Poglavlje 1

Uvod i osnove veštačke inteligencije

Svaka dovoljno napredna tehnologija ne može se razlikovati od magije.

– Artur Klark

Pre nego što zaronimo u moderne modele veštačke inteligencije, geopolitičke bitke i praktične primene koje menjaju svet, neophodno je postaviti čvrste temelje. Ovaj uvodni deo ima za cilj da demistifikuje osnovne mitove i pojmove, da pruži kontekst i odgovori na fundamentalno pitanje: šta je, zapravo, veštačka inteligencija i kako smo stigli do tačke u kojoj se danas nalazimo? Razumevanje ovih osnova je ključ za shvatanje svega što sledi.

Strahovi i mitovi

Pre sedamdeset godina, kompjuter je bio sala puna cevi i žica. Pre trideset godina, internet je bio eksperiment. Pre deset godina, veštačka inteligencija je bila naučna fantastika. Danas, ona govori, piše poeziju, dijagnostikuje bolesti i pokreće automobile, a mi se još uvek pitamo šta je zapravo.

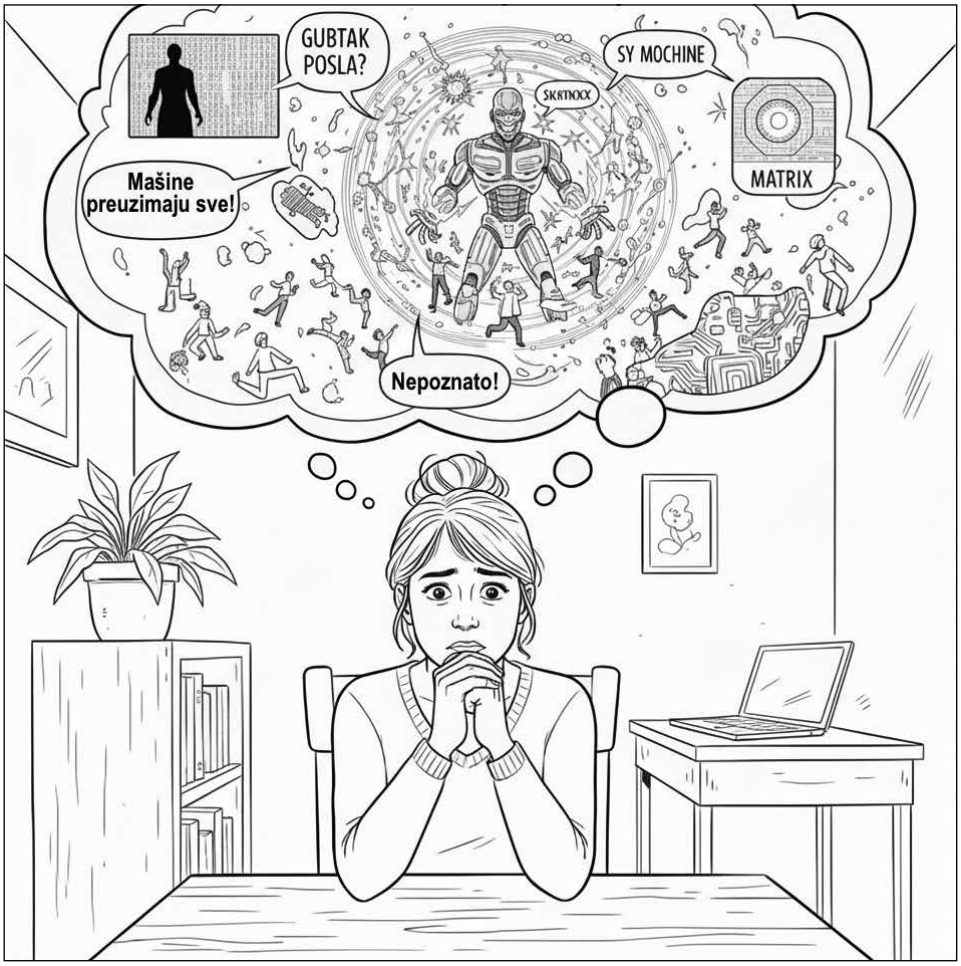
Strahovi od veštačke inteligencije

Mnogi ljudi ne žele da koriste veštačku inteligenciju (*Artificial Intelligence, AI*). Neki su umorni od novih tehnologija, kao što ima ljudi, koji ne žele ni danas da koriste mobilni telefon. Ali kada je AI u pitanju daleko češći je razlog strah.

