UDK 81’36:004
81’374:004
Pregledni rad
Rukopis primljen 1. III. 2025.
Prihvaćen za tisak 10. IV. 2025.
https://doi.org/10.31724/rihjj.52.1.1

Petra Bago
Filozofski fakulet Sveučilišta u Zagrebu
Ivana Lučića 3, HR-10000 Zagreb
https://orcid.org/0000-0002-4994-6417
pbago@ffzg.unizg.hr

RJEČNICI I GRAMATIKE U DOBA DIGITALNE TRANSFORMACIJE

Digitalna transformacija utjecala je na gotovo sve segmente života, uključujući i leksikografiju, čime se rječnici i gramatike preoblikuju kako bi zadovoljili potrebe suvremenih korisnika. Ova promjena očituje se kroz razvoj digitalnih alata i resursa koji omogućuju interaktivnije i pristupačnije metode proučavanja, učenja i upotrebe jezika.
Tradicionalni pristupi leksikografiji, koji su se oslanjali na tiskane medije, često su zbog prirode materijala rezultirali ograničenim pristupom informacijama i sporijim ažuriranjem podataka. S pojavom digitalnih tehnologija, rječnici i gramatike postali su dostupniji široj javnosti, a interaktivne platforme omogućuju korisnicima bolje razumijevanje i korištenje jezika. Razvoj *online* rječnika, gramatika i ostalih digitalnih resursa predstavlja ključnu promjenu u području e-leksikografije, primjerice omogućujući brzo pretraživanje, audioizgovor, vizualne prikaze i primjere korištenja u različitim kontekstima.
Tehnologije poput umjetne inteligencije (veliki jezični modeli) i obrade prirodnog jezika sve više se integriraju u leksikografiju, omogućujući analizu velikih skupova podataka i automatsko generiranje elemenata rječničkih natuknica, opis definicija i primjera. Ovaj napredak može značajno ubrzati proces izrade rječnika. Suočavanje s izazovima u digitalnome dobu ostaje ključno. Osiguravanje točnosti i kvalitete podataka u digitalnim rječnicima i gramatikama često je otežano, a potrebe različitih korisnika zahtijevaju specifične pristupe i resurse koji nisu uvijek dostupni. Prepoznaje se važnost interdisciplinarnih pristupa koji uključuju stručnjake iz raznih znanstvenih područja, kao što su humanističke, društvene i tehničke znanosti. Globalni trendovi ukazuju na potrebu za uključivanjem zajednica i govornika jezika u razvoj i korištenje digitalnih jezičnih resursa, čime se otvaraju mogućnosti za inovacije i nove načine razmišljanja o rječnicima i gramatikama.
Financijska potpora i istraživanja jezičnih resursa i alata ostaju ključni za održivost leksikografskih projekata, a samim time i za održavanje hrvatskog jezika relevantnim u

digitalnom dobu. Nedostatak kontinuirane financijske podrške države za istraživanje jezika, posebno u kontekstu digitalne transformacije, predstavlja ozbiljan izazov u Hrvatskoj. Također, primjetan je nedostatak središnjih institucija koje bi se posvetile prikupljanju, obradi i edukaciji u području jezičnih tehnologija. Uspostavljanje takvih centara može pomoći u jačanju suradnje kako među istraživačima, leksikografima i edukatorima tako i među stručnjacima i korisnicima jezičnih tehnologija.
Zaključno, daljnja istraživanja i razvoj u području e-leksikografije predstavljaju ključne aspekte za osiguranje održivosti i relevantnosti rječnika i gramatika u digitalnom dobu. Uz kontinuirani razvoj tehnologija i metodologija, postoji ogroman potencijal za obogaćivanje hrvatske jezične i kulturne baštine, što će doprinijeti unapređenju digitalnih resursa i alata za hrvatski jezik.

1. Uvod

Digitalna transformacija, kao globalni proces koji obuhvaća uvođenje digitalnih tehnologija u sve sfere života, značajno je utjecala na načine na koje prikupljamo, organiziramo i dijelimo znanje. Pri tome je važno razlikovati digitalnu transformaciju od digitalizacije: dok digitalizacija podrazumijeva prijenos tradicionalnih materijala iz analognog u digitalni oblik, digitalna transformacija uključuje daleko složenije promjene, poput uvođenja novih metodoloških pristupa, izmjene procesa rada, te korištenja naprednih tehnologija poput umjetne inteligencije i automatizirane obrade prirodnog jezika. U području leksikografije i gramatičkih opisa to znači ne samo dostupnost rječnika i gramatika u digitalnom obliku već i temeljitu promjenu načina njihove izrade, distribucije i korištenja. Tako tradicionalni pristupi, koji su desetljećima počivali na intenzivnom radu jezikoslovaca, danas sve više ustupaju mjesto digitalnim alatima, korpusnim analizama i automatskoj obradi podataka. Ključne prednosti koje pritom donosi e-leksikografija jesu brža ažuriranja sadržaja, mogućnost uključivanja multimedijskih elemenata te znatno šira dostupnost korisnicima. Isti trendovi sve su vidljiviji i kod digitalno dostupnih gramatičkih opisa, koji danas poprimaju oblike interaktivnih gramatičkih platformi, baza povijesnih i suvremenih gramatičkih izvora te aplikacija za automatsku gramatičku analizu teksta. Digitalni alati poput hrvatskoga projekta RETROGRAM[1], posvećenoga digitalizaciji i analizi starijih hrvatskih gramatičkih djela, te *online* gramatičkih

[1] https://retrogram.jezik.hr/

provjernika (npr. Hascheck[2]) pokazuju kako digitalna transformacija utječe ne samo na dostupnost već i na metodološki pristup obradi gramatičkog znanja, čime se tradicionalni pristupi upotpunjuju novim mogućnostima.

