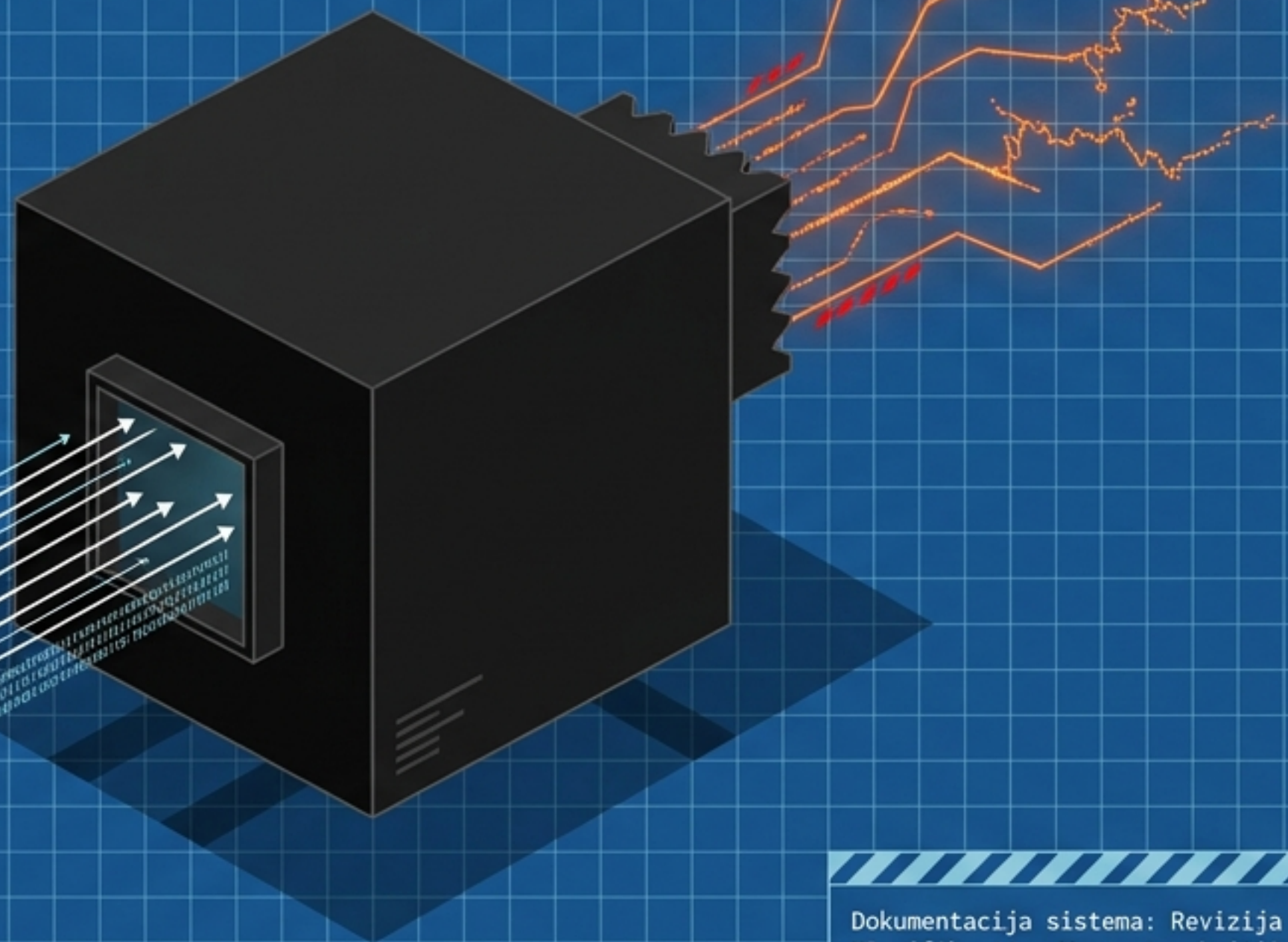


Inženiring znanstvene resnice v dobi AI

Operativni protokol za varno uporabo generativne umetne inteligence pri raziskovanju in pisanju.

