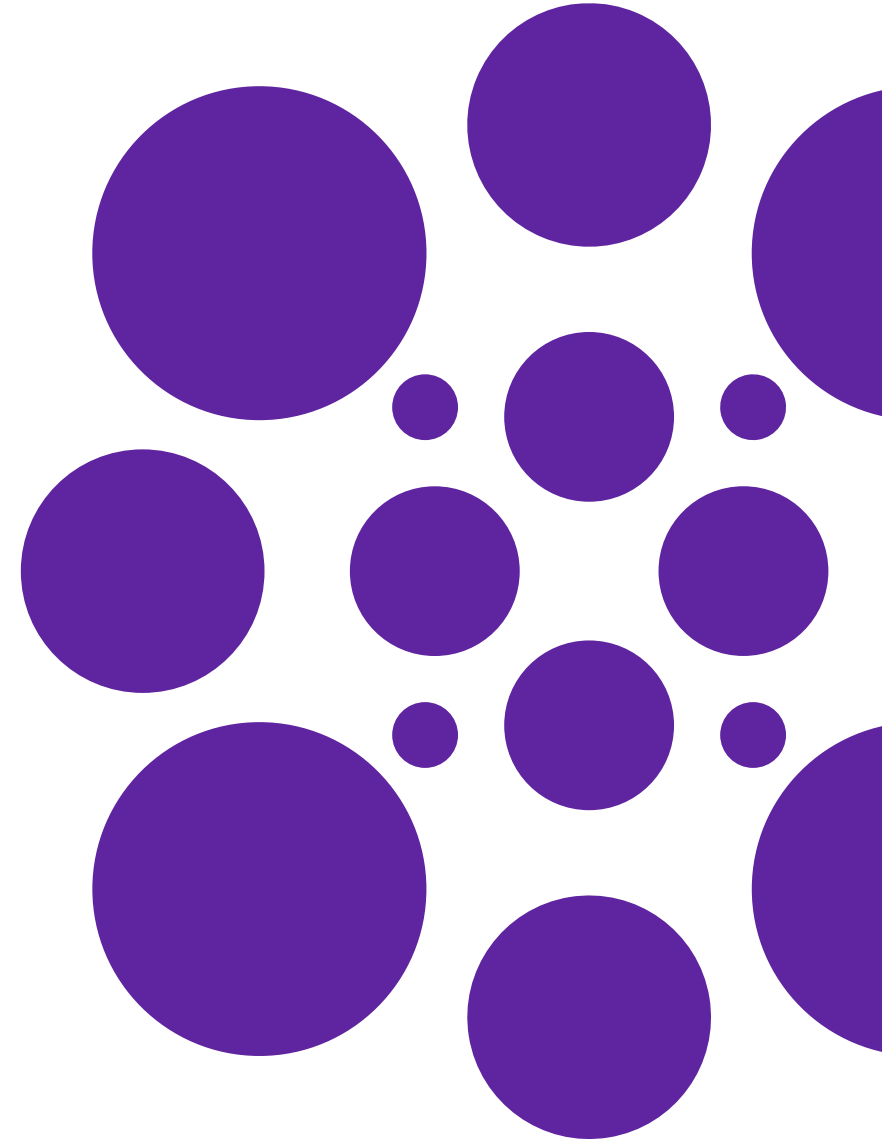


Kako da isprogramiram svog „AI“ agenta?

