Umetna inteligenca v MSP

Lastnosti, primeri uporabe in implikacije AI akta

Maja Škrjanc · Odsek za umetno inteligenco · Institut Jožef Stefan ·

Kaj boste odnesli

Tri vsebinski bloki + zaključek

01

Kaj UI je in kakšne so njene lastnosti

Klasično strojno učenje proti generativni UI. Kaj določajo podatki in kaj poziv.

02

Primeri uporabe in implikacije AI akta

Plastični primeri iz MSP. Kdaj ste samo uvajalec in kdaj ponudnik. Štiri ravni tveganja.

03

Izzivi sedanjih implementacij

Kako izbrati rešitev (make/buy/API), strojna oprema, kalibracija, hitro spreminjanje področja.

04

Ključna sporočila in razprava

Kaj odnesti v podjetje že jutri.

01

Kaj umetna inteligenca je

in kakšne so njene ključne lastnosti

Tri družine sistemov umetne inteligence

Pametni, inteligentni in kognitivni — vsak rešuje drugačno vrsto problema

PAMETNI

Znanjsko intenzivne metode

Sistem zna, ker smo znanje vanj zapisali — eksplicitna pravila, ontologije, grafi znanja.

Primer: Pravila za odobritev plačila, ekspertni sistemi, diagnostični vodiči.

INTELIGENTNI

Podatkovno intenzivne metode

Sistem se uči iz preteklih podatkov in jih posploši — klasično strojno učenje, klasifikatorji.

Primer: Napoved kreditne sposobnosti, klasifikacija prevarantskih transakcij.

KOGNITIVNI

Sistemi, ki razumejo in sklepajo

Sistem se uči, ima pomnjenje in kontekst, razume strukturo problema in zna sklepati. Kombinira znanje in učenje.

Primer: Digitalni dvojčki, hibridni odločevalni sistemi, agentni sistemi s planiranjem.

Glavno sporočilo: *vsaka družina drugače uporablja podatke, drugače kreira model.*

Sodobna slika: dva pristopa, dve filozofiji

Strojno učenje (in digitalni dvojčki) na eni strani, generativna UI na drugi

Strojno učenje · digitalni dvojčki

Kaj počne

Strojno učenje se nauči vzorcev iz preteklih podatkov in posploši pravila.

+ razumevanje strukture

Če k strojnemu učenju dodamo razumevanje, kako sistem dejansko deluje (fizikalni model, domensko znanje), dobimo **digitalne dvojčke**.

Lastnosti

- Razumemo, zakaj model reče, kar reče
- Lahko revidiramo, popravimo, interpretiramo
- Strukturiran prostor značilk in jasna tarča
- Pri digitalnem dvojčku: simulacija vzročno-posledičnih razmerij, ne le korelacij

Generativna UI · LLM

Kaj počne

Gigantske nevronske mreže, ki napovedujejo **najverjetnejšo naslednjo besedo**. Brez razumevanja — zelo zmogljivo statistično napovedovanje.

Brez razumevanja: *model ne ve, kaj govori — le napoveduje, kateri tok besed je v učnih podatkih najverjetnejši.*

Lastnosti

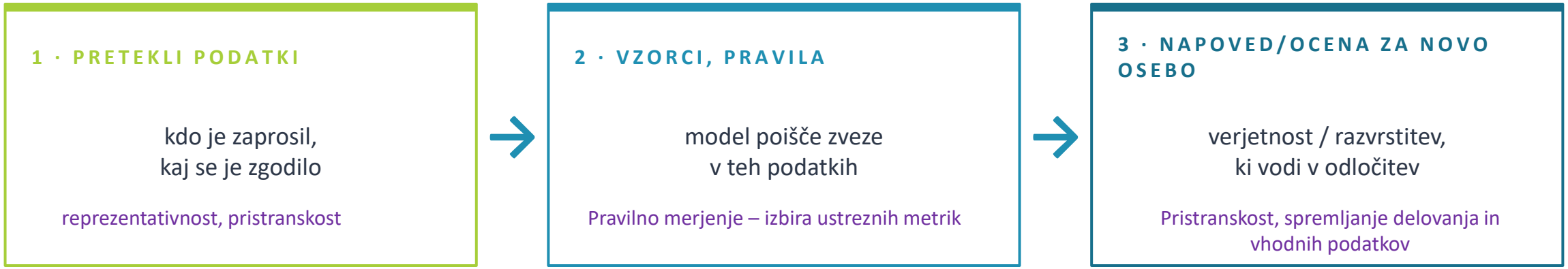
- Notranjost je »črna škatla« — milijarde parametrov, neproniknjivih
- Odgovori zvenijo samozavestno tudi, ko so napačni (halucinacije)
- Močna pri jeziku in formi, šibka pri vzročnem sklepanju
- Razumevanja v pravem pomenu ni — je le posnemanje vzorcev

Vprašanje za MSP: ***Izberite primerno vrstu UI***

Kaj model UI sploh je

Brez matematike — ena slika, ki velja za vse, kar sledi

Model se iz preteklih podatkov nauči vzorcev in jih uporabi, da za novo osebo oceni ali napove izid. Nič več in nič manj.



Zakaj je to pomembno za MSP:
Kakovost napovedi je odvisna od preteklih podatkov. Če v podatkih določene skupine strank ni bilo — modelu ne moremo zaupati na tej skupini. Če so v podatkih napake — model jih ponovi.