Ironično, mnogi koji se plaše AI-ja već koriste njegove oblike u svakodnevnom životu – preporuke na društvenim mrežama, GPS navigaciju i prepoznavanje lica i glasa. Strah se često usmerava na personifikovane oblike AI-ja poput četbotova (*chat-boot*). Četbot je računarski program koji simulira konverzaciju sa ljudima kroz tekst ili govor, koristeći unapred programirane odgovore ili veštačku inteligenciju.

Strah od veštačke inteligencije je duboko ukorenjen u ljudskoj prirodi i ima nekoliko različitih slojeva. Ljudi se plaše nepoznatog, a AI predstavlja tehnologiju koju većina ne razume u potpunosti. Ova nepoznanica stvara prostor za strahove koji često nadmašuju realnost. Nedostatak transparentnosti oko toga kako AI funkcioniše pojačava strahove. Kada nešto deluje magično ili neobjašnjivo, lakše je projektovati na to naše najveće strahove.

Jedan od glavnih uzroka strahova je naša tendencija da AI-u pripišemo ljudske karakteristike poput svesti, namera ili emocija. Ljudi prirodno pokušavaju da razumeju nepoznato pomoću poznatog, pa AI doživljavaju kao neku vrstu digitalnog bića sa vlastitom voljom. Ovo vodi ka strahovima da AI može da odluči da nam naudi ili da se okrene protiv nas.



Strahovi od AI nastaju iz mešavine filmskih stereotipa i legitimnih pitanja o tome kako će tehnologija uticati na naše živote

Popularna kultura je značajno doprinela ovim strahovima. Strah od AI je uglavnom holivudska zasluga. Čim se pomene AI, prva asocijacija nam je neki od filmskih scenarija: *Terminator serija*: Skynet postaje svestan i odlučuje da uništi čovečanstvo, *Matrix*: Mašine koriste ljude kao baterije, *Ex Machina*: AI manipuliše i vara ljude da pobegne, *HAL 9000 (2001 Odyssey)*: Kompjuter ubija astronaute „za dobro misije." Ovi filmovi su vrlo zabavni i vizuelno impresivni i duboko su ukorenili strah u mnoge ljude. Iako, realno gledano, nemaju dodirnih tačaka sa stvarnošću.

Dodajmo na to medije, koji se sa svojom mamiklik (*clickbait*) (ne)kulturom utrkuju ko će da smisli dramatičniji naslov. Naslovi poput „AI će preuzeti sve poslove do 2030. godine!", zvuče senzacionalno i privlače pažnju, ali retko su potkrepljene činjenicama. Ako se tome dodaju intervjui sa stručnjacima bez balansiranog pristupa imamo recept za osećaj ugroženosti i straha, a ne za objektivnu diskusiju o tome šta veštačka inteligencija zaista jeste i šta može da postigne i koliko novih radnih mesta može da otvori.

A tu su i stručnjaci koji upozoravaju. Ilon Mask (Elon Musk): „AI je veća pretnja od nuklearnog oružja", Stephen Hawking*:* „AI bi mogao biti poslednji izum čovečanstva", Nick Bostrom: *Superintelligence* knjiga o egzistencijalnim rizicima, Džefri Hinton (Geoffrey Hinton)*:* „otac dubokog učenja" upozorava na buduće rizike.