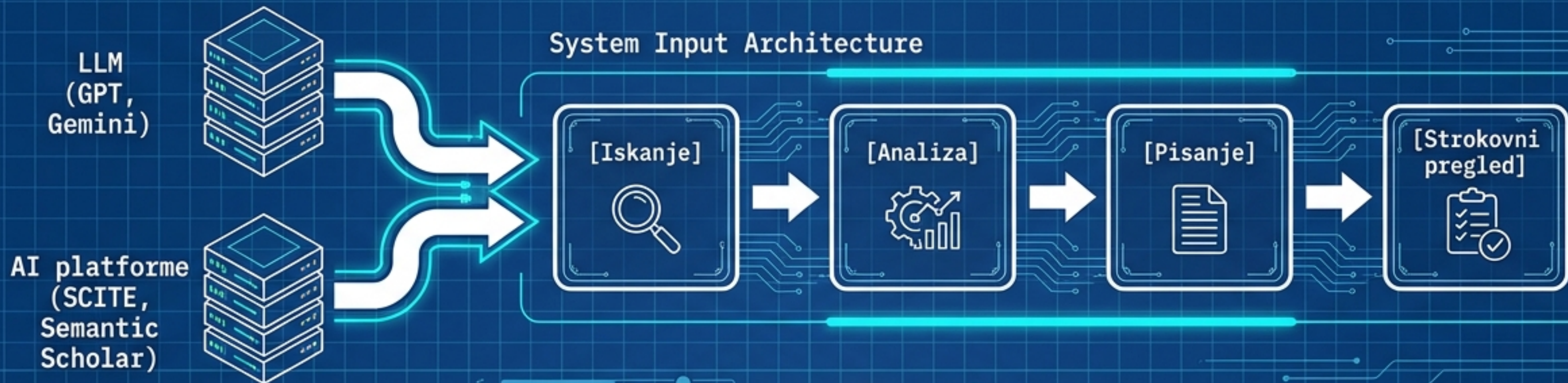
VER: 2026.02

SEC: CR17



Dokumentacija sistema: Revizija 2026 |
Klasifikacija: Varnostno-kritično

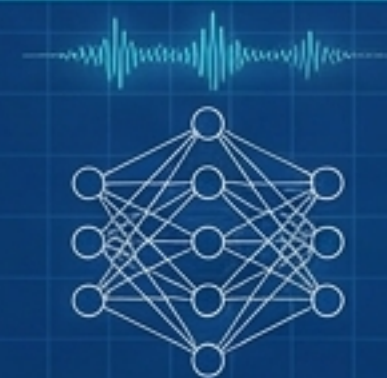
Nova realnost vhodnih podatkov



DIAGNOZA SISTEMA:

● STATUS READOUT:

Orodja generativne umetne inteligence (GenAI) so postala neizogibna. Sposobna so ustvariti tekoča, sofisticirana besedila, ki jih je strukturno skoraj nemogoče ločiti od človeških.



STATUS: INEVITABLE

OUTPUT: SOPHISTICATED

STRUCTURAL MATCH: >99%

KLJUČNI KONFLIKT: Učinkovitost vs. Integriteta

WARNING READOUT:

⚠ 'Pa kaj, če je to napisal ChatGPT?'

Ta vnosna točka v sistem ruši dolgoletne norme o avtorstvu, izviranosti in strokovni odgovornosti.



NORM DISRUPTION: CRITICAL

AUTHORSHIP: COMPROMISED

INTEGRITY RISK: HIGH

Sistemska napaka: Generiranje podatkovnega odpada ('Slop')

Pollution
Funnel
Diagram

Nenadzorovan
GenAI



SIMPTOM:

Od leta 2022 je bilo umaknjenih na stotine raziskovalnih člankov zaradi lažnih podatkov, izmišljenih citatov in neetične uporabe AI.