Klasično strojno učenje proti LLM

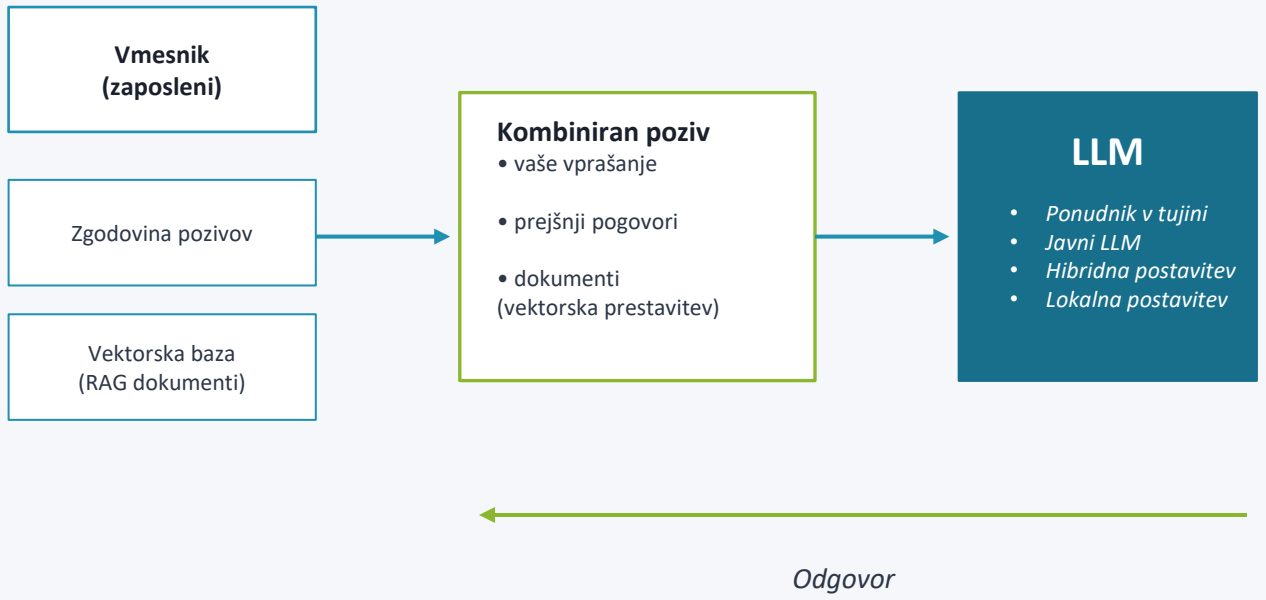
Za odločevalce v MSP se razlikuje skoraj vse, kar šteje

	Klasično strojno učenje	LLM / generativna UI
Kdo nadzira podatke	Mi sami — učno množico in značilke izberemo	Ponudnik — z zajemom z odprtega spleta
Naša tipična vloga po AI aktu	Pogosto ponudnik ali uvajalec	Skoraj vedno uvajalec
Naša vloga po GDPR	Upravljavec osebnih podatkov	Upravljavec, hkrati pa nastopajo obdelovalci in tretje države
Kaj lahko revidiramo	Značilke, definicija tarče, metrike, učna populacija	Le vhod in izhod; notranjost ostaja zaprta
Prevladujoče tveganje	Pristranskost: kdo je v podatkih, kdo ne	Izvor podatkov, izmišljanje, prenos podatkov k obdelovalcu

EDPS (oktober 2025): *pojmi »ponudnik / uvajalec« iz AI akta se NE prekrivajo z »upravljavec / obdelovalec« iz varstva podatkov.*

Veliki jezikovni modeli: kaj se zgodi z vašim pozivom

Kako se podatki gibljejo, ko zaposleni napiše prompt



Vsako besedilo, ki ga vpišete, gre obdelovalcu. Lahko tudi v tretjo državo.

4 stvari, ki jih morate vedeti

Izmišljanje o osebah

Model trdi neresnico o stranki → udari ob točnost in pravico do popravka.

Izvor učnih podatkov

Spletni zajem podatkov, ki ga niste ustvarili vi.

Poziv = prenos podatkov

Besedilo gre obdelovalcu, lahko tudi v tretjo državo.

RAG in zaupnost

Kodirani dokumenti potegnjeni v poziv; bolj varno. Kontrolirano.

02

Primeri uporabe

in implikacije evropskega Akta o umetni inteligenci

Umetna inteligenca v MSP danes

Kje jo dejansko srečate v poslovnih procesih

Strojno učenje (klasične) metode

- Napovedovanje povpraševanja (planiranje zalog, kadrovanje)
- Zaznava goljufij in nepravilnih transakcij
- Napovedna analitika za vzdrževanje strojev
- Optimizacija dobavne verige in razporejanje virov
- Razvrščanje strank (segmentacija, churn)
- Kontrola kakovosti slik in proizvodnih izdelkov