Čaprić Nina, Govt. Rel. Mgr. EMEA
Nešić Andreja, PS Engineer I

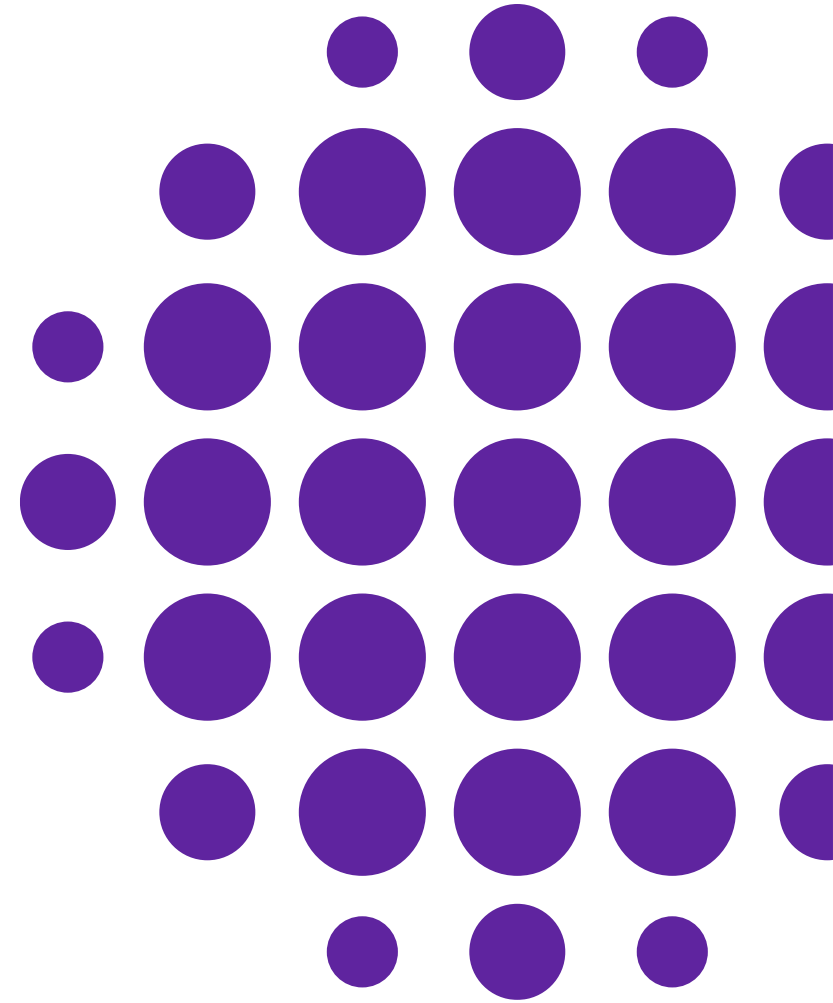


Sadržaj prezentacije

Šta ćemo danas preći?

1. O predavačima
2. Šta je veštačka inteligencija?
3. Kako rade generativni modeli?
4. Konstruišemo ideju za svoj „AI“
5. Pređimo na kucanje!
6. Demo: naša aplikacija, uživo
7. Zaključak

O predavačima





Nina Čaprić

Government Relations Manager EMEA

Master politikolog za međunarodne poslove



nina.capric@ncrvoyix.com






in/nina-capric



Andreja Nešić

PS Engineer I

Master student @ Računarski fakultet

-  andreja.nesic@ncrvoyix.com
-  anesic3023m@raf.rs
-  [in/andreja-nesic](https://www.linkedin.com/in/andreja-nesic)

NCR Voyix

Mi kreiramo budućnost trgovine

- NCR Voyix korporacija je vodeća globalna kompanija u oblasti pružanja digitalnih usluga u sektoru prodaje, ugostiteljstva i digitalnog bankarstva



16.000+

Zaposlenih



35

Zemalja sveta

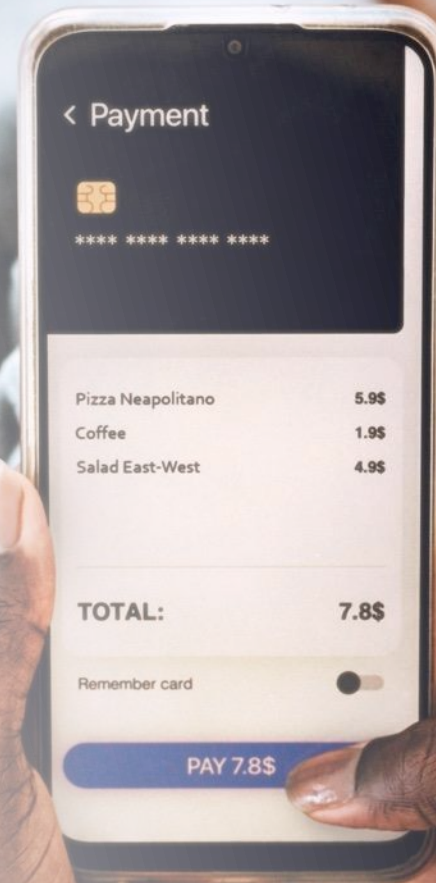
- U Srbiji smo prisutni od 2011. godine
- Naš beogradski kampus je mesto na kojem 3400 inovatora unapređuje sposobnost naših klijenata da naprave razliku kroz iskustva koja pružaju



Iz prve ruke: AI u NCR Voyix-u

- Dinamična, personalizovana i digitalna iskustva za **krajnje korisnike**
 - Krajnji korisnici: prodavac u radnji, mušterija, konobar ili menadžer restorana, itd.
- Više od podataka—jedinstvena saznanja i saveti za **poslovne korisnike**
 - Menadžeri ugostiteljskih, maloprodajnih objekata
- Produktivniji rad i jednostavnije procedure za **zaposlene širom sveta**
 - Inženjeri, instalateri, podrška, prodavci, menadžment, itd.

Veštačka inteligencija



Šta je veštačka inteligencija?

Kako je definišemo i šta obuhvata?

- Veštačka inteligencija je inteligencija mašina, računara, itd.
- Podoblast računarskih nauka
- Obuhvata mašinsko učenje, obradu jezika, ekspertne sisteme, robotiku, i slično
- Srećemo je na svakom koraku: pri veb pretrazi, pisanju poruka, fotografisanju, igranju igrica...
- U kući, u prodavnici, u restoranu



Kako rade GPT modeli?

Osnova moderne „veštačke inteligencije“

- U glavnom govorimo o generativnoj veštačkoj inteligenciji (GenAI)
- Generiše tekst, kod, sliku, čak i video!
- **Transformer model** je u osnovi većine najpopularnijih GenAI modela
- Modeli koji generišu tekst: veliki jezički modeli (large language models, LLM); klasifikacija po veličini na manje, srednje, velike
- Popularni modeli: BARD, ChatGPT, CodeLlama, DALL-E, itd.

