

# Uvodnik

Strateško načrtovan in neprekinjen razvoj jezikovnih virov, tehnologij in storitev je ključnega pomena za vsak jezik oz. jezikovno skupnost – je temeljni pogoj, da se lahko posameznice in posamezniki nemoteno vključujemo v nove načine komunikacije, dela in preživljanja prostega časa v sodobni družbi. Za slovenščino sta premišljenost in usklajenost še toliko pomembnejši, saj je razvojna naloga enaka, raziskovalno-razvojna skupnost, ki se ji posveča, pa manjša kot pri jezikih z več govorcji.

Načrtovanje digitalne infrastrukture za sodobno slovenščino v tem trenutku še ni optimalno, pozitivno pa je, da se problematiki na nacionalni ravni posveča vedno več pozornosti. Primer dobre prakse raziskovalno-razvojnega projekta, ki je povezal deležnike v slovenskem prostoru in združil znanja različnih raziskovalnih inštitucij ter jezikovnotehnoloških podjetij, je Razvoj slovenščine v digitalnem okolju (RSDO), ki sta ga med leti 2020 in 2023 financirala Ministrstvo za kulturo Republike Slovenije in Evropski sklad za regionalni razvoj.

Na projektu smo odpravili nekatere pereče vrzeli na področju odprto dostopnih virov, tehnologij in storitev za sodobno slovenščino. Prenovili smo učne množice in postopke za strojno označevanje sodobne slovenščine, osvežili in povečali temeljne jezikovne vire in nadgradili metodologijo za njihovo bodočo gradnjo. Velik del projekta je bil posvečen razvoju govorne baze in govornih tehnologij, zlasti razpoznave govora za slovenščino, ter semantičnim virom in tehnologijam, kjer so bile aktivnosti izdelava osrednje digitalne slovarske baze, baze znanja ter virov in postopkov za različne semantične naloge. Nadgradili smo metodologijo strojnega prevajanja in zasnovali ter vzpostavili portal za urejanje slovenske terminologije.

Monografija, ki je pred vami, vključuje osem poglavij, ki jih je pripravilo dvajset avtorjev in avtoric s petih različnih inštitucij: Univerze

v Ljubljani (Filozofska fakulteta in Fakulteta za računalništvo in informatiko), Univerze v Mariboru (Fakulteta za elektrotehniko, računalništvo in informatiko), ZRC SAZU (Inštitut za slovenski jezik Frana Ramovša), Inštituta »Jožef Stefan« in jezikovnotehnološkega podjetja Amebis, d. o. o., Kamnik. Prva prispevka se posvečata pripravi virov za razvoj govornih tehnologij: **Darinka Verdonik** se osredotoči na prihodnji razvoj govornih korpusov, zlasti z vidika njihove karse-da učinkovite gradnje za različne uporabniške potrebe, **Mitja Trojar** in **Andreja Bizjak** pa predstavita načela za zapis govora in izvedbo transkribiranja pri izdelavi govorne baze Artur. Sledita poglavji, ki se ukvarjata s korpusi, ki vsebujejo jezikovne popravke in so zlasti pomembni za področje jezikovnega izobraževanja: **Špela Arhar Holdt**, **Eva Pori** in **Iztok Kosem** predstavijo strategijo za prihodnost korpusa Šolar, ki vsebuje besedila osnovnošolskih in srednješolskih učencev; **Mojca Stritar Kučuk** pa prvi korpus slovenščine kot drugega oz. tujega jezika KOST 1.0. Zadnji predstavljeni jezikovni vir je slovenski učni korpus SUK, katerega kompleksno sestavo in ročno pregledane jezikoslovne oznake predstavljajo **Špela Arhar Holdt**, **Jaka Čibej**, **Kaja Dobrovoljc**, **Tomaž Erjavec**, **Polona Gantar**, **Simon Krek**, **Tina Munda**, **Nejc Robida**, **Luka Terčon** in **Slavko Žitnik**. **Slavko Žitnik** je tudi avtor prvega od dveh poglavij, ki se posvečata orodjem za obdelavo naravnega jezika – predstavi ogrodje za demokratizacijo obdelave naravnega jezika ANGLEr, vključno s podatkovnim modelom, **Aleš Žagar** in **Marko Robnik-Šikonja** pa meta-povzemalnik, ki izbira med štirimi različnimi modeli povzemanja, razvitimi za slovenščino. Monografijo zaključita **Mateja Jemec Tomazin** in **Miro Romih** z opisom Slovenskega terminološkega portala, ki ponuja uporabniku prijazen načine za urejanje slovenske terminologije.

Monografija je namenjena študentom in študentkam, predavateljem in predavateljicam, raziskovalcem in raziskovalkam, razvijalcem in razvijalkam ter vsem, ki bi radi bolje razumeli namen, sestavo in način gradnje predstavljenih projektnih rezultatov. Pomemben doprinos dela je, da pogled usmerja v prihodnost in opredeljuje korake, ki so pred nami. Zato bo uporabno branje tudi za pripravljavce in pripravljavke nacionalnih razvojnih strategij in druge področne

odločevalce in odločevalke. Urednika se prijazno zahvaljujeva vsem sodelujočim za kakovostne prispevke, recenzentoma **Moniki Kalin Golob** in **Simonu Šustru** pa za hitro branje in konstruktivne komentarje. Lepo vabljeni k branju!

*Špela Arhar Holdt in Simon Krek*