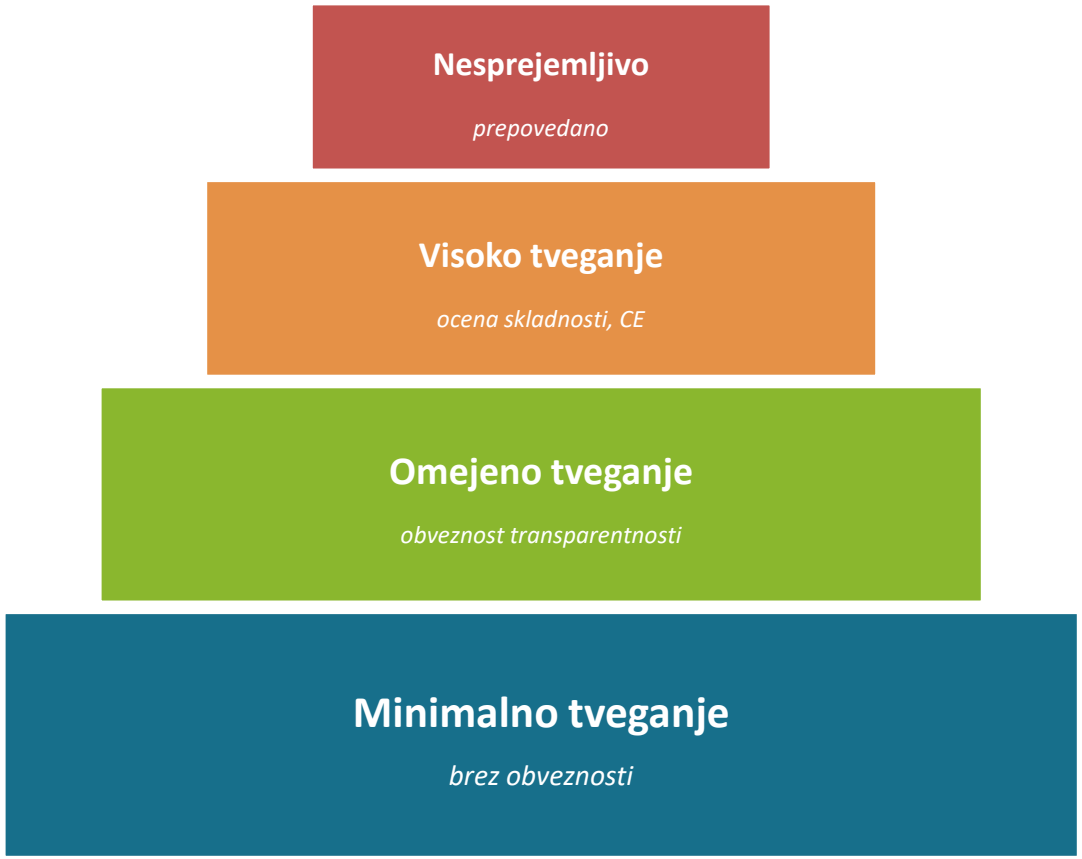
Generativna UI in LLM

- Asistenti za pisanje (osnutki, povzetki, prevodi)
- Klepetalniki za pomoč strankam (FAQ, podpora)
- Pomoč pri pisanju kode in dokumentaciji
- Iskanje po notranjih dokumentih (RAG)
- Generiranje marketinških besedil in slik
- Avtomatizacija odgovorov na elektronska sporočila

Skoraj vsako podjetje že uporablja UI — pogosto, ne da bi se tega zavedal. *Prvi korak je popis: kje, katero orodje, kdo odgovarja.*

Akt EU o umetni inteligenci: pristop po tveganju

Štiri ravni — vsaka drugačne obveznosti



Tipični primeri za MSP

Nesprejemljivo: Socialno točkovanje strank na podlagi družbenih omrežij. Biometrija v realnem času v javnih prostorih brez izrecne pravne podlage.

Visoko: Model za kreditno točkovanje. Profiliranje uspešnosti zaposlenih za HR odločitve. Sistem za triažo prošenj za zaposlitev.

Omejeno: Klepetalnik za pomoč strankam (LLM). Generirane slike in videoposnetki (»sintetični mediji«). Povzetki dokumentov za uporabnika.

Minimalno: Filter za neželeno pošto. Preprosti priporočilni sistem za novice. Avtomatska kategorizacija notranjih dokumentov.

Visoko tveganje: kdaj zadeva tudi vaš MSP

Trije primeri, ki sprožijo polni režim skladnosti

Kreditno točkovanje

Model, ki ocenjuje kreditno sposobnost posameznika ali odloča o pogojih najema. Ne glede na velikost podjetja: če ga uvedete, ste uvajalec visoko-tveganega sistema.

Triaža prošenj za zaposlitev

Sistem, ki razvršča kandidate, predlaga ožji izbor ali ocenjuje primernost. Tipično pri kadrovskih agencijah in srednjih podjetjih z veliko prijavami.

Profiliranje uspešnosti zaposlenih

Sistem, ki spremlja vedenje, produktivnost ali dodeljuje naloge na podlagi vedenjskih značilk. Pogosto skrit v »produktivnostnih« nadzornih orodjih.

Kaj to praktično pomeni — ne glede na to, ali ste ponudnik ali samo uvajalec:

- Ocena skladnosti pred uvedbo (interno ali pri priglašenem organu) in tehnična dokumentacija po prilogi IV
- Sistem upravljanja tveganj, kakovostni nadzor podatkov in zapis o izvoru podatkov (data provenance)
- Človeški nadzor na ključnih točkah odločanja in obveznost evidence — kdo je odločil, kdaj, na kakšni podlagi
- Spremljanje po dajanju na trg: nadzorne plošče za odstopanje natančnosti, pravičnosti in varnostne incidente