U samim počecima, digitalizacija rječničke i gramatičke građe uglavnom se svodila na jednostavan prijenos tiskanoga materijala na optičke medije ili *web*-stranice. Međutim, daljnji razvoj računalne lingvistike, informacijsko-komunikacijskih tehnologija i umjetne inteligencije (velikih jezičnih modela) ubrzao je proces i otvorio put inovacijama koje nadilaze običnu pretvorbu teksta. Danas govorimo o interaktivnim platformama, *online* sučeljima za traženje definicija i gramatičkih pravila, personaliziranim vježbama i aplikacijama koje nude zvučne zapise, vizualne primjere ili analizu jezičnih uzoraka. Na taj način, korisnicima su dostupni alati koji nisu samo sadržajno bogati nego i metodički raznoliki, prilagođeni različitim potrebama – od jezikoslovnih istraživanja do školskih programa i samostalnog učenja.

U središtu ove transformacije nalaze se i promjene u ulozi samih stručnjaka, odnosno leksikografa i gramatičara. Tehnologije umjetne inteligencije mogu automatizirati dio procesa, primjerice pri pretraživanju korpusa ili generiranju početnih definicija i gramatičkih objašnjenja. No, uloga stručne procjene i dalje ostaje ključna za konačnu verifikaciju, posebno kad je riječ o jezicima s manje dostupnim digitalnim resursima, poput hrvatskoga. U tom slučaju, ljudski je faktor nezamjenjiv u uočavanju finesa značenja, varijanata izgovora ili specifičnih gramatičkih fenomena. Stoga, prijelaz na digitalno ne znači nestanak tradicionalnih pristupa, nego njihovu nadogradnju i osnaživanje kroz interdisciplinarnu suradnju: jezikoslovci, informacijski i komunikacijski stručnjaci, dizajneri sučelja i pedagozi zajednički oblikuju nove alate i metodologije, zadržavajući standarde kvalitete.

Cilj ovoga rada jest istražiti i predočiti glavne trendove u razvoju e-leksikografije, s naglaskom na izazove i mogućnosti koje se otvaraju za hrvatski jezik. Analizirat će se i uloga umjetne inteligencije u izradi i ažuriranju rječničkih i gramatičkih sadržaja te ukazati na važnost održivog financiranja i uspostave središnjih institucija kao preduvjeta za daljnju afirmaciju hrvatskog jezika u digitalnome okruženju. Na temelju pregledane literature i dosadašnjih projekata, razmotrit

[2] https://ispravi.me/

će se i nužnost interdisciplinarnoga pristupa kao jednoga od ključnih faktora koji osiguravaju kvalitetu i relevantnost suvremenih e-rječnika i digitalnih gramatika.

2. Leksikografija na raskrižju: između tradicije i digitalne budućnosti

Razvoj leksikografije kroz povijest obilježila je dugotrajna i pomna izrada tiskanih rječnika, koji su stoljećima bili glavni izvor jezičnih informacija i pouzdane norme za pisanu riječ. Međutim, pojavljivanjem računala i interneta leksikografija se našla na prekretnici, postupno prelazeći s tradicionalno tiskanih formata na digitalne platforme. U ranoj fazi digitalne ere izdavači su eksperimentirali s optičkim medijima (poput CD-ROM-ova) i zatim s *web*--izdanjima, čime su otvorili nove mogućnosti za pohranu i ažuriranje podataka. Ipak, tiskani rječnici, unatoč svojstvu statičnosti, mnogima su i dalje simbol pouzdanosti i autoriteta, što upućuje na važnost uravnoteženog pristupa pri integraciji digitalnih tehnologija. Kako navode Ferrett i Dollinger (2021), gubitak leksikografskih detalja može biti posljedica prelaska na digitalne formate, pa je ključno osigurati kvalitetu i točnost podataka.

Danas je e-leksikografija u središtu pozornosti zbog eksponencijalnoga razvoja mobilnih aplikacija, interaktivnih platformi i *web*-resursa. U tom kontekstu *web*-tehnologije omogućuju neprestano ažuriranje sadržaja, brzu pretragu i integraciju multimedijskih elemenata, kao što su audio izgovor, videozapisi ili poveznice na druge *web*-izvore. Prema rezultatima opsežnoga istraživanja europskih leksikografskih praksi (Kallas i dr. 2019), korisnici sve više preferiraju digitalne formate zbog jednostavne pretrage, redovitoga ažuriranja i multimedijskih funkcionalnosti. Međutim, istraživanje je istaknulo i potrebu za očuvanjem leksikografske detaljnosti i kvalitete sadržaja, što upućuje na važnost kombinacije digitalnih inovacija s tradicionalnom stručnošću leksikografa. Upravo te interaktivne funkcionalnosti pomažu korisnicima da dublje istražuju značenja riječi, analiziraju frazeme ili etimološke podatke. Međutim, prijelaz na digitalno nije jednoznačno rješenje za sve; kako pokazuju istraživanja Kosema i suradnika (2018) o računalno potpomognutoj izradi rječnika, veličina

i raznolikost digitalnih korpusnih materijala povećavaju zahtjevnost procesa obrade i organizacije leksičkog znanja.