Ilustracija preuzeta sa [CIVITAI](#)

70mlrd

Najveći LLM modeli

13mlrd

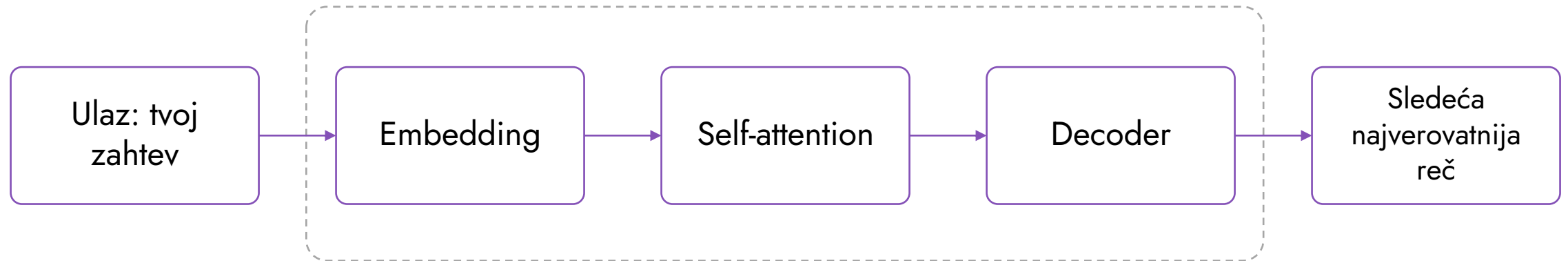
Modeli srednje veličine

<7mlrd

Manji modeli

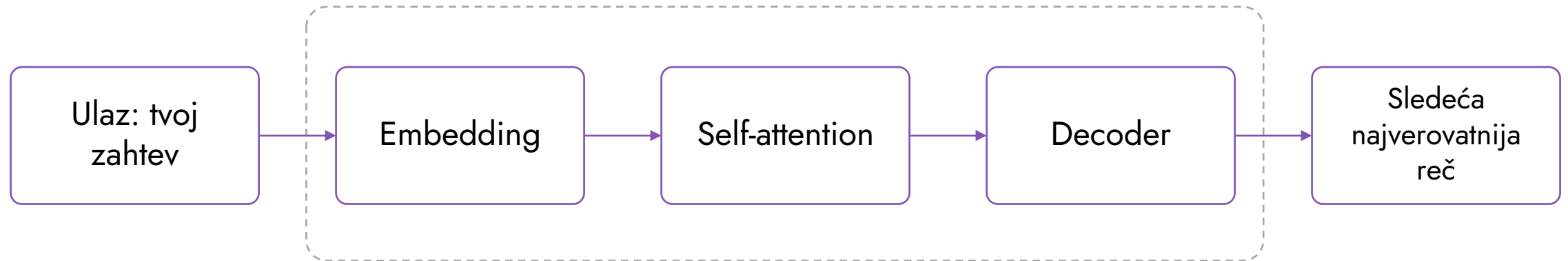
Kako rade GPT modeli?

Osnova moderne „veštačke inteligencije“



Kako rade GPT modeli?

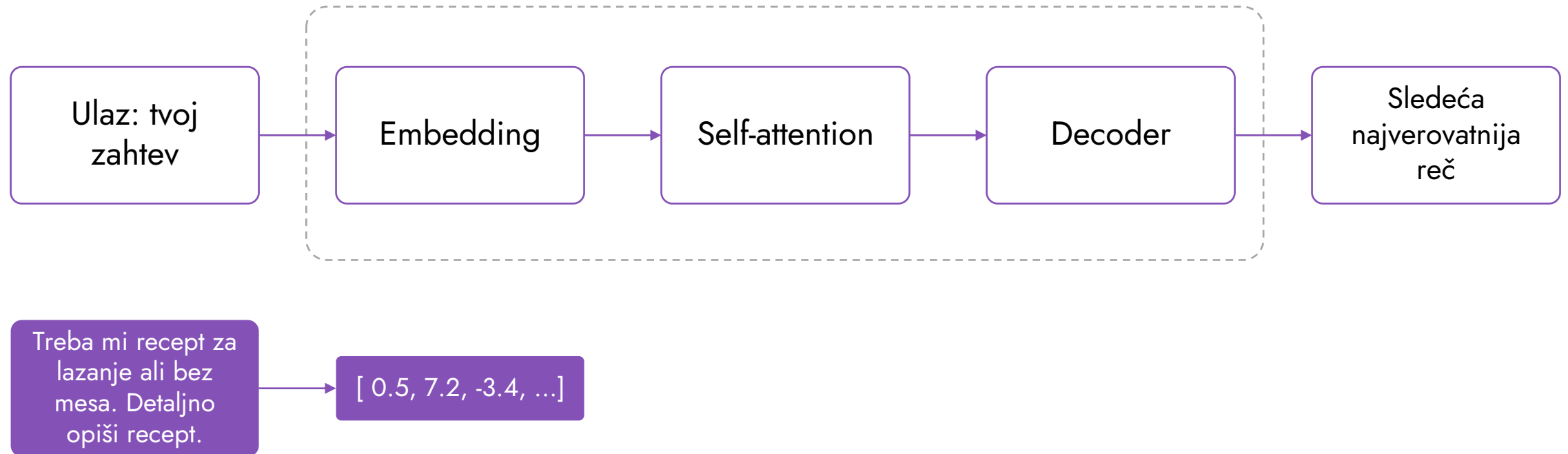
Osnova moderne „veštačke inteligencije“



Treba mi recept za
lazanje ali bez
mesa. Detaljno
opiši recept.