Postoji i dublja egzistencijalna anksioznost – strah da AI može da zameni ljude čak i u kreativnosti. Ovo nije potpuno neosnovano, ali često je preuveličano. Ljudi se plaše da će postati nepotrebni u svetu gde mašine mogu da rade sve bolje od njih.

Neosnovani strahovi (kratkoročno)

- AI će **preuzeti kontrolu**: Trenutni AI nema ciljeve, svest ili želje. Zavisi od ljudskih instrukcija i nadzora.
- AI će masovno **uništiti poslove**: Istorijski, tehnologija kreira nove poslove dok ukida stare. Proces je postupan, daje vremena za prilagođavanje.
- AI je **zao ili manipulativan**: Nema emocije, motive ili ličnost. Greške su tehnički problemi, ne zlonameran čin.

Opravdani strahovI (kratkoročno i srednjoročno)

- **Pristrasnost i diskriminacija:** AI sistemi nasleđuju predsrasude iz podataka. Primer: AI za zapošljavanje koji diskriminiše žene
- **Dipfejk (*deepfakes*) i dezinformacije:** Lažni video i audio zapisi sve teže se razlikuju od stvarnih. Ugrožava društvo i poverenje u medije
- **Privatnost i nadzor:** AI omogućava masovnu analizu ličnih podataka. Rizik od „države koja sve nadgleda". Mada takvu situaciju imamo i pre AI-ja, posle 11. septembra i rušenja dve kule u Njujorku, kako je opisano u knjizi *Data i*

Golijat (Mikro knjiga, 2019.) da je u SAD I ostalim delovima sveta, svaki pojedinac nadgledan.

- **Ekonomska nejednakost:** Koristi od AI mogu da povećaju jaz između bogatih i siromašnih. Koncentracija moći u rukama nekoliko kompanija.

Spekulativni strahovi (dugoročno)

- **Opšta veštačka inteligencija** (*Artificial General Intelligence, AGI*): Ako/kada AI postigne ljudski nivo inteligencije. Problem „usklađenosti" – kako garantovati da AI čini ono što želimo?
- **Superinteligencija:** AI koji nadmašuje ljudsku inteligenciju u svim oblastima. Teoretski rizik, ali vremenski okvir je nepoznat.

Ne treba paničiti, već treba odgovorno pristupiti. Razvijati AI sigurnost paralelno sa mogućnostima. Edukovati javnosti o realnim rizicima i koristima. Etički pristupati razvoju AI sistema.

Razumevanje da je AI alat, bez obzira na to koliko sofisticiran, može pomoći u prevazilaženju ovih strahova. Ali potrebno je vreme i edukacija da se promeni percepcija koja je godinama oblikovana kroz kulturu i medije.

Mitovi o AI

- Mit 1: **AI će postati svestan i preuzeti kontrolu:** Stvarnost je da AI nema više svesti od kalkulatora. AI je sofisticiran papagaj, odličan u prepoznavanju obrazaca, ali bez ijedne vlastite želje ili cilja. Kad AI „odgovori" na pitanje, ne razmišlja već statistički predviđa šta bi trebalo da bude sledeća reč na osnovu milijardi tekstova koje je video tokom treninga.
- Mit 2: **AI nikad ne greši:** Ovaj mit je možda gori od prethodnog. AI sistemi prave greške, haluciniraju (izmišljaju činjenice), mogu biti pristrasni. AI će sa potpunom sigurnošću reći da je Napoleon umro 1825. (umesto 1821.) ili izmisliti nepostojeću naučnu studiju. Zato uvek treba proveriti važne informacije.
- Mit 3: **AI će sutra zameniti sve poslove:** Istorijski gledano, tehnologija uglavnom stvara više poslova nego što uništava. Kad su se pojavili računari, ljudi su mislili da će svi biti nezaposleni. Umesto toga nastala je cela IT industrija. AI će zameniti neke rutinske zadatke, ali će isto tako stvoriti potrebu za ljudima koji razumeju kako da ga koriste, kontrolišu i poboljšaju.