Upravo se stoga nameće pitanje je li digitalno uvijek bolje (Ferrett i Dollinger 2021). Neosporno je da digitalni rječnici donose mnoge prednosti: jednostavno i brzo ažuriranje, veću dostupnost korisnicima, mogućnost personalizacije te interaktivno sučelje koje nudi igre ili kvizove radi olakšavanja učenja. Međutim, digitalni formati ponekad mogu patiti od površnosti ili nedostatne kontrole kvalitete unosa. Neki *online* rječnici i leksikografske platforme ovise o algoritmima umjetne inteligencije te o gomilizaciji (engl. *crowdsourcing*), što može dovesti do manjka provjere izvornosti podataka. U radu o igrifikaciji u e-leksikografiji Mihaljević (2019) upozorava i na opasnost mehaničkog „skupljanja bodova" bez istinskoga produbljivanja jezičnoga znanja, što znači da se i u digitalnom okruženju mora paziti na stvarnu vrijednost interakcije i učenja.

Paralelno s razvojem e-leksikografije odvijaju se slične promjene i na području gramatičkih opisa. Tradicionalne tiskane gramatike postupno zamjenjuju digitalni resursi koji omogućuju ne samo jednostavan pristup već i interaktivno pretraživanje, povezivanje s korpusima te automatsku provjeru gramatičke ispravnosti teksta. Primjerice, hrvatski projekt RETROGRAM jasno ilustrira potencijal digitalizacije gramatičke baštine, omogućujući korisnicima da pretražuju, analiziraju i uspoređuju povijesne gramatičke opise. Poput rječnika, i gramatike se nalaze pred izazovom očuvanja detaljnosti i kvalitete tradicionalnih pristupa u novom, digitalnom okruženju.

Drugi su izazov raznolike potrebe korisnika, jer jedni trebaju brzu provjeru značenja, a drugi detaljnu analizu jezičnih struktura. Nielsen (2022) ističe kako se suvremeni leksikografski alati trebaju osmisliti kao referentni sustavi prilagođeni potrebama korisnika i različitim načinima uporabe: od jezikoslovaca i nastavnika do učenika i šire javnosti. E-leksikografija stoga nije samo digitalizirana inačica klasičnih rječnika, nego inovativan pristup koji uključuje interdisciplinarnu suradnju stručnjaka iz različitih područja: jezikoslovaca, stručnjaka iz područja informacijskih i komunikacijskih znanosti, programera, dizajnera sučelja pa i samih korisnika.

Konačno, prelazak na digitalno otvara i pitanje buduće uloge leksikografa. Rundell (2024) upućuje na to da umjetna inteligencija i automatizacija u izradi

rječnika mijenjaju dosadašnji radni proces, ali i naglašava kako ljudska stručnost ostaje neizostavan faktor pri vrednovanju jezične relevantnosti, semantičkoj analizi i etičkim pitanjima poput autorskog prava ili zaštite privatnosti. Upravo ova kombinacija ljudskoga znanja i tehnoloških mogućnosti sugerira da će se leksikografija dalje razvijati na raskrižju tradicije i digitalne budućnosti: čuvajući leksikografske standarde (posebno pouzdanost), a istovremeno koristeći inovativne metode analize i prezentacije jezičnih podataka. U tome je ključno kontinuirano vrednovanje korisničkih potreba kako bi se zadržale najbolje strane klasične leksikografije, a istodobno iskoristili resursi i potencijali digitalnog doba.

3. Umjetna inteligencija i e-leksikografija

Umjetna inteligencija i strojno učenje u posljednjih su nekoliko godina postali ključni pokretači digitalne transformacije, a njihov je utjecaj sve vidljiviji i u području leksikografije (Lew 2023, 2024; Rundell 2024; De Schryver 2023; Tarp 2023). Dok je tradicionalna izrada rječnika podrazumijevala dugotrajnu ljudsku angažiranost – od prikupljanja korpusa i analize upotrebe riječi, do pisanja definicija i primjera – moderni sustavi umjetne inteligencije sada mogu ubrzati i automatizirati velik dio tih procesa. Ipak, takav tehnološki napredak neizbježno otvara pitanja o točnosti, autentičnosti i etici dobivenih rezultata.

Jedan od glavnih potencijala sustava umjetne inteligencije očituje se u njihovoj sposobnosti obrade velikih količina podataka u relativno kratkom vremenu (Naveed i dr. 2024; Zhao i dr. 2025). Na taj se način mogu automatski analizirati jezični korpusi, prepoznavati obrasci upotrebe, učestalost pojedinih riječi te kontekst u kojem se određeni izrazi pojavljuju. Takav pristup omogućuje brže generiranje nacrta definicija ili popisa sinonima i antonima, što značajno skraćuje vrijeme potrebno za prikupljanje osnovnih materijala. Prema Lewu (2024), time se leksikografe rasterećuje repetitivnih zadataka, a ostavlja im se više prostora za analizu, reviziju i konačnu doradu rječničkih natuknica. Time uloga leksikografa ne zastarjeva, već se transformira.

Osim u izradi rječnika, umjetna inteligencija ključna je i za razvoj suvremenih gramatičkih alata. Platforme poput *Grammarlyja*[3] koriste umjetnu inteligenciju za automatsku provjeru gramatike, pravopisa i stila, nudeći brze ispravke i prijedloge za poboljšanje teksta. Istraživanja pokazuju da ovakvi pomoćnici u pisanju, pogonjeni umjetnom inteligencijom, mogu služiti kao učinkovita alternativa za unapređenje pisanih vještina (Fitria 2021; Barrot 2021). Ipak, unatoč njihovoj sofisticiranosti, ljudska ekspertiza ostaje neizostavna za finu prilagodbu i konačnu provjeru, posebno kod jezika s kompleksnijom morfologijom ili ograničenim digitalnim resursima poput hrvatskoga.