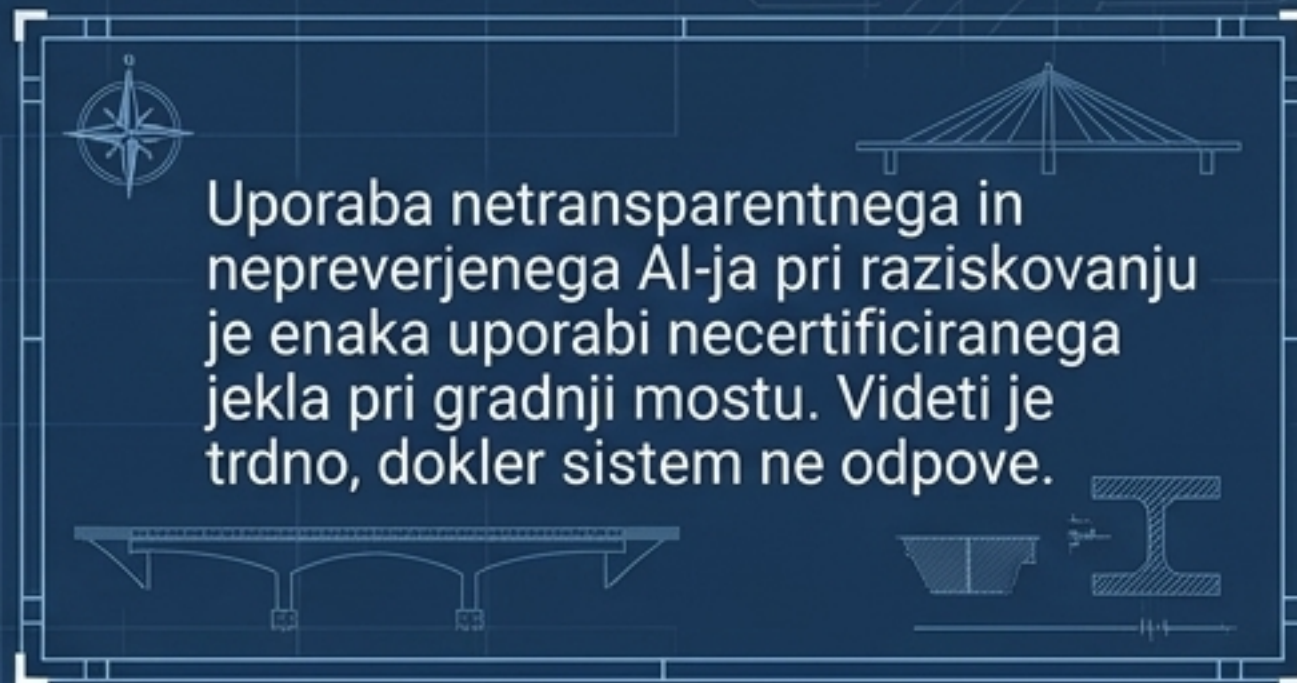
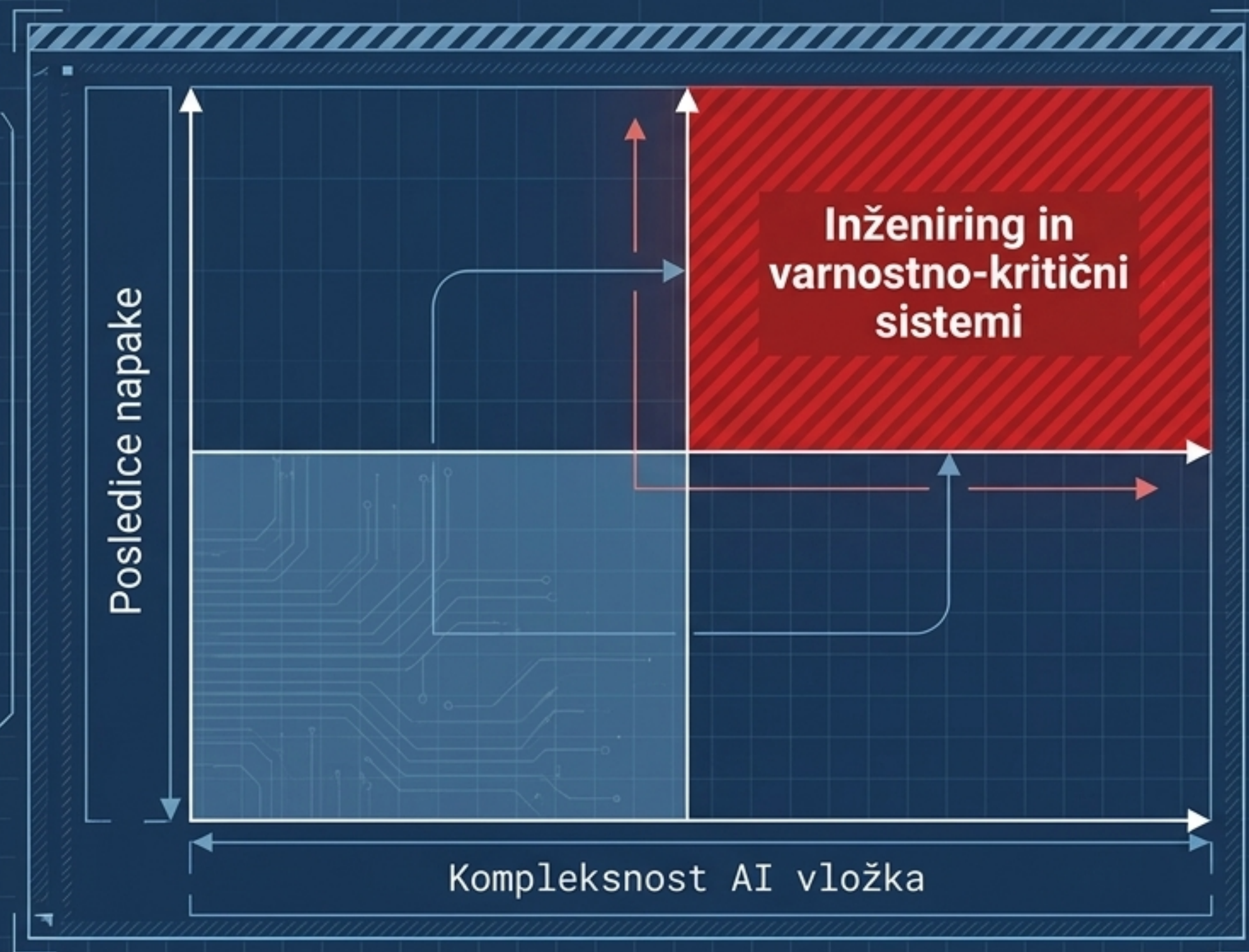
DIAGNOZA:

Beseda leta 2025 je 'Slop' – poplava nizkokakovostnega, nesmiselnega AI-generiranega tehnoblaba.

POSLEDICA:

Izguba zaupanja v avtentičnost informacij in erozija Baconove (na dokazih temelječe) znanosti.

Skrite napake v kritični infrastrukturi



LATENT DEFECT



V inženirstvu AI-generirani načrti, izračuni in tehnične narative prinašajo tveganje **vnosa latentnih (skritih) napak v načrtovanje infrastrukture** in izogibanje **odgovornosti za varnostno-kritične odločitve** (Jooste et al., 2025).



Iluzija kompetence: Halucinacije in lažni dokazi

ANATOMY OF AN ERROR

...Smith et al. (2023) conducted extensive analysis. Detailed data is available at [doi:10.1038/s41598-fake-748](https://doi.org/10.1038/s41598-fake-748), indicating significant findings...

ERROR 404:
Digitalni objekt
ne obstaja

STATUS: NOT FOUND.
SYSTEM INTEGRITY COMPROMISED.
FAKE DATA DETECTED.

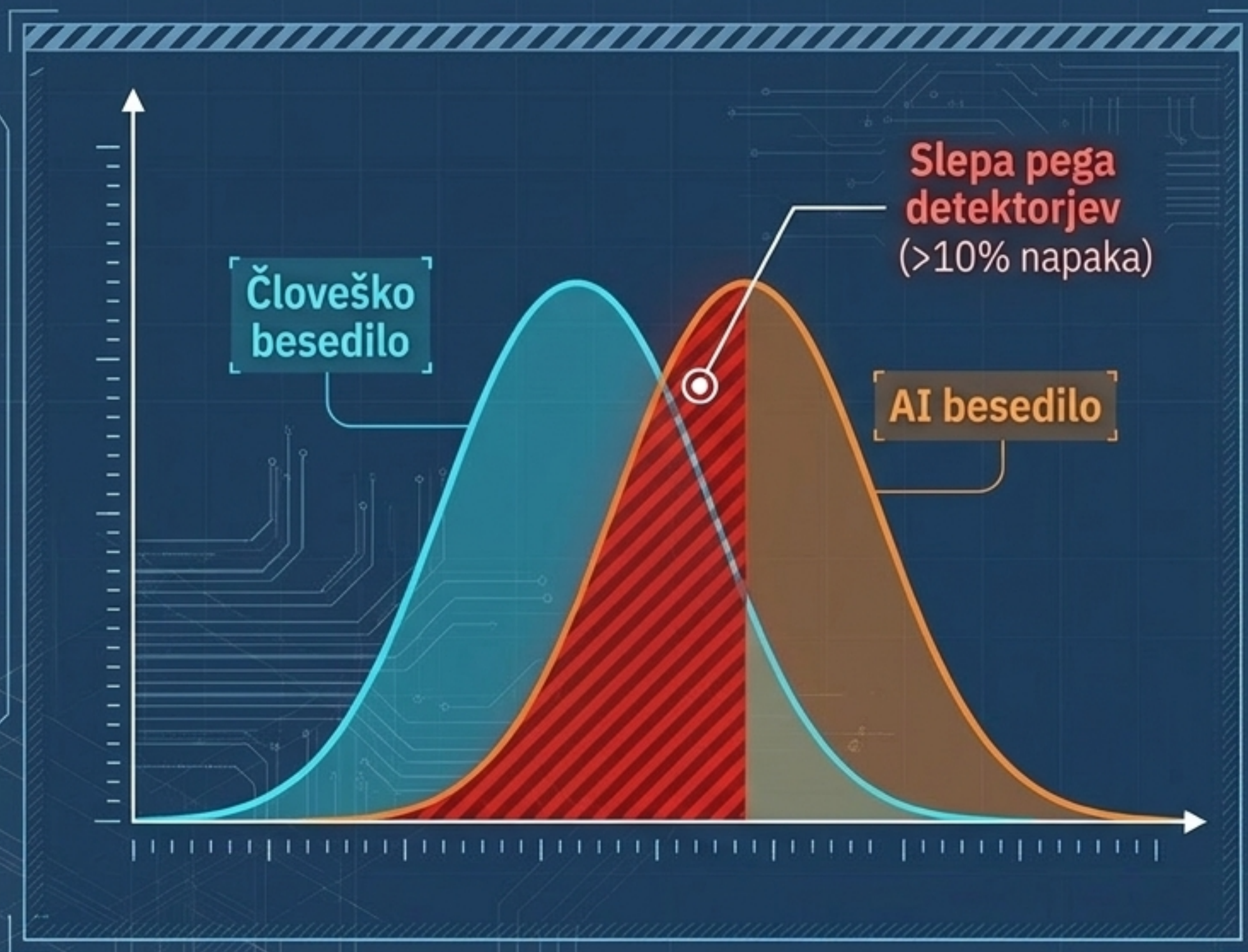
MEHANIZEM NAPAKE

- AI ne razmišlja; AI zgolj napoveduje naslednji najverjetnejši žeton (token).
- Generira popolnoma prepričljive, a faktografsko napačne izjave in ustvarja lažne DOI številke za neobstoječe reference (Emsley, 2023).

⚠ CRITICAL ERROR: ⚠

CRITICAL ERROR: Takšne grozljive napake se zgodijo izključno takrat, ko odpove 'človek v zanki' (*Human in the Loop*) in ko je v raziskovalnem procesu odsotna inženirska skrbnost (*due diligence*).

