

KOMUNIKACIJSKI NAČRT PROJEKTA

“Majhni jezikovni modeli za brezkontekstno pridobivanje informacij iz evropskih jezikov”

Ljubljana, 30.07.2026

Osnovne informacije o projektu

Oznaka in naziv projekta:	Z2-70067 Majhni jezikovni modeli za brezkontekstno pridobivanje informacij iz evropskih jezikov
Akronim projekta:	SLM4IE
Institucija:	Institut »Jožef Stefan«, IJS
Trajanje projekta:	01.03.2026 - 29.02.2028
Vodja projekta:	dr. Erik Novak
Projektni partnerji:	/
Verzija tega dokumenta:	1.0
Datum oddaje komunikacijskega načrta V1:	31.07.2026
Dokument pripravil:	dr. Erik Novak

Kazalo

1. Namen in cilji komuniciranja	2
2. Ciljne skupine	2
3. Ključna sporočila.....	3
4. Komunikacijski kanali in orodja.....	3
5. Časovni načrt komuniciranja.....	4
6. Odgovornosti	4
7. Kazalniki uspešnosti	4
8. Skladnost z zahtevami ARIS.....	5



Vsebina komunikacijskega načrta

1. Namen in cilji komuniciranja

Projekt SLM4IE razvija majhne jezikovne modele za samodejno pridobivanje informacij iz besedil v evropskih jezikih, s poudarkom na slovenščini. Ker so ključni rezultati projekta, tj. modeli, podatkovne množice in programska koda, namenjeni odprti objavi in nadaljnji uporabi v raziskovalni skupnosti ter praksi, je premišljeno komuniciranje aktivnosti in rezultatov nujen del projekta: rezultati lahko dosežejo svoj namen le, če ciljne skupine zanje vedo, jih razumejo in jih znajo uporabiti. Komuniciranje tako neposredno podpira raziskovalne cilje projekta (odprta in pregledna objava rezultatov, prenos znanja v uporabo) ter cilje ARIS glede diseminacije rezultatov javno financiranih raziskav.

Glavni cilji:

- povečanje vidnosti raziskave, raziskovalnega področja in Instituta »Jožef Stefan«,
- prenos raziskovalnih rezultatov v uporabo v raziskovalni skupnosti in praksi,
- krepitev zaupanja v znanost in prikaz družbene koristnosti javno financiranih raziskav.

Specifični cilji:

- obveščanje znanstvene skupnosti o poteku in rezultatih raziskave prek objav, konferenčnih predstavitev in odprto dostopnih virov,
- omogočanje uporabe razvitih modelov in podatkovnih množic organizacijam z omejenimi računskimi viri in občutljivimi podatki,
- izobraževanje strokovne javnosti in študentov o zmožnostih ter omejitvah majhnih jezikovnih modelov.

2. Ciljne skupine

Primarni ciljni skupini projekta sta:

- **znanstvena skupnost:** domači in mednarodni raziskovalci s področij obdelave naravnega jezika, umetne inteligence in jezikovnih tehnologij, vključno z raziskovalci sorodnih projektov (npr. PoVeJMo) in skupnostjo na platformi Hugging Face;
- **strokovna javnost oziroma uporabniki rezultatov:** organizacije in razvijalci, ki obdelujejo občutljive podatke (npr. v zdravstvu, financah in pravu) ali imajo omejene računske vire (mala in srednja podjetja, javne institucije) in bi razvite modele uporabili v lastnih okoljih.

Sekundarna ciljna skupina so **študenti**, predvsem študenti Mednarodne podiplomske šole Jožefa Stefana in drugih univerz, ki jih vodja projekta dosega prek pedagoškega dela in mentorstev.

Splošne javnosti in medijev projekt zaradi svoje ozko strokovne narave ne nagovarja sistematično; morebitne priložnosti za poljudno predstavitev rezultatov bodo obravnavane po presoji vodje projekta.

Namen komuniciranja glede na ciljno skupino:



Ciljna skupina	Namen komuniciranja
Znanstvena skupnost (primarna)	Izmenjava znanja, znanstvene objave, odprta objava modelov in podatkovnih množic, sodelovanje
Strokovna javnost / uporabniki (primarna)	Prenos znanja v prakso, dokumentacija in navodila za uporabo razvitih virov
Študenti (sekundarna)	Izobraževanje, seznanjanje z jezikovnimi tehnologijami, mentorstva

3. Ključna sporočila

Osrednja sporočila projekta so prilagojena posameznim ciljnim skupinam in zapisana okvirno; končna besedila bodo oblikovana ob posameznih objavah.

Za znanstveno skupnost:

Razvijamo majhne jezikovne modele, ki iz besedil v evropskih jezikih, s poudarkom na slovenščini, brez predhodnega učenja na označenih podatkih izluščijo strukturirane informacije, pri čemer dosegajo primerljivo kakovost kot veliki jezikovni modeli ob bistveno manjši porabi računskih virov. Vsi modeli, podatkovne množice in programska koda bodo prosto dostopni, kadar bo to mogoče.