Omejeno tveganje: tisto, kar uporablja večina MSP

Klepetalniki in sintetične vsebine — obveznost transparentnosti

Kaj morate narediti

- **Razkriti, da uporabnik komunicira z UI**
- Klepetalniki: jasna oznaka že na vstopu
- Pomočniki, ki sestavljajo dopise: oznaka v podpisu
- **Označiti sintetično vsebino**
- Slike, videi, glasovni posnetki, ustvarjeni z UI
- Pri globokih ponaredkih: izrecna in vidna oznaka
- **Obvestiti o uporabi za prepoznavanje čustev**
- Če sistem analizira ton glasu ali izraz obraza

Praktični kontrolni seznam

- Stavek »ta klepetalnik poganja umetna inteligenca« vidno pri prvem stiku
- Disclaimer pod e-poštnimi odgovori, ki jih sestavlja UI
- Vodnik za zaposlene: »pravne in mnenjske izjave gredo skozi človeški pregled«
- Watermark ali metapodatki na vseh sintetičnih medijih
- Pravila ravnanja, ko sistem zaznava občutljive ali pravne teme
- Mini »model card«: namen, meje, znane slabosti, dos & don'ts (1 stran)

Past: *če klepetalnik začne svetovati o pravicah, terjatvah ali zdravju — preverite ali ste še v »omejenem« ali ste že padli v »visoko tvegano« kategorijo.*

Pristranskost in pravičnost

Najpogostejši izvor težav v MSP — pogosto nehoten

Direktna pristranskost

Učna populacija ne odraža resnične: model za kreditiranje, treniran na strankah iz enega manjšega mesta, slabo deluje za stranke iz drugih okolij.

Implicitna pristranskost

Značilka, ki nadomesti zaščiteno lastnost: geolokacija telefona kot »sled« socialnega statusa. Tudi če ne uporabljate spola, model uporabi posredne signale.

Metodološka napaka

Korelacija, ki ni vzročnost: »otroci z večjimi stopali so bolj inteligentni« (so pač starejši). V poslovanju: »dolgi pogovori dajejo več prodaje« (ne; večji posli pač terjajo več pogovora).

Kaj lahko v MSP-ju storite že danes — brez specializiranih kadrov

- Popis učnih podatkov: katere skupine strank so prevladovale, katerih je bilo malo ali sploh ni bilo
- Testiranje na segmentih: za vsako pomembno skupino preverite natančnost ločeno, ne le skupne metrike
- Tehnike maskiranja občutljivih podatkov pri delitvi z zunanjim ponudnikom: psevdo-anonimizacija, delno zakrivanje
- Pravilo »človek v zanki« pri odločitvah, ki vplivajo na posameznika (kredit, zaposlitev, dostop do storitve)

Vaš primer uporabe → kategorija po AI aktu

Kratka samocheck-lista za odločevalce

Primer uporabe v MSP	Tipična kategorija	Vaša glavna obveznost
Klepetalnik za pomoč strankam (LLM)	**Omejeno tveganje**	Razkritje, da gre za UI; mini model card; pravila za občutljive teme
Asistent za pisanje dopisov in povzetkov	**Omejeno tveganje**	Označevanje sintetičnih izvozov; človeški pregled pri pravnih/poslovnih dopisih
Iskanje po notranjih dokumentih (RAG)	**Omejeno tveganje***	Nadzor dostopa do zaupnih dokumentov; preprečitev iznosa preko poziva
Sistem za predizbor kandidatov za zaposlitev	**Visoko tveganje**	Polni režim: tehnična dokumentacija, človeški nadzor, monitoring
Model za odločanje o kreditni sposobnosti	**Visoko tveganje**	Polni režim + ocena pristranskosti po segmentih + post-market spremljanje
Generiranje slik za marketing	**Omejeno tveganje**	Označevanje sintetične vsebine; izogibanje globokim ponaredkom
Filter za neželeno pošto	**Minimalno tveganje**	Brez specifičnih obveznosti po AI aktu — GDPR vseeno velja

** Kategorizacija je odvisna od dejanske uporabe — če RAG asistent v praksi odloča o pravicah, lahko pade v višjo kategorijo.*

Kaj naslavlja AI akt — minimalni red v MSP

Štiri stvari, ki jih po novem morate imeti urejene

1

Popis UI v podjetju

Seznam vseh sistemov, ki uporabljajo UI — kupljenih, najetih ali razvitih. Z lastnikom, namenom in podatki, ki jih obdeluje.

Brez tega ne morete vedeti, kaj sploh ureja AI akt.

2

Razvrstitev po tveganju

Za vsak sistem določite kategorijo: minimalno, omejeno, visoko, nesprejemljivo. Klasifikacijo dokumentirajte in posodabljajte.

Klasifikacija je živ dokument, ne enkratna naloga.

3

Vloge in odgovornosti

Kdo odloča o uvedbi, kdo nadzira delovanje, kdo poroča o incidentih. Pri visoko tveganih sistemih: imenovani človeški nadzor.

Brez jasnih vlog je odgovornost vseh = odgovornost nikogar.

4

Evidenca in spremljanje

Dnevniki delovanja, redno spremljanje natančnosti in pravičnosti, postopek poročanja o resnih incidentih.

Pri visoko tveganih: tudi technical documentation in CE oznaka.

