, a katkad i grčki alfabet) te su one pisane na dvadesetak različitih rukopisa, rezultati zasad nisu iskoristivi pri izradi rječnika. Treniranje novoga modela za očitavanje kartica bilo bi izrazito zahtjevno jer su razlike između pisama i rukopisa velike te bi trebalo osmisliti mnogo različitih modela koje bi trebalo kombinirati. Na *Slici 1* prikazan je primjer kartice iz korpusa na kojoj su tri pisma.



Slika 1. Primjer kartice iz korpusa  
Figure 1. An example of a corpus card

U sklopu različitih projekata u Staroslavenskome institutu digitaliziraju se ili digitalno pripremaju različiti tekstovi koji su i dio korpusa za izradu RCJHR-a. Tako se u okviru Znanstvenoga centra izvrsnosti za hrvatsko glagoljaštvo<sup>10</sup> izrađuje baza *beram.stin.hr*, koja će pružiti digitalni pristup misalima i brevijarima iz Berma u Istri. Usporedno s tim, u tijeku su projekti *Interdisciplinarni pristup hrvatskoglagoljskomu misalu na primjeru Misala kneza Novaka*<sup>11</sup>, u sklopu ko-

Transkribusa u sklopu Staroslavenskoga instituta za tisak pripremljena transliteracija *Misala po zakonu rimskog dvora* te se s pomoću njega pripremaju izdanja beramskih misala. O treniranju modela za čitanje hrvatske ustavne glagoljice v. RABUS 2022.

<sup>10</sup> Više o Znanstvenome centru izvrsnosti za hrvatsko glagoljaštvo v. <https://zci.stin.hr/>.

<sup>11</sup> Više o projektu *Interdisciplinarni pristup hrvatskoglagoljskomu misalu na primjeru Misala kneza Novaka* v. <https://stin.hr/nomis/>.

jega se razvija digitalno izdanje *Misala kneza Novaka*, i *Istraživanje starije hrvatskoglagoljske zborničke baštine*<sup>12</sup> te istraživanje, preslovljavanje i priprema za tisak *Brevijara Vida Omišljanina* (v. KAPPEL; KAPPEL; VELA 2023). Rezultati tih projekata, i kad je riječ o bazama podataka i kad je riječ o digitalnim izdanjima, planiraju se iskoristiti u daljnjemu radu na RCJHR-u. Većina se suvremenih korpusa izrađuje u alatima kao što je Sketch Engine<sup>13</sup>, koji omogućuje lematizaciju i prikaz konteksta pojedine riječi s pomoću skica riječi.<sup>14</sup> Međutim, zbog specifičnosti našega korpusa za njega se, barem zasad, ne mogu upotrebljavati takvi programi.

### 3. O POTREBI IZRADE MREŽNE INAČICE RCJHR-A

Posljednjih je nekoliko desetljeća e-leksikografija doživjela veliki uzlet i e-inačice rječnika postale su svojevrsnim standardom te se čak i rječnici koji su izvorno objavljeni u tiskanome izdanju sve više objavljuju i na mreži. Već su 2015. Štrkalj Despot i Möhrs napisale da je područje e-leksikografije toliko razvijeno da se smatra zasebnim istraživačkim područjem (ŠTRKALJ DESPOT; MÖHRS 2015: 330), a to je područje otad ubrzano napredovalo. Uzlet je e-leksikografije posljedica svojevrsne promjene paradigme, tj. činjenice da se korisnici rječnika učestalo koriste internetom.<sup>15</sup> Filipović-Petrović i Parizoska donose pregled područja e-leksikografije i e-rječnika (FILIPOVIĆ-PETROVIĆ; PARIZOSKA 2022). De Schryver donosi popis prednosti koje donosi e-leksikografija za korisnika, leksikografa i društvo (DE SCHRYVER 2003: 152–187).

Ključna je prednost koja nas je potakla na izradu e-inačice RCJHR-a brz pristup informacijama te besplatna uporaba bilo gdje u svijetu. Ako postoji e-inačica rječnika, svaki korisnik koji ima pristup internetu našem bi rječniku mogao pristupiti bilo od kud u svijetu bilo u kojemu trenutku bez potrebe imanja fizičkoga rječnika. Iz leksikografske perspektive najvažnija nam je prednost to što se rječnik može slobodno posuvremenjivati, što se pogreške mogu ispraviti te što se mogu dodavati nove riječi pod već obrađena slova. To

<sup>12</sup> Više o projektu *Istraživanje starije hrvatskoglagoljske zborničke baštine* v. [http://elte2.fsb.hr:8000/staroslavenski\\_institut](http://elte2.fsb.hr:8000/staroslavenski_institut).

<sup>13</sup> Više o Sketch Engineu v. <https://www.sketchengine.eu/> i npr. KILGARRIFF; RYCHLÝ; SMRZ; TUGWELL 2004.

<sup>14</sup> Više o uporabi toga programa u izradi hrvatskih mrežnih rječnika v. MIHALJEVIĆ, J. 2022: 59–63.

<sup>15</sup> Usp. FUENTES-OLIVERA; BERGENHOLTZ 2011: 2.

u tiskanoj inačici nije moguće, tj. dosad su povremeno dodavane riječi na kraju novijih sveščića, međutim takva praksa otežava korisniku pronalazak riječi.

Štrkalj Despot i Möhrs donose i pregled vrsta rječnika prema dotad relevantnoj literaturi (ŠTRKALJ DESPOT; MÖHRS 2015: 332–333). Kao vrste e-rječnika navode se rječnici na samostojećemu računalu kojima se može služiti jedan korisnik na jednome računalu (više korisnika ne može se njime koristiti istodobno, nego se korisnici moraju izmjenjivati) i rječnik na mreži, kojim se može koristiti više korisnika istodobno. E-izdanje RCJHR-a zamišljeno je kao rječnik na mreži kojim se istodobno može koristiti neograničeni broj korisnika. Zatim se rječnici dijele na: unutamrežne (intranetske) rječnike, kojima se najčešće koristi na mreži unutar određene ustanove, i mrežne (ili internetske) rječnike. Naš je rječnik zamišljen kao mrežni, tj. dostupan i korisnicima izvan naših ustanova, ali je trenutačna demoinačica dostupna samo odabranim korisnicima dok ne prođe testnu fazu. Mrežni rječnici dijele se na statične i dinamične. Statični su rječnici gotovi rječnici koji se ne dopunjuju i najčešće su digitalna preslika rječnika bez izmjena u tekstu. Dinamični rječnici mogu nastati iz izvorno tiskanoga rječnika koji je prvo pretvoren u digitalni oblik, a zatim se digitalna inačica proširuje i obogaćuje (npr. *Oxford English Dictionary*), ili mogu biti izvorno digitalni (engl. *born digital*), kao npr. rječnik Instituta za hrvatski jezik i jezikoslovlje *Mrežnik*,<sup>16</sup> koji nije prethodno bio izdan kao tiskana knjiga, nego je izvorno nastao na mreži. Prema toj podjeli RCJHR pripada u skupinu rječnika koji su izvorno objavljeni u tiskanu obliku i svesci koji su dosad objavljeni zasad se uglavnom pri mrežnome objavljivanju ne dorađuju iako se ispravljaju pogreške koje se pri digitalizaciji uoče i u budućnosti postoji mogućnost daljnjega dorađivanja ili proširivanja već obrađenih riječi. Rječnik se ipak može definirati i kao dinamičan jer se i dalje izrađuje.

Prema kriteriju hipertekstualizacije RCJHR je hipertekstualiziran jer pojedine natuknice upućuju na druge, npr. u rječniku se navode sinonimi pod oznakom Cf. (od lat. *confer* ‘uspoređi’), pojedini alternativni oblici riječi upućuju na normalizirani oblik, navode se različite inačice pojedinih riječi i sl. RCJHR također pripada skupini rječnika vođenih korpusom.