Ima još mitova o AI. Jedan od njih je mit koji se često provlači, nalazi se u dodatku knjige, jeste *Apokalipsa spajalicama*. To je ideja da bi AI, u pokušaju da ispuni naizgled bezazlen cilj kao što je proizvodnja spajalica, mogla da uništi svet.

Tako se stvara slika da je AI neka sila koja je tu da nas uništi ili zameni, a ne moćan alat koji može da nam pomogne u rešavanju složenih problema. Zato, kad sledeći put čujete neku priču o AI apokalipsi, setite se da je verovatno začeta u nekom holivudskom studiju, a dodatno je začinjena bombastičnim naslovom iz nekog tabloida. Prava priča o AI je mnogo manje dramatična, ali zato i mnogo zanimljivija.

Naravno, treba da budemo budni i trezvenoumni kada je AI u pitanju, o čemu nam govori lično Džefri Hinton (u Dodatku *Otac veštačke inteligencije o AI*).

Zašto je prethodno izneto važno?

Ako ne možete da se oslobodite straha ili verujete u mitove, propustićete priliku. AI je danas fenomenalan alat za pisanje, prevođenje, kodiranje, izradu audio i video sadržaja, učenje, kreativnost, ali samo ako ga razumete realno. Nije ni strašno čudovište iz filmova, niti čarobni štapić koji će rešiti sve probleme. Jednostavno je vrlo moćan alat koji zahteva pametno korišćenje. A ako stvari krenu u pravcu najgorih snova, bolje je da budemo u toku sa znanjem šta je AI i njegovim korišćenjem da bismo bili spremniji da mu se suprotstavimo.

Budite skeptični prema velikim tvrdnjama (i pozitivnim i negativnim), eksperimentišite sa AI alatima lično, i uvek imajte na umu da je to tehnologija u razvoju, korisna, ali nesavršena.

Šta je veštačka inteligencija

Veštačka inteligencija je grana računarstva koja se bavi stvaranjem mašina i softvera sposobnih da obavljaju zadatke koji, kada ih obavlja čovek, zahtevaju inteligenciju. Ovi zadaci uključuju učenje, rezonovanje, rešavanje problema, percepciju i razumevanje jezika.

Ključna reč ovde je „zahtevaju inteligenciju“. Ne radi se o tome da mašine misle ili osećaju na isti način kao ljudi. Radi se o simulaciji inteligentnog ponašanja. Kalkulator može da računa brže od bilo kog čoveka, ali ga ne smatramo inteligentnim jer samo sledi fiksna, unapred programirana pravila. AI, s druge strane, poseduje sposobnost adaptacije i učenja iz novih informacija.

Da bismo razumeli modernu AI, moramo se upoznati sa dva fundamentalna koncepta koja su pokretačka snaga današnje revolucije: *mašinsko učenje* i *duboko učenje*.

Mašinsko učenje

Zamislite da želite da naučite kompjuter da prepozna slike mačaka. Tradicionalni pristup bi bio da pokušate da napišete beskonačan niz pravila: „Ako ima šiljate uši i brkove i krzno... onda je mačka." Ovaj pristup je krut, neefikasan i osuđen na propast zbog ogromne varijabilnosti u izgledu mačaka.

Mašinsko učenje (*Machine Learning, ML*) okreće ovaj proces naglavačke. Umesto da mu dajemo eksplicitna pravila, mi mu dajemo podatke. Pokažemo mu hiljade i hiljade slika sa oznakom „mačka" i hiljade slika sa oznakom „nije mačka". Algoritam mašinskog učenja zatim samostalno, kroz statističku analizu, uči da identifikuje obrasce (*patterns*) koji definišu mačku. On „uči iz iskustva" (podataka), umesto da bude eksplicitno programiran.

