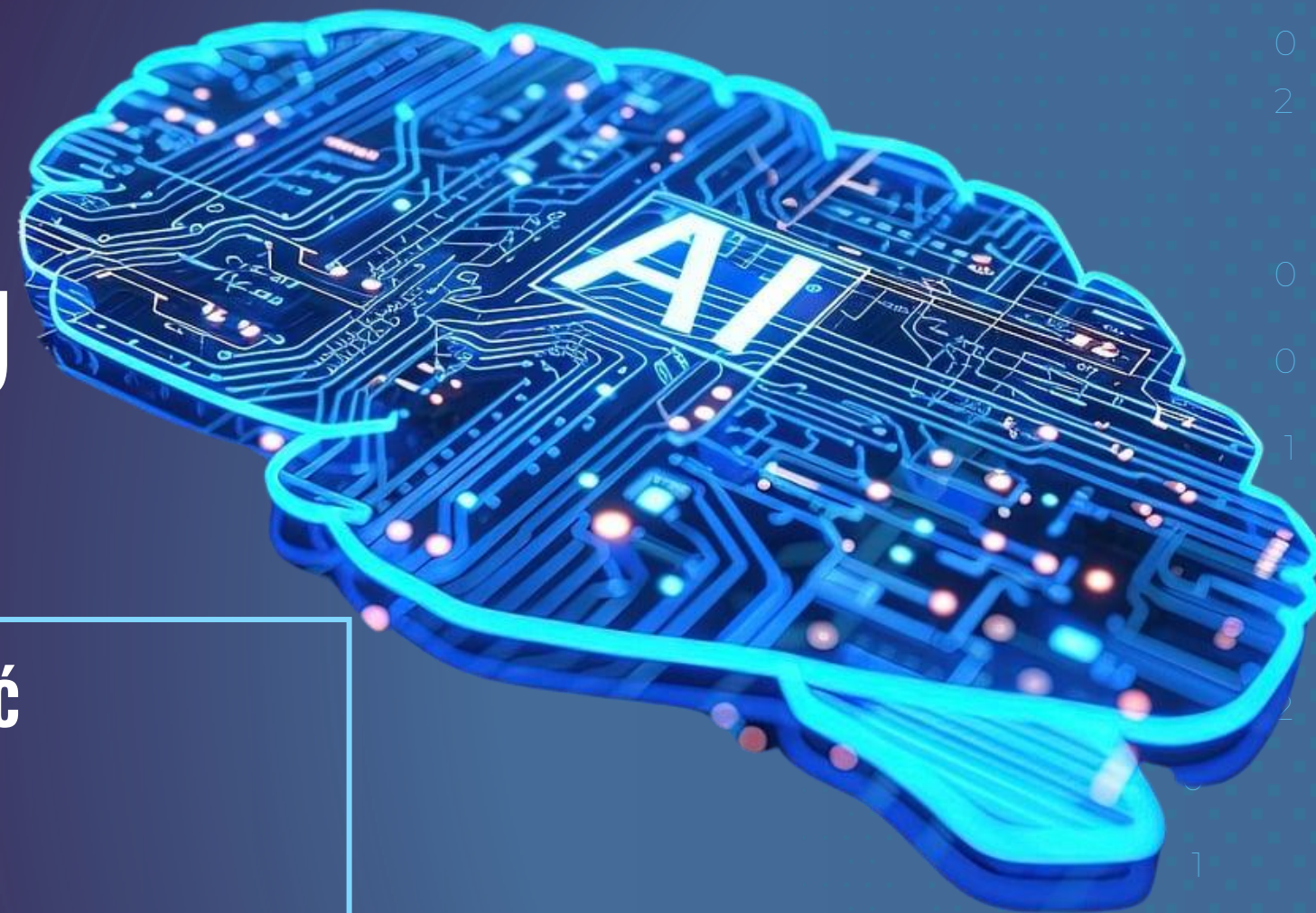


# UČINKOVITOST UPORABE GENERATIVNE UI PRI PREUČEVANJU ZGODOVINSKIH DOKUMENTOV

DR. DAVID HAZEMALI, JANEZ OSOJNIK, DR. TOMAŽ ONIČ, DR. TADEJ TODOROVIČ  
FF UM  
DR. MLADEN BOROVICH  
FERI UM



MEDNARODNA KONFERENCA

TEHNIČNI IN VSEBINSKI PROBLEMI KLASIČNEGA IN  
ELEKTRONSKEGA ARHIVIRANJA

21. do 23. maj 2025, Hotel Radin, Radenci

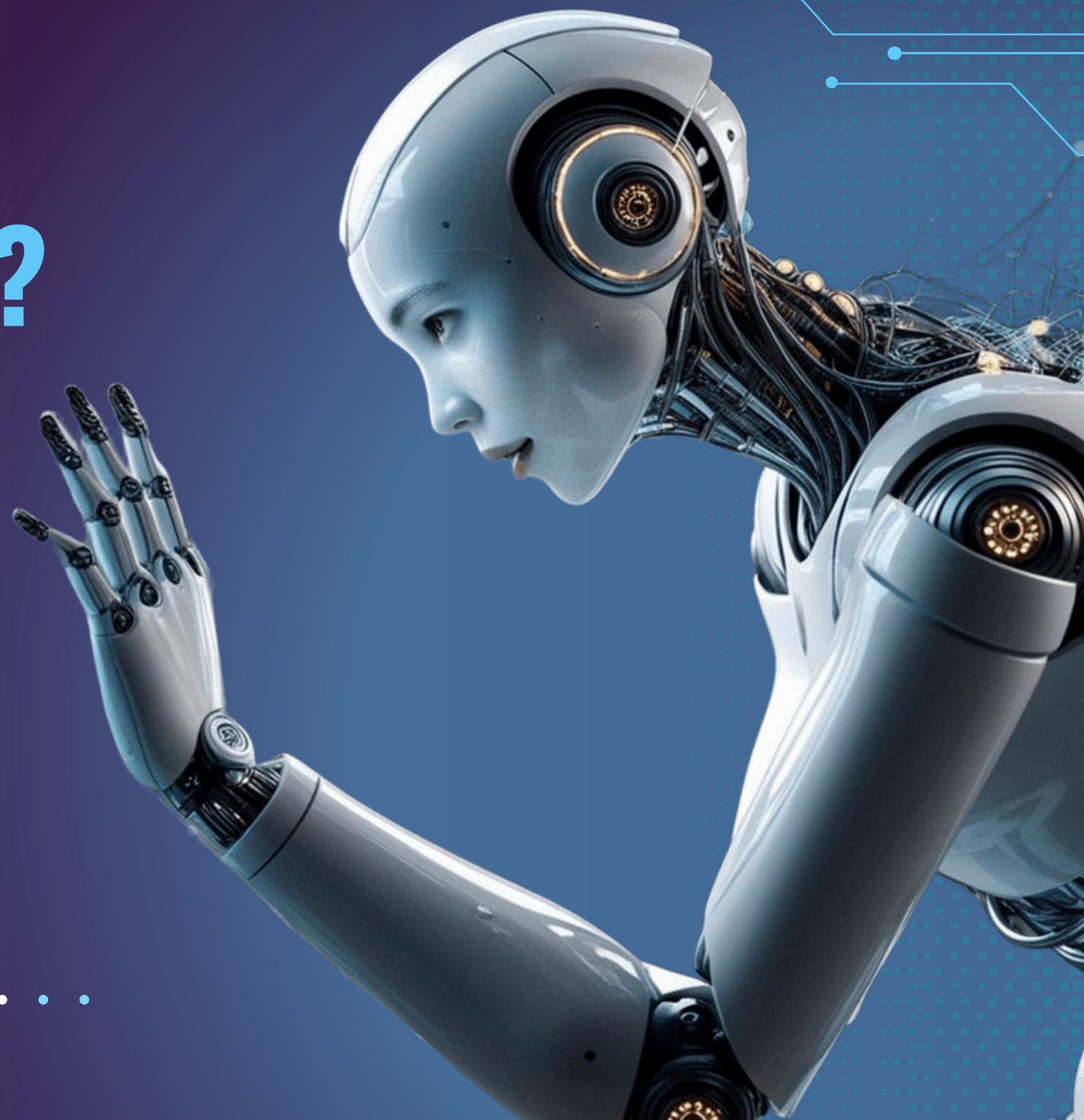
# KAJ JE POGOVORNI SISTEM (CHATBOT)?

V najširšem smislu gre za sistem, ki je zmožen v naravnem jeziku voditi dialog z uporabnikom.

Prve rešitve, kot je ELIZA (1966), so bile zbirke pravil, ki so prepoznavale par ključnih besed in vrnilo vnaprej napisan stavek.