Međutim, stručnjaci upozoravaju da kvaliteta uvelike ovisi o podacima na kojima se obučavaju sustavi umjetne inteligencije (Lew 2024). Ako korpusi nisu reprezentativni, raznovrsni ili dostatno uravnoteženi, konačni rezultati mogu odražavati te nedostatke. Posebno to vrijedi za jezike s ograničenim resursima, uključujući hrvatski, gdje postoji opasnost da se u automatski generiranim rječničkim sadržajima ne zabilježe sve stilske nijanse ili regionalne varijante. Stoga je i dalje neizostavna uloga ljudskih stručnjaka koji mogu uočiti greške i ispraviti pristranosti proizašle iz neuravnoteženih baza podataka.

Uz to, ostaje izazov autentičnosti (Lew 2024; de Schryver 2023). Iako automatizirani sustavi mogu brzo izraditi nacrte definicija, često izostavljaju različite registre i dijalekte, ključne za dublje razumijevanje jezika. Etika je još jedna važna dimenzija: pitanje vlasništva i distribucije korpusnih podataka, kao i transparentnosti algoritama, otvara rasprave o mogućem kršenju autorskih prava ili narušavanju privatnosti korisnika.

Osim izrade rječnika, umjetna je inteligencija sve zastupljenija u razvoju platformi za učenje jezika (Bonner, Lege, i Frazier 2023; Cong 2024; Kasneci i dr. 2023), pri čemu se personalizirane aktivnosti prilagođavaju zahtjevima korisnika. Ipak, kao što primjećuju Ferrett i Dollinger (2021), ključno je izbjeći površno klikanje ili gomilizaciju koja ne pridonosi stvarnom jezičnom napretku.

Na kraju, vrijedno je prisjetiti se misli Garija Kasparova: „Ne treba se bojati inteligentnih strojeva, već surađivati s njima" (*Don't fear intelligent machines. Work with them* 2017). Umjetna inteligencija može ubrzati leksikografske procese, ali ne može zamijeniti stručnu ljudsku prosudbu pri procjeni jezične

[3] https://www.grammarly.com/

relevantnosti i semantičkih finesa, kao ni pri rješavanju etičkih i pravnih pitanja. Takav je pristup zapravo logičan nastavak dugogodišnje prakse suradnje u leksikografiji, gdje inovativni digitalni alati nadopunjuju, a ne zamjenjuju ljudsku kreativnost i kritičko razmišljanje. U konačnici, upravo će ta sinergija između čovjeka i stroja usmjeravati daljnji razvoj e-leksikografije i doprinijeti kvaliteti jezičnih resursa u budućnosti.

4. Interdisciplinarnost kao budućnost leksikografije

Razvoj e-leksikografije tijekom posljednjih dvadesetak godina jasno pokazuje koliko je interdisciplinarnost važna za razumijevanje i stvaranje suvremenih rječnika. Dok su rani leksikografski projekti počivali mahom na radu jezikoslovaca, današnji pristup neizbježno uključuje niz stručnjaka iz područja informacijskih i komunikacijskih znanosti, (računalne) lingvistike, dizajna sučelja, psihologije učenja te, sve češće, samih korisnika (Klosa-Kückelhaus i Michaelis 2022). Naime, izrada digitalnih rječnika više nije tek prebacivanje sadržaja iz tiskanoga formata na *web*-platforme, nego sveobuhvatni proces koji zahtijeva uvažavanje principa strukture podataka, algoritamskog pretraživanja, interaktivnih elemenata i vizualnog dizajna.

Jedna od ključnih prednosti interdisciplinarnosti jest mogućnost rješavanja kompleksnih izazova koji se javljaju pri izradi rječnika, od prikupljanja i obrade velikih korpusa do oblikovanja sučelja prilagođenih različitim skupinama i potrebama korisnika (Klosa-Kückelhaus i Michaelis 2022). Primjerice, stručnjaci iz područja računalne lingvistike mogu se pobrinuti za optimizaciju algoritama obrade jezika, dok dizajneri sučelja osiguravaju da te algoritamski dobivene informacije postanu jednostavno i učinkovito dostupne korisnicima. S druge strane, jezikoslovci i leksikografi zadržavaju ključno mjesto u jamčenju znanstvene točnosti i jezične relevantnosti. Na taj se način znanje iz različitih disciplina integrira u jedinstveni sustav, omogućujući učinkovitije rješavanje složenih problema.

Razlog zašto je takav pristup postao nužan može se objasniti tempom tehnoloških i društvenih promjena. Današnji korisnici od rječnika očekuju mnogo više od samih definicija: traže personalizirane pristupe, interaktivne načine učenja i

kontinuirano ažuriranje jezičnih podataka (Müller-Spitzer i Koplenig 2014). Iste tendencije vrijede i za gramatičke opise, gdje se digitalna transformacija očituje kroz interaktivne platforme za učenje gramatike, automatizirane provjere teksta i analize jezičnih korpusa za gramatičke svrhe. Uvođenje umjetne inteligencije i pametnih algoritama dodatno ubrzava ovaj proces, ali istovremeno zahtijeva etički i sigurnosni nadzor, posebno u pitanjima autorskih prava i zaštite privatnosti (Wu, Duan, i Ni 2024). Upravo u tim segmentima stručnjaci iz pravnih te informacijskih i komunikacijskih znanosti postaju nezaobilazni suradnici.