Za strokovno javnost in uporabnike:

Organizacije z občutljivimi podatki (zdravstvo, finance, pravo) ne morejo uporabljati oblačnih storitev oziroma jih lahko uporabljajo le pod omejenimi pogoji. Naši modeli tečejo na običajni računalniški opremi v prostorih organizacije, kar omogoča varno in cenovno dostopno samodejno pridobivanje informacij iz besedil.

Za študente:

Jezikovne tehnologije niso omejene na velike modele svetovnih podjetij; tudi z manjšimi, skrbno načrtovanimi modeli je mogoče doseči vrhunske rezultate, jih razumeti in razvijati doma – tudi za slovenščino.

4. Komunikacijski kanali in orodja

Izbor kanalov je prilagojen ciljnim skupinam in obsegu podoktorskega projekta ter je realno izvedljiv:

- *znanstvene objave*: članki v mednarodnih revijah s področja (npr. Computational Linguistics, Transactions of the Association for Computational Linguistics, Computer Speech & Language, Journal of Machine Learning Research);
- *konference in delavnice*: predstavitve na mednarodnih konferencah s področja obdelave naravnega jezika in strojnega učenja (npr. ACL, EMNLP, EACL, ICML, IJCAI, NeurIPS) ter domačih dogodkih (npr. Slovenska konferenca o odkrivanju znanja v podatkih, SiKDD);
- *spletna stran projekta*: GitHub stran projekta (<https://github.com/eriknovak/SLM4IE>) z dokumentacijo, kodo in navodili za uporabo ter podstran projekta na spletni strani Instituta »Jožef Stefan« (<https://ijs.si/ijsw/ARRSProjekti/2026>), kjer je objavljen tudi ta komunikacijski načrt;



- *Hugging Face Hub*: odprta objava razvitih modelov in podatkovnih množic ter spremljajoče strokovne objave;
- *družbena omrežja*: objave o pomembnejših rezultatih in mejnikih projekta prek profila vodje projekta na omrežju LinkedIn;
- *blog objave*: poljudno-strokovne predstavitve rezultatov na spletni strani projekta oziroma platformi Hugging Face.

Pri pomembnejših mejnikih se lahko projekt po presoji opre tudi na obstoječe komunikacijske kanale Instituta »Jožef Stefan« (spletna stran ter profili na omrežjih Facebook, Instagram, LinkedIn in TikTok).

5. Časovni načrt komuniciranja

Časovni načrt je okvirni in usklajen z vsebinskimi fazami projekta ter njegovimi rezultati (poročila D1–D4). Aktivnosti so predvidene šele v fazah, ko so vsebinsko mogoče.

Faza projekta	Obdobje	Komunikacijske aktivnosti
Začetek projekta	M1-M6 (03/2026 - 08/2026)	Vzpostavitev GitHub strani projekta in podstrani na spletni strani IJS; objava komunikacijskega načrta; predstavitev ciljev projekta in raziskovalnega pristopa (blog, LinkedIn)
Vmesne faze	M7-M18 (09/2026 - 08/2027)	Objava prvih podatkovnih množic in vmesnih rezultatov (Hugging Face, blog); predstavitve na konferencah in delavnicah; oddaja znanstvenih člankov
Zaključek projekta	M19-M24 (09/2027 - 02/2028)	Odprta objava končnih modelov, podatkovnih množic in kode z dokumentacijo; zaključne znanstvene objave; strnjena predstavitve glavnih rezultatov projekta (blog, LinkedIn)

6. Odgovornosti

Projekt je doktorski raziskovalni projekt z enim raziskovalcem. Vse komunikacijske aktivnosti načrtuje, koordinira in izvaja vodja projekta, ki je odgovoren tudi za posodabljanje spletnih vsebin, objavo rezultatov v odprtih repozitorijih in za skladnost komuniciranja z zahtevami ARIS.

7. Kazalniki uspešnosti

Uspešnost komuniciranja se spremlja s preprostimi, merljivimi kazalniki, sorazmernimi z obsegom projekta. Ciljne vrednosti veljajo za celotno trajanje projekta.

Kazalnik	Ciljna vrednost
Znanstvene objave (revije oziroma konference z recenzijo)	Najmanj 2
Predstavitve na mednarodnih konferencah ali delavnicah	Najmanj 2



Objavljeni modeli na platformi Hugging Face	Najmanj 1
Objavljene podatkovne množice na platformi Hugging Face	Najmanj 1
Blog in LinkedIn objave o projektu	Najmanj 4
Odziv skupnosti (prenos modelov in množic, GitHub zvezdice)	Spremljanje brez fiksne ciljne vrednosti

8. Skladnost z zahtevami ARIS

Komunikacijski načrt in njegovo izvajanje bosta realizirana skladno z zahtevami ARIS, zlasti glede navajanja sofinanciranja ARIS pri objavah in dogodkih, diseminacije rezultatov raziskav, upoštevanja pravil odprtega dostopa ter etičnih načel raziskovalnega in javnega delovanja.