03

Izzivi sedanjih implementacij

Kaj izbrati, kako uvesti, čemu se izogniti

Kako sploh izbrati primerno rešitev

Tri poti — naredi sam, kupi gotovo, uporabi storitev

USE

Uporabi storitev / API

Najpogostejša izbira za MSP

+ prednosti

- Najhitrejši zagon
- Brez naložbe v strojno opremo
- Velika kakovost vodilnih modelov

– pasti

- Vaši podatki gredo k obdelovalcu in lahko v tretjo državo
- Cena raste s količino
- Odvisnost od enega ponudnika

BUY

Kupi gotov produkt

Standardni primeri (CRM AI, marketing)

+ prednosti

- Domensko izpiljen produkt
- Vključen podporni okvir
- Pogosto manj integracije

– pasti

- Manjša prilagodljivost
- Težko vidimo, kaj se zares dogaja
- Cena licence + vendor lock-in

MAKE

Naredi sam (in-house ali z odprtokodnim modelom)

Smiselno pri zaupnih podatkih

+ prednosti

- Podatki ostanejo doma
- Polna kontrola nad modelom
- Dolgoročno najnižji marginalni stroški

– pasti

- Najvišja začetna investicija
- Potreba po kompetenci v ekipi
- Trajno breme vzdrževanja

Ovrednotenje zaupnosti informacij

Najprej klasifikacija podatkov — šele potem izbira rešitve

T1	**Javno objavljive informacije** Marketinška besedila, blog osnutki, splošni opisi izdelkov	*nizko*
T2	**Notranje, a ne občutljivo** Povzetki srečanj, splošne notranje komunikacije, procesni opisi	*nizko*
T3	**Poslovno občutljivo** Cenovne strategije, intelektualna lastnina, načrti za stranke	*srednje*
T4	**Osebni podatki (GDPR)** Podatki strank, zaposlenih, kandidatov; vse, kar je osebno	*visoko*
T5	**Regulirani podatki** Zdravstveni zapisi, finančni profili, podatki, ki so predmet sektorske zakonodaje	*kritično*

Standardi, ki vam pri tem pomagajo

ISO 9001

Sistem vodenja kakovosti — okvir za postopke, vloge in dokumentacijo

ISO 27001

Upravljanje informacijske varnosti — klasifikacija informacij, dostop, zaščita

ISO 27701

Razširitev 27001 za varstvo osebnih podatkov — usklajenost z GDPR

ISO 42001

Sistem upravljanja UI — politike, vloge, evidence za uporabo UI v podjetju

GDPR / ZVOP-2

Pravna podlaga za obdelavo osebnih podatkov — kaj sploh smete pošiljati k zunanjemu LLM-u

Pravilo: *ne začnite z izbiro modela — začnite s klasifikacijo podatkov. Tip podatkov določa, katere rešitve UI sploh smete uporabiti.*

Tipi GenAI rešitev: kam gredo vaši podatki

Javna storitev, hibridna rešitev v oblaku, lokalna inštalacija

JAVNI

Javna storitev (API)

ChatGPT, Claude.ai, Gemini, Copilot

Podatki: Podatki gredo k ponudniku
Strošek: Plačilo po porabi (tokeni)

+ prednosti

- Najhitrejši zagon, brez infrastrukture
- Vedno najnovejši modeli
- Pogosto najboljša kakovost

– pasti

- Vaši podatki zapustijo podjetje
- Pogosto tretja država (ZDA)
- Vendor lock-in pri količinski rabi

Kdaj se obnese: *Za T1–T2 informacije (javno, notranje, ne občutljivo).*

HIBRIDNI

Lasten LLM v oblaku

Azure OpenAI Service, AWS Bedrock, dedicated tenancy

Podatki: Podatki v vašem oblačnem najemniku
Strošek: Plačilo po porabi + premija za zasebnost

+ prednosti

- Podatki ostanejo v vašem oblačnem prostoru
- Možno EU-bivanje (data residency)
- Skalabilnost brez svojega železa

– pasti

- Konfiguracija je netrivialna
- Cena višja kot javni API
- Še vedno odvisnost od ponudnika oblaka

Kdaj se obnese: *Za T3–T4 informacije (poslovno občutljivo, osebni podatki).*

LOKALNI

Lokalna inštalacija

Llama, Mistral, Qwen na svojem strežniku

Podatki: Podatki nikoli ne zapustijo podjetja
Strošek: Naložba v strojno opremo + vzdrževanje

+ prednosti

- Polna kontrola in zasebnost
- Nobeni tokeni se ne plačujejo
- Možno tudi v zaprtem omrežju (air-gapped)

– pasti

- Najvišja začetna investicija
- Pogosto manjša kakovost kot vodilni javni modeli
- Potreba po tehnični ekipi za vzdrževanje

Kdaj se obnese: *Za T5 informacije in regulirane sektorje.*

Strojna oprema in računske zahteve

Najpogosteje podcenjen del — še posebej pri LLM-jih

Tip rabe	Tipična strojna oprema
Klepetalnik prek API	Brez svoje opreme — plačujete tokene
Klasični ML model (npr. churn)	Strežnik ali oblak; CPU zadošča
Manjši LLM lokalno (7–13B parametrov)	1× GPU 24 GB VRAM (npr. RTX 4090 razreda)
Srednji LLM lokalno (70B parametrov)	2× GPU s skupno 80+ GB VRAM
Učenje (fine-tuning) lastnega modela	Več GPU; pogosto v oblaku, na uro plačila
Vektorska baza za RAG	Pomnilnik in disk; CPU zadošča do nekaj 10 mio dokumentov

Kar vas bo presenetilo

Inferenca ni brezplačna

Vsak poziv stane — pri API jasno, pri lastni opremi prikrito v elektriki in amortizaciji.

Latenca raste s kontekstom

Bolj zapleten poziv → daljši odzivni čas. Pri integracijah s strankami to opazijo.