Važno je naglasiti i ulogu samih korisnika, koji se sve češće vide kao suistraživači ili suradnici pri metodi gomilizacije (Sajous i Josselin-Leray 2022). Korisnici mogu predlagati nove natuknice, slati povratne informacije o postojećim definicijama ili upozoravati na specifične registre i dijalektalne varijante. Međutim, takva otvorenost nije održiva bez razvijenih mehanizama provjere kvalitete, što ponovno traži usku suradnju računalnih stručnjaka i jezikoslovaca.

Nadalje, interdisciplinarnost se širi i u obrazovno okruženje. Suradnja s nastavnicima, pedagozima i psiholozima posebno je važna za razvoj platformi koje poboljšavaju proces učenja jezika, primjerice nudeći interaktivne vježbe, prilagodljive testove te personalizirane povratne informacije (Shadiev i Yang 2020). Time e-leksikografija ne ostaje zatvorena u akademskim okvirima, već se povezuje s konkretnim potrebama šire društvene zajednice.

Konačno, interdisciplinarnost nije samo instrument za tehnička i tehnološka poboljšanja nego i temelj za daljnji znanstveni i društveni napredak. Kada stručnjaci iz različitih područja zajedno rade na izradi rječnika i gramatika, dobivamo sveobuhvatan pogled na jezik kao živi sustav te kreiramo alate koji mogu pridonijeti boljem razumijevanju i očuvanju jezične raznolikosti (Nielsen 2017). Drugim riječima, zajednički rad nije samo nužno rješenje nego i strateški korak koji e-leksikografiji jamči razvoj i relevantnost u godinama koje dolaze.

Upravo stoga budućnost leksikografije leži u još čvršćoj i otvorenijoj suradnji: jezikoslovaca, računalnih stručnjaka, informacijskih i komunikacijskih stručnjaka, dizajnera, pedagoga, korisnika i drugih ključnih dionika. Dokaz tomu su i aktualni projekti koji pokazuju kako razmjena znanja među disciplinama itekako pomaže da digitalni rječnici i gramatike budu sadržajno bogati, metodički

pouzdani i tehnološki inovativni (Blagus Bartolec, Duplančić Rogošić, i Ordulj 2024; Horvat, Kramarić, i Mihaljević 2024; Hudeček, Mihaljević, i Jozić 2024). Na tim će temeljima e-leksikografija i ubuduće nastaviti graditi nove modele u kojima spoj ljudske stručnosti, analitičkih alata i raznolikih znanstvenih perspektiva čini srž inovacija i kvalitete.

5. Kontekst hrvatskog jezika: izazovi i mogućnosti

Specifičnosti hrvatskoga jezika i ograničeni broj njegovih govornika predstavljaju poseban izazov pri izradi i održavanju suvremenih leksikografskih resursa. Iako je globalni trend razvoja e-leksikografije jasno usmjeren prema sve snažnijoj primjeni informacijsko-komunikacijske tehnologije, ostaje pitanje kako takve inovacije održivo integrirati u sredinu u kojoj je jezična zajednica relativno mala i resursi ograničeni (Lignos i dr. 2022). Pritom se ključno ističe važnost kontinuiteta i financijske potpore kako bi se osiguralo da hrvatski rječnici i gramatike, kao i drugi digitalni jezični resursi, budu aktualni i dostupni svim korisnicima.

Upravo je financijska potpora često najkritičnija točka. Nedostatak održivih modela financiranja ili dugoročnih strategija ulaganja u jezične resurse (Europski parlament 2019) rezultira sporadičnim i povremenim projektima koji se teško mogu sustavno nadograđivati (Ide i Pustejovsky 2010). Posljedica toga je i izostanak središnjih institucija koje bi koordinirale prikupljanje i obradu jezičnih podataka, a time i uspostavile jasne metodološke standarde i smjernice za hrvatsku e-leksikografiju. U većim, globalnim jezicima taj je problem rjeđe izražen jer se financiraju projekti velikoga opsega i postoji čitava mreža izdavača, sveučilišta te privatnih poduzeća spremnih ulagati u inovativne lingvističke alate (Pasikowska-Schnass 2020).

Drugi suvremeni izazov odnosi se na nerazvijenu infrastrukturu za masovno prikupljanje i obradu podataka. U Hrvatskoj nedostaje koordinirana platforma za kontinuirano ažuriranje i dijeljenje jezičnih resursa i alata, što otežava pravovremeno praćenje promjena u jeziku, kao i brzu prilagodbu novim tehnologijama. To se podjednako odnosi na razvoj opsežnih leksičkih baza za rječnike i na stvaranje detaljnih gramatičkih korpusa neophodnih za napredne

analize i automatsku obradu gramatike. Primjerice, dok su veći jezici obično uključeni u komercijalnim platformama za strojno učenje i umjetnu inteligenciju od samih početaka njihova razvoja, hrvatski se u pravilu uvodi naknadno ili s ograničenim resursima. Nerijetko se pritom koristi engleski kao metajezik, što odražava nedostatak sveobuhvatnih i kvalitetnih jezičnih baza za hrvatski jezik. Takav pristup zahtijeva dodatno prilagođavanje, kao i povećanu financijsku i tehničku podršku, a u konačnici utječe na dostupnost i kvalitetu rješenja za korisnike hrvatskog jezika. Uz to, bez primjerene infrastrukture gotovo je nemoguće provoditi redovite analize upotrebe novih riječi, frazema i žargonizama, kao ni pratiti evoluciju gramatičkih konstrukcija ili razvoj kolokvijalnih jezičnih tendencija, čime leksikografija gubi na ažurnosti i sveobuhvatnosti.

