
RSDO-5 SLOVENSKI TERMINOLOŠKI PORTAL NAČRT RAZVOJA IN VZDRŽEVANJA

Pripravili: Mateja Jemec Tomazin, Miro Romih, Mojca Žagar Karer, Tanja Fajfar, Simon Atelšek, Senja Pollak, Milan Ojsteršek, Marko Ferme

Verzija: 3

Datum: 24. 2. 2023

1 Izhodišča

Slovenski terminološki portal (<https://terminoloski.slovenscina.eu/>) je bil predstavljen javnosti decembra 2022. Zasnovan in pripravljen je kot skupno spletno mesto, namenjeno vsem uporabnikom terminologije, na katerem lahko iščejo najustreznejši termin, preverjajo definicije v slovarjih in iščejo ustreznike v tujih jezikih. Registrirani uporabniki pa lahko poleg tega sestavljajo sami ali skupaj s sodelavci tudi svoje virove, luščijo terminološke kandidate iz lastnih korpusov in zastavljajo vprašanja v Terminološki svetovalnici, ki je povezana s Terminološko svetovalnico ZRC SAZU.

Izdelavo portala so vodili sodelavci podjetja Amebis in Inštituta za slovenski jezik Frana Ramovša ZRC SAZU, pri razvoju luščilnika in pripravi vzorčnih zbirk pa še sodelavci Fakultete za elektrotehniko, računalništvo in informatiko Univerze v Mariboru, sodelavci Inštituta Jožef Stefan, področni strokovnjaki s Fakultete za upravo in Fakultete za računalništvo in informatiko Univerze v Ljubljani.

Sestavni deli portala so **iskalnik** z odprto dostopnimi terminološkimi slovarji, **luščilnik** terminoloških kandidatov in primerov rabe iz uporabnikovih izbranih besedil in izbranih besedil z Nacionalnega portala odprte znanosti, **urejevalnik** terminoloških virov, **terminološka svetovalnica**, ki je povezana s Terminološko svetovalnico na spletišču Terminologišče ter **administratorski del**, ki omogoča upravljanje portala in aktiven razvoj virov. Slovenski terminološki portal je rezultat projekta *Razvoj slovenščine v digitalnem okolju*.

Da bi se lahko portal vsebinsko in programsko razvijal, predlagamo naslednji načrt vsebinskega dopolnjevanja terminološkega portala, razdeljenega po posameznih sestavnih delih portala:

1. TERMINOLOŠKI PORTAL
2. ORODJE ZA LUŠČENJE
3. SMERNICE IN NAVODILA ZA IZDELAVO TERMINOLOŠKIH VIROV
4. UREJEVALNIK ZA IZDELAVO TERMINOLOŠKIH VIROV
5. TERMINOLOŠKA SVETOVALNICA

1. TERMINOLOŠKI PORTAL

Terminologija je ključni del specializirane komunikacije, ki jo uporabljajo področni strokovnjaki, prevajalci, lektorji in drugi, ki se pri svojem vsakodnevem delu srečujejo s terminologijo. Prav tako je terminologija potrebna za razvoj različnih jezikovnotehnoloških orodij in storitev. V ta namen je bil v okviru projekta RSDO izdelan terminološki portal, ki na enem mestu ponuja urejene in standardizirane eno- in večjezične terminološke vire ter odprto dostopno strežniško infrastrukturo za izdelavo novih terminoloških virov in orodja za luščenje terminologije iz specializiranih korpusov. Na portalu je na voljo tudi terminološka svetovalnica za svetovanje ob terminoloških zagatah različnih uporabnikov terminologije.

Za nemoteno delovanje in razvoj terminološkega portala je potrebno naslednje:

- Zagotavljanje ustrezne strojne opreme ter njeno vzdrževanje in nadgrajevanje (npr. 2 grafični kartici Tesla V100)
- Vzdrževanje in nadgrajevanje systemske in aplikativne programske opreme, ki jo portal uporablja
- Zagotavljanje varnostnega kopiranja podatkov
- Skrb za povezovanje portala z drugimi instancami portali
- Zagotavljanje tehnične in vsebinske podpore uporabnikom portala
- Spremljanje dejavnosti na portalu zaradi zagotavljanja nemotenega delovanja
- Skrb za kakovostno vsebino objavljenih terminoloških virov
- Skrb za nemoteno delovanje in morebitno prilagajanje programskih vmesnikov
- Zbiranje pripomb in predlogov uporabnikov za izboljševanje funkcionalnosti portala
- Odpravljanje napak in pomanjkljivosti ter nadgradnja portala z novimi funkcionalnostmi
- Izboljševanje uporabniškega vmesnika za zagotavljanje čim boljše uporabniške izkušnje, predvsem za osebe s posebnimi potrebami
- Dodatno izboljšanje »backend validacije«
- Pretvorba dodatnih obstoječih odprto dostopnih terminoloških virov v standardni format in njihova vključitev na portal

2. ORODJE ZA LUŠČENJE

S pomočjo specializiranih korpusov, ki so na voljo za vse več strokovnih področij, je s pomočjo orodja za luščenje terminov mogoče s samodejnimi metodami hitro izluščiti

terminološke kandidate in jih dati v nadaljnjo obdelavo terminologom in strokovnjakom. S tem se prva faza izdelave terminološkega slovarja poenostavi in skrajša.