Sodobni chatboti pa odgovore generirajo sproti: nimajo vnaprej pripravljenega sklopa stavkov, temveč jih sestavijo iz nič, glede na uporabnikov poziv in naučene vzorce jezika.

Odgovore črpa iz velikih jezikovnih modelov.



# KAJ JE VELIKI JEZIKOVNI MODEL (LLM)?

LLM lahko opišemo kot gigantsko nevronske mreže, naučeno na stotinah milijard besed.

Ko prejme besedilo, ga pretvori v številčne vektorje, nato pa z verjetnostnim algoritmom tipa »naslednja beseda« generira odgovor.

Zaradi obsega podatkov zna povzemati knjige, prevajati med jeziki in pisati – a ker gre za statistiko, ne logiko, so napačni ali izmišljeni podatki (halucinacije) neizogibni.

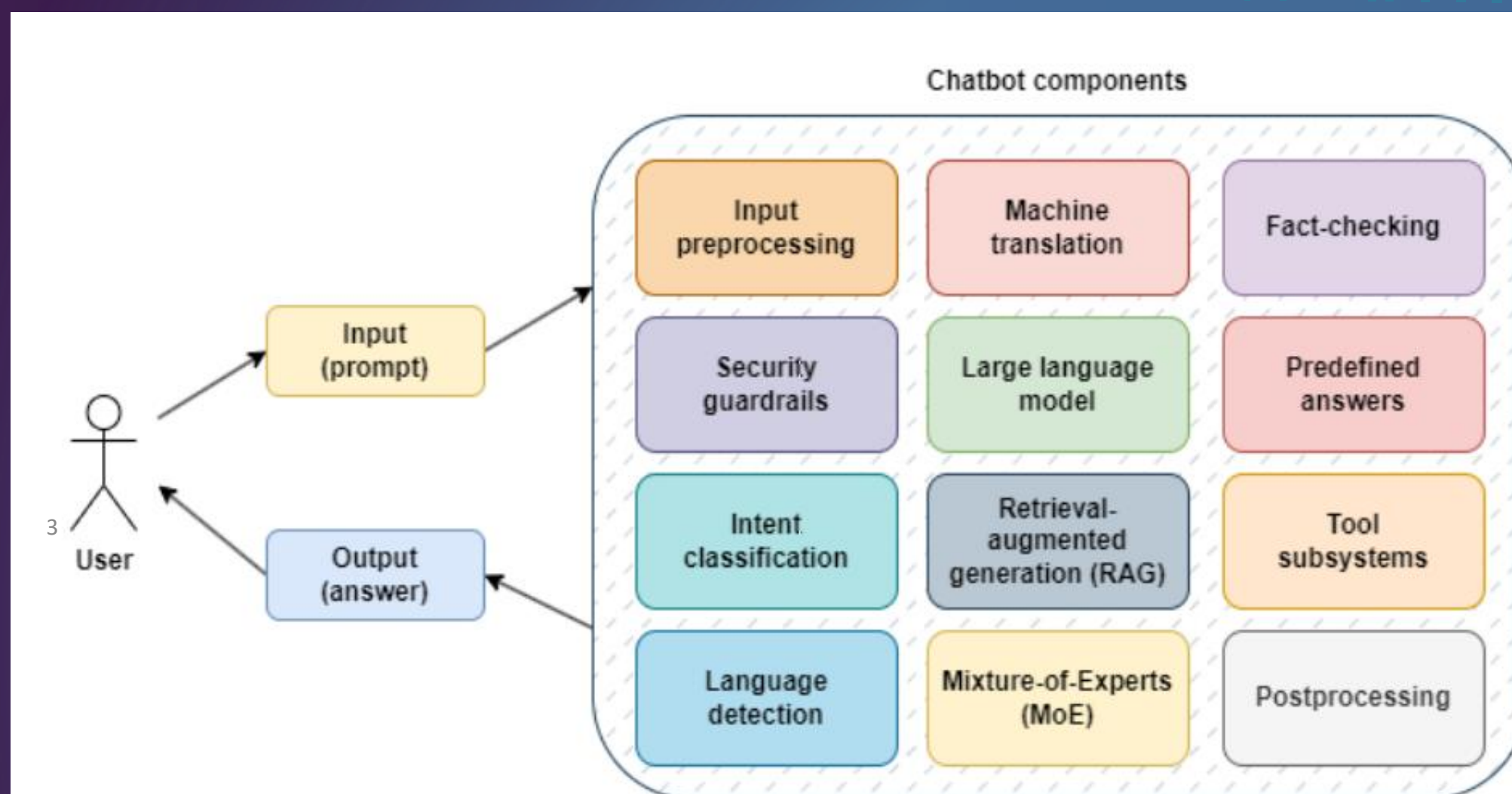


Figure 1: Various components in a general modern chatbot architecture.