Hrvatski jezik dio je kulturne baštine koju treba čuvati i promicati, osobito u suvremenom digitalnom okruženju. Jezične inovacije ne smiju se svesti samo na komercijalne proizvode ili kratkoročne projekte, već bi trebale uključivati širu viziju razvoja znanosti, obrazovanja, kulture i društva. Pri tome je ključno priznati da budućnost jezičnoga znanja značajno oblikuju nove tehnologije, mijenjajući način na koji učimo, razumijemo i zajednički gradimo spoznaje o jeziku (Chun, Kern, i Smith 2016). Alati za automatizirano generiranje definicija, interaktivne platforme za izgovor ili prilagodljivi didaktički materijali samo su dio rastućega digitalnog prostora, čije će širenje ovisiti o upravo spremnosti da kontinuirano ulažemo u jezičnu infrastrukturu i razvijamo održive institucionalne okvire.

Na taj se način dugoročno može osigurati da hrvatski jezik, usprkos svojoj veličini, pa onda i ekonomskim ograničenjima, zadrži vidljivost i relevantnost u digitalnome dobu. U konačnici, to nije samo pitanje tehnologije već i strateške odluke da se jezični identitet Hrvatske očuva i obogati u svim sferama društvenoga života.

6. Zaključak

Promjene koje je donijela digitalna transformacija u području leksikografije dalekosežne su i predstavljaju tek početak jednoga šireg procesa. E-rječnici i e-gramatike više se ne doživljavaju samo kao praktični alati već i kao dinamični sustavi znanja koji se stalno razvijaju u skladu s tehnološkim napretkom i

potrebama korisnika. Primjeri interaktivnih *web*-platformi, mogućnost brzog ažuriranja građe te uvođenje multimedijskih elemenata jasno pokazuju da je digitalni format daleko nadmašio početne koncepte jednostavnog prijenosa tiskane građe u *web*-prostor. Istovremeno, praksa je pokazala da prelazak na digitalno nije uvijek bez izazova: nedostatak kontrole kvalitete, povremeno izostajanje leksikografskih ili gramatičkih detalja te ovisnost o sve složenijoj tehnologiji zahtijevaju kontinuirani rad stručnjaka i uspostavu odgovarajuće infrastrukture.

Središnji naglasak u suvremenoj e-leksikografiji svakako leži u interdisciplinarnoj suradnji. Jezikoslovci, informacijsko-komunikacijski stručnjaci, dizajneri korisničkog sučelja, pedagozi i pravnici trebaju djelovati zajednički kako bi se ostvarila ravnoteža između tehnoloških mogućnosti i jezičnog realiteta. Umjetna inteligencija (veliki jezični modeli) osobito postaje nezaobilazan saveznik: od preliminarne analize korpusa i prepoznavanja trendova u uporabi riječi do predlaganja pravila i primjera u interaktivnim leksikografskim djelima. No, upravo uloga ljudskoga stručnjaka ostaje ključna za konačno vrednovanje i prilagodbu takvih sustava kako bi se očuvale stilske, dijalektalne i druge specifične značajke jezika.

Hrvatski jezik, kao jezik s relativno malom zajednicom govornika, posebno je osjetljiv na opasnost zaostajanja za naprednijim tehnologijama koje se razvijaju uglavnom za globalno rasprostranjene jezike. Ovdje se jasno uviđa važnost dugoročne i stabilne financijske podrške, ali i strateške posvećenosti javnih i privatnih institucija, kako bi se izgradila održiva infrastruktura za prikupljanje, obradu i dijeljenje jezičnih podataka. Uspostava središnjih ustanova ili koordiniranih mreža istraživača mogla bi pospješiti razmjenu znanja i resursa, olakšati izradu suvremenih e-rječnika i e-gramatika te potaknuti inovacije u svim segmentima jezičnoga planiranja i obrazovanja.

Zaključno, daljnji razvoj e-leksikografije ne predstavlja samo tehnički i tehnološki izazov već i priliku za širu društvenu i kulturnu nadogradnju. Kvalitetno utemeljeni digitalni resursi izravno utječu na obrazovanje, znanstvena istraživanja, ali i na način na koji percipiramo i vrednujemo vlastiti jezik u doba globalizacije i intenzivne razmjene informacija. Ako želimo da hrvatski jezik ostane relevantan i ravnopravno zastupljen u svjetskome digitalnom okruženju, nužno je nastaviti s ulaganjima u jezične resurse i alate, poticati interdisciplinarni

dijalog i osnaživati institucionalne okvire. Samo tako možemo odgovoriti na sve izazove digitalnoga doba, a istovremeno zadržati jezičnu raznolikost i bogatstvo koje nas, kao govornike hrvatskoga jezika, čini jedinstvenima.

Napomena: Ovaj je rad nastao u sklopu istraživanja koje je sufinancirala Europska unija kroz program Next Generation EU projektom *Socijalna konstrukcija apstraktnih značenja čitanjem* – SKAZ (581-990-1036).

Literatura

Anatomija Rječnika: Hrvatski Mrežni Rječnik-Mrežnik. 2024. Ur. Hudeček, Lana; Mihaljević, Milica; Jozić, Željko. Institut za hrvatski jezik i jezikoslovlje. Zagreb.

Barrot, Jessie Saraza. 2022. Integrating Technology into ESL/EFL Writing through Grammarly. *RELC Journal* 53(3). 764–768. https://doi.org/10.1177/0033688220966632

Blagus Bartolec, Goranka; Duplančić Rogošić, Gorana; Ordulj, Antonia. 2024. INIKOL - Collocational Database for Learning Croatian as a Foreign Language. *Proceedings of the 2024 CLASP Conference on Multimodality and Interaction in Language Learning*. Ur. Qiu, Amy i dr. Association for Computational Linguistics. Gothenburg. 8–12. https://aclanthology.org/2024.clasp-1.2/ (pristupljeno 1. ožujka 2025.).