ANALOGIJA ZA MAŠINSKO UČENJE

Mašinsko učenje je slično učenju deteta da prepozna životinje. Ne objašnjavate mu teoriju o sisarima ili anatomiju psa. Jednostavno mu pokazujete slike i govorite: „Ovo je kuca, ovo je maca." Posle dovoljno primera, dete samo počinje da prepoznaje nove kuce i mace koje nikada pre nije videlo.

Duboko učenje

Duboko učenje (*Deep Learning, DL*) je podskup mašinskog učenja, ali toliko moćan da je postao gotovo sinonim za današnji AI. Njegova snaga leži u korišćenju struktura koje se nazivaju Veštačke neuronske mreže *(Artificial Neural Networks, ANNs*), koje su labavo inspirisane strukturom i funkcijom ljudskog mozga.

Ove mreže se sastoje od slojeva međusobno povezanih neurona (matematičkih funkcija). Kada podatak, poput slike, uđe u mrežu, on prolazi kroz ove slojeve. Svaki sloj prepoznaje sve apstraktnije i kompleksnije karakteristike.

- **Prvi sloj** može prepoznati osnovne elemente poput ivica i boja.
- **Sledeći sloj** može kombinovati te ivice da prepozna oblike poput očiju ili ušiju.
- **Još dublji sloj** može kombinovati te oblike da prepozna lice mačke.
- **Poslednji sloj** donosi konačnu odluku: „Verovatnoća 98% da je ovo mačka."

Reč „duboko“ u dubokom učenju odnosi se upravo na veliki broj ovih slojeva (desetine, stotine, pa i hiljade). Što je mreža dublja, to kompleksnije obrasce može da nauči. Upravo je duboko učenje omogućilo proboje u prepoznavanju govora, generisanju slika i razumevanju jezika koje vidimo danas.

Vrste veštačke inteligencije

Kada govorimo o AI-ju, važno je razlikovati ono što imamo danas od onoga što je još uvek u domenu teorije i naučne fantastike. AI se obično kategoriše na osnovu svojih sposobnosti.

- **Uska veštačka inteligencija** (*Artificial Narrow Intelligence, ANI*): Ovo je jedini tip AI-ja koji trenutno postoji. ANI, ili samo AI, je dizajniran i treniran za jedan specifičan zadatak. AI koji pobeđuje svetskog prvaka u šahu ne može da vozi automobil. AI koji preporučuje filmove na Netflixu ne može da dijagnostikuje bolesti. AI koji generiše slike ne može da vodi smislen razgovor (iako se granice zamagljuju sa multimodalnim modelima). 99.9% AI sistema u upotrebi danas su primeri uskih AI-ja.
- **Opšta veštačka inteligencija** (*Artificial General Intelligence, AGI*): Ovo je hipotetički tip AI-ja koji bi posedovao inteligenciju i kognitivne sposobnosti ravne ljudskim. AGI bi mogao da uči, razume, prilagođava se i primenjuje znanje na širok spektar zadataka, baš kao čovek. Postizanje AGI je „sveti gral“ AI istraživanja, ali smo još uvek daleko od toga. Ne postoji konsenzus o tome da li je AGI uopšte moguć, a kamoli kada bi se mogao pojaviti.
- **Superinteligencija** (*Artificial Superintelligence, ASI*): Takođe hipotetički koncept, koji je popularizovao filozof Nik Bostrom. ASI bi bila inteligencija koja daleko nadmašuje najpametnije ljudske umove u praktično svakom domenu, uključujući naučnu kreativnost, opštu mudrost i socijalne veštine. Diskusije o ASI se često fokusiraju na egzistencijalne rizike i etičke dileme koje bi takva tehnologija donela.

Za potrebe ove knjige, fokusiraćemo se isključivo na usku veštačku inteligenciju (ANI), jer je to tehnologija koja oblikuje naš svet ovde i sada.

Sada kada smo prošli strahove i mitove, i pojasnili šta je to veštačka inteligencija, možemo da pogledamo zanimljivu priču o njenom nastanku. Da bismo razumeli gde se nalazi danas, moramo znati odakle je došla veštačka inteligencija.