Prema ovdje osmišljenome modelu RCJHR je zasad zamišljen kao rječnik bez mogućnosti interakcije s korisnicima i bez multimedijских sadržaja. Međutim, s obzirom na to da je riječ o digitalnome rječniku, koncepciju je moguće i kasnije mijenjati i dorađivati te se u budućnosti mogu dodati i drugi sadržaji ili se može omogućiti interakcija s korisnicima. Rječničkim se član-

<sup>16</sup> Više o *Mrežniku* v. <http://ihjj.hr/mreznik/>; HUDEČEK; MIHALJEVIĆ 2020.

cima na portalu može pristupiti kombinirano – s pomoću popisa natuknica s hiperpoveznicama, ali i s pomoću tražilice. U sklopu portala moguće je pristupiti i inačici rječnika u PDF-u, u kojoj se rječničkim člancima pristupa s pomoću pomicanja prema dolje kroz niz rječničkih članaka.

Kao uzor pri izradi poslužila su dva portala na kojima su dostupni rječnici starocrkvenoslavenskoga jezika – *Gorazd* i *Cyrillomethodiana*. Pri izradi mrežnoga izdanja SLOVNÍKA na portalu *Gorazd*, osim ispravljavanja pogrešaka, dodane su i engleske istovrijednice kojih u tiskanome izdanju nema, tj. rađene su nešto veće preinake od preinaka u izradi našega mrežnog rječnika. Na portalu *Gorazd* riječi se također mogu pretraživati s pomoću virtualne tipkovnice, ali se, za razliku od našega rječnika, ne nudi mogućnost prikaza svih natuknica pod pojedinim slovom. Na portalu *Cyrillomethodiana* mogu se prikazati sve natuknice koje počinju određenim slovom ili se pojedina natuknica može upisati, ali ne nudi se mogućnost virtualne tipkovnice. Na obama se portalima može pristupiti više rječnika te oba osim digitaliziranih rječnika donose i druge sadržaje – primjerice korpuse, kartičnu građu i sl.

#### 4. IZRADA MREŽNE INAČICE RCJHR-A

Pri istraživanju mogućnosti stvaranja mrežne inačice rječnika koja bi zadovoljila sve zahtjeve korisnika, obrađivača i urednika iskristalizirali su se ovi kriteriji za odabir programa za izradu e-inačice RCJHR-a:

1. mogućnost implementiranja starih pisama
2. lakoća uporabe za suradnike na rječniku koji nisu informatički stručnjaci
3. mogućnost povezivanja s PDF dokumentom skeniranoga tiskanog izdanja
4. mogućnost lakoga unosa novih rječničkih članaka s obzirom na to da se rječnik još izrađuje
5. mogućnost usporednoga rada više suradnika
6. mogućnost dodjeljivanja različitih uloga suradnicima (razlikovanje obrađivača od urednika)
7. mogućnost upravljanja kriterijima pretrage
8. laka dostupnost programa (ako je moguće, besplatan program).

Nakon toga pokušali smo pronaći program koji bi na najbolji način mogao zadovoljiti najveći broj postavljenih kriterija. Pri analizi krenuli smo od

programa kao što su *Lexonomy*<sup>17</sup>, *iLex*<sup>18</sup> i *TshwaneLex*<sup>19</sup>. Prednost je programa *TshwaneLex* što je osmišljen upravo za sastavljanje digitalnih rječnika te omogućuje detaljnu izradu poželjne strukture rječnika, no nedostatak mu je što je komercijalan (njegovo se korištenje plaća) te što su njegove mogućnosti izvoza podataka iz rječnika za pregled na mreži ograničene. U svojoj osnovnoj inačici omogućuje izvoz rječničkih članaka rječnika kao niza statičnih HTML ili XML dokumenata koji se, kako bi se pravilno prikazali na mreži, naknadno moraju prilagoditi i prenijeti u drugi sustav ili programsko rješenje za izradu mrežne stranice (na projektu *Mrežnik* uvezilo se stvorene HTML datoteke iz *TshwaneLexa* u sustav *WordPress* s pomoću programskih dodatka)<sup>20</sup>. Zato je poželjnija opcija da se odabere programsko rješenje za sastavljanje rječnika u kojemu se nakon izrade strukture rječničkoga članka i upisivanja podataka, unutar istoga rješenja ti rječnički članci mogu objaviti na mreži. O radu u *iLexu* razgovaralo se s kolegama s Inštituta za slovenski jezik Frana Ramovša iz Slovenije te se proučio njihov portal *Fran*<sup>21</sup>. U suradnji s kolegama iz Češke s instituta Slovanský ústav Akademie věd ČR i Bugarske s instituta Institut "t za b"lgarski ezik »Profesor Lûbomir Andrejčin« isprobavana su njihova programska rješenja kojima su se oni koristili za mrežno izdanje rječnika *Slovník jazyka staroslověnského* (dalje SLOVNÍK), koji je mrežno dostupan u sklopu portala *Gorazd*<sup>22</sup> i za mrežno izdanje rječnika *Rečnik na b"lgarskiâ ezik*<sup>23</sup> te za druge rječnike u tim institutima.<sup>24</sup> Međutim, utvrđeno je da njihova rješenja, premda iznimno dobro rade za njihove rječnike, ne mogu biti dovoljno dobro prilagođena našem rječniku da bi omogućila učinkovitu izradu e-izdanja RCJHR-a. Iako se nismo mogli poslužiti njihovim rješenjima, njihova su nam iskustva i pristupi bili iznimno dragocjeni u procesu razvijanja vlastitih rješenja. U radu nam je iznimno korisna bila i mreža *Elexis*<sup>25</sup>, čijim smo se goto-

<sup>17</sup> Više o *Lexonomyju* v. <https://lexonomy.elex.is/>.

<sup>18</sup> Više o *iLexu* v. [https://issuu.com/jens.erlandsen/docs/ilex\\_brochure\\_120dpi](https://issuu.com/jens.erlandsen/docs/ilex_brochure_120dpi).

<sup>19</sup> Više o *TshwaneLexu* v. <https://tshwanedje.com/tshwanelex/>.

<sup>20</sup> Više v. MIHALJEVIĆ, J. 2022: 64–67.

<sup>21</sup> Portal je dostupan na <https://www.fran.si/>.

<sup>22</sup> Portal je dostupan na <http://gorazd.org/?q=cs>. Više o projektu *Gorazd* i programskim rješenjima za taj portal v. PILÁT 2017; PILÁT 2020; PILÁT; KNOLL 2017, također ŠIMIĆ 2022. Za programska rješenja v. TUČEK; ČIPEROVÁ; MAJER; PILÁT 2018.a; TUČEK; ČIPEROVÁ; MAJER; PILÁT 2018.b.

<sup>23</sup> Rječnik je dostupan na <https://ibl.bas.bg/rbe/>.

<sup>24</sup> Više o bugarskim i češkim mrežnim izvorima v. MIHALJEVIĆ, A. 2022.

<sup>25</sup> Više o *Elexisu* v. <https://elex.is/>.



vim rješenjima (kao što su *Lexonomy* i *Elexifier*<sup>26</sup>) također služili kao idejnim rješenjima, a i koja nam je omogućila suradnju sa stručnjacima iz Bugarske. Također, kao ciljni uzor služio nam je i portal *Cyrrillomethodiana*<sup>27</sup> te radovi GANEVA 2018; TOTOMANOVA 2012; TOTOMANOVA 2021.