Bonner, Euan; Frazier, Erin; Lege, Ryan. 2023. Large Language Model-Based Artificial Intelligence in the Language Classroom: Practical Ideas for Teaching. *Teaching English with Technology* 23(1). 23–41. http://dx.doi.org/10.56297/BKAM1691/WIEO1749

Chun, Dorothy; Kern, Richard; Smith, Bryan. 2016. Technology in Language Use, Language Teaching, and Language Learning. *The Modern Language Journal* 100(S1). 64–80. http://dx.doi.org/10.1111/modl.12302

Cong, Yan. 2024. AI Language Models: An Opportunity to Enhance Language Learning. *Informatics* 11(3): 49. http://dx.doi.org/10.3390/informatics11030049

Don't Fear Intelligent Machines. Work with Them. 2017. https://www.ted.com/talks/garry_kasparov_don_t_fear_intelligent_machines_work_with_them (pristupljeno 17. veljače 2025).

Europski parlament. 2019. Jezična ravnopravnost u digitalnom dobu: Rezolucija Europskog parlamenta od 11. rujna 2018. o jezičnoj ravnopravnosti u digitalnom dobu (2018/2028(INI)). *Službeni list Europske unije* 62(C 433). 42–49.

Ferrett, Emma; Dollinger, Stefan. 2021. Is Digital Always Better? Comparing Two English Print Dictionaries with Their Digital Counterparts. *International Journal of Lexicography* 34(1). 66–91. http://dx.doi.org/10.1093/ijl/ecaa016

Fitria, Tira Nur. 2021. Grammarly as AI-powered English Writing Assistant: Students' Alternative for Writing English. *Metathesis: Journal of English Language, Literature, and Teaching* 5(1). 65–78. https://doi.org/10.31002/metathesis.v5i1.3519

Horvat, Marijana; Kramarić, Martina; Mihaljević, Ana. 2024. Challenges in the Process of Retro-Digitisation of Croatian Grammar Books before Illyrism. *Structuring Lexical Data and Digitising Dictionaries*. Ur. Martín Arista, Javier; Ojanguren López, Ana Elvira. Leiden. Brill. 63–94. http://dx.doi.org/10.1163/9789004702660_005

Ide, Nancy; Pustejovsky, James. 2010. What Does Interoperability Mean, Anyway? Toward an Operational Definition of Interoperability for Language Technology. *Proceedings of the Second International Conference on Global Interoperability for Language Resources*. Ur. Fang, Alex; Ide, Nancy; Webster, Jonathan. City University of Hong Kong. Hong Kong.

Kallas, Jelena i dr. 2019. *Lexicographic Practices in Europe: A Survey of User Needs.* https://elex.is/wp-content/uploads/2020/06/Revised-ELEXIS_D1.1_Lexicographic_Practices_in_Europe_A_Survey_of_User_Needs.pdf (pristupljeno 15. rujna 2025.).

Kasneci, Enkelejda i dr. 2023. ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. *Learning and Individual Differences* 103. Elsevier. Amsterdam. 102274. https://www.sciencedirect.com/science/article/abs/pii/S1041608023000195?via%3Dihub (pristupljeno 15. rujna 2025.). http://dx.doi.org/10.1016/j.lindif.2023.102274

Klosa-Kückelhaus, Annette; Michaelis, Frank. 2022. The Design of Internet Dictionaries. *The Bloomsbury Handbook of Lexicography.* Ur. Jackson, Howard. Bloomsbury Academic. London. 405–421. http://dx.doi.org/10.5040/9781350181731.ch-024

Kosem, Iztok i dr. 2018. Collocations Dictionary of Modern Slovene. *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts*. Ur. Čibej, Jaka i dr. Znanstvena založba Filozofske fakultete Univerze v Ljubljani. Ljubljana. 989–997. https://www.researchgate.net/profile/Spela-Arhar-Holdt/publication/347848752_Collocations_Dictionary_of_Modern_Slovene/links/5fe4e604299bf140883f39a9/Collocations-Dictionary-of-Modern-Slovene.pdf (pristupljeno 17. veljače 2025).

Lew, Robert. 2023. ChatGPT as a COBUILD Lexicographer. *Humanities and Social Sciences Communications* 10(1). 1–10. http://dx.doi.org/10.1057/s41599-023-02119-6

Lew, Robert. 2024. Dictionaries and Lexicography in the AI Era. *Humanities and Social Sciences Communications* 11(1): 1–8. http://dx.doi.org/10.1057/s41599-024-02889-7

Lignos, Constantine i dr. 2022. Toward More Meaningful Resources for Lower-Resourced Languages. *Findings of the Association for Computational Linguistics: ACL 2022*. Ur. Muresan, Smaranda; Nakov, Preslav; Villavicencio, Aline. Association for Computational Linguistics. Dublin. 523–532. http://dx.doi.org/10.18653/v1/2022.findings-acl.44

Mihaljević, Josip. 2019. Gamification in E-Lexicography. *Proceedings of the 7th International Conference The Future of Information Sciences (INFuture 2019)*. Ur. Bago, Petra i dr. Odsjek za informacijske i komunikacijske znanosti Filozofskog fakulteta Sveučilišta u Zagrebu. Zagreb. 155–164. http://dx.doi.org/10.17234/INFUTURE.2019.19