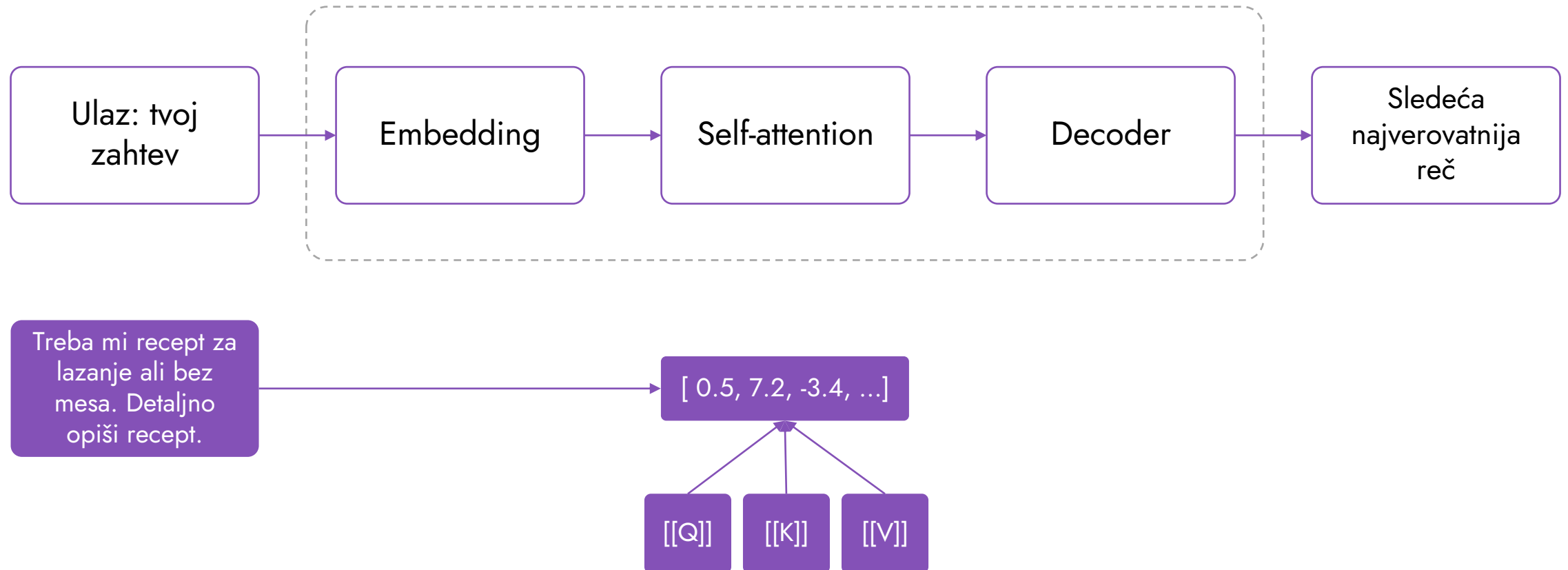
Kako rade GPT modeli?

Osnova moderne „veštačke inteligencije“



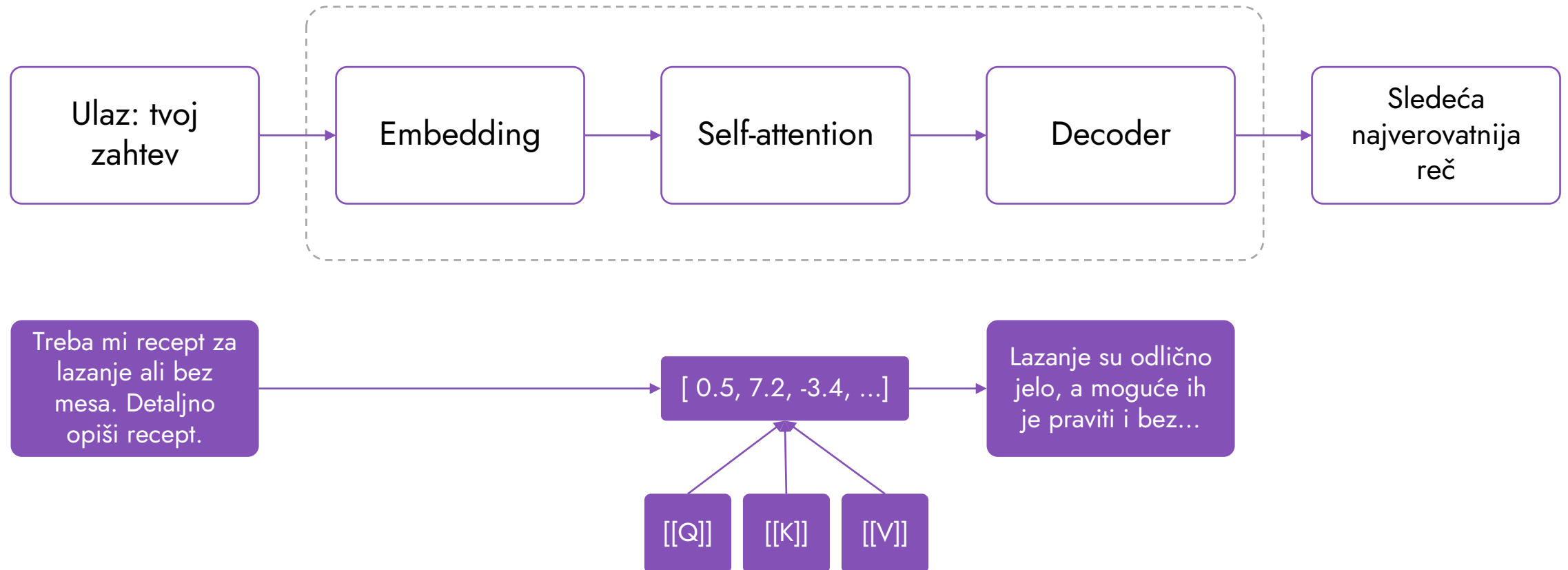
Kako rade GPT modeli?

