

Komunikacijski načrt projekta Z2-70070

Ljubljana, 31. 8. 2026

Osnovne informacije o projektu:

- Oznaka in naziv projekta: Z2-70070, BRIDGE: Bridging Reasoning and Informational Deficiencies in Generative Engines (Odprava halucinacij in izboljšanje razumskega sklepanja v generativnih jezikovnih modelih)
- Akronim projekta: BRIDGE
- Institucija: Institut »Jožef Stefan«, IJS
- Trajanje projekta: 24 mesecev
- Vodja projekta: Matej Martinc
- Projektni partnerji: /
- Verzija tega dokumenta: 1.0
- Datum oddaje komunikacijskega načrta V1: 15. 7. 2026
- Dokument pripravil: Matej Martinc

Kazalo

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

1. Namen in cilji komuniciranja

Umetna inteligenca (UI) in veliki jezikovni modeli (LLM) izrazito preoblikujejo sodobno družbo. Čeprav so komercialni modeli zmogljivi, so zaradi svoje zaprtosti, stroškov in vprašanj zasebnosti za mnoga področja in manjše jezike (kot je slovenščina) neustrezni. Odprtokodni modeli so pomembna alternativa, a so trenutno bolj nagnjeni k t. i. halucinacijam in napakam pri logičnem sklepanju.

Namen komunikacijskega načrta v okviru projekta BRIDGE je obveščati strokovno in širšo javnost o razvoju novih, inovativnih metod usposabljanja, ki bodo odprtokodne modele naučile prepoznati lastne vrzeli v znanju. S transparentnim komuniciranjem želimo graditi zaupanje v zanesljivo in varno umetno inteligenco ter spodbujati uporabo razvitih slovenskih in angleških modelov v kritičnih praksah (zdravstvo, pravo, znanost).

Glavni cilji:

- Povečanje vidnosti raziskave, razvitih modelov in Instituta »Jožef Stefan«, krepitev zaupanja v transparentno in odprtokodno znanost ter vpliv na pravičnejši razvoj UI tehnologij.

Specifični cilji:

- Diseminacija novih metodologij za bogatenje podatkov in usposabljanje (DPO, RLHF) znotraj globalne znanstvene NLP skupnosti.
- Omogočanje neposrednega prenosa razvitih modelov v aplikativna področja in ozaveščanje domenskih strokovnjakov (v medicini, pravi in znanosti) o prednostih in omejitvah generativne UI.
- Informiranje splošne javnosti in odločevalcev o nujnosti razvoja in ohranjanja slovenskega jezika v digitalni dobi skozi razvoj suverenih jezikovnih tehnologij.

2. Ciljne skupine

Namen komuniciranja glede na ciljno skupino:

Ciljna skupina	Namen komuniciranja
Znanstvena javnost (raziskovalci s področja obdelave naravnega jezika (NLP) in UI doma in v tujini)	Izmenjava znanja, znanstvene objave metodologij (WP1, WP2), spodbujanje uporabe in nadgradnje naše odprtokodne kode ter sintetičnih naborov podatkov.
Strokovna in laična javnost (pravniki, zdravstveno osebje, nevladne organizacije, organizacije za socialna vprašanja)	Omogočanje uporabe raziskovalnih rezultatov v praksi (orodja za analizo medicinskih dialogov, orodja za poenostavitev pravnih besedil o socialnih pravicah).
Odločevalci in oblikovalci javnih politik (Ministrstvo za digitalno preobrazbo, Ministrstvo za zdravje idr.)	Podpora pri strateškem odločanju o uvajanju varne in domače (slovenske) umetne inteligence v državne sisteme ter poudarjanje pomena razvoja digitalnih virov za slovenščino.
Gospodarstvo (IT in AI podjetja)	Prenos tehnologij v gospodarstvo, spodbujanje uporabe razvitih odprtokodnih modelov (z manj halucinacijami) v komercialnih in storitvenih rešitvah.

Mediji (novinarji s področja znanosti in tehnologije)

Širjenje informacij o prebojih, demistifikacija umetne inteligence (kako modeli "razmišljajo" in zakaj "halucinirajo") ter promocija slovenske znanosti.

Financerji (ARIS)

Poročanje, pregled napredka, potrditev smotrnosti vlaganja v bazične tehnološke projekte.

3. Ključna sporočila

Da bi učinkovito nagovorili izbrane ciljne skupine, so oblikovana naslednja osrednja sporočila:

- **Kaj raziskujemo:** V projektu BRIDGE razvijamo inovativne postopke učenja, s katerimi manjšim, prosto dostopnim jezikovnim modelom (umetni inteligenci) pomagamo, da se začnejo zavedati svojih omejitev. Namesto da bi si model odgovor izmislil (haluciniral), ga učimo, da samostojno prepozna pomanjkanje znanja in odgovori z "Ne vem".
- **Zakaj je to pomembno:** Najzmogljivejši modeli umetne inteligence (kot je ChatGPT) so v rokah velikih globalnih korporacij, kar prinaša visoke stroške in tveganja za zasebnost podatkov. Z izboljšanjem odprtokodnih modelov krepimo tehnološko neodvisnost ter zagotavljamo, da jeziki z manj viri, predvsem slovenščina, v tehnološkem razvoju ne zaostanejo.
- **Kakšni so (ali bodo) glavni rezultati:** Ustvarili bomo nove odprtokodne algoritme, zbirke podatkov ter nove različice modelov, ki bodo neposredno in brezplačno dostopni. Njihovo zanesljivost bomo praktično dokazali na zahtevnih primerih iz prakse – pri pomoči detekcije demence iz zdravstvenih pogovorov, pri prevajanju zapletenih pravnih besedil v ljudem razumljiv jezik in pri analizi strokovne literature.
- **Kakšen je širši pomen raziskave:** Projekt prispeva k varnejši, pravičnejši in bolj trajnostni umetni inteligenci. Zagotavlja orodja, na katera se uporabniki (od zdravnikov do običajnih državljanov) lahko zanesejo brez strahu pred dezinformacijami.

