

RRI PoVejMo

Rezultat **D3.9: Aplikacija za strojno ekstrakcijo entitet in izvedbo anonimizacije dokumentov**

Vodilni partner: **Semantika, d.o.o.**

Sodelujoči partnerji: **Fakulteta za računalništvo in informatiko UL Inštitut za novejšo zgodovino**

Februar 2026

## Uvod

Programska oprema GalisOnline je celovit dokumentacijski sistem za muzeje in arhive, v katerem so o predmetih shranjena obsežna besedila — opisi, popisi, prepisi arhivskega gradiva in druge proste vsebine. V teh besedilih so pogosto omenjene imenske entitete: osebe, kraji, organizacije in datumi. Te entitete so dvojno pomembne. Po eni strani so ključ za povezovanje zapisov z že obstoječimi šifranti in registri (avtorji, kraji, ustanove), kar je sicer zamudno ročno opravilo. Po drugi strani lahko ista besedila vsebujejo osebne podatke živih oseb, ki jih je treba pred javno objavo ali deljenjem z raziskovalci ustrezno zakriti (anonimizirati) v skladu z varstvom osebnih podatkov. V okviru raziskovalno-razvojnega projekta PoVejMo (»Prilagodljiva obdelava naravnega jezika z velikimi jezikovnimi modeli«) smo zato razvili novo storitev, ki iz besedila zapisa samodejno izlušči imenske entitete ter na isti tehnološki osnovi omogoči zakrivanje osebnih podatkov. S tem nadgrajujemo obstoječo dokumentacijsko infrastrukturo — kustosu olajšamo povezovanje zapisov in pripravo gradiva za varno objavo. Nova rešitev, ki jo opisujemo v tem poročilu, je bila v sklopu projekta preizkušena v okolju GalisOnline (dosežena tehnološka zrelost TRL 6) in predstavlja pomemben korak k boljše povezanim in varno objavljenim slovenskim zbirkam kulturne dediščine.

## CILJ C3.9 – ekstrakcija entitet in anonimizacija

Projekt PoVejMo je usmerjen v raziskave in razvoj novih rešitev na področju umetne inteligence in jezikovnih tehnologij s poudarkom na slovenščini. Znotraj projekta poteka več delovnih sklopov, med katerimi je delovni sklop C3 osredotočen na praktične demonstratorje uporabe velikih jezikovnih modelov nad slovenskimi besedilnimi viri. Glavni cilj tega rezultata je razviti in v GalisOnline integrirati storitev, ki iz besedila zapisa samodejno prepozna imenske entitete (osebe, kraje, organizacije, datume) in jih predstavi v strukturirani obliki, na isti osnovi pa omogoči zakrivanje osebnih podatkov. To pomeni, da se besedilo zapisa posreduje jezikovnemu modelu, ta pa vrne seznam prepoznanih entitet z njihovimi vrstami; za anonimizacijo se podmnožica osebnih entitet v besedilu nadomesti z oznakami (npr. [OSEBA], [NASLOV]). Prepoznane entitete služijo dvema namenoma: kot predlogi za povezavo zapisa s šifranti in registri zbirke ter kot osnova za pripravo anonimizirane različice besedila, primerne za javno objavo ali deljenje. Rezultat neposredno podpira širši cilj projekta PoVejMo, da z uporabo velikih

jezikovnih modelov izboljša dostopnost, povezanost in varno uporabo slovenskih besedilnih virov v digitalni obliki.

## Razvoj rešitve in uporabljene tehnologije

Pri razvoju rešitve smo izhajali iz pomočnika z umetno inteligenco, ki smo ga vgradili v GalisOnline kot stransko ploščo ob zapisu predmeta. Plošča ponuja klepetalni vmesnik in hitre akcije (med njimi »Izlušči entitete« in »Anonimiziraj«), odzivi pa se uporabniku izpisujejo pregledno, tako da je delo z njim tekoče tudi pri daljših besedilih.

Ko kustos odpre zapis predmeta, aplikacija samodejno zbere besedilne vsebine zapisa — opise, popise in druge proste vsebine iz glavnih in podrejenih sekcij forme — in jih skupaj z imenom forme in identifikatorjem zapisa pripravi kot strukturiran kontekst. Ob sprožitvi ekstrakcije se ta kontekst posreduje strežniški storitvi, ki sestavi sistemski poziv z navodili, naj model iz besedila prepozna imenske entitete in jih razvrsti po vrstah (osebe, kraji, organizacije, datumi). Tako pripravljeno zahtevo obdela veliki jezikovni model GaMS, razvit v sklopu PoVejMo in prilagojen za slovenski jezik; model vrne strnjen, strukturiran seznam entitet. Za anonimizacijo se isti postopek omeji na osebne entitete, ki jih nato deterministični strežniški modul v besedilu nadomesti z oznakami vrste entitete — tako iz istega prepoznavanja nastane zakrita različica besedila.