Müller-Spitzer, Carolin; Koplenig, Alexander. 2014. Online Dictionaries: Expectations and Demands. *Using Online Dictionaries*. Ur. Müller-Spitzer, Carolin. Walter de Gruyter. Berlin – Boston. 143–188. http://dx.doi.org/10.1515/9783110341287.143

Naveed, Humza i dr. 2024. A Comprehensive Overview of Large Language Models. *ACM Transactions on Intelligent Systems and Technology* 16(5). 1–72. http://dx.doi.org/10.1145/3744746

Nielsen, Sandro. 2017. Lexicography and Interdisciplinarity. *The Routledge Handbook of Lexicography*, Ur. Fuertes-Olivera, Pedro A. Routledge. London. 93–104. http://dx.doi.org/10.4324/9781315104942-7

Nielsen, Sandro. 2022. The Future of Dictionaries, Dictionaries of the Future. *The Bloomsbury Handbook of Lexicography*. Ur. Jackson, Howard. Bloomsbury Academic. London. 389–403. http://dx.doi.org/10.5040/9781350181731.ch-023

Pasikowska-Schnass, Magdalena. 2020. *European Day of Languages: Digital Survival of Lesser-Used Languages.* European parliamentary Research Service. Belgium. https://policycommons.net/artifacts/1426568/european-day-of-languages/2041013/ (pristupljeno 1. ožujka 2025).

Rundell, Michael. 2024. Automating the Creation of Dictionaries: Are We Nearly There? *Humanising Language Teaching* 26(1). Instill Education Limited. Oxford. https://www.hltmag.co.uk/feb24/automating-the-creation-of-dictionaries (pristupljeno 17. veljače 2025).

Sajous, Franck; Josselin-Leray, Amélie. 2022. Issues in Collaborative and Crowdsourced Lexicography. *The Bloomsbury Handbook of Lexicography*. Ur. Jackson, Howard. Bloomsbury Academic. London. 343–358. http://dx.doi.org/10.5040/9781350181731.ch-020

de Schryver, Gilles-Maurice. 2023. Generative AI and Lexicography: The Current State of the Art Using ChatGPT. *International Journal of Lexicography* 36(4). 355–387. http://dx.doi.org/10.1093/ijl/ecad021

Shadiev, Rustam; Yang, Mengke. 2020. Review of Studies on Technology-Enhanced Language Learning and Teaching. *Sustainability* 12(2). 524. http://dx.doi.org/10.3390/su12020524

Tarp, Sven. 2023. Eppur Si Muove: Lexicography Is Becoming Intelligent! *Lexikos* 33(2): 107–131. http://dx.doi.org/10.5788/33-2-1841

Wu, Xiaodong; Duan, Ran; Ni, Jianbing. 2024. Unveiling Security, Privacy, and Ethical Concerns of ChatGPT. *Journal of Information and Intelligence* 2(2): 102–115. http://dx.doi.org/10.1016/j.jiixd.2023.10.007

Zhao, Wayne Xin i dr. 2025. A Survey of Large Language Models. *arXiv.* Cornell University. Ithaca. 2303.12223v16. https://arxiv.org/abs/2303.18223 (pristupljeno 15. rujna 2025.). http://dx.doi.org/10.48550/arXiv.2303.18223

Dictionaries and Grammars in the Age of Digital Transformation

Abstract

Digital transformation has impacted nearly every aspect of life, including lexicography, reshaping dictionaries and grammars to meet the needs of modern users. This change is evident in the development of digital tools and resources that enable more interactive and accessible methods for studying, learning, and using languages.

Traditional approaches to lexicography, which relied on print media, often resulted in limited access to information and slower updates due to the nature of the medium. With the advent of digital technologies, dictionaries and grammars have become more widely available to the public, and interactive platforms allow users to better understand and use languages. The development of online dictionaries, grammars, and other digital resources represents a key shift in the field of e-lexicography, offering features like quick searches, audio pronunciations, visual representations, and examples of usage in various contexts.

Technologies such as artificial intelligence (large language models) and natural language processing are increasingly being integrated into lexicography, enabling the analysis of large datasets and the automatic generation of dictionary entry elements, such as definitions and examples. This progress can significantly accelerate the dictionary creation process. Addressing challenges in the digital age remains critical. Ensuring the accuracy and quality of data in digital dictionaries and grammars is often difficult, and the diverse needs of users require specific approaches and resources that are not always readily available. The importance of interdisciplinary approaches is recognized, involving experts from various fields, such as the humanities, social sciences, and technical sciences. Global trends highlight the necessity of involving communities and language speakers in the development and use of digital language resources, opening opportunities for innovation and new ways of thinking about dictionaries and grammars.

Financial support and research into language resources and tools remain crucial for the sustainability of lexicographic projects, and consequently for maintaining the relevance of the Croatian language in the digital age. A lack of continuous financial support from the state for language research, particularly in the context of digital transformation, presents a serious challenge in Croatia. Additionally, there is a noticeable lack of central institutions dedicated to the collection, processing, and education in the field of language

technologies. Establishing such centers could enhance collaboration among researchers, lexicographers, educators, and experts as well as users of language technologies.

In conclusion, further research and development in the field of e-lexicography represent key aspects of ensuring the sustainability and relevance of dictionaries and grammars in the digital age. With the continuous development of technologies and methodologies, there is tremendous potential to enrich Croatian linguistic and cultural heritage, contributing to the advancement of digital resources and tools for the Croatian language.

***Ključne riječi*:** e-leksikografija, digitalna transformacija, hrvatski jezik, interdisciplinarnost, veliki jezični modeli

***Keywords*:** e-lexicography, digital transformation, Croatian language, interdisciplinarity, large language models