Osnova moderne „veštačke inteligencije“



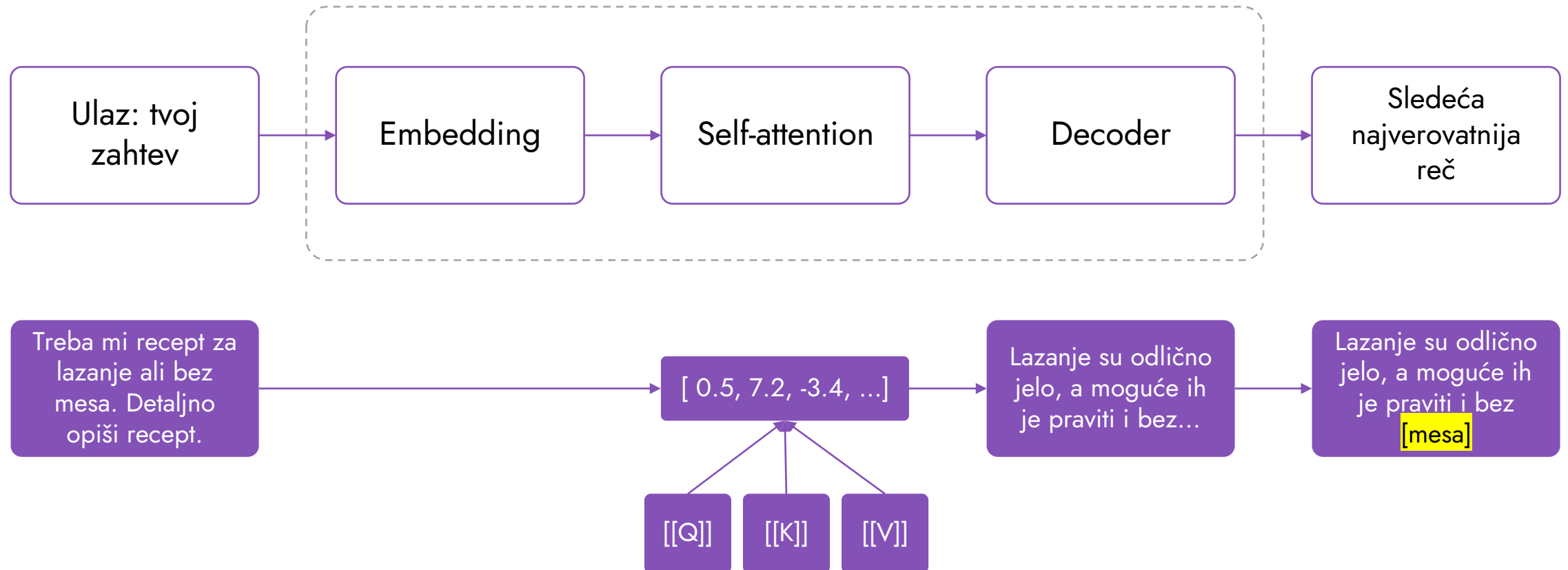
Kako rade GPT modeli?