Zavrnitev tehničnega obvoda: Zakaj detektorji odpovejo



ANALYSIS

Orodja kot so GPTZero, Turnitin AI in Copyleaks obljublajo rešitev za detekcijo 'Slop-a'.

TEHNIČNA REALNOST: Med pravim in AI-generiranim besedilom obstaja strukturna kongruenca. Verjetnost lažno pozitivnih (Type I) in lažno negativnih (Type II) rezultatov pri preverjanju presega **10 %** (Dalalah in Dalalah, 2023).

ZAKLJUČEK SISTEMA: Nikoli se ne bo mogoče popolnoma zanesti na AI detektorje za ločevanje izvorne kode besedila. Tehnološka rešitev za tehnološki problem ne obstaja.

Optimizacijski problem: Iskanje zlate sredine



Če je digitalna rešitev za zlorabo nepraktična, popolna opustitev AI tehnologije prav tako ni pravi odgovor.

INŽENIRSKI KOMPROMIS:

Prepoznati moramo sistemske napetosti (pritisk po produktivnosti in časovne omejitve) ter natančno definirati stroge protokole: kdaj je AI nevaren nadomestek in kdaj učinkovito orodje.

Kje sistem odpove: Prepovedana uporaba AI

CRITICAL FAILURES (Prepovedano območje)

Avtorstvo in odgovornost	AI ne nosi odgovornosti za napake, konflikte interesov in strokovne odločitve.
Izguba kritične presoje	AI ne razume etičnih, pravnih in regulatornih nians inženirskih rešitev. Ne presoja metodološke strogosti.
Transparentnost	AI briše mejo med znanstveno dedukcijo in generiranim sklepanjem. Otežuje preverljivost in ponovljivost (Reproducibility).