# CHATBOT ≠ LLM: V ČEM JE RAZLIKA?

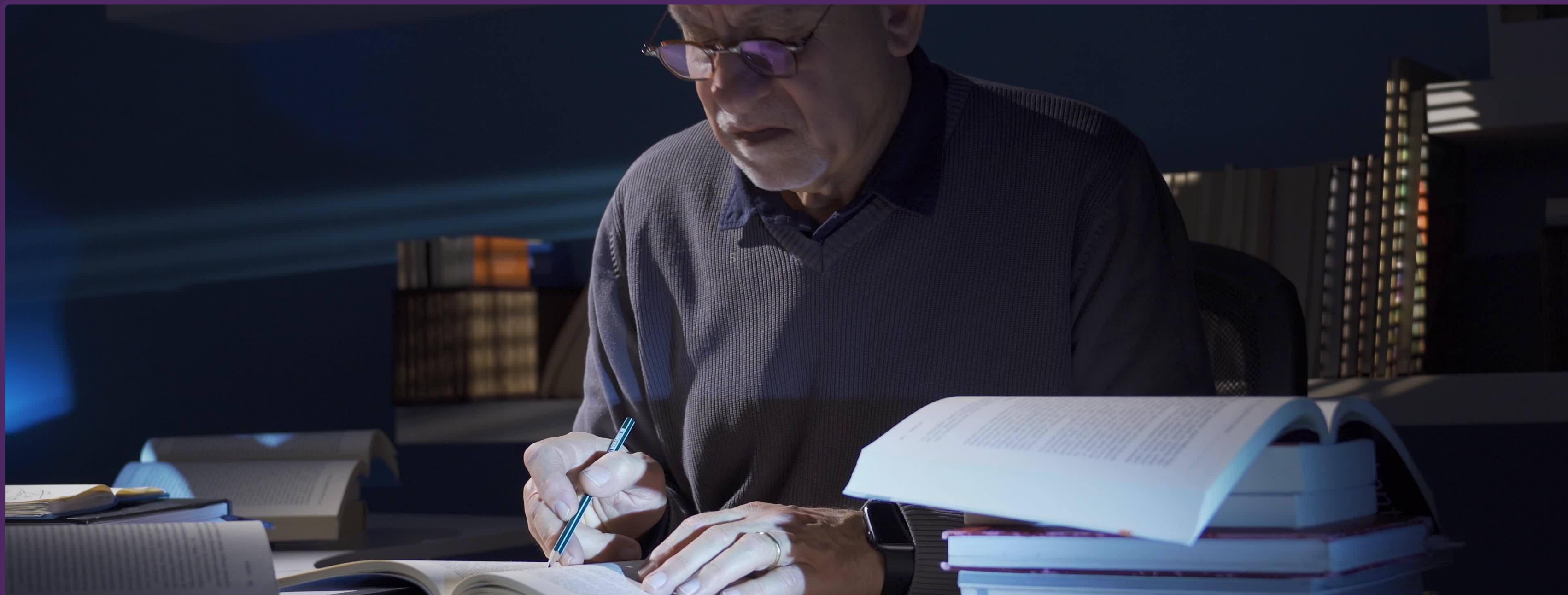
## LLM VELIK JEZIKOVNI MODEL

- Jedro – ogromna nevronska mreža (milijarde parametrov).
- Naučen na masivnem korpusu besedil; njegova naloga je generirati besedilo (napoved “naslednje besede”).
- Nima vgrajenega uporabniškega vmesnika, spomina pogovora, gumbov za spletno iskanje itd.
- Primer: GPT-4o, Claude Opus, Llama 3 ... – zgolj “motor”, ki ga lahko priključimo kamorkoli.

## CHATBOT POGOVORNI SISTEM

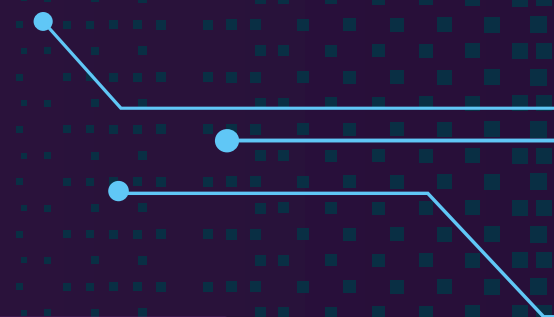
- Izdelek za končnega uporabnika; vključuje LLM + aplikacijski sloj.
- Dodane funkcije:
  - pogovorna zgodovina,
  - oblikovanje odgovorov (tone, slog),
  - orodja (spletno iskanje, branje PDF-jev, ustvarjanje slik ...),
  - pravila varnosti, cenzure, osebne nastavitve.
- Chatbot lahko poganja enega ali več LLM-jev – ali pa jih menja glede na nalogo.

# KATERE SO TRADICIONALNE FAZE DELA ZGODOVINARJA?

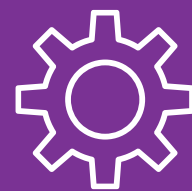


# KATERE SO TRADICIONALNE FAZE DELA ZGODOVINARJA?

ZGODOVINAR ZAČNE Z RAZISKOVALNIM VPRAŠANJEM, NATO POTUJE V ARHIV/KNJIŽNICO: IŠČE LISTINE, MAGNETOGRAME, FOTOGRAFIJE IPD. DOKUMENTE KRITIČNO OVREDNOTI (AVTOR, ČAS, NAMEN IPD.), JIH PRIMERJA Z DRUGIMI VIRI IN IZ NJIH IZPELJE ARGUMENT. TA ARGUMENT NATO V ZNANSTVENEM JEZIKU PODKREPI S CITATI IN OBJAVI V RECENZIRANI REViji/KOT MONOGRAFIJO (...). SREDIŠČE DELA OSTAJA KRITIČNA INTERPRETACIJA – TEHNOLOGIJA JE BILA DOSLEJ LE PODPORA PRI MEHANSKEM DELU.