Osnova moderne „veštačke inteligencije“



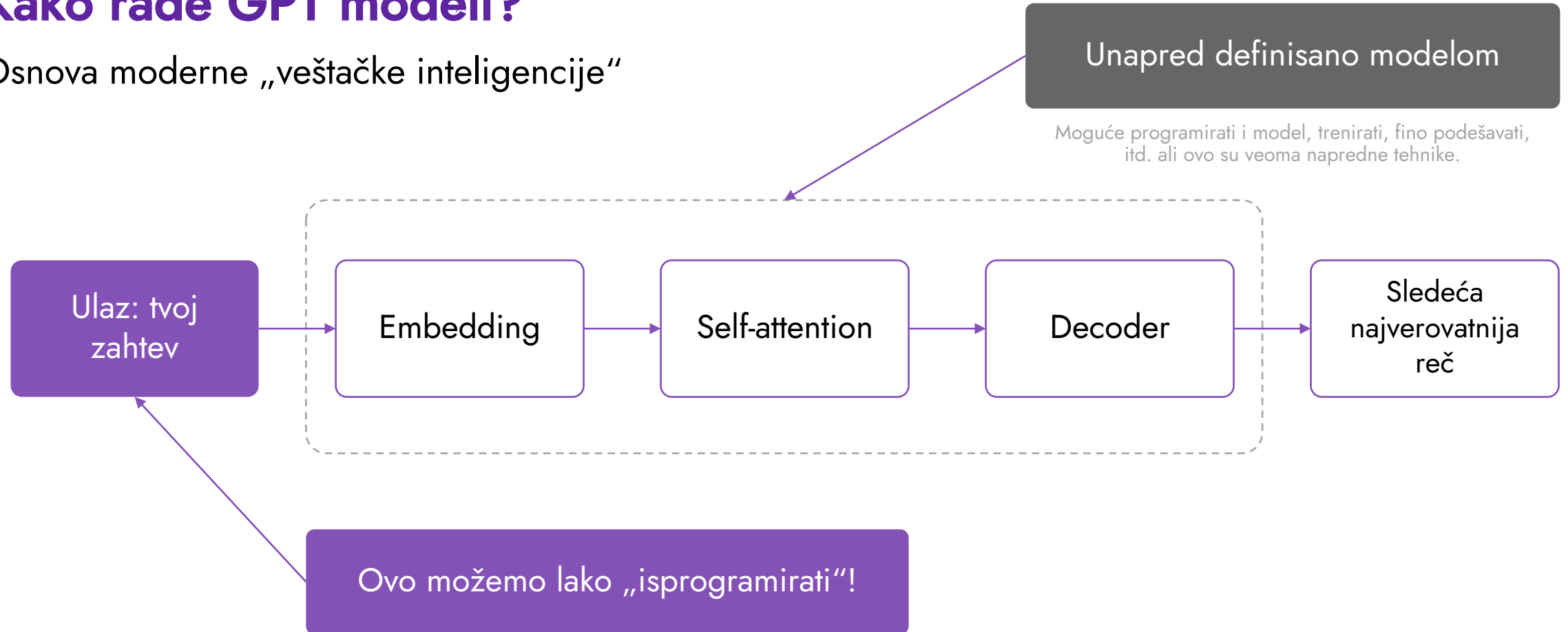
Kako rade GPT modeli?

Osnova moderne „veštačke inteligencije“

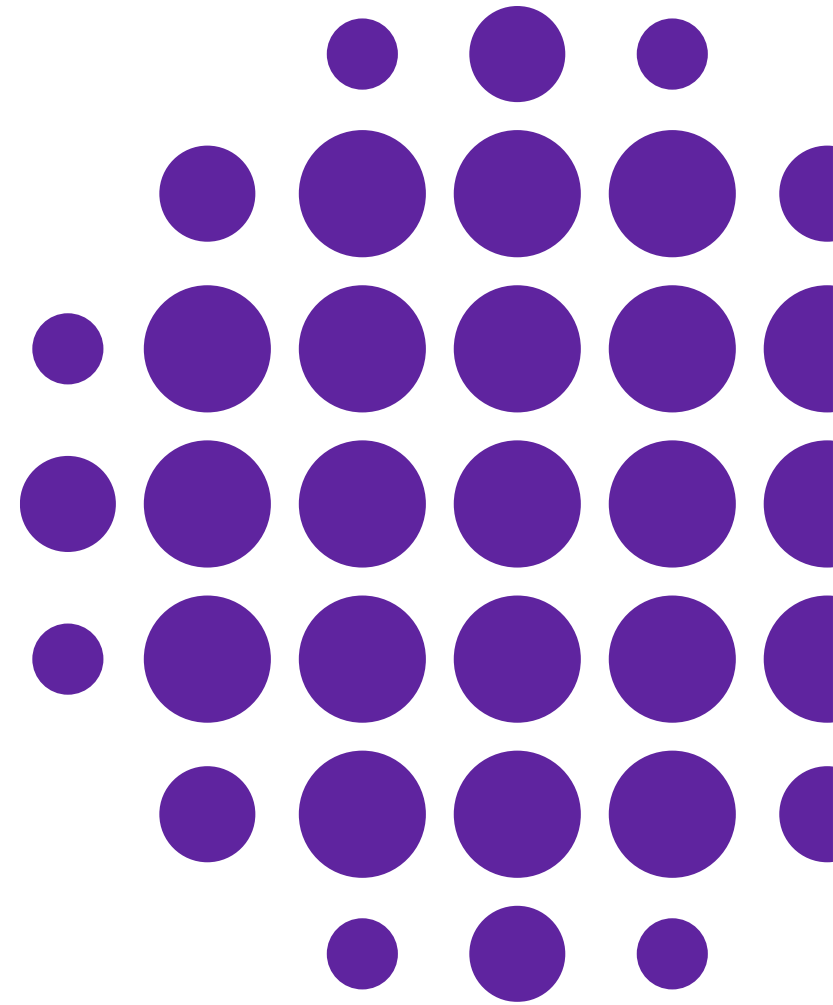


Kako rade GPT modeli?

Osnova moderne „veštačke inteligencije“



Napravimo
sopstveni „AI“!