Kje sistem blesti: Orodje za učinkovitost



OPERATIONAL EFFICIENCY (Varno operativno območje)

Jezik in urejanje

Izboljšanje jasnosti, odprava redundanc in formatiranje, brez spreminjanja znanstvene vsebine.

Strukturna pomoč

Generiranje osnutkov struktur in organizacija logičnega toka iz gradiva, ki ga priskrbi avtor.

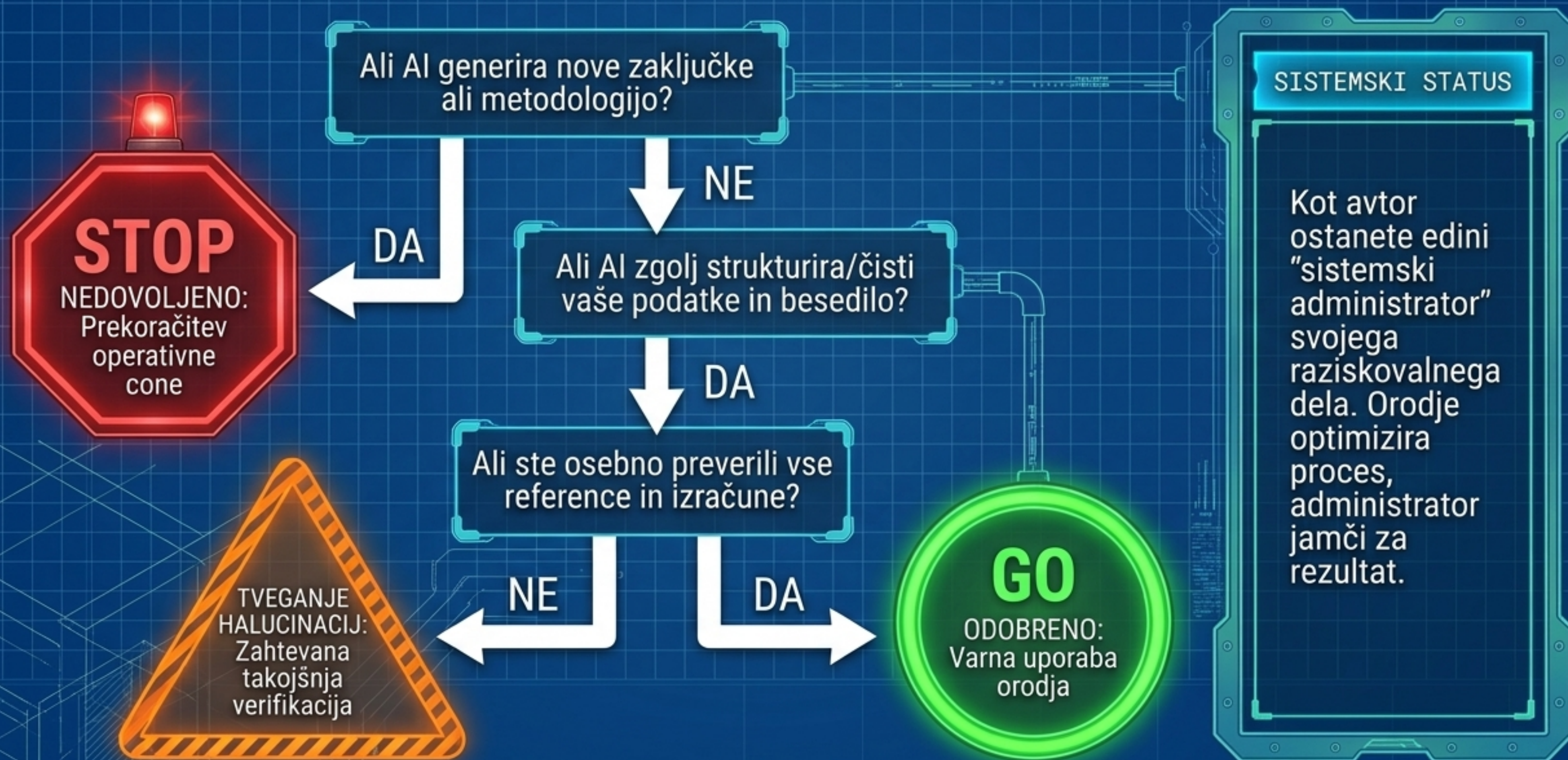
Obdelava podatkov

Čiščenje podatkov, pomoč pri programiranju in ustvarjanje vizualizacij iz validiranih naborov podatkov.