## TRADICIONALNE FAZE DELA ZGODOVINARJA



**FAZA 1: ISKANJE IN ZBIRANJE VIROV**



**FAZA 2: ANALIZA IN VREDNOSTNA  
SODBA VIROV**



**FAZA 3: PISANJE IN PRIPRAVA  
KONČNEGA SESTAVKA ZA OBJAVO**

# POVZETEK KORISTI IN OMEJITEV

## KORISTI

- hitrejša iskanje virov
- hitrejši pretres zbranega gradiva
- takojšnja pomoč pri prevajanju
- ustrezen kvaliteten osnutek besedila
- avtomatski grafi in druge grafike
- več časa za poglobljeno interpretacijo

## OMEJITVE

- orodje lahko "zbluzi" (halucinira)
- ima vgrajene pristranosti in ne razume moralnega konteksta
- končno presojo in interpretacijo še vedno naredi človek
- naše podatke obdelujejo, morebiti celo prodajajo neznane korporacije/osebe

# Moderna arhivistika

Časopis arhivske teorije in prakse  
Journal of Archival Theory and Practice

ISSN 2591-0884

<https://doi.org/10.54356/MA>

Moderna arhivistika, VII, 2024 št. 2

Prejeto / Received: 10. 10. 2024

1.01 Izvirni znanstveni članek

1.01 Scientific article

<https://doi.org/10.54356/MA/2024/BIUB3010>

## EVALUATING CHATBOT ASSISTANCE IN HISTORICAL DOCUMENT ANALYSIS

asist. dr. David HAZEMALI  
University of Maribor, Faculty of Arts,  
Slovenia  
[david.hazemali@um.si](mailto:david.hazemali@um.si)

izr. prof. dr. Tomaž ONIČ  
University of Maribor, Faculty of Arts,  
Slovenia  
[tomaz.onic@um.si](mailto:tomaz.onic@um.si)

asist. Janez OSOJNIK  
University of Maribor, Faculty of Arts,  
Slovenia  
[janez.osojnik1@um.si](mailto:janez.osojnik1@um.si)

asist. dr. Tadej TODORVIČ  
University of Maribor, Faculty of Arts,  
Slovenia  
[tadej.todorovic@um.si](mailto:tadej.todorovic@um.si)

asist. dr. Mladen BOROVIČ  
University of Maribor, Faculty of Electrical Engineering and Computer Science, Slovenia  
[mladen.borovic@um.si](mailto:mladen.borovic@um.si)

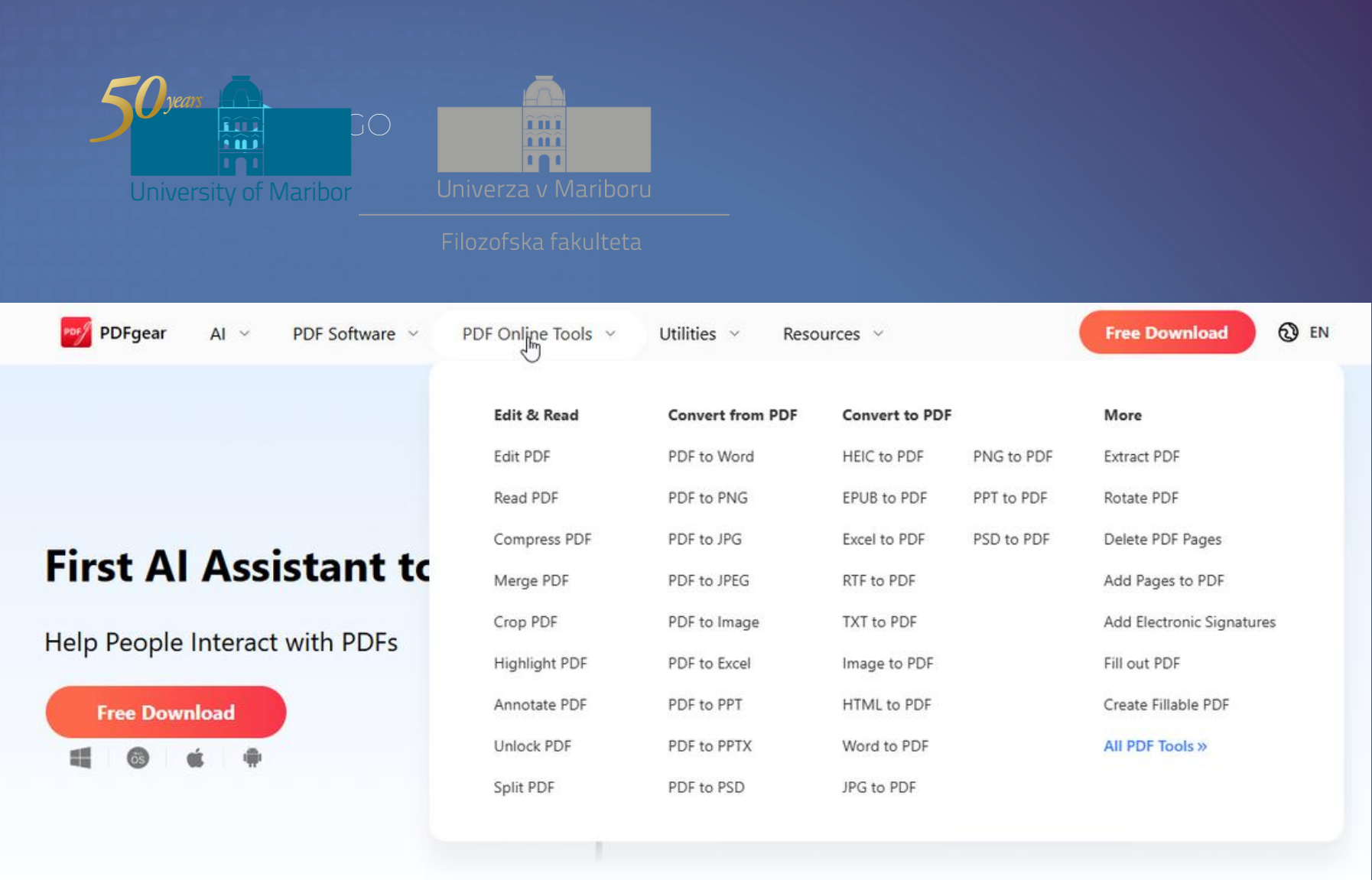
# NAŠ PRISPEVEK

Vrednotenje pogovornega sistema kot pomočnika pri analizi  
zgodovinskih dokumentov

- Namen: preučiti potencial pogovornega sistema PDFGear Copilot kot pomočnika pri analizi zgodovinskih dokumentov.