Nakon analize različitih sustava i programa, koji osim prethodno spomenutih programa uključuju sustave *Joomla*<sup>28</sup>, *Publii*<sup>29</sup> i *Wix*<sup>30</sup>, odabran je sustav *WordPress*<sup>31</sup>. *WordPress* je odabran iz sljedećih razloga:

- besplatan je i dostupan bez potrebe za registracijom
- kao sustav za upravljanje sadržajem (engl. CMS – *Content management system*) omogućuje da više korisnika radi na rječniku koristeći se grafičkim sučeljem
- riječ je o sustavu otvorenoga koda koji ima mnogo dostupnih besplatnih dodataka koji proširuju njegove funkcionalnosti
- sustav ima jednostavno grafičko sučelje za rad te se lako održava redovitim posuvremenjivanjem.

*Joomla* također ima sve navedene funkcionalnosti kao *WordPress*, ali ima ograničenije mogućnosti za rad u odnosu na *WordPress*. To ograničenje proizlazi iz manjega broja dostupnih dodataka, manje zajednice ljudi koja se bavi unaprjeđenjem i održavanjem sustava te složenijega grafičkog sučelja posebice za nove korisnike.<sup>32</sup> *Publii* funkcionira kao sustav koji izrađuje statične mrežne stranice te, unatoč tomu što ima velike mogućnosti prilagodbe i ima jednostavno grafičko sučelje, ne omogućuje da više korisnika istodobno uređuje sadržaje, nego je u prvome redu namijenjen za osobnu uporabu. *Wix*, unatoč tomu što je jednostavniji od *WordPresa* kad je riječ o izradi dizajna, nije u potpunosti besplatan i otvorenoga koda te se ne može jednostavno prilagoditi za različite sadržaje kao *WordPress*.

Pri radu u sustavu *WordPress* trebalo je riješiti sljedeće probleme povezane s integracijom rječnika na mrežu:

- omogućiti integraciju latiničnih, glagoljičnih i ćiriličnih slova te grčkoga alfabeta za pisanje teksta

<sup>26</sup> Više o *Elexifieru* v. <https://elexifier.elex.is/>.

<sup>27</sup> Portal je dostupan na [histdict.uni-sofia.bg](http://histdict.uni-sofia.bg).

<sup>28</sup> Sustav je dostupan na <https://www.joomla.org/>.

<sup>29</sup> Sustav je dostupan na <https://getpublii.com/>.

<sup>30</sup> Sustav je dostupan na <https://www.wix.com/>.

<sup>31</sup> Sustav je dostupan na <https://wordpress.org/>.

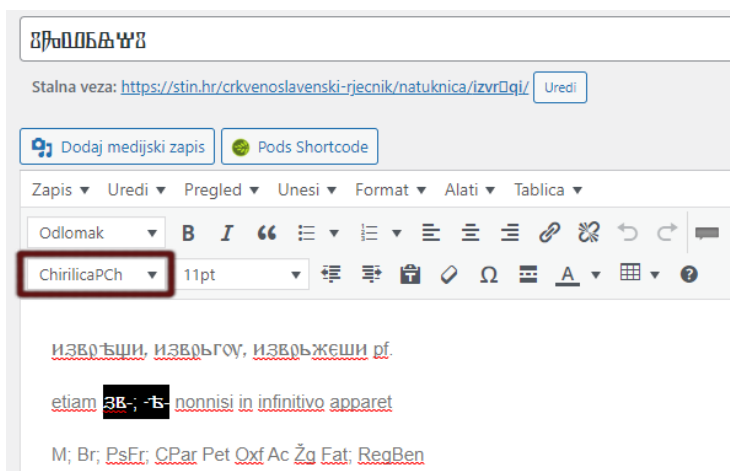
<sup>32</sup> Više o usporedbi na <https://www.hostinger.com/tutorials/joomla-vs-wordpress>.

- odvojiti rječničke članke od ostalih sadržaja koji se mogu nalaziti na stranici rječnika
- odrediti strukturu za unos podataka u sučelju sustava *WordPress*.

Rješenja tih problema prikazat će se u sljedećim poglavljima.

## 5. PODRŠKA ZA STARA PISMA

Jedna je od glavnih izmjena koja se morala napraviti u sustavu podrška za prikaz glagoljičnoga i ćiriličnoga fonta, koji se pojavljuju kao fontovi u kojima se pojavljuje natuknica. Fontovi FSGLA i ChirilicaPCh, od kojih je FSGLA font koji se može preuzeti na internetu,<sup>33</sup> integrirani su u sustav *WordPress* te se prikazuju administratorima, urednicima i obrađivačima tijekom uređivanja rječničkih članaka u sučelju za rad. Obradivači za promjenu pisma moraju samo označiti određeni tekst i promijeniti font u sučelju (v. *Sliku 2*). U sustavu je također određeno da je natuknica automatski pisana glagoljicom kako bi se ubrzao unos.

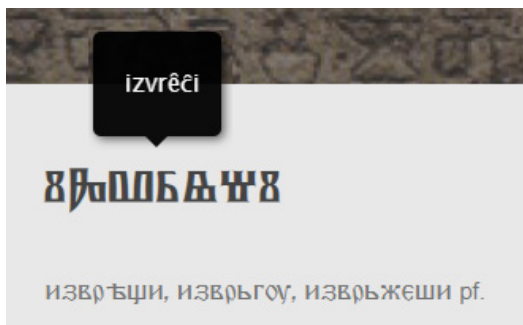


Slika 2. Izgled okvira za unos glavnoga teksta rječničkoga članka (kvadratom je označen dio sučelja u kojemu se mijenjaju fontovi)

Figure 2. The frame for the insertion of the main entry text (the square marks the interface for changing fonts)

<sup>33</sup> Font je dostupan na stranici <https://stin.hr/glagoljica/>.

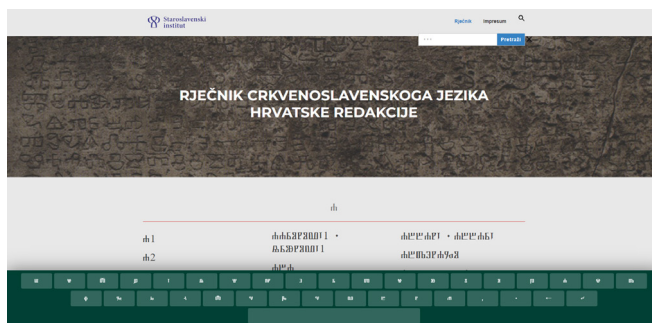
Kako bi se pomoglo korisnicima u prepoznavanju slova glagoljične azbuke, u sustavu je omogućen prikaz latinične transliteracije glagoljične natuknice. Transliteracija glagoljičnih slova prikazuje se kad se pokazivačem miša prijeđe preko natuknice ili, ako korisnik rječnik pregledava na mobitelu, pritiskom na riječ. Na *Slici 3* prikazan je primjer prikaza transliteracije pri pregledavanju rječnika na računalu.



Slika 3. Primjer prikaza transliteracije glagoljične natuknice

Figure 3. An example of the representation of transliteration of the Glagolitic headword

Dodatna je pomoć korisnicima u pretraživanju riječi da se, kad korisnik otvori polje tražilice rječnika, na dnu ekrana stvara prilagođena virtualna tipkovnica koja sadržava sva glagoljična slova, od kojih je neka teže utipkati standardnom QWERTY tipkovnicom. Na *Slici 4* prikazana je virtualna tipkovnica.