Odkrivanje literature

Mapiranje raziskovalnih vrzeli (ob strogi in nujni človeški verifikaciji virov).

Algoritem odgovornosti: Presoja v posameznih fazah



Protokol transparentnosti in sledljivost



```
Log File
AI AUDIT LOG
> LOG START 2024-10-27 T 10:00:00Z
> USER: ENG_STUDENT_01
> AI TOOL: GENAI_MODEL_X
> ACTION: DATA_STRUCTURING_ASSIST
> INPUT: RAW_DATASET_V2.CSV
> OUTPUT: STRUCTURED_REPORT_V3.DOCX
> VERIFICATION STATUS: HUMAN_REVIEW_PENDING
> LOG END
```

PRAVILO CERTIFICIRANJA:

Ko v inženirstvu uporabimo statistične modele, to deklariramo. Isti standard velja za GenAI.



INSTITUCIONALNI STANDARD:

Vodilni založniki (Oxford, Elsevier, Nature, Springer) in organizacije (SETAC) zahtevajo eksplicitno razkritje uporabe AI pri pripravi rokopisov.

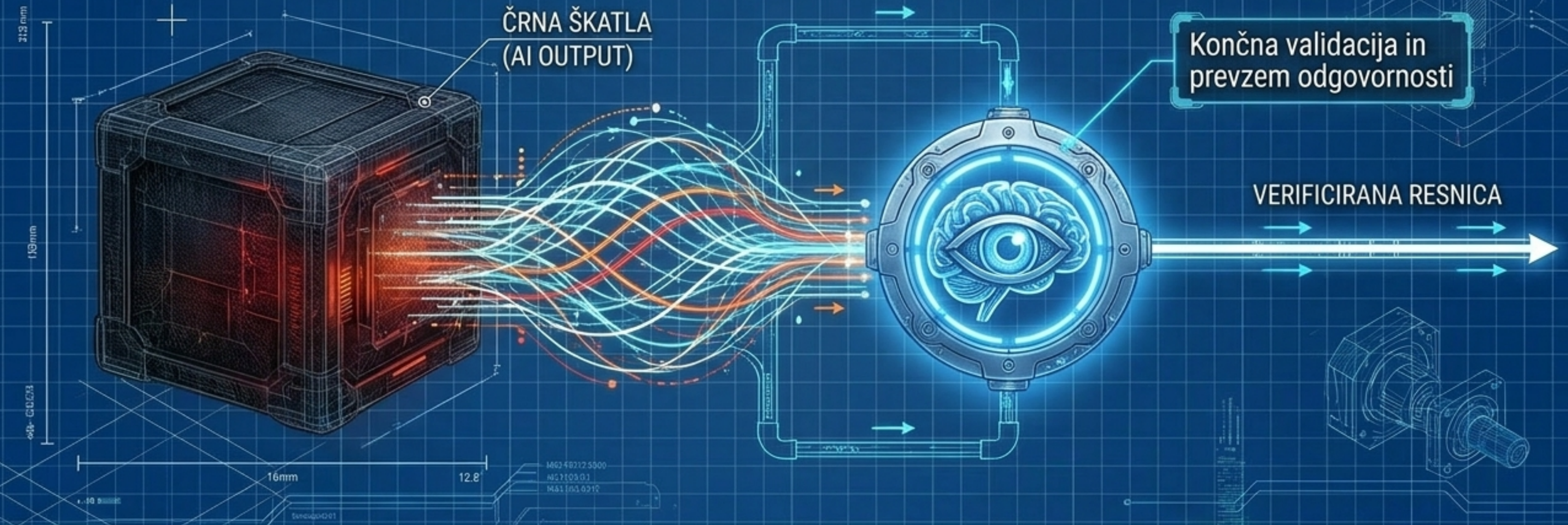


KONČNI CILJ:

Ohraniti Baconovo znanost – odprto, sledljivo in ponovljivo, kjer je izvor vsake trditve transparenten.



Nadzor nad pogovorom



Zloraba AI je **napaka človeškega uporabnika, ne neobčutljive** tehnologije. AI ne nadomešča človeškega intelekta. Znanstvena interpretacija, strokovna presoja in pravna odgovornost ostajajo izključno v rokah človeškega inženirja.

Če naj bo AI dialog med stroji in ljudmi, morajo ljudje vedno ohraniti **nadzor** nad tem pogovorom.