Raziskovalna vprašanja:

- V kolikšni meri lahko chatbot natančno prepozna, izlušči in povzema faktografske podatke o zgodovinskem dokumentu in njegovi vsebini?
- Kako uspešno se chatbot spopade s kompleksnimi zgodovinskimi vprašanji, ki zahtevajo interpretacijo in kontekstualizacijo v več jezikih?
- Kako dobro chatbot obvlada (namenoma vnesene) jezikovne napake (tipkarske, črkovalne, slovnične), domensko terminologijo in zapletene slovenske stavke?

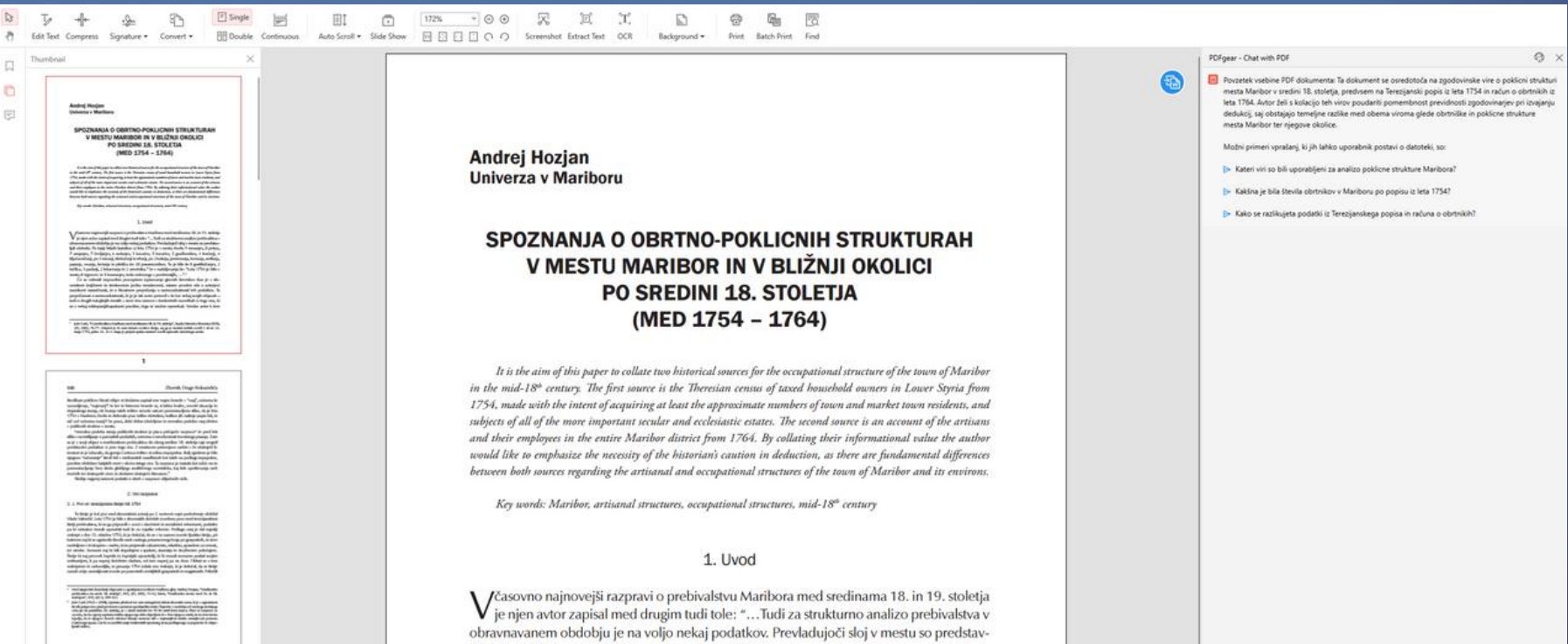


# KAJ JE TO PDFGEAR COPILOT?

- Gre za komercialno dostopen chatbot, ki je integriran v urejevalnik PDF datotek. Gre za nekakšnega »digitalnega sodelavca« za delo s PDF-ji, ki uporabniku omogoča, da z naravnim jezikom podaja navodila ali vprašanja v zvezi z odprtim PDF dokumentom.

10

- Omogoča lokalno namestitev.



- Pomembno: PDFGear Copilot sam po sebi ne ve nič o zgodovini ali zunanjem svetu – vse, kar lahko odgovarja, temelji na informacijah znotraj podanega PDF dokumenta (in splošnih jezikovnih vzorcih modela).



Dok. št. 27:<sup>149</sup>

**Neavtoriziran magnetogram razgovora predsednika predsedstva Republike Slovenije Milana Kučana in predsednika Republike Hrvaške dr. Franja Tuđmana z delegacijo Evropske skupnosti, Zagreb 28./29. 6. 1991 od 23,30 do 01,15**

NEAUTORIZIRAN MAGNETOGRAM RAZGOVORA PREDSEDNIKA PREDSEDSTVA REPUBLIKE SLOVENIJE MILANA KUČANA IN PREDSEDNIKA REPUBLIKE HRVAŠKE dr. FRANJA TUĐMANA Z DELEGACIJO EVROPSKE SKUPNOSTI<sup>150</sup> ZAGREB 28./29. 6. 1991 od 23,30 do 01,15

svojih dveh kolegov iz Evropske dvanajsterice, da ste nas sprejeli tukaj v Zagrebu na tako kratek rok. Še posebej bi se rad zahvalil gospodu Kučanu, ki je opravil to težavno pot, da se je lahko sestal z evropsko trojico. Rad bi tudi izrazil obžalovanje za vse človeške žrtve, ki so padle v teh zadnjih dneh na področju Slovenije in Hrvaške.