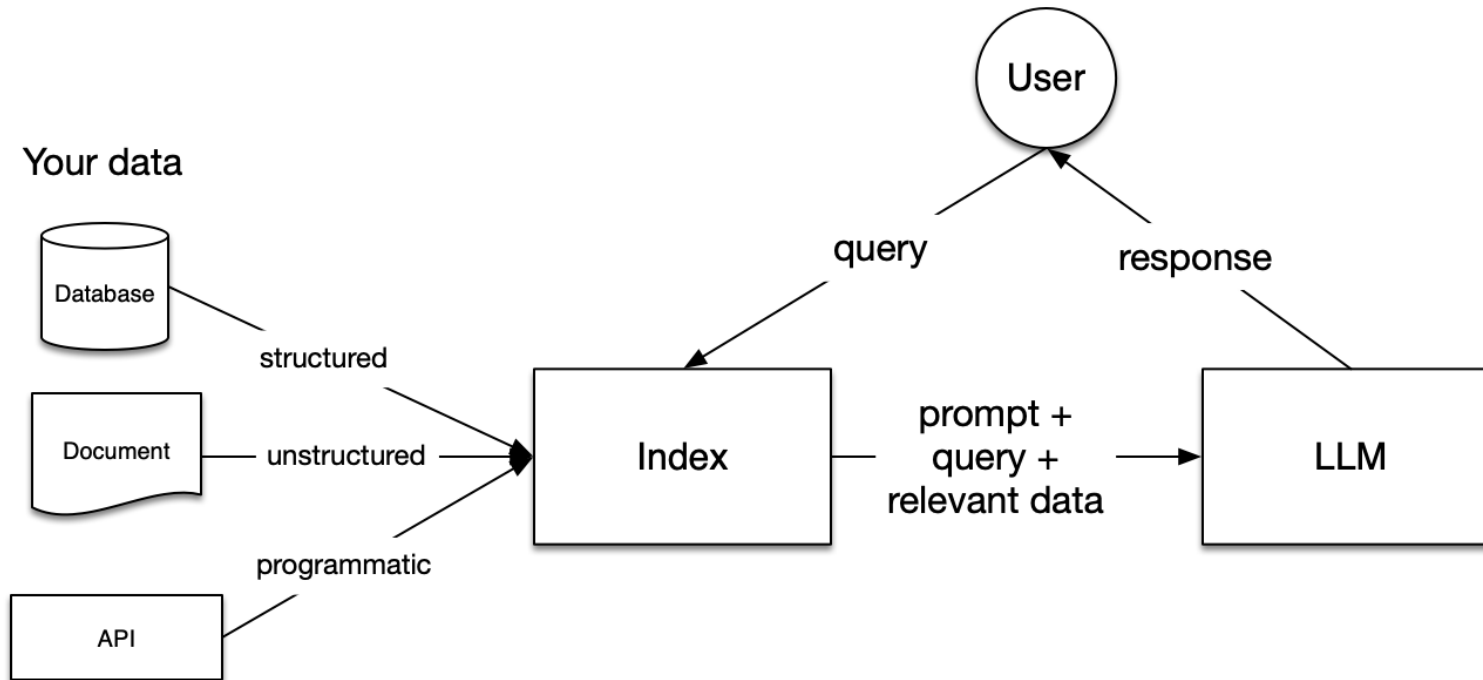


Napravimo sopstveni „AI“!