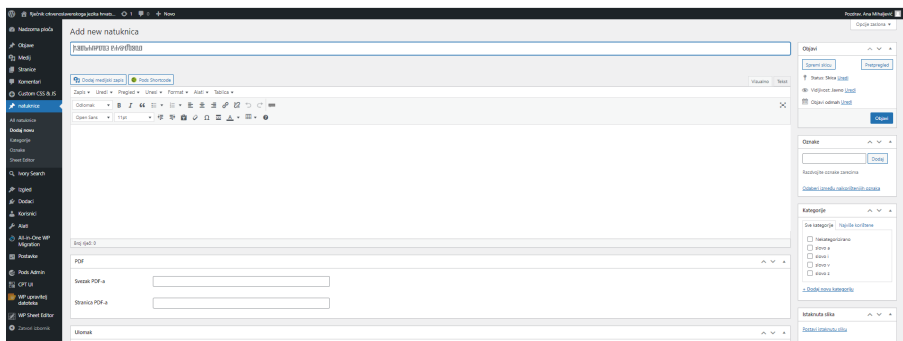
Slika 4. Prikaz virtualne tipkovnice koja se otvara pritiskom na polje tražilice te sadržava slova koja je teže unijeti s pomoću standardne fizičke tipkovnice

Figure 4. The virtual keyboard which opens by pressing the field of the search engine and contains letters difficult to insert by a standard keyboard

Problem u pretraživanju mogu biti i dijakritički znakovi te činjenica da se određeni znakovi zbog postavka fonta FSGLA ne pišu kao na latinici. Stoga su opcije pretraživanja izmijenjene te je omogućeno pretraživanje riječi pri kojemu se zanemaruju dijakritički znakovi, npr. pretraga riječi sa slovom *s* obuhvatit će i riječi koje sadržavaju slovo *š*, pretraga riječi sa slovom *c* obuhvatit će riječi *s ĉ* i *ĉ* i sl.

## 6. PODJELA SADRŽAJA RJEČNIKA

U *WordPressu*, koji je izvorno nastao kao sustav za pisanje bloga, glavni su tip sadržaja objave (engl. *post*). Budući da izrađujemo rječnik u kojemu su glavni sadržaj rječnički članci, odlučili smo da ćemo stvoriti novi prilagođenih tip sadržaja unutar sustava koji smo nazvali *natuknice*. Taj tip sadržava vlastita polja za ispunjavanje koja su određena rječničkom strukturom (za osnovnu strukturu za unos imenica vidi *Sliku 7* u sljedećemu poglavlju). Također, objave smo postavili kao odvojene sadržaje od rječničkih članaka jer se njima može ubuduće koristiti za pisanje obavijesti ili drugih tipova podataka na stranici koji nisu rječnički članci ili dijelovi rječničkih članaka. Na *Slici 5* prikazan je prvi dio polja prilagođenoga tipa sadržaja koji smo nazvali *natuknice*.



Slika 5. Polje za dodavanje novoga rječničkog članka  
Figure 5. The field for the insertion of a new entry

Rječnički su članci kategorizirani i raspoređeni na stranici u pododjeljke koji su određeni prema početnome slovu riječi (npr. *natuknice* koje počinju slovom *a* nalaze se u kategoriji *a*). Obrađivač pri unosu riječi određuje kojoj





Slika 7. Prikaz prvoga dijela popisa slova  
Figure 7. The first part of the list of letters

## 7. STRUKTURA ZA UNOS PODATAKA

U početnim fazama izrade e-inačice rječnika u sustavu *WordPress* rječnički članci pisali su se unutar glavnoga tekstnog okvira sustava. Samo oblikovanje i strukturiranje podatka radilo se u tekstnome okviru na sličan način kao oblikovanje teksta unutar *Wordovih* dokumenta. To je prvotno provedeno za nekoliko rječničkih članaka, ali nakon što se precizno odredila struktura rječničkoga članka i našao način unutar sustava da se u odvojenim poljima unose različiti podatci s pomoću dodatka *Pods*<sup>34</sup>, rječnički su se članci počeli unositi i obrađivati u sklopu polja. Razlozi su za obradu natuknica s pomoću unaprijed određenih polja sljedeći:

1. strukturiranje podataka
2. ujednačavanje izgleda prikaza rječničkoga članka u rječniku
3. mogućnost pretraživanja te uvoza i izvoza strukturiranih podatka.

Dodatak *Pods* omogućuje da se definira više polja za pojedine dijelove rječničkoga članka te se za ta polja mogu odrediti tipovi vrijednosti koji se unose u polja (tekst, slike, mrežne poveznice itd.). Strukturiranje podataka unutar rječničkoga članka rječnika važno je kako bi se moglo lakše rukova-

<sup>34</sup> Dodatak je dostupan na <https://wordpress.org/plugins/pods/>.

ti podacima. Urednici lakše mogu ispravljati informacije unutar rječničkoga članka rječnika kad znaju unutar kojih se polja nalaze. Kod unosa rječničkoga članka ne trebaju paziti na njihovo grafičko oblikovanje kao kad se sve radilo unutar pojedinoga tekstnog okvira. To je omogućeno postavljanjem sustava kako bi određeni podatci bili prikazani pri pregledu mrežne stranice. Određena polja mogu se kopirati jer postoji mogućnost da se trebaju više puta ponoviti. Jedan je od glavnih razloga za strukturiranje podataka i činjenica da se podatci iz rječnika, ako se strukturiraju, mogu izvesti u različitim drugim formatima poput TEI-ja<sup>35</sup>. Sva polja rječničkoga članka nalaze se unutar baze sustava *WordPress* te se mogu izvesti iz tablica u strukturiranome formatu (npr. kao CSV, XML JSON ili XLSX). Taj strukturirani format dalje se može prebaciti u drugu shemu (npr. spomenuti TEI) ili se njime može koristiti za uvoz podataka u druge sustave. Na *Slici 8* prikazana je osnovna struktura za unos imenica.

---

<sup>35</sup> TEI (Text Encoding Initiative) standardiziran je XML format koji se upotrebljava za označivanje tekstova u digitalnome obliku. Često se njime koristi pri označivanju starih izdanja i suvremenih leksikografskih izdanja. O standardu se brine TEI konzorcij koji je i odredio oznake kojima se treba koristiti za označivanje dijelova teksta u XML-u. Za razliku od drugih jezika za označivanje poput HTML-a, TEI oznake ne služe u prvome redu za grafički prikaz podataka, nego za njihovo semantičko označivanje. Oznakama unutar TEI-ja možemo detaljno opisati svaki dio teksta ili dodati metapodatke koji mogu biti povezani s autorima, urednicima, datumima i podrijetlom teksta. To mogu biti podatci koji uključuju fizičku i digitalnu inačicu (BOWERS; ROMARY 2016: 2–3).

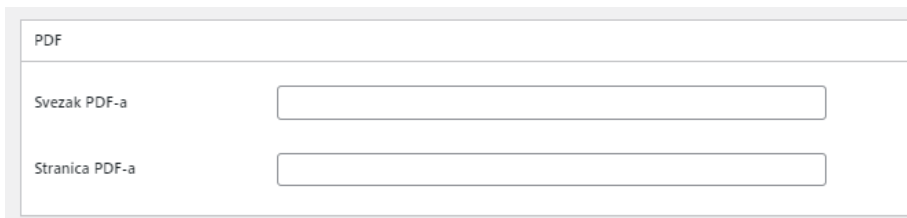




## 8. POVEZIVANJE SA STRANICOM U TISKANOME RJEČNIKU

U strukturi rječničkoga članka dodano je polje u sustavu koje stvara poveznicu između rječničkoga članka i stranice rječnika u PDF-u. PDF rječnika izrađen je u programu ABBYY FineReader, unutar kojega su se učitale skenirane stranice rječnika koje su se dodatno grafički uredile (izbrisalo se crne rubove i izravvalo stranice). Pokušali smo u programu napraviti i optičko prepoznavanje znakova (OCR), ali program ne može prepoznati određene glagoljične i ćirilične simbole i OCR ne daje zadovoljavajuće rezultate, tj. očitavanje nije dovoljno precizno da bi omogućilo učinkovito pretraživanje dokumenta. Rječnički članak i stranica tiskanoga rječnika povezane su, osim toga da se može pronaći natuknica u izvornome rječniku, kako bi autori rječnika korisnicima omogućili što brži pristup natuknicama rječnika prije nego što se one u potpunosti digitalno obrade i prikažu jer je proces obrade određenih natuknica za mrežni rječnik dugotrajan te je brže napraviti poveznicu na stranicu PDF-a tiskane inačice rječnika. Poveznica na dokument u PDF-u osobito je bila potrebna za rječničke članke iz prvih deset sveščića RCJHR-a. Naime, riječ je o starijim sveščićima koji nisu dostupni u obliku teksta, nego je jedini njihov oblik PDF lošije kvalitete. Taj pak format nije dovoljno precizno čitljiv programima za OCR kako bi se tekst rječničkoga članka mogao kopirati u polja unutar sustava.