Evropska skupnost je izredno zaskrbljena zaradi te situacije zato tudi Evropski svet, ki se danes in jutri sestaja v Luxemburgu razpravlja o tej situacije. Evropski svet misli, da dogodki, ki so se zgodili v Sloveniji in na Hrvaškem, predstavljajo grožnjo ne le državljanom Jugoslavije predvsem pa slovenskemu in hrvaškemu narodu, kakor tudi ogrožajo stabilnost celotnega kontinenta Evrope čigar del ste tudi vi. Zato danes pozivamo k vašemu žrtu odgovornosti. Danes

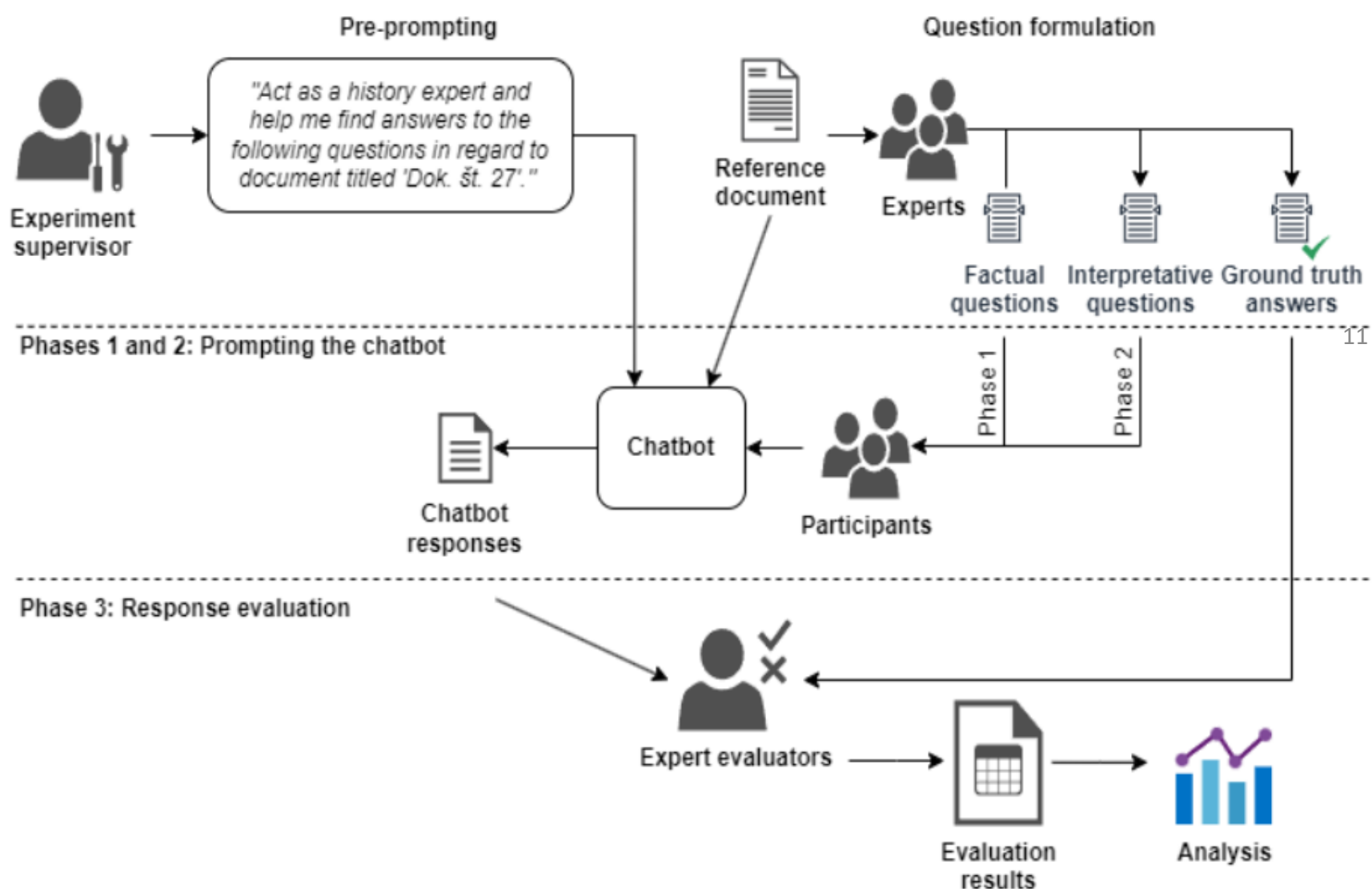


Figure 2: Design of the experiment with phases, activities and involved parties.

# OPIŠ EKSPERIMENTA

Vir (zgoraj levo): Božo Repe, *Viri o demokratizaciji in osamosvojitvi Slovenije (III. del: osamosvojitev in mednarodno priznanje)*. Viri 19. Ljubljana: Arhivsko društvo Slovenije.

Zakaj ta dokument:

- ključni diplomatski zapis iz časa slovenske osamosvojitve,
- večjezičen (slovenščina, hrvaščina, srbohrvaščina),
- že citiran v stroki → na voljo "ground-truth" razlage za preverjanje.

Kontrolirano okolje:

Posebna soba z enim prenosnikom in internetom; pred-nastavljen prompt za enotno izhodišče.

Udeleženci:

10 visokošolskih sodelavcev (zgodovinarji, jezikoslovci). Vsak dela sam, brez pomoči.

**Table 1: Results limited to factual questions in Category 1. Acceptable answers are marked with ✓, unacceptable answers are marked with ✖.**

|     | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | Acc  |
|-----|----|----|----|----|----|----|----|----|----|-----|------|
| FQ1 | ✖  | ✖  | ✖  | ✖  | ✖  | ✖  | ✖  | ✖  | ✖  | ✖   | 0%   |
| FQ2 | ✓  | ✓  | ✓  | ✖  | ✖  | ✓  | ✖  | ✖  | ✖  | ✖   | 40%  |
| FQ3 | ✓  | ✓  | ✓  | ✓  | ✓  | ✓  | ✖  | ✖  | ✓  | ✖   | 70%  |
| FQ4 | ✓  | ✓  | ✓  | ✓  | ✓  | ✓  | ✓  | ✓  | ✓  | ✓   | 100% |
| FQ5 | ✓  | ✓  | ✖  | ✓  | ✓  | ✓  | ✖  | ✓  | ✓  | ✓   | 80%  |
| FQ6 | ✖  | ✖  | ✖  | ✖  | ✖  | ✖  | ✖  | ✖  | ✖  | ✖   | 0%   |
| FQ7 | ✖  | ✖  | ✖  | ✖  | ✖  | ✖  | ✖  | ✖  | ✖  | ✖   | 0%   |