RAG potrebuje pripravo dokumentov

Pred prvo poizvedbo: razrez, vektorizacija, indeksiranje. Ni izvedeno samo od sebe.

Tržne cene padajo, a hitro

Cena tokena vodilnih modelov je padla cca 10× v 18 mesecih — zaklenjene večletne pogodbe so tvegane.

Pravilo palca: *preden kupite GPU, najprej preverite, ali bi lahko isti rezultat dosegli s storitvijo. Marsikateri MSP najprej zazna pravo potrebo šele po 3–6 mesecih dejanske rabe.*

Kalibracija modelov in tipične napake

Zakaj iste tehnologije v dveh podjetjih dajo različne rezultate

KALIBRACIJA

Prilagajanje modela na vaš kontekst

- **Prompt engineering**
- Sistemski poziv, ki opiše vlogo, ton in meje
- Primeri želenih odgovorov (few-shot)
- **Fine-tuning**
- Doučevanje na ozkem domenskem korpusu
- Dražje, smiselno šele po jasni rabi
- **Evaluacija**
- Vnaprej določen nabor primerov in metrike

RAG

Retrieval-augmented generation

- Model išče v vaših dokumentih, preden odgovori
- **Kje gre pogosto narobe**
- Slabo razrezani dokumenti — odgovor ne najde konteksta
- Neevaluiran retrieval — vrne napačne dokumente
- Manjkajoča svežina — stara verzija pravilnika
- Nadzor dostopa: model vidi tudi tisto, česar uporabnik ne bi smel

HALUCINACIJE

Ko model trdi neresnico s prepričanjem

- **Tipični vzroki**
- Vprašanje izven podatkov v treningu
- Premalo konteksta v pozivu
- Prevod izmišljotin med jeziki
- **Kaj pomaga**
- Sklicevanje na vire (citiranje v izhodu)
- Strukturiran izhod (JSON s polji)
- Človeški pregled pri pravnih in mnenjskih izjavah

Kako se soočati s hitro spreminjajočim področjem

Vsakih 6 mesecev nov »najboljši« model — kako to upravljati

Ne lovite zadnjega modela. Lovite svoj problem in dovolj dober model zanj.

01

Locitev modela in aplikacije

Postavite arhitekturo tako, da model lahko zamenjate v nekaj dneh, ne v nekaj mesecih. Vmesnik (interface) je ločen od modela.

02

Avtomatska evaluacija, ki pa ni dovolj

Nabor 50–200 testnih primerov za vaš primer uporabe. Ko se pojavi nov model, ga lahko v eni uri primerjate s trenutnim — brez ugibanja.

03

Realna meritev vrednosti

Kateri kazalnik se izboljša, ko vključite UI? Čas obravnave, stopnja reševanja na prvi stopnji, NPS. Brez tega ne veste, ali je sploh smiselno menjati.