Na *Slici 9* prikazano je polje za unos podatka o svesku i stranici u tiskanome izdanju.



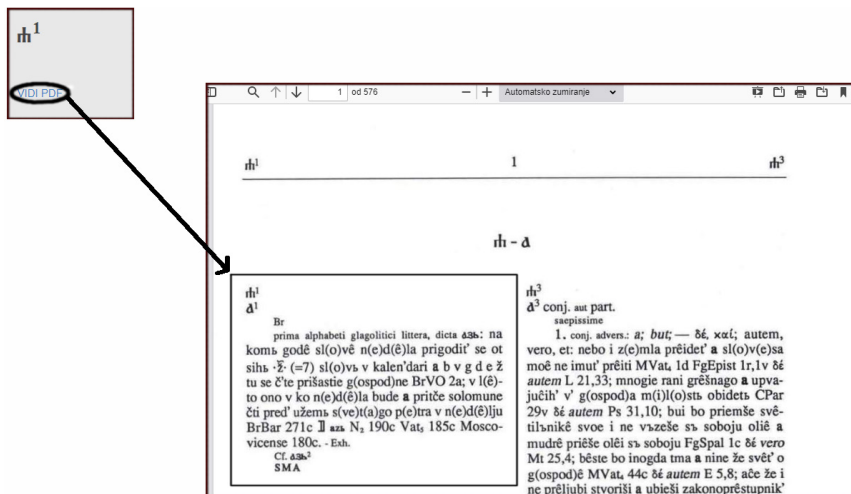
PDF	
Svezak PDF-a	<input type="text"/>
Stranica PDF-a	<input type="text"/>

Slika 9. Polje za unos podatka o svesku i stranici u tiskanome izdanju.

Figure 9. The field for the insertion of the data on the volume and page of the printed edition

Na *Slici 10* vidljiv je prikaz poveznice na PDF tiskane inačice rječnika i stranica koja se otvara pritiskom mišem na poveznicu. U prvoj inačici mrežnoga rječnika hiperveze na rječničke članke koji trenutačno sadržavaju samo

poveznicu na PDF bez obrade u sustavu izravno otvaraju PDF stranicu dokumenta na kojoj se nalazi rječnički članak.

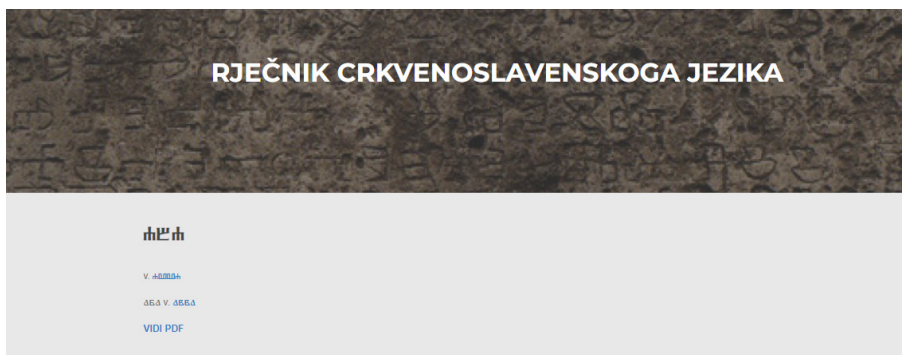


Slika 10. Prikaz poveznice koja vodi na stranicu na kojoj se nalazi rječnički članak u PDF dokumentu

Figure 10. A link to the page of the entry in PDF format

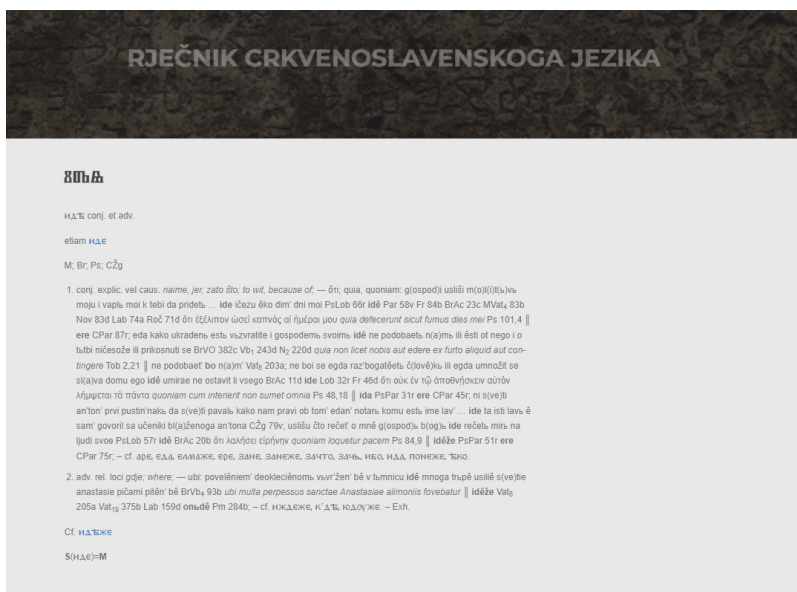
## 9. POVEZIVANJE RIJEČI

Prednost je e-rječnika pred tiskanima lakoća prelaska s natuknice na natuknicu. U RCJHR-u mnogo je natuknica s kojih se upućuje na drugu natuknicu (rjeđi oblici riječi upućuju na češći oblik, nenormalizirani oblici upućuju na normalizirane, upućuje se na sinonime itd.). Kad se susretne s natuknicom koja upućuje na drugu natuknicu, korisnik tiskanoga izdanja rječnika mora fizički tražiti drugu natuknicu okretanjem stranica. E-izdanje rječnika omogućuje da se riječi povezuju poveznicama te da korisnik pritiskom mišem na poveznicu automatski bude preusmjeren na drugu riječ. Na *Slici 11* prikazan je primjer natuknice koja upućuje na drugu natuknicu. Pritiskom mišem na riječ *abba* korisnik je preusmjeren na rječnički članak s tom natuknicom.



Slika 11. Natuknica koja upućuje na drugu natuknicu  
Figure 11. A headword linked to another headword

Takvo je povezivanje moguće i za navođenje sinonima, koji su u RCJHR-u navedeni kao Cf. Na *Slici 12* prikazan je primjer natuknice za koji je unesen sinonim na koji se upućuje s pomoću hiperpoveznice. Pritiskom mišem na riječ *идѣже* korisnik je preusmjeren na rječnički članak s tom natuknicom.



Slika 12. Primjer povezivanja natuknice i njezina sinonima s pomoću hiperpoveznice  
Figure 12. An example of linking the headword and its synonyms by a hyperlink

E-izdanje otvara i mogućnost daljnjega povezivanja. Na kraju se rječničkoga članka u tiskanome RCJHR-u pojavljuje podatak o tome navodi li se riječ u trima rječnicima: SLOVNÍKU; MIKLOSICH 1963 i ARj 1880–1976. SLOVNÍK i MIKLOSICH 1963 u doba početaka izrade RCJHR-a bila su dva najpotpunija i najvažnija rječnika crkvenoslavenskoga jezika, a ARj 1880–1976 povijesni je rječnik hrvatskoga jezika.<sup>36</sup> Budući da postoji mrežno izdanje SLOVNÍKA, u e-izdanju rječnika moguće je izravno povezivanje pojedine natuknice s natuknicom u SLOVNÍKU. Primjer takvoga povezivanja prikazan je na *Slici 13*. Na kraju se rječničkoga članka natuknice *igo* oznakom S potvrđuje da je riječ zabilježena i u SLOVNÍKU. Pritiskom mišem na oznaku S korisnik se automatski prebacuje na rječnički članak iz SLOVNÍKA na portalu *Gorazd* u kojemu je obrađena riječ *igo*.