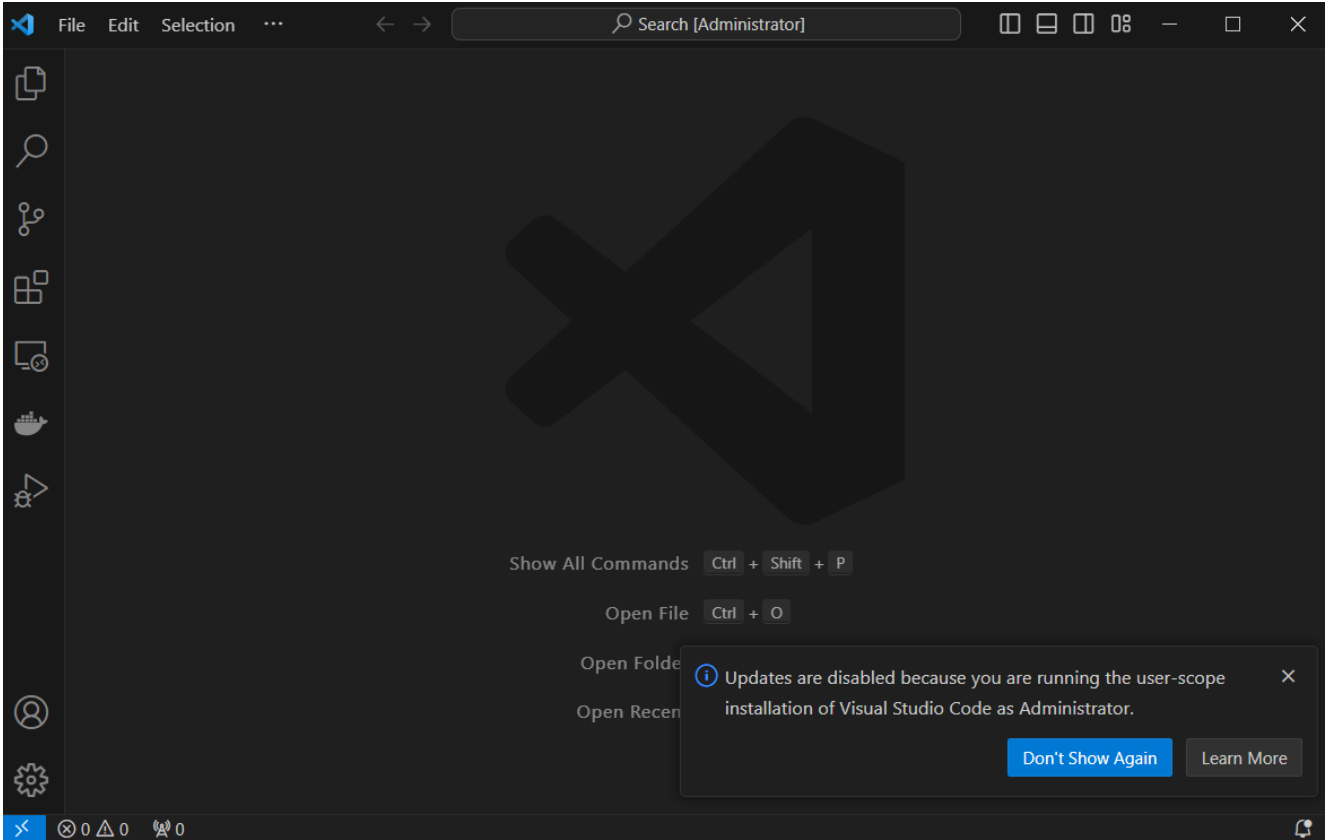
Naša ideja

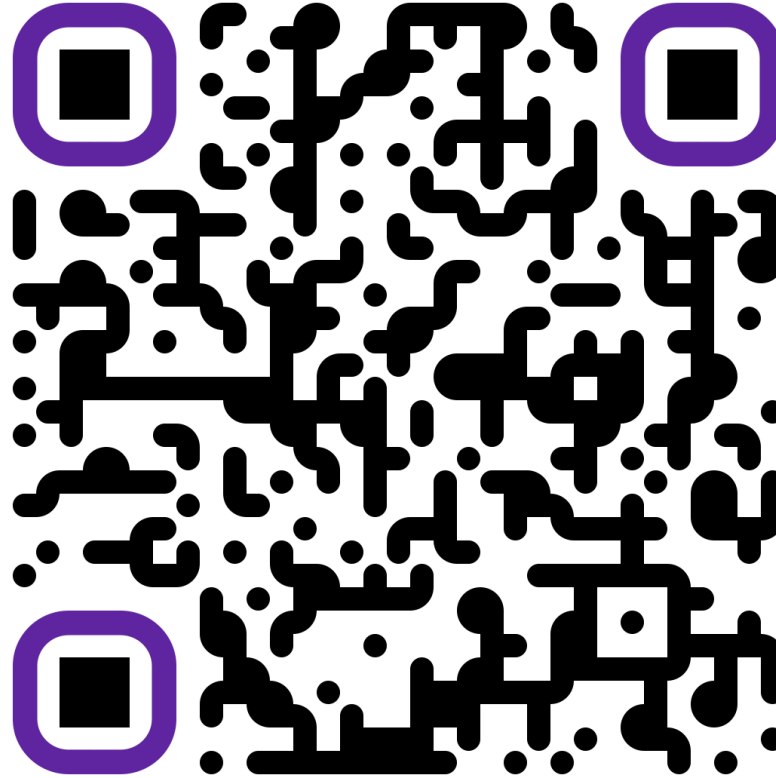
- Programiramo rešenje za e-commerce radnju
 - Kreiramo **čebot asistenta**
- Zadaci: pomoć klijentima u kupovini i reklamaciji
- Asistent mora ispuniti uslove:
 - Pruža korisne i tačne informacije
 - Odgovara isključivo ukoliko zna tačne informacije
 - Posедуje znanje o našim proizvodima
- Softverski alati: LlamaIndex, Gradio, OpenAI API (koristimo jezik Python)



LlamaIndex: Osnovni Koncept

Slika preuzeta sa https://docs.llamaindex.ai/en/stable/getting_started/concepts/





Prezentacija + Demo

<https://github.com/andrejanesic/raf-it-days-2024>

Dodatni resursi

Saznaj još ili napravi ovo kod kuće!

- [GitHub repozitorijum demonstriranog koda](#)
- [LlamaIndex dokumentacija](#): okvir za razvoj LLM-ova sa kontekstom tvojih podataka
- [LangChain dokumentacija](#): još jedan alat za razvoj agenata na LLM-ovima
- [Gradio dokumentacija](#): za rapidan razvoj veb aplikacija sa LLM-ovima
- [DeepLearning.AI kursevi](#): za sticanje osnovnog znanja o radu sa/razvoju AI agenata
- [Open LLM Leaderboard](#): tabela najnaprednijih otvorenih LLM modela
- [Attention is all you need](#): naučni rad koji je prvi uveo Transformer model
- [Coursera Neural Networks and Deep Learning](#): kurs o osnovama mašinskog učenja i dubokih mreža

Hvala na pažnji!

Pitanja, komentari? Više o NCR Voyix-u:

- nina.capric@ncrvoyix.com
- andreja.nesic@ncrvoyix.com
- ncrvoyix.com