4. Komunikacijski kanali in orodja

Za doseganje komunikacijskih ciljev bomo uporabili naslednje kanale:

- **Znanstvene objave:** Vrhunske odprtodostopne revije in zborniki konferenc s področja računalniškega jezikoslovja ter strojnega učenja (npr. *Transactions of the Association for Computational Linguistics (TACL)*, *ACL*, *ICLR*, *NeurIPS*, *LREC*).
- **Spletna stran in repozitoriji:**
 - Projektna spletna stran znotraj domene Instituta »Jožef Stefan« (za splošno in strokovno javnost).
 - Repozitoriji za kodo (GitHub) in modele (Hugging Face, Clarin.si) – namenjeno znanstvenikom in gospodarstvu. Pripravljen bo tudi javni API za preizkušanje modelov.
- **Konference, simpoziji in delavnice:** Udeležba na mednarodnih konferencah za predstavitev dosežkov (WP1, WP2, WP3) in morebitna predstavitev rezultatov na domači konferenci JT-DH (Jezikovne tehnologije in digitalna humanistika).
- **Družbena omrežja:** Komuniciranje napredka in objav preko uradnih LinkedIn profilov

Instituta »Jožef Stefan« in Odseka za tehnologije znanja (LinkedIn) ter Instagram profila Centra za Jezikovne Vire in Tehnologije (CJVT).

- **Medijski prispevki:** Intervju, poljudni članek ali radijski prispevek.
- **Javni dogodki:** Sodelovanje na dogodkih popularizacije znanosti (npr. Dnevi Instituta »Jožef Stefan«, Evropska noč raziskovalcev).

5. Časovni načrt komuniciranja

Dinamika komuniciranja je neposredno vezana na potek delovnih sklopov (WP).

- **Začetek projekta (M1–M6):**
 - Postavitev spletne strani projekta (M1).
 - Vzpostavitev notranjih in javnih odprtokodnih repozitorijev (GitHub).
 - Objava o začetku projekta in ciljih na institucionalnih družbenih omrežjih IJS.
- **Vmesne faze (M7–M18):**
 - Objava prvega znanstvenega članka s področja ustvarjanja sintetičnih podatkov in detekcije vrzeli v znanju (D1.1, okvirno M12).
 - Udeležba in predstavitve na prvi mednarodni konferenci.
 - Objava drugega znanstvenega članka o metodologijah usposabljanja (D2.1, okvirno M18).
- **Zaključek projekta (M19–M24):**
 - Objava in predstavitve zadnjih rezultatov s poudarkom na evalvacijah v specifičnih domenah – pravo, zdravstvo, znanost (D3.1 in D3.2).
 - Javna lansiranje (Release): Široka objava celotnega odprtokodnega koda, sintetičnih naborov in LLM modelov preko API-ja in Hugging Face okolja.
 - Priprava sporočila za medije in predstavitve rešitev strokovni laični javnosti.

6. Odgovornosti

Za izvajanje komunikacijskih aktivnosti so predvidene naslednje vloge:

- **Vodja projekta:** Glavni nosilec komunikacijskega načrta. Odgovoren je za vodenje in pripravo znanstvenih publikacij, udeležbo in predstavitve na znanstvenih konferencah, objavljanje orodij, modelov in kode v repozitorije (GitHub, Hugging Face) ter ažuriranje vsebin na projektni spletni strani.
- **Služba za stike z javnostjo IJS:** Podpora pri diseminaciji širši javnosti. Skrbi za distribucijo sporočil za javnost medijem ter za objavljanje ključnih informacij in prebojev projekta na uradnih družbenih omrežjih in osrednji spletni strani Instituta »Jožef Stefan«.

7. Kazalniki uspešnosti

Za spremljanje uspešnosti zastavljenih ciljev smo določili naslednje merljive kazalnike (za obdobje celotnega projekta):

- **Znanstvene objave:** Objava vsaj 3 znanstvenih člankov (1 v reviji in 2 konferenčni članka) v odprtem dostopu v vrhunskih mednarodnih edicijah.
- **Dogodki:** Aktivna udeležba z ustno ali poster predstavitvijo na vsaj 2 mednarodnih konferencah s področja NLP/AI.
- **Digitalni doseg (repozitoriji):** Vzpostavitev vsaj 1 repozitorija na platformi Hugging Face/GitHub in doseganje vsaj 100 prenosov (downloads) javno dostopnih modelov ali

naborov podatkov s strani širše razvojne skupnosti.

- **Spletna in medijska prisotnost:** Postavitev projektne spletne strani do konca M1 ter objava vsaj 1 prispevka (ali omemb) v medijih, namenjenih splošni javnosti (intervju, poljudni članek, radijski prispevek ipd.).

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. Vse znanstvene objave bodo sledile načelom politike odprtega dostopa («as open as possible, as closed as necessary»), vsa razvita koda pa bo deljena pod odprtokodnimi licencami.