**ИГО**

ИГО, –а. п.

in PsFr dat. sg. игѡѣ

M; Br; PsFr(Comm); RegBen

*jaram. teret, yoke, burden: — ζυγός; iugum; — fig.: otímeť se iġo ležeće na nih' BrVO 38c Vb<sub>1</sub> 29a Pm 26a Vat<sub>2</sub> 21c Vat<sub>19</sub> 22b N<sub>2</sub> 21c Bar 100b ὀφειλῆται ὁ ζυγός ὁ ἐπὶ αὐτῶν κειμενος iugum enim oneris eius ... superasti Is 9,4; ot toga dne nemu nist' digъ ot molstira izlisti ni iġa ili ġrma regule s'voje šie z'vríci RegBen 38b nec collum excutere de sub iugo regulae; prideľ že vréme vneġda razoríši i razdréšiši iġo ego ot vie tvoee BrVO 166c τὸν ζυγὸν αὐτοῦ iugum eius Gn 27,40; predast' e b'ogъ iġs tebě i povinuti buduť. pod' iġo kréposti tvoee BrVO 389b Vb<sub>1</sub> 248c N<sub>2</sub> 228b sub iugo potentiae tuae Jdt 5,24; podai m(o)ġim le vs(e)m(o)ġi b'ozě da iže pod grěha iġom' vethoju rabotoju pognátem se gredućimi' edinočedago s(i)ġna tvoego novimi' rožd'stvomi' da izb(a)vim se MVal<sub>4</sub> 5b Nov 6c Roč 4c sub peccati iugo; sim. BrVO 38b Pm 25d Vat<sub>2</sub> 21b N<sub>2</sub> 21b Bar 100a; ne ibo povinuše se iġové v(ě)řě iġuděisci ġudi vsudu že se protivlahu PsFr 1d τὸ ζυγὸν τῆς πτοστεως CommPs 2,3; egda s(ve)ġi moisei s(i)ġni iz(dra)ġ(e)vi ot iġa eġupčans. plěna izb(a)vi BrVal<sub>4</sub> 76a N<sub>2</sub> 74c Bar 164b de iugo Aegyptiae captivitatis ¶ ot raboti VO 156c Vb<sub>1</sub> 103c; prosim' ġ(ospod)ġi priš'stvě tvoego da skoro prideš i otrěšiši iġo plěnenié n(a)šego BrVO 24b Vb<sub>1</sub> 15b Pm 10b Vat<sub>2</sub> 9a Bar 84c;*

*fig. de praecceptis Euangelii, de Christum sequendo: vřzmitě iġo moe na vi ... iġo bo moe bl(a)ġo est' MVal<sub>4</sub> 243b/c Nov 245d/246a Roč 197a Lab 207c τὸν ζυγὸν μου ... ὁ γὰρ ζυγός μου iugum meum ... iugum enim meum Mt 11,29; sim. BrN<sub>2</sub> 325b/c Vb<sub>1</sub> 103d Vat<sub>2</sub> 36b Ac 40a Vat<sub>19</sub> 228a Bar 466a; d'ěvġju nečistotu i blġenie otvrġoše i priěše iġo ġ(ospod)ne slatġoe i brěme ego laġkoe BrVal<sub>4</sub> 114b N<sub>2</sub> 425c Lab 56b Vat<sub>19</sub> 295b iugum domini suave; c(ěsa)ġi z(e)m(l)bsci ... h(r)sta) radi idoli potřebě i povinut' se vse vlasti pod' iġo h(r)sto)vo BrN<sub>2</sub> 330d Vb<sub>1</sub> 112a Vat<sub>2</sub> 41d Bar 471c iugo Christi; raz'boťěv še beš'tinoju vřizrav še se iġo h(r)sto)vo otvr'ġše ot s(e)be PsFr 119a τὸν ζυγὸν τοῦ Χριστοῦ ἀπέρριψαν ὡφ' ἑαυτῶν CommPs 143,14.*

*Erronee pro cγχο: o m(ġ)žu dġv'ni (sc. frančisko) v zn(aj)m(e)ni i čġdesěh' nedugi i bėsi progone iġo dae vs(ě)m' pġcat'm dubrav'ni'm prop(o)v(ě)dae ġ(ago) i rasmotrenié BrVal<sub>4</sub> 360d ¶ uho N<sub>2</sub> 493a Vat<sub>2</sub> 185c dat aurem suis avium praedicans silvestrium verbis intendentem.*

Cf. ѣбъиѣ

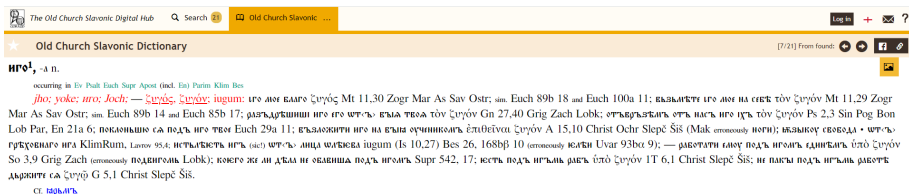
SMA

Slika 13. Mogućnost izravnoga povezivanja s natuknicom iz SLOVNÍKA na portalu Gorazd

Figure 13. The possibility of direct linking to headwords in SLOVNÍK on the Gorazd portal

<sup>36</sup> V. i Šimić 2017: 108.

Na *Slici 14* prikazana je stranica iz SLOVNÍKA s portala *Gorazd* koja se otvara pritiskom mišem na poveznicu S na prethodnoj slici.



Slika 14. Rječnički članak iz SLOVNÍKA koja se otvara pritiskom mišem na poveznicu S na prethodnoj slici

Figure 14. An entry from SLOVNÍK which is opened by pressing the link S on the previous page

## 10. ZAKLJUČAK I DALJNI PLANOVI

Nakon što su analizirani dostupni rječnički programi i sustavi te nakon što je zaključeno da nijedan sustav ili program nije prikladan za obradu našega rječnika, odabrano je programsko rješenje koji nije izvorno namijenjeno za izradu rječnika, ali koje je uz niz prilagodba ocijenjeno kao najprikladnije rješenje za izradu e-inačice RCJHR-a. Napravljene su potrebne prilagodbe i struktura za unos rječničkih članaka, a sad u sustav treba unijeti sve obrađene rječničke članke. Dosad su uneseni svi rječnički članci za natuknice pod slovima *a* i *i* te dio natuknica pod slovom *z*. Slova *a* i *i* odabrana su jer su natuknice koje započinju slovom *a* najstarije rječničke obrade dostupne samo u fizičkome i skeniranome obliku, a natuknice koje započinju slovom *i* posljednje obrade koje su dostupne i u DOCX obliku, stoga se različito pristupalo njihovom prenošenju na mrežu i za svaku od tih skupina izradio se poseban tip pravila i posebna struktura. Tako je osmišljen model za izradu e-inačice RCJHR-a te za računalnu potporu daljnjoj obradi RCJHR-a.