04

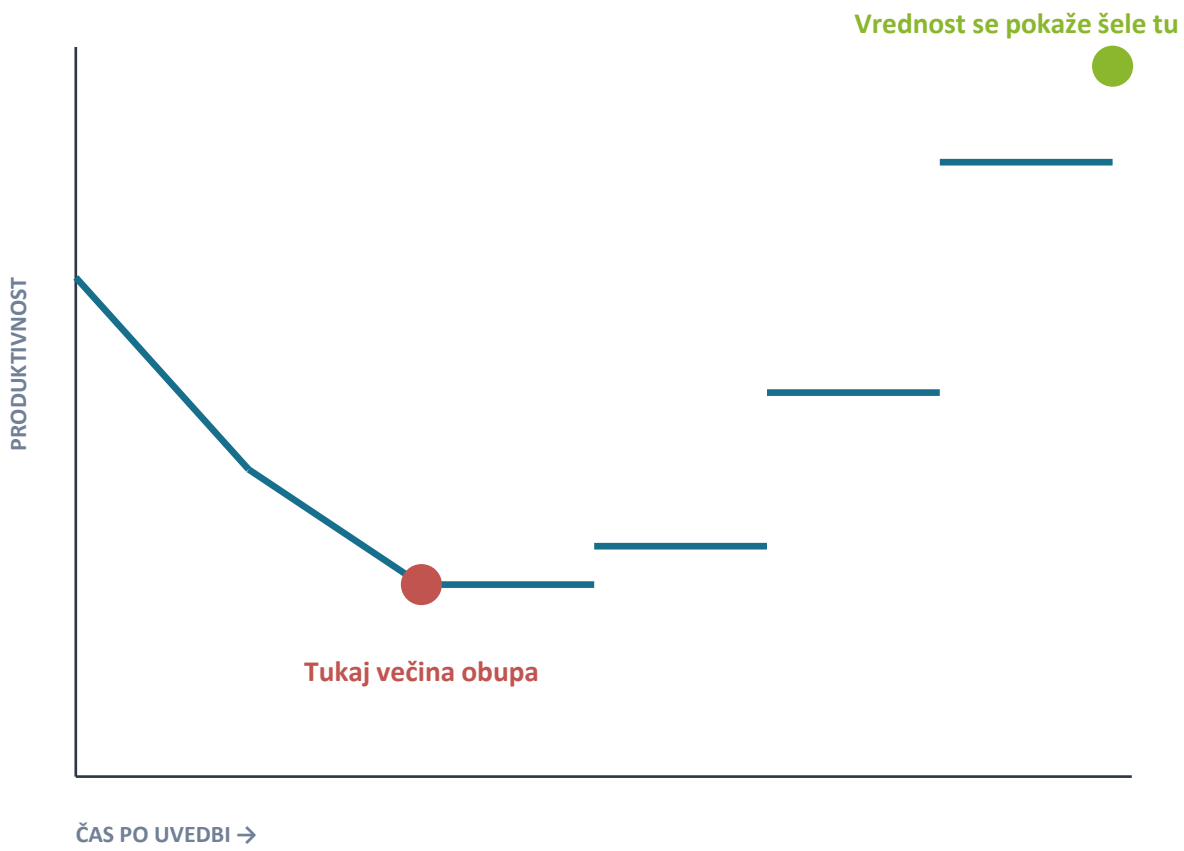
Spremljanje regulatornih sprememb

AI akt se uveljavlja postopno. Določite osebo, ki kvartalno pregleda novosti. Ne tehnik — pravnik s tehničnim razumevanjem.

Praktično: *če imate evaluacijski nabor in modularno arhitekturo, je »področje se hitro spreminja« vaša prednost, ne grožnja.*

Pričakovanja in realnost: J-krivulja uvedbe

Zakaj večina projektov zataji pri tretjem mesecu



Zakaj pride do padca

- Stari procesi se rušijo, novi še niso utečeni
- Zaposleni se učijo orodja in se mu upirajo
- Modeli se kalibrirajo za realno rabo, ne le za demo
- Vodstvo pričakuje takojšen donos, ki ga ni

Kako preprečiti odpoved

- Načelo 10 / 20 / 70: algoritmi 10 %, tehnologija 20 %, ljudje in procesi 70 %
- Začnite z enim primerom, ne s petimi naenkrat
- Vnaprej določen kazalnik uspeha — in vnaprej določen trenutek odločitve
- Komunikacija z zaposlenimi od prvega dne, ne ob krizi

Validacija LLM-rešitev: kaj kupujete, če ne merite

Stotine ponudnikov, podobne obljube, različne kakovosti — brez metodologije ne vidite razlik

Pomembno: LLM-i odgovore vedno podajo samozavestno — tudi ko so napačni. Brez evalvacijske metodologije ne morete vedeti, ali rešitev dela dobro, ali bi lahko delala bistveno bolje.

Zakaj brez evalvacije ne vidite razlik

- **Trg je preplavljen**
- Stotine ponudnikov, podobne obljube, isti modeli pod pokrovom — drugačni promptni okviri in kakovost.
- **Demo ≠ vaša raba**
- Prodajni primeri so izbrani; vaši dokumenti, jezik in robni primeri delujejo drugače.
- **Samozavest brez pravilnosti**
- LLM nikoli ne reče »ne vem«. 30 % napak v izhodu zveni enako kot 0 %.
- **Halo učinek**
- Prvi nekaj dobrih odgovorov zgradi zaupanje, ki ga slabši odgovori kasneje ne razgradijo.

Kaj zajema prava metodologija evalvacije

- **Vaš lasten testni nabor**
- 50–200 reprezentativnih primerov iz dejanske rabe — ne demo primeri ponudnika.
- **Vnaprej definirana merila uspeha**
- Kaj je »dovolj dobro«? Točnost, čas, robustnost — določiti pred testom, ne po njem.
- **Slepa primerjava ponudnikov**
- Anonimiziran izhod, da ocena ni pristranska zaradi blagovne znamke ali cene.
- **Robni primeri, ne le enostavni**
- Težki, dvoumni in redki primeri pokažejo, kdaj rešitev odpove.
- **Periodična ponovitev**
- Modeli se spreminjajo — evalvacija je proces, ne enkratna naloga.

Past: *brez metodologije se vsaka rešitev zdi v redu. To je v dolgoročnih stroških najdražja možna odločitev — plačujete za inferiorni rezultat in tega ne veste.*

Kaj odnesete v svoj MSP

Pet stvari, ki jih lahko naredite že jutri

1

Naredite popis UI v podjetju

Tabela: sistem, namen, ponudnik, podatki, ki gredo vanj. To je temelj za vse, kar sledi.

2

Vsakemu sistemu pripišite kategorijo tveganja

Minimalno / omejeno / visoko / nesprejemljivo. Klasifikacija je živa — vrnite se k njej vsakih 6 mesecev.

3

Pri klepetalnikih in sintetičnih medijih: razkrijte

Stavek »ta klepetalnik poganja UI« in oznaka na sintetičnih izvozih. Mali vložek, velik regulatorni prihranek.

4

Imejte evaluacijski nabor za vsak primer rabe

50–200 testnih primerov. Brez tega ne morete reči, ali je nov model boljši — niti regulatorju ne morete dokazati skrbnosti.

5

Določite eno osebo za »UI in skladnost«

Ne nujno polni delovni čas. Lahko obstoječi zaposleni — pravnik s tehničnim razumevanjem ali tehnik s pravnim posluhom.

Najpogostejši strahovi okoli umetne inteligence

Kaj skrbi družbo na splošno in kaj posebej MSP-jevca

Družbeni in osebni strahovi (globalno)

Izguba delovnih mest

UI prevzame opravila, ki so jih dosedaj delali ljudje — predvsem pisarniška in kognitivna.

Koncentracija moči

Najmočnejši modeli so v rokah peščice ameriških in kitajskih korporacij — odvisnost regij in držav od njih raste.

Manipulacija in dezinformacije

Deepfaki, sintetični mediji in personalizirana prepričevanja vplivajo na volitve, mnenja in zaupanje v informacije.

