

RRI PoVejMo

Rezultat **D3.5: Demonstracijska aplikacija za samodejno generiranje opisov zbirk**

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, ki kustosom in dokumentalistom omogoča strukturiran vnos podatkov o predmetih — naziv, avtor, datacija, material in tehnika, dimenzije, provenienca, hramba in podobno. Ti podatki so praviloma zabeleženi v ločenih, dobro strukturiranih poljih, manjka pa povezano besedilo — vsebinski opis predmeta ali zbirke, ki ga obiskovalec ali raziskovalec dejansko prebere. Tak opis kustosi danes pišejo ročno; delo je zamudno, pri večjih zbirkah pa pogosto ostane nedokončano ali neenotno po slogu in obsegu. V okviru raziskovalno-razvojnega projekta PoVejMo (»Prilagodljiva obdelava naravnega jezika z velikimi jezikovnimi modeli«) smo zato razvili novo storitev za samodejno generiranje opisov, ki iz obstoječih strukturiranih podatkov zapisa tvori koherenten, slovnično pravilen opis v slovenščini in je vgrajena neposredno v GalisOnline. S tem nadgrajujemo obstoječo dokumentacijsko infrastrukturo — kustosu ponudimo kakovosten osnutek opisa, ki ga le še pregleda in po potrebi uredi, namesto da bi ga pisal od začetka. Nova rešitev, ki jo opisujemo v tem poročilu, je bila v sklopu projekta že preizkušena v realnem okolju GalisOnline (dosežena tehnološka zrelost TRL 6) in predstavlja pomemben korak k hitrejšemu in bolj enotnemu opisovanju slovenske kulturne dediščine.

CILJ C3.5 – samodejno generiranje opisov

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 na podlagi strukturiranih kataložnih podatkov samodejno tvori vsebinski opis predmeta ali zbirke v slovenščini. To pomeni, da se polja zapisa (npr. naziv, avtor, datacija, material, tehnika, provenienca) posredujejo jezikovnemu modelu, ta pa iz njih zgradi tekoče, naravno

slovensko besedilo, prilagojeno muzejski katalogizacijski praksi — ob čemer se strogo drži danih podatkov in ne dodaja izmišljenih dejstev. Rezultat neposredno podpira širši cilj projekta PoVejMo, da z uporabo velikih jezikovnih modelov izboljša dostopnost in uporabnost 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 »Generiraj opis«), odzivi pa se uporabniku izpisujejo pretočno (sprotno), tako da je delo z njim tekoče tudi pri daljših besedilih.

Ko kustos odpre zapis predmeta, aplikacija samodejno zbere vrednosti vseh prikazanih polj zapisa — tako iz glavnih kakor iz podrejenih sekcij forme — in jih skupaj z imenom forme in identifikatorjem zapisa pripravi kot strukturiran kontekst. Ob sprožitvi generiranja se ta kontekst posreduje strežniški storitvi, ta pa sestavi sistemski poziv, v katerega vključi podatke zapisa v obliki JSON in navodila o želenem slogu in registru opisa (muzejska katalogizacijska praksa, slovenščina, brez dodajanja podatkov, ki jih v zapisu ni). Tako pripravljeno zahtevo obdela veliki jezikovni model **GaMS**, razvit v sklopu PoVejMo in prilagojen za slovenski jezik. Model na podlagi posredovanih polj tvori povezano besedilo: pri posameznem predmetu opis enega zapisa, pri zbirki pa agregiran opis, ki povzame skupne značilnosti več zapisov (obdobje, avtorstvo, tematske ali materialne sklope).

Pomemben del rešitve je oblikovanje pozivov (prompt engineering), s katerim model usmerimo v ustrezen muzejski register in obenem omejimo tveganje halucinacij — model opisuje izključno iz danih polj in ob pomanjkanju podatka tega ne nadomesti z izmišljenim dejstvom. S tem ostane generirani opis preverljiv glede na izvorni zapis, kustos pa nad rezultatom obdrži nadzor: predlog pregleda, po potrebi popravi in šele nato shrani v zapis.

Tehnološko gledano je storitev zasnovana kot neodvisen REST API znotraj infrastrukture GalisOnline (/api/ai/chat). Strežniški modul prejme podatke zapisa, sestavi poziv in zahtevo posreduje jezikovnemu modelu, odgovor pa pretočno (SSE) vrača odjemalcu. Celoten proces je dovolj hiter, da deluje v interaktivnem času — kustos za posamezen predmet predlog opisa 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 opisov 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 »Generiraj opis« pošlje strukturirane podatke zapisa na storitev za generiranje besedila. Storitev vrne predlog opisa, ki ga kustos pregleda, po potrebi uredi in shrani kot del podatkov o predmetu. S tem lahko opis sodeluje v nadaljnji katalogizaciji, prikazu predmeta in iskanju, kustos pa namesto pisanja od začetka opravi le še uredniški pregled.

Ena od prednosti razvite rešitve je, da je bila neposredno preizkušena in integrirana v obstoječo aplikacijo GalisOnline — tehnično smo GalisOnline nadgradili z novim modulom za samodejno generiranje opisov.

The screenshot displays the GalisOnline interface for editing a record titled "Rumena hiša (ulica)". The main panel shows a form with various fields for metadata and description. On the right, an "AI Pomočnik" (AI Assistant) panel is open, displaying a generated description for the record.

Record Details:

- Inventory Number (SM):** 2014025
- Accession Number (SM):** [Empty]
- Subject:** slika
- Subject Additional:** test
- Title:** Rumena hiša (ulica)
- Title Additional:** [Empty]
- Category:** slikarstvo
- Author:** Vincent van Gogh
- Creation Date:** september 1888 do januar 1889
- Creation Date Range:** 1888 / 1889
- Location:** Arles, Francija
- Signature:** Vincent
- Location:** [Empty]

AI Pomočnik - Poglobljen opis predmeta:

Dela *Rumena hiša* spada med najpomembnejša van Goghova dela iz obdobja v Arlesu. Slika prikazuje hišo na kolu ulice, v kateri je van Gogh nastanil svoje atelje in v kateri je nameraval ustanoviti umetniško kolonijo za druge umetnike.

****Pomen dela:****

- Hiša je predstavljala simbol za van Goghovo sanjar o skupnosti umetnikov
- Delo odraža van Goghovo zanimanje za arhitekturo in urbanistične motive
- Barvna paleta je značilna za van Goghovo obdobje v Arlesu – tople, svetle barve

****Krajevidi:****

Slika prikazuje arhitekturo v južni Franciji, z značilnimi rumenimi fasadami, ki so doble ime po delu. Umetnik je s svojim značilnim, energičnim čopičnim premikom dočrтал arhitekturne detaje in perspektivo ulice.

Predlogi za dopolnitev zapisa

<actions>

Priporočilo za restavratorsko dokumentacijo

Zaradi starosti dela (1888-1889) in specifičnosti [Empty]

Vprašaj kaj o tem zapisu...