U sljedećemu razdoblju planira se unijeti i ostala obrađena rječnička građa. Probna inačica e-rječnika trenutačno je dostupna na stranici <https://stin.hr/cr-kvenoslavenski-rjecnik/>, ali samo uz unos lozinke. Kad se unese još koje slovo, otvorit će se pristup korisnicima bez lozinke. Sustav WordPress redovito se posuvremenjuje i nadograđuje novim funkcionalnostima te će se u buduć-

nosti vjerojatno moći proširiti funkcionalnosti rječnika te unaprijediti dizajn bez potrebe da se izrađuje nova stranica. Mogućnost ostavljanja komentara na natuknice jedna je od mogućnosti koja već postoji u sustavu te se planira ugraditi u kasnijoj fazi radi poboljšanja usluge. Nakon što se izradila osnovna struktura za imenice, pristupilo se izradi strukture za druge vrste riječi. Nakon početnoga unosa svih rječničkih članaka, sve će se natuknice povezati s natuknicama na koje upućuju te sa sinonimima. Postoji mogućnosti i povezivanja sa SLOVNÍKOM, na koji i tiskano izdanje RCJHR-a upućuje.

## DIGITALNE REFERENCIJE

- ABBY® FlexiCapture®*. URL: <https://www.abby.com/flexicapture/> (13. 4. 2023.).
- Cyrrilomethodiana*. URL: [histdict.uni-sofia.bg](http://histdict.uni-sofia.bg) (13. 4. 2023.).
- Elexifier*. URL: <https://elexifier.elex.is/> (19. 12. 2022.).
- Elexis – European lexicographic infrastructure*. URL: <https://elex.is/> (19. 12. 2022.).
- Fran – Slovarji Inštituta za slovenski jezik Frana Ramovša ZRC SAZU*. URL: <https://www.fran.si/> (13. 4. 2023.).
- GORAZD: Digitální portál staroslověnštiny*. URL: <http://gorazd.org/?q=cs> (19. 12. 2022.).
- Hostinger*. URL: <https://www.hostinger.com/tutorials/joomla-vs-wordpress> (13. 11. 2023.).
- Hrvatski mrežni rječnik – Mrežnik*. URL: <http://ihjj.hr/mreznik/> (22. 11. 2022.).
- iLEX integrated Lexicography, Editing and XML*. URL: [https://issuu.com/jens.erlandsen/docs/ilex\\_brochure\\_120dpi](https://issuu.com/jens.erlandsen/docs/ilex_brochure_120dpi) (13. 4. 2023.).
- Interdisciplinarni pristup hrvatskoglagoljskomu misalu na primjeru Misala kneza Novaka*. URL: <https://stin.hr/nomis/> (13. 4. 2023.).
- Istraživanje starije hrvatskoglagoljske zborničke baštine*. URL: [http://elte2.fsb.hr:8000/staroslavenski\\_institut](http://elte2.fsb.hr:8000/staroslavenski_institut) (13. 4. 2023.).
- Joomla*. URL: <https://www.joomla.org/> (13. 11. 2023.).
- Lexonomy*. URL: <https://lexonomy.elex.is/> (19. 12. 2022.).
- Mrežna inačica Rječnika crkvenoslavenskoga jezika hrvatske redakcije*. URL: <https://stin.hr/crkvenoslavenski-rjecnik/> (13. 4. 2023.).
- Publii*. URL: <https://getpublii.com/> (13. 11. 2023.).
- Rečnik na b"lgarskiâ ezik. = Речник на българския език*. URL: <https://ibl.bas.bg/rbe/> (13. 4. 2023.).
- Sketch Engine*. URL: <https://www.sketchengine.eu/> (19. 12. 2022.).
- Transkribus*. URL: <https://readcoop.eu/transkribus/> (19. 12. 2022.).
- TshwaneDJe Software and Consulting*. URL: <https://tshwanedje.com/tshwanelex/> (19. 12. 2022.).
- Wix*. URL: <https://www.wix.com/> (13. 11. 2023.).
- WordPress*. URL: <https://wordpress.org/> (19. 12. 2022.).
- Znanstveni centar izvrsnosti za hrvatsko glagoljaštvo*. URL: <https://zci.stin.hr/> (13. 4. 2023.).

## LITERATURA

- ARj 1880–1976. = *Rječnik hrvatskoga ili srpskoga jezika*. Zagreb: Jugoslavenska akademija znanosti i umjetnosti.
- BOWERS, J.; L. ROMARY. 2016. Deep Encoding of Etymological Information in TEI. *Journal of the Text Encoding Initiative 10*: 1–44.
- ČAVAR, D.; I. JAZBEC; B. NAHOD. 2009. Struktura i razvoj baze podataka za potrebe projekta Hrvatsko strukovno nazivlje (STRUNA) – projekt koordinacije. N. Ledinek; M. Žagar Karer; M. Humar. (ur.). *Terminologija in sodobna terminografija*. Ljubljana: Založba ZRC, ZRC SAZU, 311–317.
- DE SCHRYVER, G.-M. 2003. Lexicographers' dreams in the electronic-dictionary age. *International Journal of Lexicography 16/2*: 143–199.
- FILIPOVIĆ PETROVIĆ, I.; J. PARIZOSKA. 2022. Elektronički rječnici u informacijsko doba: primjeri rješenja u rječnicima hrvatskoga i drugih jezika. *Studia lexicographica 31/16*: 111–134.
- FUERTES-OLIVERA, P. A.; H. BERGENHOLTZ. 2011. *e-Lexicography: The Internet, Digital Initiatives and Lexicography*. London – New York: Continuum.
- GANEVA, G. 2018. Electronic Diachronic Corpus and Dictionaries of Old Bulgarian. *Studia Ceranea 8*: 111–119.
- HUDEČEK, L.; M. MIHALJEVIĆ. 2020. The Croatian Web Dictionary – Mrežnik Project – Goals and Achievements. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje 46/2*: 645–667.
- KAPPEL, G.; J. KAPPEL; J. VELA (ur.). 2023. *Brevijar Vida Omišljanina (1396.) – preslovljeni tekst / transliteracija*. Zagreb: Staroslavenski institut.
- KILGARRIFF, A.; P. RYCHLÝ; P. SMRZ; D. TUGWELL. 2004. The Sketch Engine. G. Williams; S. Vessier (eds.). *Proceedings of the 11th EURALEX International Congress*. Lorient: Université de Bretagne Sud, 105–116.
- MAGDIĆ, A. 2015. Kako je digitalizirana građa Staroslavenskoga instituta. V. Badurina Stipčević; S. Požar; F. Velčić (ur.). *Hrvatsko glagoljaštvo u europskom okružju*. Zagreb: Staroslavenski institut, 167–209.
- MAREŠ, F. V. 2007. [1962]. Návrh přípravných prací pro slovník jazyka církevněslovanského. E. Bláhová; E. Šlaufová; V. Čermák (ur.). *Církevněslovanská lexikografie 2006*. Praha: Slovanský ústav AV ČR, Euroslavica, 64–84.
- MIHALJEVIĆ, A. 2022. Bugarski i češki mrežni jezikoslovni izvori. *Hrvatski jezik 9/4*: 20–24.
- MIHALJEVIĆ, J. 2022. *Igrifikacija hrvatskoga mrežnog rječnika*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- MIKLOSICH, F. 1963. *Lexicon Palaeoslovenico-Graeco-Latinum*. Wien: Scientia Verlag Aalen.
- NAZOR, A. 1997. Napomene o izvorima *Rječnika crkvenoslavenskoga jezika hrvatske redakcije*. M. Миовски; Д. Пандев; Е. Црвенковска (ур.). *Зборник во чест на Радмила Угринова-Скаловска: по повод седамдесетгодишнината*. Скопје: Филолошки факултет »Блаже Конески«, 195–200. [M. Miovski; D. Pandev; E. Cr-