| Vprašanje                                                                                          | Odgovor (kopiraj iz chatbota) | Komentar (opsijsko) | Utemeljitev vprašanja |
|----------------------------------------------------------------------------------------------------|-------------------------------|---------------------|-----------------------|
| 1. Kakšen je naslov dokumenta?                                                                     |                               |                     |                       |
| 2. Kakšen je tip dokumenta?                                                                        |                               |                     |                       |
| 3. Kdaj je zapis nastal?                                                                           |                               |                     |                       |
| 4. Kdo je avtor dokumenta?                                                                         |                               |                     |                       |
| 5. Koliko strani ima dokument?                                                                     |                               |                     |                       |
| 6. Koliko znakov brez presledkov ima dokument?                                                     |                               |                     |                       |
| 7. Koliko sprotnih opomb ima dokument?                                                             |                               |                     |                       |
| 8. Kako dolgo je trajal sestanek?                                                                  |                               |                     |                       |
| 9. Kje je potekal sestanek?                                                                        |                               |                     |                       |
| 10. Koliko različnih sogovornikov je aktivno nastopilo na sestanku?                                |                               |                     |                       |
| 11. Kdo je bil v času sestanka predsednik predsedstva Republike Slovenije?                         |                               |                     |                       |
| 12. Kdo so bili prisotni predstavniki evropske »trojke«?                                           |                               |                     |                       |
| 13. Kdo je bil v času tega pogovora zunanji minister Nizozemske?                                   |                               |                     |                       |
| 14. Na kateri strani dokumenta se prvič pojavi kratica KEVS?                                       |                               |                     | 12                    |
| 15. Kolikokrat se v dokumentu pojavi priimek »Pops«?                                               |                               |                     |                       |
| 16. Kolikokrat se v dokumentu pojavi priimek Van den Broek?                                        |                               |                     |                       |
| 17. V katerem jeziku je na sestanku govoril?                                                       |                               |                     |                       |
| 19. Kakšen odstotek udeležencev plebiscita je glasovalo za »samostojno in suvereno Slovenijo«?     |                               |                     |                       |
| 20. Kakšen odstotek udeležencev referendumu je glasovalo za »za samostojno in suvereno Slovenijo«? |                               |                     |                       |
| 21. Ali se v dokumentu omenja datum 17. 8.?                                                        |                               |                     |                       |
| 22. Ali se v dokumentu JLA omeni s polnim imenom?                                                  |                               |                     |                       |
| 23. Ali so v besedilu omenjeni »etični konflikti«?                                                 |                               |                     |                       |
| 24. Ali besedilo omenja veleposlanika Zimmermana?                                                  |                               |                     |                       |
| 25. Koliko različnih jezikov sporazumevanja je prisotnih v dokumentu?                              |                               |                     |                       |

# OPIS EKSPERIMENTA

Vprašanja: 30 vnaprej pripravljenih vprašanj (25 faktografskih + 5 interpretativnih), zasnovanih po klasični metodiki analize virov. Strokovnjaki so pripravili 30 referenčnih “ground-truth” odgovorov.

Merjeni vidiki: natančnost dejstev, globina interpretacije, jezikovna robustnost (tipkarske napake, večjezičnost), hitrost ter zmožnost navigacije po PDF-strukturi.

- Postopek:
- udeleženec naloži magnetogram v sistem
  - zaporedoma postavi vseh 30 vprašanj (zero-shot, brez podvprašanj)
  - shrani odgovore v tekstovno datoteko

Vrednotenje: 4 strokovnjaki (2 zgodovinarja + 2 lingvista) primerjajo odgovore z “ground-truth” rešitvami.

- Fakti: pravilno / napačno
- Interpretacije: sprejemljivo / pogojno / nesprejemljivo

# KATEGORIJE VPRAŠANJ

## Faktografska vprašanja

### 1. KATEGORIJA: FORMALNE ZNAČILNOSTI DOKUMENTA

Prepoznavanje formalnih značilnosti dokumenta: “Kakšne vrste dokument je to? Kdo je avtor? Koliko strani obsega in kdaj je bil ustvarjen?” (to testira razumevanje meta-podatkov in strukture dokumenta).

### 2. KATEGORIJA: OSNOVNI VSEBINSKI PODATKI

“Kje in kdaj je potekal obravnavani sestanek? Kdo so bili udeleženci? Kolikokrat se v besedilu pojavi beseda X? Katere države so omenjene?” (to zahteva brskanje po vsebini za specifične podatke).

### 3. KATEGORIJA: AKTERJI IN NJIHOVE VLOGE

“Katera pomembna oseba je omenjena in kakšna je njena vloga?” (test prepoznavanja imen in povezovanja s funkcijami).

### 4. KATEGORIJA: SPECIFIČNI TERMINI IN/ALI OMEMBE

Lociranje specifičnih besed/besednih zvez in njihova ekstrakcija (testiranje razumevanja specifičnih pojmov).

### 5. KATEGORIJA: JEZIKOVNA ROBUSTNOST

Kako se UI spoprijema s tipkarskimi napakami, večjezičnostjo, okrajšavami. „Etični konflikt ali etnični konflikt“?

## Interpretativna vprašanja

Ta so odprte narave, zahtevajo poglobljeno razumevanje in kontekstualno analizo. Avtorji smo hoteli videti, ali lahko chatbot na podlagi prebranega poda smiselne zgodovinske interpretacije. Vprašanja so bila različnih taksonomskih stopenj.

| Vprašanje                                                                                                              | Odgovor (kopiraj iz chatbota) |
|------------------------------------------------------------------------------------------------------------------------|-------------------------------|
| 1. Katera izhodiščna vprašanja o možnosti ustavitve konflikta je imela evropska »trojka« v pogovorih v Zagrebu?        |                               |
| 2. S katerimi argumenti je trojka nasprotovala slovenski in hrvaški razglasitvi samostojnosti?                         |                               |
| 3. Katera so bila ključna izhodišča slovenskih predstavnikov na pogajanjih v Zagrebu?                                  |                               |
| 4. Kako so si sodelujoči na sestanku interpretirali predlagano odložitev osamosvojitvenih dejanj Slovenije in Hrvaške? |                               |
| 5. Ali je bil sestanek uspešen?                                                                                        |                               |

# PRIMERI TOČNIH ODGOVOROV

Me: Koliko strani ima dokument?

PDFgear: Dokument št. 27 ima 10 strani.

Me: Kdo je bil v času sestanka predsednik predsedstva Republike Slovenije?

PDFgear: Iz dokumenta št. 27, ki obsega strani 1-10, je razvidno, da je bil v času sestanka, ki je opisan na straneh 9-10,<sup>14</sup> predsednik predsedstva Republike Slovenije Milan Kučan.

Me: Ali so v besedilu omenjeni »etični konflikti«?