Za nemoteno delovanje in nadaljnji razvoj orodja za luščenje je potrebno naslednje:

- Nadgradnja uporabniškega vmesnika za še boljšo uporabniško izkušnjo
- Dodajanje podrobnejših poročil o stanju in rezultatih posameznega luščenja
- Povečanje robustnosti in zanesljivosti luščenja
- Povečanje hitrosti luščenja
- Dodajanje funkcionalnosti za označevanje terminov v besedilih (Terminator)
- Izboljšanje rezultatov luščenja iz besedil korpusa OSS
- Zagotovitev ustreznega okolja za prenos produkcijskega luščilnika iz razvojnih strežnikov Univerze v Mariboru (dovolj zmogljiv strežnik s pripadajočim vzdrževanjem)
- Odpravljanje morebitnih napak luščenja
- Razvoj še boljših postopkov luščenja na osnovi novih postopkov
- Nadgradnja obstoječih in priprava dodatnih korpusov, posodobitev jezikovnih modelov, priprava novih učnih in testnih množic
- Izboljšanje kanonizatorja, potrebnega za izdelavo končnega seznama terminoloških kandidatov
- Izboljšanje in pohitritev luščenja dobrih primerov

3. SMERNICE IN NAVODILA ZA IZDELAVO TERMINOLOŠKIH VIROV

Za izdelavo kakovostnega terminološkega slovarja je treba upoštevati relevantno strokovno literaturo v izbranem jeziku, jo uravnotežiti glede na vrsto in časovno aktualnost ter upoštevati različne avtorje, potem pa izdelati specializirani korpus, na podlagi katerega se izdela nabor terminoloških kandidatov. Nabor terminoloških kandidatov pregledajo področni strokovnjaki in nato izberejo termine, ki jih bodo uvrstili v terminološki vir. Smernice pomagajo uporabnikom terminologije, zlasti področnim strokovnjakom, ki si želijo izdelati svoj terminološki vir, pri vseh fazah njegove izdelave, da bo tako nastal kakovosten terminološki vir.

Za lažje delo in pomoč avtorjem terminoloških virov je zato treba tudi smernice glede na odzive uporabnikov ustrezno nadgrajevati in jih sproti dopolnjevati glede na spremembe uporabniškega vmesnika in drugih funkcionalnosti.

4. UREJEVALNIK ZA IZDELAVO TERMINOLOŠKIH VIROV

Terminološki viri nastajajo v različnih okoljih. Ustvarjajo jih strokovnjaki v javnih in zasebnih ustanovah, raziskovalci, prevajalci, študenti in drugi. Da bi jim olajšali delo, smo izdelali spletni urejevalnik za izdelavo terminoloških virov, s pomočjo katerega avtorji lahko urejajo terminologijo svojega področja. Ta v enotnem formatu uporabnikom ponuja možnosti za definiranje različnih tipov terminoloških virov, npr. eno- ali večjezični slovar. Med drugimi funkcionalnostmi so omogočeni tudi uvoz podatkov, izvoz v različne formate zapisa in večuporabniško delo z različnimi vlogami uporabnikov.

Za nemoteno delovanje in nadaljnji razvoj urejevalnika za izdelavo terminoloških virov je potrebno naslednje:

- Spremljanje in odpravljanje napak v delovanju urejevalnika
- Tehnična in predvsem vsebinska podpora avtorjem terminoloških virov
- Dodajanje novih možnosti, kot npr. dodatne informacije o oblikah, naglasih in izgovoru besed v terminu, novi formati uvoza (csv, tsv) ...

5. TERMINOLOŠKA SVETOVALNICA

Uporabniki terminologije kljub pomoči obstoječih slovarjev pogosto ne najdejo odgovorov na svoja vprašanja o strokovnih poimenovanjih, zato smo terminološko svetovalnico vključili v terminološki portal, ki uporabnikom na enem mestu ponuja terminološko svetovanje.

Terminološka svetovalnica uporabnikom prek spletnega vmesnika omogoča postavljanje s terminologijo povezanih vprašanj, ob tem pa vmesnik uporabnika spodbuja k čim bolj natančni formulaciji vprašanja z navedbo vira, sobesedila, morebitnih tujejezičnih ustreznikov itd. Za učinkovito delovanje storitve smo zagotovili infrastrukturo – spletni vmesnik z arhiviranjem že odgovorjenih vprašanj. Na vprašanja bodo odgovarjali terminologi, ki se bodo po potrebi posvetovali tudi z zunanji (področnimi) strokovnjaki.

Za nemoteno delovanje in nadaljnji razvoj terminološke svetovalnice oz. terminološkega svetovanja je potrebno predvsem redno in pravočasno terminološko svetovanje za to delo usposobljenih svetovalcev (večinoma v okviru rednih delovnih aktivnosti Inštituta za slovenski jezik Frana Ramovša ZRC SAZU) ter nemoteno delovanje svetovalne infrastrukture.

TERMINSKI IN FINANČNI NAČRT ZA POSODABLJANJE IN NADALJNI RAZVOJ TERMINOLOŠKEGA PORTALA IN PRIPADAJOČIH ORODIJ

Strojna oprema

Trenutna infrastruktura, ki teče na obstoječih strežnikih CJVT in FERI, ne zadostuje za resno in dolgoročno delovanje portala. Za delovanje osrednje instance terminološkega portala (»Slovenski terminološki portal«) je potreben dovolj zmogljiv strežnik, še posebej zaradi možnosti luščanja terminoloških kandidatov iz besedil. Vrednost takega strežnika z npr. dvema grafičnima karticama ter zadostno procesno močjo, ustreznim pomnilnikom in diski ocenjujemo na okrog 100.000 EUR.

Delo

Ocenjujemo, da je za vzdrževanje in nadgradnjo terminološkega portala potrebnih vsaj 12 človek-mesecev letno (1 FTE/letno).