- venkovska (ur.). *Zbornik vo čest na Radmila Ugrinova-Skalovska: po povod sedamdesetgodišnjinata*. Skopje: Filološki fakultet »Blaže Koneski«, 195–200.]
- NAZOR, A. 2008. Rječnik crkvenoslavenskoga jezika iniciran na IV. međunarodnom slavističkom kongresu u Moskvi 1958. godine (u povodu 50. obljetnice inicijative). M. Samardžija (ur.). *Vidjeti Ohrid. Referati hrvatskih sudionika i sudionika za XIV. međunarodni slavistički kongres (Ohrid, 10. – 16. rujna 2008.)*. Zagreb: Hrvatsko filološko društvo – Hrvatska sveučilišna naklada, 65–82.
- NAZOR, A. 2017. O početcima rada na *Rječniku crkvenoslavenskoga jezika hrvatske redakcije*. *Filologija* 68: 117–125.
- PILÁT, Š. 2017. GORAZD Project: Digital Portal of Old Church Slavonic. B. Niševa; K. Skwarska; D. Blažek (eds.). *Moderní slovanská lexikografie*. Praha: Slovanský ústav AV ČR, v. v. i., 39–48.
- PILÁT, Š. 2020. Elektronický Slovník jazyka staroslověnského – perspektivy dalšího rozvoje. *Bohemica Olomucensia* 12/2: 74–90.
- PILÁT, Š.; V. KNOLL. 2017. Bericht vom Projekt GORAZD: Digitales Portal des Altkirchenslawischen. *Byzantinoslavica* 75/1–2: 338–340.
- RABUS, A. 2022. Handwritten Text Recognition for Croatian Glagolitic. *Slovo* 72: 181–192.
- RCJHR 1991 = *Rječnik crkvenoslavenskoga jezika hrvatske redakcije. Uvod*. Zagreb: Staroslavenski institut.
- RCJHR 2000 = *Rječnik crkvenoslavenskoga jezika hrvatske redakcije. A – VRĚD*. Zagreb: Staroslavenski institut.
- RCJHR 2015 = *Rječnik crkvenoslavenskoga jezika hrvatske redakcije. VRĚDbNb<sup>1</sup> – ZAPOVĚDNICA*. Zagreb: Staroslavenski institut.
- RCJHR 2016 = *Rječnik crkvenoslavenskoga jezika hrvatske redakcije. ZAPOVĚDb – ZEMLbNb*. Zagreb: Staroslavenski institut.
- RCJHR 2018 = *Rječnik crkvenoslavenskoga jezika hrvatske redakcije. ZEMLbNb – I*. Zagreb: Staroslavenski institut.
- RCJHR 2021 = *Rječnik crkvenoslavenskoga jezika hrvatske redakcije. I – IZGNILSTVO*. Zagreb: Staroslavenski institut.
- RIBAROV, K.; M. CAMUGLIA. 2003. Incorporation of Old Church Slavonic Card Files into a Corpus. *Scripta & e-Scripta* 1: 65–74
- SLOVNÍK 1959–1997. = *Slovník jazyka staroslověnského*. Praha: Academia, nakladatelství Československé akademie věd.
- Struna – Hrvatsko strukovno nazivlje*. URL: <http://struna.ihjj.hr/> (1. 3. 2023.).
- ŠIMIĆ, A. 2017. Mikrostruktura i istraživački izazovi obrade rječnika crkvenoslavenskoga jezika hrvatske redakcije. *Filologija* 69: 99–128.
- ŠIMIĆ, A. 2022. Gorazd: digitalni portal starocrkvenoslavenskoga jezika. *Slovo* 72: 473–475.
- ŠTRKALJ DESPOT, K.; C. MÖHRS. 2015. Pogled u e-leksikografiju. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje* 41/2: 329–353.
- TOGNINI-BONELLI, E. 2001. *Corpus Linguistics at Work*. Amsterdam – Philadelphia: John Benjamins.



- TOTOMANOVA, A.-M. 2012. Digital presentation of Bulgarian lexical heritage. Towards an electronic historical dictionary. *Studia Ceranea* 2: 221–232.
- TOTOMANOVA, A.-M. 2021. Electronic Research Infrastructure for Bulgarian Medieval Written Heritage: history and perspectives. *Diacronia* 14: 1–9.
- TUČEK, V.; O. ČIPEROVÁ; M. MAJER; Š. PILÁT. 2018.a. Gorazd Generator [software]. Slovanský ústav AV ČR, v. v. i. URL: <http://gorazd.org/?q=cs/node/35> (19. 12. 2022.).
- TUČEK, V.; O. ČIPEROVÁ; M. MAJER; Š. PILÁT. 2018.B. Gorazd Editor [software]. Slovanský ústav AV ČR, v. v. i. URL: <http://gorazd.org/?q=cs/node/38> (19. 12. 2022.).
- VUKOJA, V. 2012. O korpusu Rječnika crkvenoslavenskoga jezika hrvatske redakcije i njegovu odnosu prema korpusima hrvatskoga jezika. *Filologija* 59: 207–229.
- VUKOJA, V. 2014. The Corpus of the Croatian Church Slavonic Texts and the Current State of Affairs Concerning the Dictionary of the Croatian Redaction of Church Slavonic Compiling. A. Abel; C. Vettori; N. Ralli (eds.). *Proceedings of the XVI EURALEX International Congress: The User in Focus. 15–19 July 2014, Bolzano/Bozen*. Bolzano/ Bozen: EURAC Research – EURALEX, 1221–1235.
- VUKOJA, V.; L. TURKALJ. 2016. Converting a Paper Card-file Corpus of Croatian Church Slavonic Texts into a Machine Readable Corpus. V. Baranovas; T. Timčenko (ur.). *El'Manuscript – 2016. Rašytinis palikimas ir skaitmeninės technologijos: VI tarptautinė mokslinė konferencija, Vilnius, 2016 m. rugpjūčio 22–28 d. Pranešimai*. Vilnius: Iževskas, 170–173.

## Summary

Ana MIHALJEVIĆ, Josip MIHALJEVIĆ

### THE ONLINE VERSION OF THE DICTIONARY OF THE CROATIAN REDACTION OF CHURCH SLAVONIC

This paper presents the development process of an online version of the Dictionary of the Croatian Redaction of Church Slavonic. Compiling this dictionary is a long-term project of the Old Church Slavonic Institute in Zagreb. The project began in 1958, but the first fascicle was published in 1991. So far, 23 volumes have been published with entries from *a* to *izgnilstvo*. The dictionary is being compiled as a MS Word document. This paper presents the steps in the retrodigitalisation of the previously published fascicles and the problems involved in this process. A web portal for this dictionary is currently being developed, a process that is complicated due to the dictionary's complex microstructure. The entries contain five languages (Croatian, Croatian Church Slavonic, English, Greek, and Latin) and four scripts (Glagolitic, Cyrillic, Latin, and Greek). After testing and analysing many pre-existing solutions and models for creating a reliable retrodigitalization model for this dictionary, the authors decided that the online version will be built using the *WordPress* content management system (CMS). By modifying

*WordPress* into a custom dictionary writing system, previous entries will be imported into the database, and new entries will be created within the system itself. Glagolitic and Cyrillic fonts have been implemented in the system and can be used by editors through the user interface. A custom virtual keyboard was created to enable the user to type all Glagolitic characters and thus make searching easier. Dictionary entries are interlinked by hyperlink. The technical aspects of developing an online dictionary are also analysed and problems and plans for future development are outlined.

Keywords: Croatian Church Slavonic, e-dictionary, web dictionary, retodigitalisation, Dictionary of the Croatian Redaction of Church Slavonic

Ana MIHALJEVIĆ  
Josip MIHALJEVIĆ  
Old Church Slavonic Institute  
Zagreb (Croatia)  
[amihaljevic@stin.hr](mailto:amihaljevic@stin.hr)  
[jmihaljevic@stin.hr](mailto:jmihaljevic@stin.hr)