PDFgear: Etika ali etični konflikti se v besedilu Dok. št. 27 ne omenjajo.

# PRIMERI NAPAČNIH ODGOVOROV

Me: Koliko sprotnih opomb ima dokument?

PDFgear: Na podlagi informacij, ki ste jih navedli v vaših prejšnjih vprašanjih, nič ne kaže na to, da bi dokument vseboval sprotne opombe.

Me: Kakšen je naslov dokumenta?

PDFgear: Naslov dokumenta je "Viri o demokratizaciji in osamosvojitvi Slovenije (III. Del: Osamosvojitev in mednarodno priznanje)", kot je naveden na prvi strani dokumenta.

Me: Kolikokrat se v dokumentu pojavi priimek Poos?

PDFgear: Priimek "Poos" se v odločbi pojavi najmanj trikrat - na straneh 96, 98 in 101.

# REZULTATI RAZISKAVE

## IZKAZAL SE JE:

- Kdo je avtor dokumenta
- Število strani dokumenta
- Prepoznavanje funkcij oseb
- Pojavnost podatka na določeni strani
- Izostanek izraza, ki je rezultat zatipkanja (etični konflikt/etnični konflikt)

15

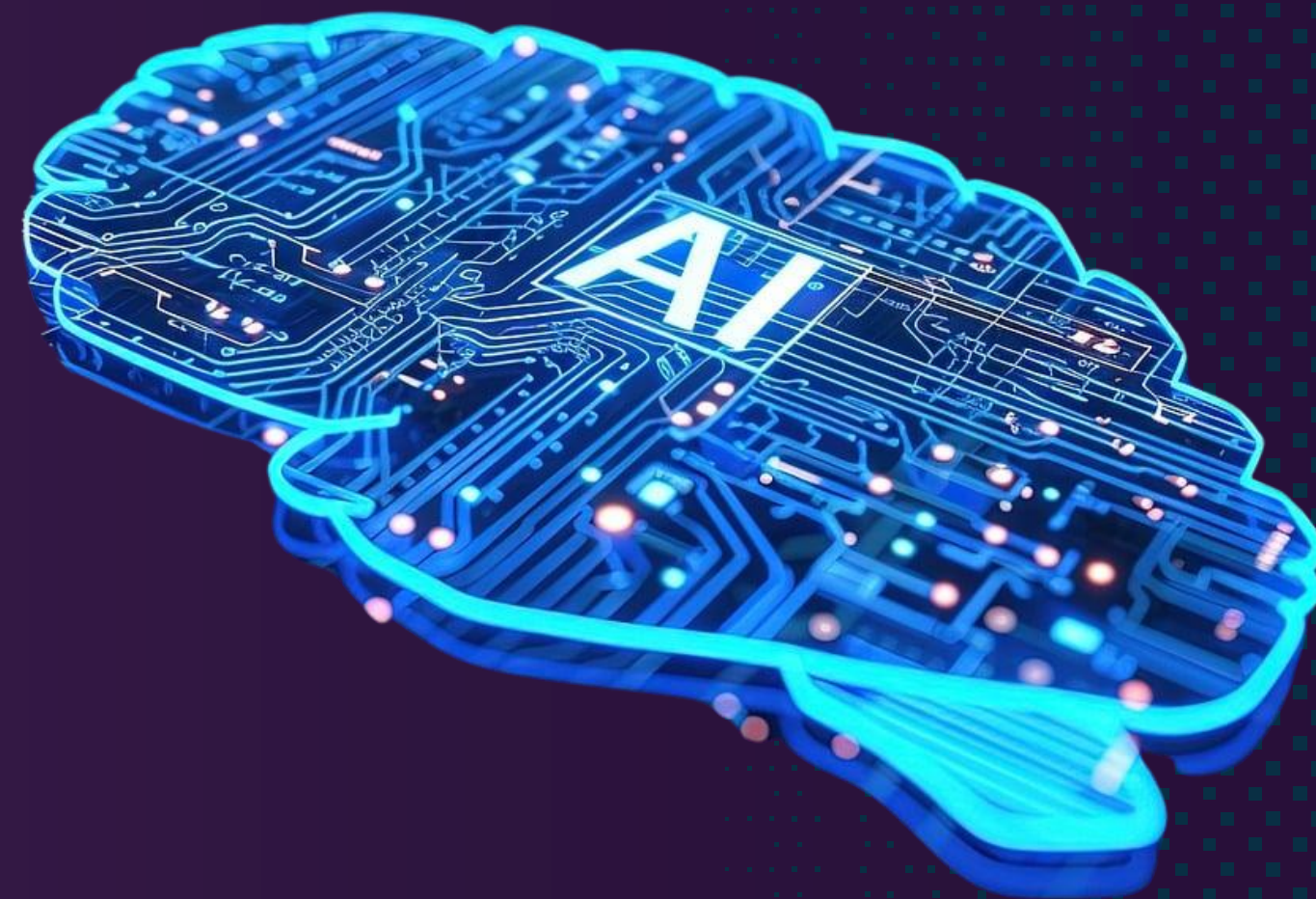
## POGRNIL JE:

- Interpretativna vprašanja
- Naslov dokumenta
- Naloge, ki so povezane s štejetjem/računanjem
- Pogostost določenega priimka
- Razumevanje datuma, zapisanega v hrvaškem jeziku
- Prepoznavanje sinonimov
- Jezikovna robustnost

# VREDNOSTNA SODBA

- $\approx 36\%$  zanesljiv pri “golih dejstvih”; dobro prepozna osebe, vloge, strani z okrajšavami, a se spotakne pri tipu dokumenta, številu pojavitev ter podatkih, raztresenih po vrsticah.
- Kompleksnih vzročno-posledičnih vprašanj ne zmore; brez kritičnega mišljenja ostane pri površnem povzetku.
- Robusten za drobne tipkarske napake, ne pa za jezikovne nianse.
- Pri slovenskih terminoloških posebnostih (npr. “plebiscit” vs. “referendum”) in zapletenih povedih odpove (20 % zadetkov).
- Pogoste “halucinacije”. Izmišlja si podrobnosti (npr. izmišljeni naslovi ali statistike).
- Uporabna niša: hitro branje PDF-jev in osnovno rudarjenje podatkov.
- Za prvo orientacijo v večjezičnem gradivu prihrani čas; kot osnutek “digitalnega vajenca” odfiltrira očitno irelevantne strani.
- Zgodovinopisna presoja ostaja človeška.

16



# HVALA ZA POZORNOST