Erozija zasebnosti

Več podatkov v modelih → manj zasebnosti. Sledenje vedenja in profiliranje sta lažje kot kadar koli.

Eksistenčno tveganje

Skrb, da bo UI presegla človeški nadzor — fokus razprav v raziskovalnih in vladnih krogih.

Strahovi pri uvajanju v MSP

Uhajanje občutljivih podatkov

Naši dokumenti, e-pošta, podatki strank gredo k ponudniku in lahko v tretjo državo.

Napačne odločitve in odgovornost

Model halucinira, klepetalnik svetuje napačno, postane to pravna težava — kdo nosi posledice?

Investicija brez merljive vrednosti

Drago orodje, pilotni projekti se ne pretvorijo v produkcijo. Več kot 80 % projektov UI v podjetjih ne preživi pilota.

Demoralizacija zaposlenih

Strah pred zamenjavo, upor pri uvajanju, atrofija veščin pri rutinski uporabi UI.

Regulatorno breme

AI akt, GDPR, sektorska pravila — bo skladnost požrla več kot dobiček iz UI?

Šest prebojev iz zadnjega časa

Kaj je danes mogoče, kar lani še ni bilo

01 Agentni sistemi

UI, ki ne le odgovarja, ampak izvaja zaporedje korakov — planiranje, izvedba, popravki. Avtomatizacija celih procesov, ne le posameznih opravil.

02 Modeli, ki »razmišljajo«

Modeli si vzamejo čas in problem razdelajo po korakih. Rešujejo naloge na ravni matematike, programiranja in znanosti, ki so bile pred letom dni nedosegljive.

03 Multimodalni modeli

En model za besedilo, sliko, zvok in video. Zaposleni pošlje fotografijo pokvarjenega stroja — model neposredno predlaga ukrep.

04 UI v znanosti

Marca 2026 je Sakana AI Scientist samostojno generiral hipotezo, izvedel poskus in napisal članek, ki je prestal recenzijo (Nature). Začetek »UI raziskovalca v ekipi«.

05 Manjši in lokalni modeli

Odprto-utežni modeli na prenosniku dosegajo zmogljivosti, ki so bile lani v dosegu samo največjih. Kakovosten LLM na svojem strežniku za delček cene.

06 UI za razvoj kode

Naloge, ki so bile dosedaj programiranje, opravljene z opisom v naravnem jeziku. Razvoj v urah, ne tednih. Velika prerazporeditev v IT sektorju.

Vzorec: *največje pridobitve niso v novih »funkcijah«, ampak v tem, da UI prehaja iz orodja za en odgovor v sodelavca, ki opravi celo nalogo.*

Zaključek

UI je orodje — odgovornost ostaja na uporabniku

Zavedanje potenciala: UI pospešuje produktivnost, učenje in inovacije – a odgovorno in preudarno uvajanje je ključno

UI je zelo uporabno orodje - zato se je razširilo — tudi neustrezna raba.

01

Ustrezno kontinuirano izobraževanje

Vpliv na razvoj posameznika: nove veščine (promptiranje, verifikacija, razlaga), kritično razmišljanje /presoja, partnerstvo človek+AI, etična presoja

Razumevanje lastnosti, klasifikacija podatkov, izbira pravilne rešitve.

02

UI kot konkurenčna prednost

Vpliv na organizacije: konkurenčna prednost iz discipliniranega upravljanja AI (standardizacija, merjenje, skladnost, varnost), evalvacija GenAI (LLM) sistemov je še v veliki meri neustrezna
.

Razumevanje, na kakšen način je potrebno uvajati UI, načelo odgovornosti – poslovni učinki

03

UI kot strateška geopolitična dobrina

Vpliv na državo: strateška suverenost, dvig digitalnih kompetenc, pospeševanje gospodarstva; konkurenčna prednost iz zgodnjega sprejema standardov in dobre prakse.

Razvijanje EU infrastrukture, EU modelov

Razumeti lastnosti. Klasificirati podatke. Izbrati pravilno rešitev. Beležiti odločitve. Ustrezna evalvacija. Način uvajanja je ključen.

Hvala za udeležbo!

Maja.Skrjanc@ijs.si

SLAIF
Slovenska tovarna UI
Slovenian AI Factory

IZUM
Institute of Information Science, Maribor, Slovenia

arnes

University of Maribor

Univerza v Novi Gorici
Scientia Vincet 1995

Gospodarska zbornica Slovenije
Chamber of Commerce and Industry of Slovenia

Jožef Stefan Institute
Ljubljana, Slovenia

UNIVERZA V LJUBLJANI
University of Ljubljana

FIS
Fakulteta za informacijske študije
Faculty of information studies

Univerzita del Litorale
Università del Litorale
Univerza na Primorskem

TP LJ01

Financerja / Financed by:

REPUBLIKA SLOVENIJA
MINISTRSTVO ZA VISOKO ŠOLSTVO, ZNANOST IN INOVACIJE

EuroHPC
Joint Undertaking

Projekt SLAIF: Slovenska tovarna umetne inteligence je finančno podprlo Ministrstvo za visoko šolstvo, znanost in inovacije. Projekt je bil na razpisu skupnega podjetja EuroHPC izbran za financiranje v okviru programov Obzorje Evropa ter Digitalna Evropa.

SLAIF: Slovenian AI Factory has been funded by the Ministry of Higher Education, Science and Innovation of Republic of Slovenia. At a call by EuroHPC JU, the project has received a positive funding decision under Horizon Europe and Digital Europe Programmes.