Pomemben del rešitve je oblikovanje pozivov (prompt engineering), s katerim model usmerimo, da entitete prepozna zanesljivo in v predvidljivi strukturirani obliki, ter omejimo tveganje halucinacij — model ne sme dodajati entitet, ki jih v besedilu ni. Razdelitev nalog, pri kateri jezikovni model prepozna entitete, samo zakrivanje pa izvede deterministična koda nad izvirnim besedilom, zagotavlja, da je rezultat preverljiv in ponovljiv. Kustos nad rezultatom obdrži nadzor: predlagane entitete in povezave pregleda in potrdi, anonimizirano besedilo pa pred uporabo preveri.

Tehnološko gledano je storitev zasnovana kot neodvisen REST API znotraj infrastrukture GalisOnline (/api/ai/chat). Strežniški modul prejme besedilo zapisa in vrsto zahteve (ekstrakcija ali anonimizacija), sestavi poziv, zahtevo posreduje jezikovnemu modelu in iz njegovega odgovora zgradi strukturiran seznam entitet oziroma zakrito besedilo. Celoten proces je dovolj hiter, da deluje v interaktivnem času — kustos rezultat prejme v nekaj sekundah na standardni strojni opremi, kar je sprejemljivo za neposredno uporabo v aplikaciji.

V okviru razvoja smo izhode modela GaMS vzporedno primerjali tudi z modeloma Qwen 3.5 in Qwen 3.6 (splošnonamenska večjezična modela), kar nam je omogočilo ovrednotenje kakovosti in primernosti generiranih rezultatov glede na alternativne pristope.

## Implementacija kot del GalisOnline

Razvito rešitev smo tesno povezali z obstoječim sistemom GalisOnline, tako da je uporabnikom na voljo brez potrebe po zunanjem orodju. Arhitektura integracije sledi načelu modularnosti: pomočnik je vgrajen kot stranska plošča ob zapisu, ki ob akciji »Izlušči entitete« pošlje besedilne

vsebine zapisa na storitev za ekstrakcijo. Prepoznane entitete se kustosom prikažejo kot predlogi za povezavo zapisa s šifranti in registri zbirke (avtorji, kraji, ustanove), ki jih ta s potrditvijo poveže z zapisom. Ob akciji »Anonimiziraj« storitev vrne zakrito različico besedila, ki jo kustos uporabi pri javni objavi predmeta oziroma pri deljenju gradiva z raziskovalci.

Ena od prednosti razvite rešitve je, da je bila zasnovana za neposredno integracijo v obstoječo aplikacijo GalisOnline — tehnično GalisOnline nadgrajujemo z novim modulom za ekstrakcijo entitet in anonimizacijo, ki ponovno uporablja že vgrajeno infrastrukturo pomočnika z umetno inteligenco.

The screenshot displays the 'AI Pomočnik' (AI Assistant) interface within GalisOnline. The interface is divided into two main sections: a left sidebar and a right main content area.

**Left Sidebar:**

- Header: **AI Pomočnik** 3-ITEMFORM
- Link: **Povezan z zapisom #22710**
- Button: **Izlušči imena**
- Section: **Najdene entitete:**
- Category: **Osebe 2** (People 2)
  - Entity: **Vincent** (Vincent van Gogh)
- Category: **Kraji 2** (Locations 2)
  - Entity: **Arles** (Arles)
  - Entity: **Francija** (France)
- Category: **Organizacije 0** (Organizations 0)
- Category: **Datumi 3** (Dates 3)
  - Entity: **september 1888 do januar 1889** (September 1888 to January 1889)
  - Entity: **1888**
  - Entity: **18**
- Buttons: **Opiši predmet** (Describe subject), **Predlagaj metapodatke** (Suggest metadata)
- Button: **Izlušči entitete** (Extract entities)

**Right Main Content Area:**

- Tab: **Mediji** (Media)
- Tab: **Metapodatki** (Metadata)
- Section: **SLIKE** (Images)
- Image: A reproduction of Vincent van Gogh's painting 'Oudestraat in Arles' (1888), showing a street scene with buildings and figures.
- Filename: **1695806791421.jpg**
- Page indicator: **1 / 1**

Slika 1: Posnetek zaslona GalisOnline — stranska plošča pomočnika s seznamom prepoznanih entitet in predogledom anonimiziranega besedila.